NEIGHBOURHOOD DETECTION AND IDENTIFICATION OF SPATIO-TEMPORAL DYNAMICAL SYSTEMS USING A COARSE-TO-FINE APPROACH

L.Z. GUO, S.S. MEI AND S.A.BILLINGS

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

Research Report No. 828
November 2002

Neighbourhood Detection and Identification of Spatio-Temporal Dynamical Systems Using a Coarse-to-Fine Approach

L.Z.Guo, S.S.Mei and S.A. Billings

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

Abstract

A novel approach to the determination of the neighbourhood and the identification of spatio-temporal dynamical systems is investigated. It is shown that thresholding to convert the pattern to a binary pattern and then applying cellular automata (CA) neighbourhood detection methods can provide an initial estimate of the neighbourhood. A coupled map lattice model can then be identified using the CA detected neighbourhood as the initial conditions. This provides a coarse-to-fine approach for neighbourhood detection and identification of coupled map lattice models. Three examples are used to demonstrate the application of the new approach.

1 Introduction

The modelling, analysis, and control of spatio-temporal dynamical systems present many computational and theoretical challenges with potential applications to physical, chemical, biological, and ecological systems (Kaneko 1993, Sóle, Valls and Bascompte 1992, Yanagita and Kaneko 1997, Tabuchi, Yakawa and Mallick et al 2002, Köhler, Reinhard and Huth 2002). Traditionally such systems have been studied using partial differential equations (PDE's). More recently Coupled Map Lattice (CML) models have emerged as an effective and powerful tool to study these systems due to the computational efficiency and ability to reproduce complex spatio-temporal behaviour. CML models were initiated in the late 80's by Kaneko (1985, 1986, 1989a) and can exhibit surprisingly rich dynamical behaviours, including spatio-temporal chaos, intermittency, traveling waves and pattern formation (Kaneko 1989b). CML's have been used to model convected temperature fluctuations in the atmosphere (Platt and Hammel 1997), boiling processes (Yanagita 1992), spatio-temporal chaos in fluid flows (He, Cao and Li 1995) and cloud dynamics

(Yanagita and Kaneko 1997). Despite the considerable attention devoted to CML's, the identification of CML models from observations or measured data is still regarded as a difficult problem. The aim of this paper is to introduce a new approach to identifying the CML equations from spatio-temporal observations.

Various methods for the identification of local CML models from spatio-temporal observations have already been proposed (Coca and Billings 2001, Billings and Coca 2002, Mandelj, Grabec and Govekar 2001, Marcos-Nikolaus, Martin-Gonzalez and Sóle 2002, Grabec and Mandeji 1997, Parlitz and Merkwirth 2000). An important step in all these modelling methods is the proper reconstruction of the local state vectors at some specified site from the measured data, that is determining the spatio-temporal region which influences the dynamics of that site. This region can be decomposed into two effects: the spatial domain and the time lag. The determination of these two regions is critical for obtaining a good approximation of the underlying spatio-temporal dynamical relationship. In Coca and Billing 2001, and Billings and Coca 2002, the CML equations were restricted to be symmetrically coupled with neighbouring lattice sites within a circular neighbourhood with some finite radius in the spatial domain. In Grabec and Mandeji 1997 and Parlitz and Merkwirth 2000, a state vector was reconstructed from a set of neighbouring values in a rectangular spatio-temporal region while a simple triangular region was used by Mandeji, Garbec and Govekar (2001). Recently, an effective approach to determining the neighbourhood for CA modelling in a binary space has been proposed by Mei and Billings (2002). This motivates our study, and instead of searching for the neighbourhood directly from the original space where real cell entities and CML modelling have to be used, an initial estimate of the neighbourhood is first extracted from an associated binary pattern. The approach involves thresholding the real valued pattern to produce a binary pattern. CA neighbourhood detection methods are then applied to produce an initial neighbourhood to prime the CML modelling. CML identification methods can then be applied to produce the final model. This is a coarse-to-fine modelling strategy which can considerably reduce the computational requirements compared to the more direct approach of searching for the entire neighbourhood in the original space. Providing the main features of the system can be extracted from the thresholded pattern, the approximate neighbourhood should provide a quick method of priming the CML neighbourhood search.

The paper is organised as follows. Section 2 introduces the CML model of spatio-temporal dynamical systems and provides an alternative derivation for the input-output representation. The identification method is presented in section 3 which includes system identification, pattern thresholding and neighbourhood detection. Section 4 illustrates the proposed approach using three examples. Finally conclusions are given in section 5.

2 The CML model

Consider a d-dimensional lattice I consisting of the set of all integer coordinate vectors $i = (i_1, \cdots, i_d) \in \mathbf{Z}^d$. The deterministic CML state-space model of spatio-temporal dynamical systems defined over I is of the following form (Coca and Billings 2001)

$$x_i(t) = f_l(x_i(t-1), u_i(t-1)) + f_c(x_i(t-1), u_i(t-1), \mathbf{s}^m x_i(t-1), \mathbf{s}^m u_i(t-1)) \tag{1}$$

where $x_i(t) \in X_i \subset \mathbf{R}^n$ and $u_i(t) \in U_i \subset \mathbf{R}^l$, X_i and U_i are open sets, n and l-dimensional vectors representing the local state and input variables respectively at the ith site in I, and f_l and f_c are piecewise differentiable maps. $\mathbf{s}^m$ is a spatial shift operator, which is defined as

$$\mathbf{s}^m = (s^{p_1}, s^{p_2}, \cdots, s^{p_m}) \tag{2}$$

such that

$$\mathbf{s}^m x_i(t) = (s^{p_1} x_i(t), s^{p_2} x_i(t), \cdots, s^{p_m} x_i(t)) = (x_{i+p_1}(t), x_{i+p_2}(t), \cdots, x_{i+p_m}(t)) \tag{3}$$

where $p_1, p_2, \cdots, p_m$ are the indices of the neighbours of the ith site - that is the region in I around the ith site, which influences the dynamics of that site.

The CML model (1) can also be written, in terms of the global state and input variables $x = \{x_i\}_{i \in I} \subset X = \prod_{i \in I} X_i$ and $u = \{u_i\}_{i \in I} \subset U = \prod_{i \in I} U_i$, as follows

$$x(t) = f(x(t-1), u(t-1)) \tag{4}$$

where $f : X \times U \to X$ is the function sequence $f = \{f_i\}_{i \in I}$ with $f_i = f_l + f_c$ and $i = \{i_1, \cdots, i_d\} \in I$.

In general, the direct measurement of the state vector x is not possible and only some observable variable y which depends on the state and input can be measured. Therefore, the state-space model of the CML is usually complemented with a measurement equation

$$y_i(t) = h_i(x(t)) \tag{5}$$

Here it is assumed that the measurement function is identical for each site, i.e., $h_i = h$. Furthermore, it is assumed that the lattice equations are spatially invariant over the observed spatial domain. This implies that the difference equations corresponding to each lattice site or location are the same for all lattice sites. Generally it is also assumed that the following input-output representation

$$y_i(t) = g(\mathbf{q}^{n_y} y_i(t), \mathbf{q}^{n_u} u_i(t), \mathbf{s}^{m'} \mathbf{q}^{n_y} y_i(t), \mathbf{s}^{m'} \mathbf{q}^{n_u} u_i(t)) \tag{6}$$

can be derived for any site from (1) and (5). In (6), $\mathbf{q}$ is a backward shift operator such that

$$\mathbf{q}^{n_y} y_i(t) = (y_i(t-1), y_i(t-2), \cdots, y_i(t-n_y)) \tag{7}$$

A number of sufficient conditions which ensure that this can be found are given in Billings and Coca 2002 for the symmetric or anti-symmetric coupling cases Here an alternative derivation of this result for a more general case is given below.

From the measurement equation (5) and state equation (1),

$$\begin{aligned} y_i(t) &= h_i(x(t)) =: \Phi_i^1(x(t)) \\ y_i(t+1) &= h_i(x(t+1)) = h_i \circ f(x(t),u(t)) =: \Phi_i^2(x(t),u(t)) \\ &\vdots \\ y_i(t+n_y-1) &= h_i(x(t+n_y-1)) = h_i \circ f^{n_y-1}(x(t),u(t)) \\ &=: \Phi_i^{n_y}(x(t),u(t),u(t+1),\cdots,u(t+n_y-2)) \end{aligned}$$

where $\circ$ denotes the composition of functions. Then

$$\begin{aligned} \mathbf{s}^m y_i(t) &= \mathbf{s}^m h_i(x(t)) = \mathbf{s}^m \Phi_i^1(x(t)) \\ \mathbf{s}^m y_i(t+1) &= \mathbf{s}^m h_i(x(t+1)) = \mathbf{s}^m \Phi_i^2(x(t),u(t)) \\ &\vdots \\ \mathbf{s}^m y_i(t+n_y-1) &= \mathbf{s}^m h_i(x(t+n_y-1)) = \mathbf{s}^m \Phi_i^{n_y}(x(t),u(t),u(t+1),\cdots,u(t+n_y-2)) \end{aligned}$$

Let $Y_i(t) = (y_i(t), y_i(t+1), \cdots, y_i(t+n_y-1), \mathbf{s}^m y_i(t), \mathbf{s}^m y_i(t+1), \cdots, \mathbf{s}^m y_i(t+n-1))^T$, $U(t) = (u(t), u(t+1), \cdots, u(t+n_y-2))^T$, and $\Phi_i = (\Phi_i^1, \Phi_i^2, \cdots, \Phi_i^n, \mathbf{s}^m\Phi_i^1, \mathbf{s}^m\Phi_i^2, \cdots, \mathbf{s}^m\Phi_i^{n_y})^T$, then it follows that

$$Y_i(t) = \Phi_i(x(t), U(t)) \tag{8}$$

If Φ_i is a C^r map (differentiable up to r), $r \geq 1$ and the partial derivatives of Φ_i in the measurement site i with respect to $x(t)$ are isomorphic with respect to $U(t)$, then from the implicit function theorem in Banach space that $x(t)$ can be expressed locally in terms of $Y_i(t)$ and $U(t)$ which implies

$$y_i(t+n_y) = h_i(x(t+n_y)) = \Phi_i^{n_y+1}(x(t),u(t),u(t+1),\cdots,u(t+n_y-1)) \tag{9}$$

that is

$$\begin{aligned} y_i(t) &= g'(y_i(t-1), y_i(t-2), \cdots, y_i(t-n_y); \\ &\quad \mathbf{s}^m y_i(t-1), \mathbf{s}^m y_i(t-2), \cdots, \mathbf{s}^m y_i(t-n_y); u(t-1), u(t-2), \cdots, u(t-n_y)) \\ &= g'(\mathbf{q}^{n_y} y_i(t), \mathbf{s}^m \mathbf{q}^{n_y} y_i(t), \mathbf{q}^{n_y} u(t)) \end{aligned} \tag{10}$$

In practice, it is reasonable to expect that the measurement function $y_i(t)$ will depend only on a finite number of states and input variables in the neighbourhood of the measurement site, then (10) can be rewritten in the following form

$$y_i(t) = g(\mathbf{q}^{n_y} y_i(t), \mathbf{q}^{n_u} u_i(t), \mathbf{s}^{m'} \mathbf{q}^{n_y} y_i(t), \mathbf{s}^{m'} \mathbf{q}^{n_u} u_i(t)) \tag{11}$$

3 CML identification

3.1 The identification procedure

From (11), the task of the identification is to reproduce the dynamical relation g from the measured data. For a specific site i, the identification procedure can be outlined as below

i) Determine the spatial neighbourhood sites (represented by $\mathbf{s}^{m'}$) of the ith site;

ii) Determine the time lags n_y and n_u;

iii) Apply the Orthogonal Least Squares (OLS) algorithm to obtain the parameters of the CML model (polynomials as regressors).

If data is available from more sensors than the minimum required to extract the CML equations, the additional measurements can be used in model validation. The CML model identified using a set of data from a given spatial site can be validated on data recorded at different spatial locations by computing the model predicted output

$$\hat{y}_j(t) = \tilde{g}(\mathbf{q}^{n_y} \hat{y}_j(t), \mathbf{q}^{n_u} u_j(t), \mathbf{s}^{m'} \mathbf{q}^{n_y} \hat{y}_j(t), \mathbf{s}^{m'} \mathbf{q}^{n_u} u_j(t)) \tag{12}$$

Model predicted output is a much more rigorous test than one step ahead predicted outputs which most authors use.

Note that in the above identification procedure, the spatial neighbourhood sites (represented by $\mathbf{s}^{m'}$) of the identified site and the time lags (n_y, n_u) need to be known *a priori*. In other words, the neighbourhood of the identified site, i.e., the region around that site which influences the dynamics of that site in the spatial domain and the time domain need to be known before starting the identification. In practice, these two factors are important in determining the spatio-temporal dynamics of the underlying system. Determining which site and what time lag should be included in the model structure is therefore very important in CML identification. Whilst Chen and Billings (1990) proposed an algorithm for time lag determination which has been used in many applications, little progress has been made in the determination of the spatial neighbourhood. As noted in section 1, existing methods tend to choose a circular, rectangular or triangular region around the identified site as the initial spatial neighbourhood. In the following

sections, a novel approach will be presented to detect the initial neighbourhood using thresholding and a CA method.

3.2 Feature extraction by thresholding

In this study, thresholding will be applied to extract the main features of a spatio-temporal dynamical system. For a real-value spatio-temporal dynamical system, snapshots of the evolution of the system both in time and space show up as a series of images which represent the features of the underlying dynamical system. The features in these images are reflected by the magnitude distribution of the variables, generally real entities, over a spatial domain of interest at some specific time instants. By thresholding these images to form binary counterparts, the main features of the original system may, depending on the dynamics of the patterns involved, be retained using only two-levels, which significantly reduces the computational requirements for searching for candidate neighbours.

The thresholding process can be accomplished using histograms. The basic procedure is as follows

i) Produce a histogram of the measured or simulated data from the underlying spatio-temporal system;

ii) Determine the threshold value by setting it as some value between the peaks or the centre of the distribution occurring in the histogram;

iii) Generate a binary spatio-temporal data set.

3.3 Initial Neighbourhood detection using CA's

The feature extraction process above produces an approximate binary representation of the spatio-temporal pattern. Applying this procedure to the input-output measurements $(u_i(t), y_i(t))$ of the underlying spatio-temporal system, therefore produces a collection of corresponding binary data $(u_i^*(t), y_i^*(t))$. The assumption is that a dynamical relationship exists among these binary input-output data which partially represents the main features of the original dynamical system (6). Assume that the spatio-temporal dynamical relationship (6) can be expressed by a CA model counterpart in binary space

$$y_i^*(t) = g(\mathbf{q}^{n_y} y_i^*(t), \mathbf{q}^{n_u} u_i^*(t), \mathbf{s}^{m'} \mathbf{q}^{n_y} y_i^*(t), \mathbf{s}^{m'} \mathbf{q}^{n_u} u_i^*(t)) \tag{13}$$

The approach developed by Mei and Billings (2002) can then be employed to detect the neighbourhood both in the spatial domain and the time domain for the thresholded data $(u_i^*(t), y_i^*(t))$. Initially, assume the CA model has the following structure

$$y_i^*(t) = g(\mathbf{q}^{n_{y0}} y_i^*(t), \mathbf{q}^{n_{u0}} u_i^*(t), \mathbf{s}^{m'0} \mathbf{q}^{n_{y0}} y_i^*(t), \mathbf{s}^{m'0} \mathbf{q}^{n_{u0}} u_i^*(t)) \tag{14}$$

where n_{y_0}, n_{u_0}, and m'_0 are the candidate time lags and the spatial range. The candidate term vector is therefore

$$(\mathbf{q}^{n_{y_0}} y_i^*(t), \mathbf{q}^{n_{u_0}} u_i^*(t), \mathbf{s}^{m'_0}\mathbf{q}^{n_{y_0}} y_i^*(t), \mathbf{s}^{m'_0}\mathbf{q}^{n_{u_0}} u_i^*(t)) \in B^{(m'_0+1)\times(n_{y_0}+n_{u_0})}$$

$B = \{0, 1\}$, and the result from Mei and Billings (2002) can be used to determine the neighbourhood for (13) if the candidate set is large enough:

Proposition 1 Let $y = f(x_1, x_2, \cdots, x_n)$ be a Boolean function from B^n to B, $B = \{0, 1\}$. For any given component x_i, there exists a Boolean function $f' : B^{n-1} \to B$, $f'(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \in B$ such that $f = f' \circ p_i$, where p_i is the natural projection from B^n to the other $n-1$ components except i if and only if

$$f(x_1, \cdots, x_{i-1}, 1 - x_i, x_{i+1}, \cdots, x_n) \equiv f(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_n) \tag{15}$$

for all $[x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n] \in B^{n-1}$.

Condition (15) implies that the ith component makes no contribution to the Boolean function f so that it can be discarded. Repeatedly applying the above proposition to (14), produces a minimal collection of components which represent the neighbourhood. This neighbourhood can then be used to prime the CML identification.

During the implementation of the neighbourhood detection algorithm, a filtering procedure is employed to reduce the possibility of problems caused by noise or other disturbances in the feature extraction process by thresholding. Typically problems may arise when noise causes a sampling point to cross the binarisation threshold because then the obtained binary pattern may not reflect the original dynamics correctly. These effects are identical to the problems which arise in Probabilistic CA (PCA) with dynamic noise where cell states can be flipped during the evolution of the CA (Billings and Yang 2003). Algorithms which reduce and in some cases mitigate these problems can be employed here. The function of the filtering procedure is to filter these erroneous data out. This can be achieved by applying the criterion (15) to the expected values instead of the variables themselves. In practice, the expected values will be approximated by the sample means computed over the data obtained from the thresholding process. For more details, refer to Mei and Billings (2002).

Obviously this strategy will work well when the dynamics of the CML pattern, can be reasonably approximated by a binary pattern. There will be cases where this is not the case. However, our simulations suggest that this approach works well in the many examples that we have simulated. The advantage is that the neighbourhood search over the binary pattern is very straightforward and relatively quick. This is then used to prime the CML identification and can be considered as a coarse-to-fine approach to neighbourhood detection. The histogram of the original pattern provides a good indicator of whether the method will work well and this can be used as the decision boundary. Little computing effort will have been wasted even if the histogram suggests the method may not be appropriate for a particular spatio-temporal pattern. If the coarse-to-fine

neighbourhood detection method is however applied model validation, based for example on the CML model predicted output, should be used as a means of testing the authenticity of the final CML model.

4 Simulation studies

4.1 Example 1 ? Non-homogeneous Wave Equation

Consider the following non-homogeneous wave equation

$$\frac{\partial^2 v(t,x)}{\partial t^2} - C\frac{\partial^2 v(t,x)}{\partial x^2} = u(t,x), x \in [0,1] \tag{16}$$

with initial conditions

$$\begin{aligned} v(0,x) &= 0 \\ \frac{dv(0,x)}{dt} &= 4\text{exp}(-x) + \text{exp}(-0.5x) \end{aligned} \tag{17}$$

where

$$u(t,x) = -13\text{exp}(-x)\text{cos}(1.5t) - 9.32\text{exp}(-0.5x)\text{cos}(2.1t) \tag{18}$$

For $C = 1.0$ the exact solution $v(t,x)$ of the initial value problem (16), (18) is

$$\begin{aligned} v(t,x) &= 4\text{exp}(-x)\text{cos}(1.5t) + 2\text{exp}(-0.5x)\text{cos}(2.1t) \\ &\quad -4\text{exp}(-x)\text{exp}(-t) - 2\text{exp}(-0.5x)\text{exp}(-0.5t) \end{aligned} \tag{19}$$

The measurement function was taken as

$$y(t,x) = v(t,x) \tag{20}$$

The reference solution was sampled at 21 equally spaced points over the spatial domain $\Omega = [0,1], x = \{x_1, \cdots, x_{21}\} = (0, 0.05, \cdots, 0.95, 1)$. From each location, 1000 input/output data points sampled at $\Delta t = \pi/100$ were generated. The data are plotted in Fig. (1).

Terms	Estimates	ERR	STD
$y_i(t-1)$	0.16799E+01	0.99734E+00	0.11860E+00
$y_i(t-2)$	-0.99911E+00	0.26578E-02	0.86569E-11
$u_i(t-1)$	0.97171E-03	0.16440E-05	0.14638E-12
$y_{i-1}(t-1)$	0.16906E+00	0.89148E-07	0.57112E-01
$y_{i+1}(t-1)$	0.15065E+00	0.83493E-15	0.61560E-01
$y_{i-1}(t-2)$	-0.79775E-03	0.35227E-15	0.77924E-11
constant	0.78551E-17	0.41257E-16	0.42323E-15
$u_i(t-2)$	0.14685E-04	0.24471E-17	0.14312E-12

Table 1: Example 1: The terms and parameters of the final CML model

The histogram of the input/output data is shown in Fig. (2) which exhibits a normal-like distribution. Therefore, the threshold value was set to be zero to obtain a binary counterpart of the original pattern. The obtained neighbourhood from cellular automata was $i-1, i+1$, and $i+2$ in the spatial domain and $t-1, t-2$ in the time domain. This estimated neighbourhood comprises the neighbourhood, $i-1$ and $i+1$, used in Billings and Coca 2002, which indicates the rationale for using the proposed strategy. Using the neighbourhood as an initial estimation, different combinations of the sites within this were tested for the CML modelling. The case, which the neighbourhood was selected to be $i-1$ and $i+1$ in the spatial domain and $t-1, t-2$ in the time domain without the symmetrical restrictions, is presented here.

The identification data consisted of 1000 data points of input/output data $u_i(t), y_i(t)$ at site $i = 3$ corresponding to $x = x_3 = 0.1$. In addition, 1000 input and output data $u_i(t), y_{i-1}(t)$ and $y_{i+1}(t)$ from neighbouring locations $x = x_2 = 0.05$ and $x = x_4 = 0.15$ acted as inputs during the identification. The identified model is listed in Table (1), where ERR denotes the Error Reduction Ratio (Chen and Billings 1989) and STD denotes the standard deviations. It is shown that the number of terms in the identified model is 8 which is less then that (9 terms) from Billing and Coca 2002.

The model predicted output and error are plotted in Figs.(3) and (4), which show very good agreement between the exact solution and the CML model output.

To test the possible effects from noise upon the neighbourhood detection algorithm, white noise with variance 0.3 was superimposed on the data set described above. The noisy data for site x_3 is shown in Fig. (5). Fig. (6) shows the binary counterparts of $y(t, x_3)$ with and without noise. By applying the filtering procedure based on sample means (Mei and Billings 2003) and the CA algorithm, exactly the same neighbourhood was obtained.

4.2 Example 2 - Sine-Gordon Equation

Consider the two-dimensional Sine-Gordon Equation (Hirota 1973)

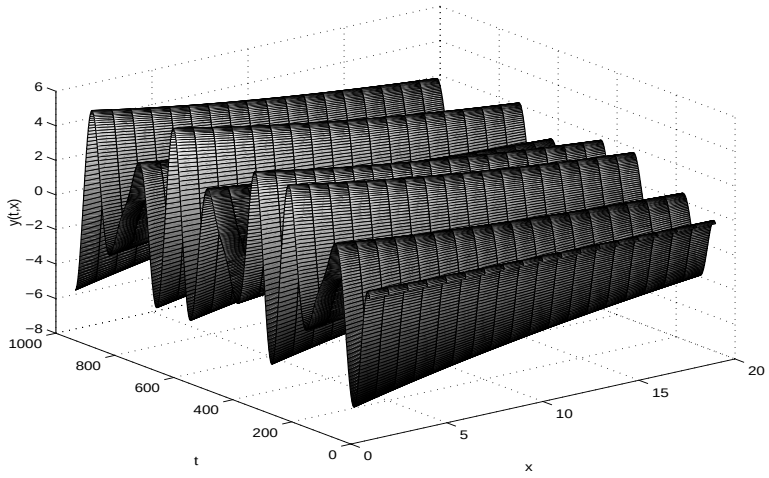

Figure 1: Example 1: System output

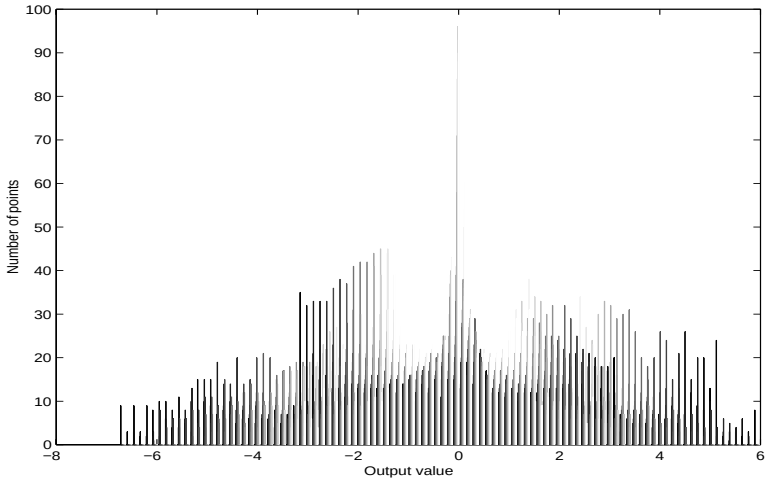

Figure 2: Example 1: Histogram of the data

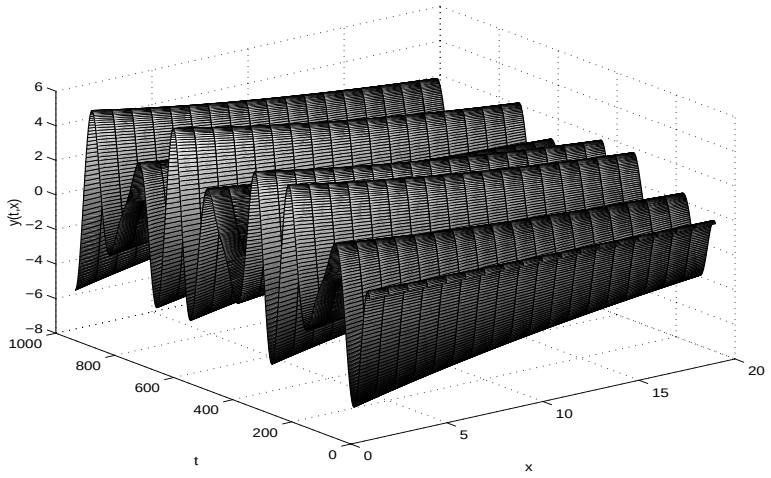

Figure 3: Example 1: Model predicted output

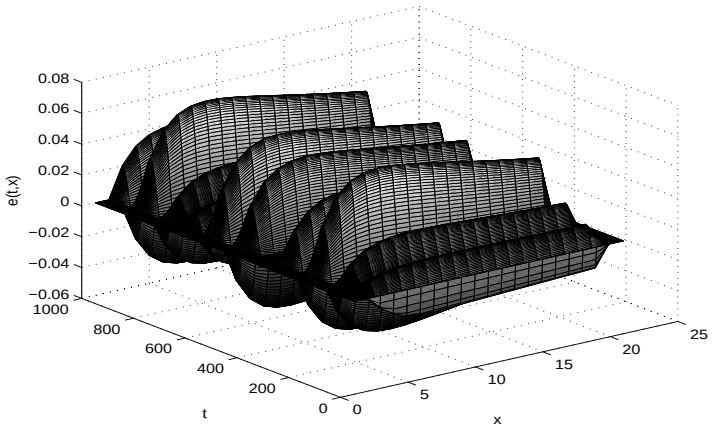

Figure 4: Example 1: Model predicted error

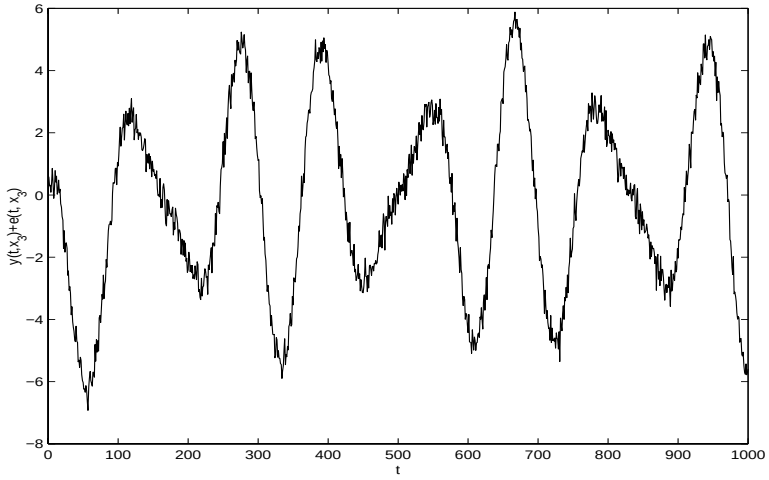

Figure 5: Example 1: $y(t, x_3)$ with white noise

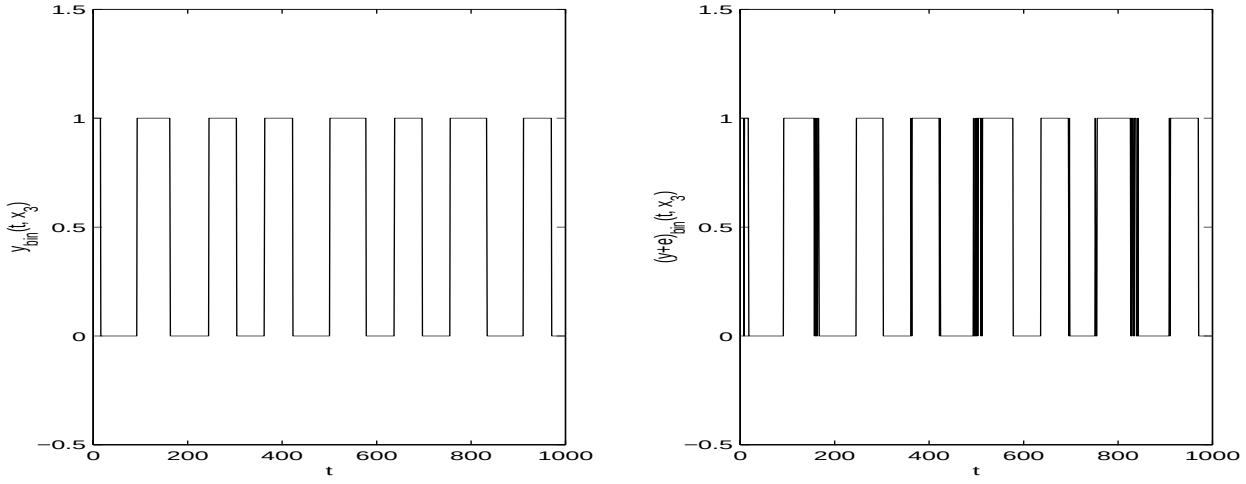

Figure 6: Example 1: Binary counterparts of $y(t, x_3)$ and $(y + e)(t, x_3)$

$$\frac{\partial^2 u(t,x,y)}{\partial x^2} + \frac{\partial^2 u(t,x,y)}{\partial y^2} - \frac{\partial^2 u(t,x,y)}{\partial t^2} = \sin(u(t,x,y)) \tag{21}$$

which describes the motion of the magnetic flux quanta on a Josephson junction transmission line.

The exact three-soliton solution of (21) can be expressed in the following form

$$u(t,x,y) = 4\tan^{-1}(\frac{g(t,x,y)}{f(t,x,y)}) \tag{22}$$

where

$$\begin{aligned} f(t,x,y) &= 1 + a(1,2)\exp(\eta_1+\eta_2) + a(1,3)\exp(\eta_1+\eta_3) + a(2,3)\exp(\eta_2+\eta_3) \\ g(t,x,y) &= \exp(\eta_1) + \exp(\eta_2) + \exp(\eta_3) + a(1,2)a(1,3)a(2,3)\exp(\eta_1+\eta_2+\eta_3)) \end{aligned} \tag{23}$$

in which

$$\begin{aligned} a(i,j) &= \frac{(P_i-P_j)^2 + (q_i-q_j)^2 - (w_i-w_j)^2}{(P_i+P_j)^2 + (q_i+q_j)^2 - (w_i+w_j)^2} \\ \eta_i &= P_i x + q_i y - w_i t - \eta_i^0, (\eta_i^0 \text{ is constant}) \\ P_i^2 + q_i^2 - w_i^2 &= 1, \text{for } i,j = 1,2,3 \end{aligned}$$

provided that the parameters P_i, q_i, and $w_i, i = 1,2,3$ satisfy the condition

$$det \begin{pmatrix} P_1 & q_1 & w_1 \\ P_2 & q_2 & w_2 \\ P_3 & q_3 & w_3 \end{pmatrix} = 0$$

The solution (22) was sampled at 50×50 spaced points with $\Delta h = 0.05$ with the following parameter values, $P_1 = 1.1, P_2 = P_3 = 0.3; q_1 = 0.0, q_2 = q_3 = 1.2$ and $w_1 = 0.4583, w_2 = w_3 = 0.6633$, and initial conditions

$$u(0,x,y) = 4\tan^{-1}(\frac{g(0,x,y)}{f(0,x,y)}) \tag{24}$$

The measurement function was taken as

Terms	Estimates	ERR	STD
$y_{i,j-2}(t-1)$	0.85718E+00	0.99999E+00	0.11960E+01
$y_{i-2,j}(t-1)$	0.38689E+00	0.11733E-05	0.52748E+00
$y_{i,j-3}(t-1)$	-0.27956E+00	0.28690E-07	0.81747E+00
constant	-0.14333E-04	0.10377E-10	0.25698E-04
$y_{i-1,j}(t-1)$	0.60874E-01	0.47631E-11	0.10475E+01
$y_{i-1,j}(t-1)y_{i-1,j-1}(t-1)$	-0.12208E-02	0.12146E-13	0.13893E-01
$y^2_{i-2,j}(t-1)$	-0.14189E-01	0.11647E-13	0.11793E+00
$y_{i-1,j}(t-1)y^2_{i-2,j}(t-1)$	0.78280E-03	0.10020E-13	0.26307E-01
$y_{i-1,j-1}(t-1)y_{i,j-2}(t-1)$	0.65649E-01	0.91298E-14	0.49675E+00
$y^2_{i,j-2}(t-1)$	-0.49985E-01	0.72695E-14	0.37229E+00
$y_{i,j}(t-1)y_{i-2,j}(t-1)y_{i,j-3}(t-1)$	0.24243E-01	0.53862E-14	0.77486E-01
$y_{i-1,j}(t-1)y_{i,j-2}(t-1)y_{i,j-3}(t-1)$	0.18671E-02	0.45060E-14	0.16395E-01
$y_{i,j}(t-1)y_{i-1,j}(t-1)y_{i,j-2}(t-1)$	0.21671E-01	0.29832E-14	0.40881E-01
$y_{i,j}(t-1)$	-0.26475E-01	0.16857E-14	0.23102E+00
$y_{i,j}(t-1)y_{i-1,j-1}(t-1)y_{i-2,j}(t-1)$	0.39445E-01	0.14168E-14	0.10067E+00
$y_{i-1,j}(t-1)y_{i-2,j}(t-1)y_{i,j-3}(t-1)$	-0.10455E-01	0.12683E-14	0.11613E-01
$y_{i,j}(t-1)y_{i,j-2}(t-1)y_{i-2,j}(t-1)$	-0.66183E-01	0.12633E-14	0.14130E+00
$y_{i,j-2}(t-1)y^2_{i,j-3}(t-1)$	-0.13504E-02	0.12183E-14	0.10186E-01
$y_{i,j}(t-1)y_{i-1,j}(t-1)y_{i-2,j}(t-1)$	-0.18752E-01	0.11756E-14	0.42723E-01
$y^2_{i-1,j}(t-1)y_{i-1,j-1}(t-1)$	0.87364E-02	0.10906E-15	0.37293E-01

Table 2: Example 2: The terms and parameters of the final CML model

$$y(t,x,y) = u(t,x,y) \tag{25}$$

From each location, 100 input/output data points sampled at $\Delta t = 0.1$ were generated. Fig. (7) shows four snapshots of $y(t,x,y)$ at $t = 0, t = 0.1 \times 10 = 1, t = 0.1 \times 20 = 2$ and $t = 0.1 \times 30 = 3$, respectively. Fig. (8) shows the time series of the system outputs.

The histogram of all input/output data is shown in Fig. (9) which clearly exhibits two peaks which suggests a suitable threshold value would lie between these two peaks. The threshold value was therefore set to be 2 in this simulation. The obtained neighbourhood from cellular automata was $(i-1,j),(i-1,j-1),(i,j-2),(i-2,j)$, and $(i,j-3)$ in the spatial domain and $t-1$ in the time domain.

The identification data consisted of 30 data points of input/output data $u_i(t), y_i(t)$ at the node $(i,j) = (25,25)$. The identified model is listed in Table (2)

The model predicted outputs and errors are plotted in Figs. (10) and (11), which show that the identified CML model can reproduce the spatio-temporal patterns of the original system very well.

4.3 Example 3 ? A Chaotic Spatio-temporal System

Consider the following two-dimensional deterministic CML with symmetrical nearest neighbour coupling

$$x_{i,j}(t) = (1-\varepsilon)f(x_{i,j}(t-1)) + \frac{\varepsilon}{4}(f(x_{i,j-1}(t-1)) + f(x_{i,j+1}(t-1)) + f(x_{i-1,j}(t-1)) + f(x_{i+1,j}(t-1))) \tag{26}$$

where $x_{i,j}(t) i, j = 1, \cdots, N$ is the state of the CML located at site (i, j) at discrete time t, ε is the coupling strength, and N is the size of lattice. Periodic boundary conditions, that is $x_{1,j}(t) = x_{N,j}(t), x_{i,1}(t) = x_{i,N}(t)$ for all i, j and t, are used throughout this simulation. Let $y_{i,j} = x_{i,j}$ is the observation variable of the CML at site (i, j). The evolution of the CML on the lattice sites is governed by the local map f, which is chosen to be the logistic map

$$f(x) = 1 - ax^2 \tag{27}$$

This model has been extensively studied. It has been observed that for small ε (< 0.3) the system evolves from a frozen random state to pattern selection and to fully developed spatio-temporal chaos via spatio-temporal intermittency. For stronger coupling $\varepsilon > 0.3$ neither a frozen random pattern nor a pattern selection regime is formed which implies there are no pattern changes in this case (Kaneko 1989a).

In order to apply the proposed method, the model (26) with (27) was simulated for a lattice of the size 50×50 with random initial conditions, periodic boundary conditions, and parameters $\varepsilon = 0.4, a = 1.55$. The observation variable was set to be $y_{i,j} = x_{i,j}$. All data were normalised to the interval [0, 1]. Some snapshot patterns are shown in Fig. (12). With these parameters, the system is actually in a chaotic regime with Lyapunov exponents $\lambda_1 = 0.0648, \lambda_2 = 0.0622, \lambda_3 = 0.0158, \lambda_4 = -0.0014, \lambda_5 = -0.0106, \lambda_6 = -0.0275, \lambda_7 = -0.0478, \lambda_8 = -0.0811, \lambda_9 = -0.1360$. The Lyapunov exponents were calculated through the product of Jacobians of time steps 1 to 100 for a sub-lattice of the size 3×3 (the site $(25, 25)$ as the centre point), where the boundary effect is neglected. It follows that the KS entropy is 0.1428. In order to be able to calculate the largest positive Lyapunov exponent from the data, a numerical algorithm proposed by Rosenstein, Collins, and De Luca (1993) was employed. For the data from site $(25, 25)$, the slope of the curve obtained by the algorithm was found to converge towards a common value for the choice of embedding dimensions m and provided a value of $\lambda_1 \cong 0.0644$ for the largest Lyapunov exponent which is very close to the value of 0.0648 obtained by the product of Jacobians. The correlation dimension was also estimated by Rosenstein's method to be around 0.495.

In the identification, a set of 100 observation pairs randomly selected among the data set was used for identification. The threshold value was set to be 0.5 in this simulation and the obtained neighbourhood from cellular automata was, $(i, j-1), (i, j+1), (i-1, j)$, and $(i+1, j)$ and the time lag was 1.

During the identification, the maximal number of selected terms in the OFR selection algorithm

Terms	Estimates	ERR	STD
Constant	7.4046e-01	8.1197e-01	2.9845e-01
$y_{i,j}(t-1)^3$	5.0098e-01	1.8189e-01	5.3908e-02
$y_{i,j}(t-1)$	1.4612e+00	2.3249e-03	4.2481e-02
$y_{i,j-1}(t-1)y_{i,j+1}(t-1)y_{i-1,j}(t-1)$	-2.1694e-01	1.3244e-03	3.4310e-02
$y_{i,j}(t-1)^2$	-2.3338e+00	7.6475e-04	2.8547e-02
$y_{i+1,j}(t-1)^3$	-9.8188e-02	7.0280e-04	2.1954e-02
$y_{i,j+1}(t-1)^3$	-3.1194e-02	9.9214e-05	2.0856e-02

Table 3: Example3: The terms and parameters for the final model

was set to be 10 and the tolerance ρ was chosen as 10^{-3}. The final model is listed in Table (3). The model predicted outputs and errors from the obtain model are shown in Figs. (13) and (14). By using Rosenstein's method to the data from site $(25, 25)$, a positive value of $\lambda_1 = 0.0635$ and an estimated correlation dimension 0.494 were found.

The identification results clearly show that the proposed method can provide satisfactory prediction performance for this chaotic dynamical system. Both estimated largest Lyapunov exponents and correlation dimensions are quite close to the values calculated using the correct model.

5 Conclusions

A novel approach to the identification of spatio-temporal dynamical systems has been introduced. It has been demonstrated that determining the neighbourhood for a real-value CML model can be significantly simplified by employing a binary counterpart of the original data. Computing a histogram of the data provides a good indication of whether this coarse-to-fine neighbourhood detection algorithm will be worthwhile. Simulation results were included to demonstrate that the coarse-to-fine procedure can produce excellent final CML models with a significant reduction in computational resource.

6 Acknowledgement

The authors gratefully acknowledge financial support from EPSRC (UK).

References

[1] Billings, S. A. and Coca, D., (2002) Identification of coupled map lattice models of deterministic distributed parameter systems, *Int. J. Syst. Sci.*, In Print.

[2] Billings, S. A. and Yang, Y. Y., (2003) Identification of probabilistic cellular automata, *IEEE Trans. Systems, Men and Cyber.*, Vol. 33, No. 8, pp. 225-236.

[3] Chen, S. and Billings, S. A., (1989) Modelling and analysis of non-linear time series. *International Journal of Control*, Vol. 50, No. 6, pp. 2151-2171.

[4] Coca,D. and Billings, S. A., (2001) Identification of coupled map lattice models of complex spatio-temporal pattern, *Phys. Lett.*, A287, pp. 65-73.

[5] Grabec, I. and Mandelj, S., (1997) Continuation of chaotic fields By RBFNN, in *Biological and Artificial Computation: From Neuroscience to Technology: Proc.*, Mira, J. et al. eds., Lecture Notes in Computer Science, Springer-Verlag, Vol. 1240, pp. 597-606.

[6] He, G., Cao, L., and Li, J., (1995) Convective coupled map for simulating spatiotemporal chaos in flows, *Acta Mechanica Sinica*, Vol. 11, pp. 1-7.

[7] Hirota, R., (1973) Exact three-soliton of the two-dimensional Sine-Gordon equation, *J. Phys. Soc. Japan*, Vol. 35, pp. 1566.

[8] Kaneko, K., (1985) Spatio-temporal intermittency in couple map lattices, *Progress of Theoretical Physics*, Vol. 74, No. 5, pp. 1033-1044.

[9] Kaneko, K., (1986) Tuebulence in coupled map lattices, *Physica*, D18, pp. 475-476.

[10] Kaneko, K., (1989a) Spatiotemporal chaos in one- and two-dimensional coupled map lattices, *Physica*, D37, pp. 60-82.

[11] Kaneko, K., (1989b) Pattern dynamics in spatiotemporal chaos: pattern selection, diffusion of defect and pattern competition intermittency, *Physica*, D34, pp. 1-41.

[12] Kaneko, K. (eds.), (1993) *Coupled map lattice: theory and experiment*, World Scientific, Singapore.

[13] Köhler, P., Reinhard, K., and Huth, A., (2002) Simulating anthropogenic impacts to bird communities in tropical rain forests, *Biological Conservation*, Vol. 108, pp. 35-47.

[14] Mandelj, S., Grabec, I., and Govekar, E, (2001) Statistical approach to modeling of spatiotemporal dynamics, *Int. J. Bifurcation and Chaos*, Vol. 11, No. 11, pp. 2731-2738.

[15] Marcos-Nikolaus, P., Martin-Gonzalez, J. M. and Sóle, R. V., (2002) Spatial forcasting: detecting determinism from single snapshots, *Int. J. Bifurcation and Chaos*, Vol. 12, No. 2, pp. 369-376.

[16] Mei, S. S. and Billings, S. A., (2002) Neighbourhood determination and transition rule identification for deterministic cellular automata, *Int. J. Bifurcation and Chaos (submitted for publication)*.

[17] Parlitz, U. and Merkwirth, C., (2000) Prediction of spatiotemporal teim series based on reconstructed local states, *Phys. Rev. Lett.* , Vol. 84, No. 9, pp. 2820-2823.

[18] Platt, N. and Hammel, S., (1997) Pattern formation in driven coupled map lattices, *Physica*, A239, No. 1-3, pp. 296-303.

[19] Rosenstein, M. T., Collins, J. J., and De Luca, C. J., (1993) A practical method for calculating largest Lyapunov exponents from small data sets, *Physica*, D65, pp. 117-134.

[20] Sóle, R. V., Valls, J. and Bascompte, J., (1992) Spiral waves, chaos and multiple attactors in lattice models of interacting populations, *Phys. Lett.*, A166, No. 2, pp. 123.

[21] Tabuchi, E., Yakawa, T., Mallick, H., Inubushi, T., Kondoh, T., Ono, T., and Torii, K., (2002) Spatio-temporal dynamics of brain activated regions during drinking behaviour in rats, *Brain Research*, Vol. 951, pp. 270-279.

[22] Yanagita, T., (1992) Phenomenology of boiling: a coupled map lattice model, *Chaos*, Vol. 2, No. 3, pp. 343-350.

[23] Yanagita, T. and Kaneko, K, (1997) Modeling and characterisation of cloud dynamics, *Phys. Rev. Lett.*, Vol. 78, No. 22, pp. 4297-4300.

Lingzhong Guo received the B.Sc. degree in mathematics from Bohai University, Jinzhou, China, in 1986. After that he received the M.Sc. and Ph.D. degrees from Northeastern University, China, and University of the West of England, UK. He is currently a research associate in the Department of Automatic Control and Systems Engineering of the University of Sheffield, UK. His research interests include identification, analysis and control of nonlinear dynamic systems.

Shengsong Mei received the Beng degree in Automatic Control in 1990, the Msc Degree in Computer Engineering in 1997, from Wuhan University of Science and Technology, China, and the degree of PhD in Automatic Control and Systems Engineering from the University of Sheffield in 2005. He is currently a Research Associate in the Psychology Department, University of Sheffield, UK. Previously, he has worked as a Research Associate/fMRI Physicist in the Department of Ophthalmology, University of Leicester, UK, Lecturer in the Department of Automatic Control, Wuhan University of Science and Technology, China, Design and Development Researcher, in the Neural Network Centre, and System Analyst in Computer Based Learning Centre , Ngee Ann Polytechnic, Singapore. His current research interests include nonlinear system modelling, narmax methods, model validation, prediction, wavelets, cellular automata, spatio-temporal systems, fMRI and optical imagery of the brain, and medical data analysis.

Stephen A Billings received the BEng degree in Electrical Engineering with first class honours from the University of Liverpool in 1972, the degree of PhD in Control Systems Engineering from the University of Sheffield in 1976, and the degree of DEng from the University of Liverpool in 1990. He is a Chartered Engineer [CEng] , Chartered Mathematician [CMath], Fellow of the IEE [UK] and Fellow of the Institute of Mathematics and its Applications. He was appointed

as Professor in the Department of Automatic Control and Systems Engineering, University of Sheffield, UK in 1990 and leads the Signal Processing and Complex Systems research group. His research interests include system identification and information processing for nonlinear systems, narmax methods, model validation, prediction, spectral analysis, adaptive systems, nonlinear systems analysis and design, neural networks, wavelets, fractals, machine vision, cellular automata, spatio-temporal systems, fMRI and optical imagery of the brain, metabolic systems engineering, systems biology and related fields.

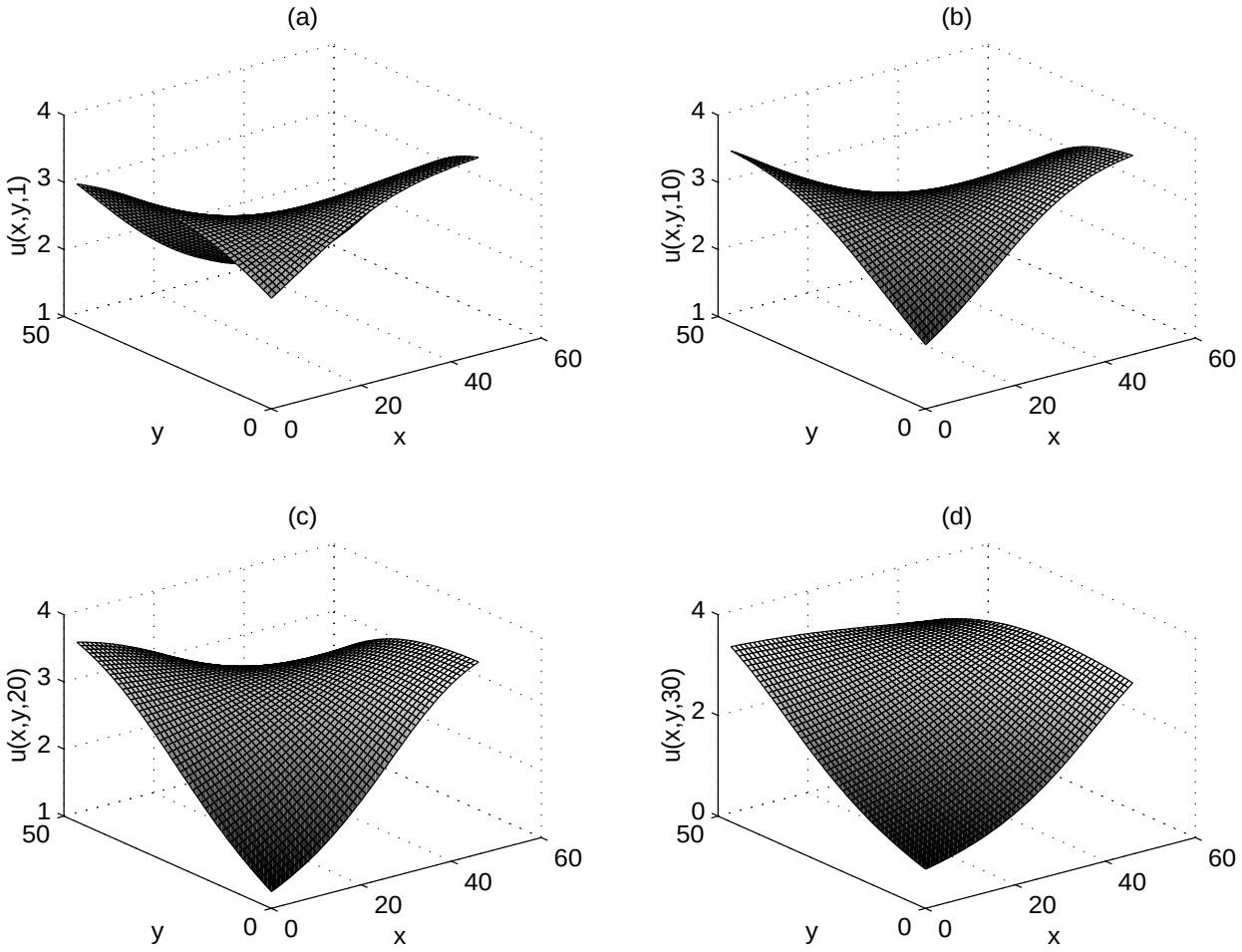

Figure 7: Example 2: Sytem outputs at (a) $t = 0$, (b) $t = 0.1 \times 10 = 1$, (c) $t = 0.1 \times 20 = 2$, (d) $t = 0.1 \times 30 = 3$

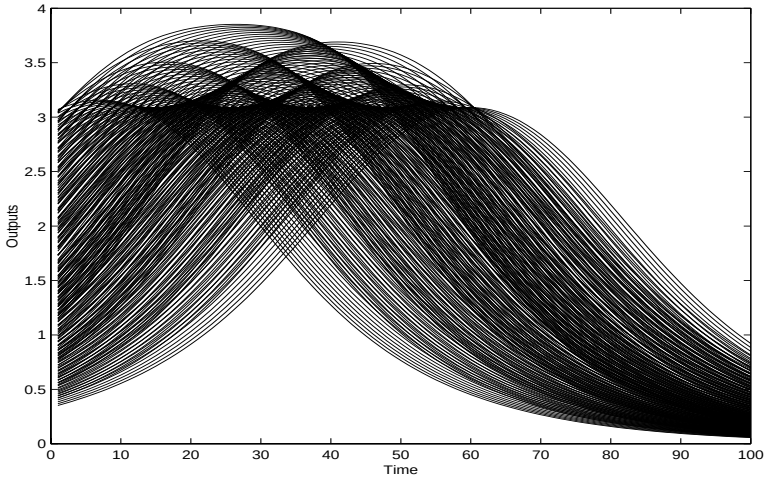

Figure 8: Example 2: Time series of the system output

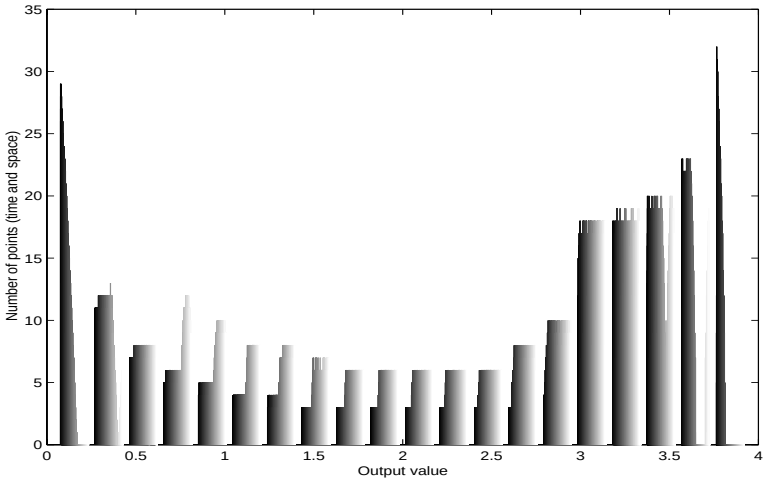

Figure 9: Example 2: Histogram of the data

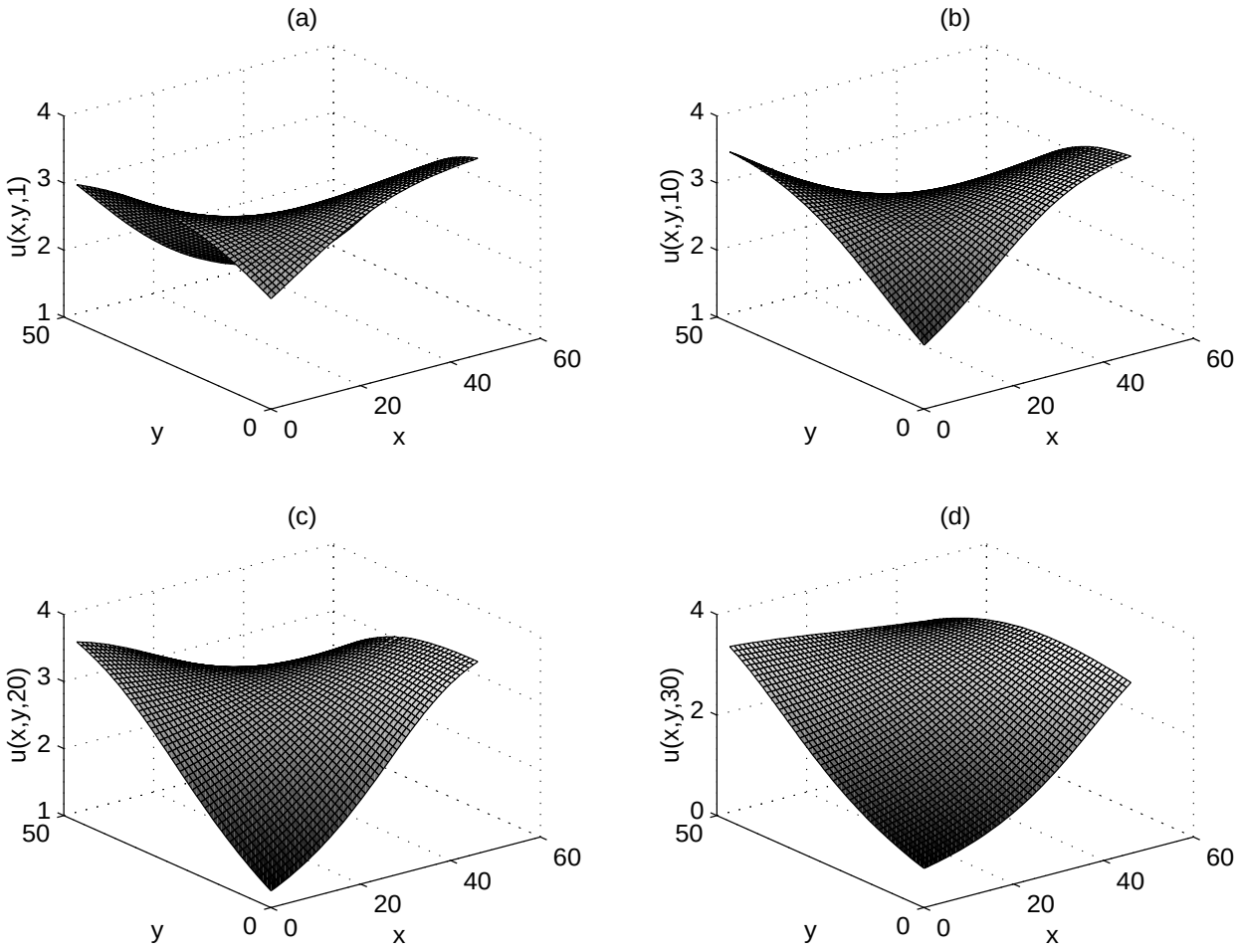

Figure 10: Example 2: Model predicted output at (a) $t = 0$, (b) $t = 0.1 \times 10 = 1$, (c) $t = 0.1 \times 20 = 2$, (d) $t = 0.1 \times 30 = 3$

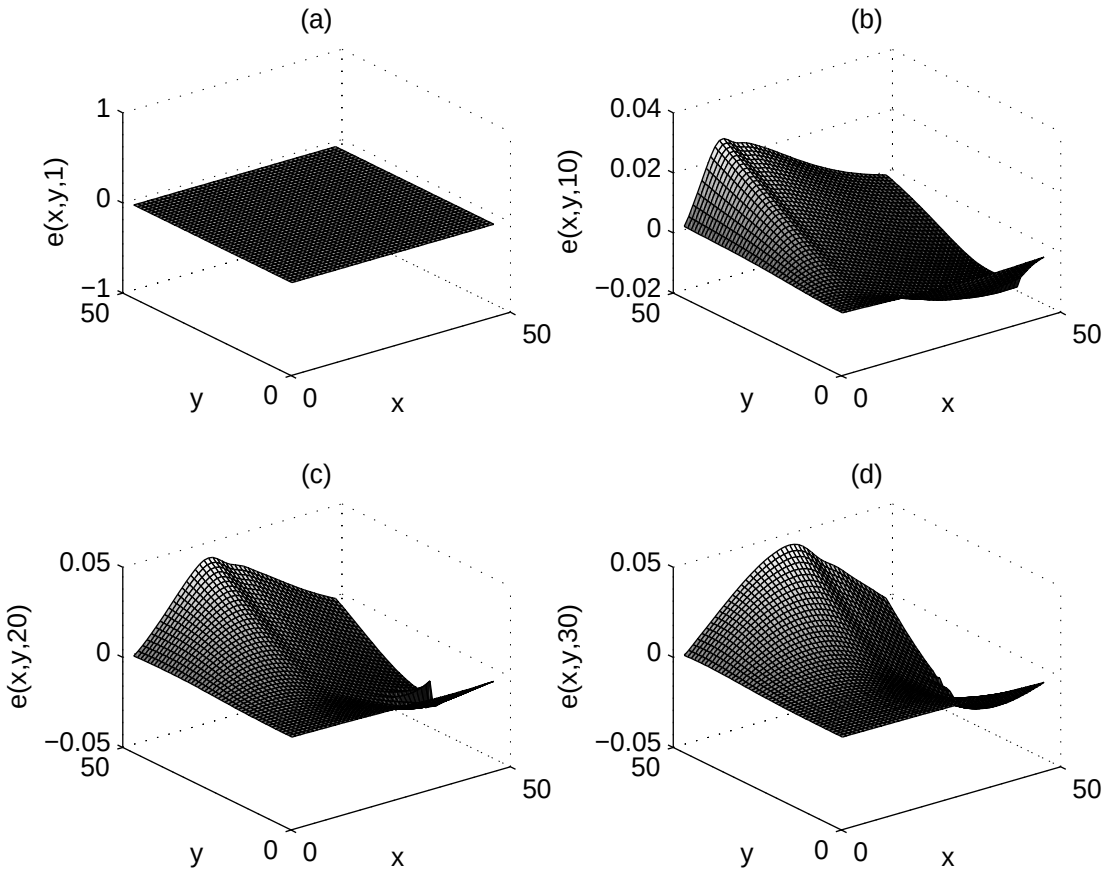

Figure 11: Example 2: Model predicted error at (a) $t = 0$, (b) $t = 0.1 \times 10 = 1$, (c) $t = 0.1 \times 20 = 2$, (d) $t = 0.1 \times 30 = 3$

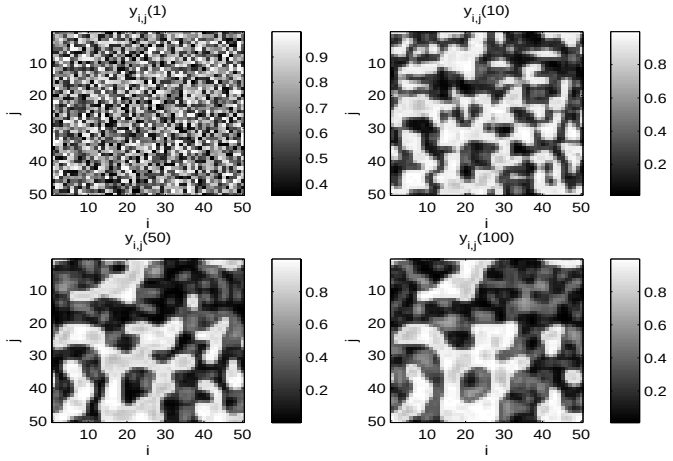

Figure 12: Example3: Some snapshots (at $t = 1, 10, 50$, and 100) from simulated data

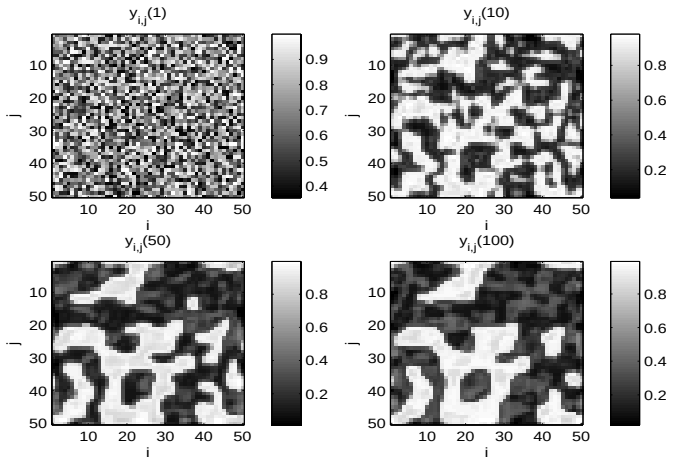

Figure 13: Example 3: Snapshots (at $t = 1, 10, 50$, and 100) from the estimated polynomial model predictive output

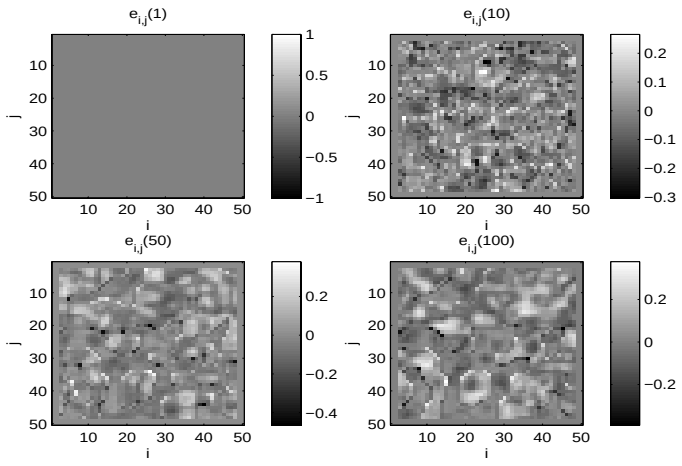

Figure 14: Example 3: Snapshots (at $t = 1, 10, 50$, and 100) from the estimated polynomial model predictive output