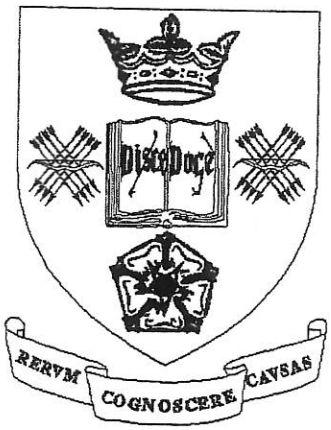
Monograph:
Billings, S.A., Wei, H.L., Mei, S.S. et al. (2003) Identification of Spatio-Temporal Systems Using Multi-resolution Wavelet Models. Research Report. ACSE Research Report 847 . Department of Automatic Control and Systems Engineering

Identification of Spatio-Temporal Systems Using Multiresolution Wavelet Models

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



Research Report No. 847

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

October 2003

Identification of Spatio-Temporal Systems Using Multiresolution Wavelet Models

S.A. Billings, H.L. Wei, S.S. Mei and L.Z. Guo
Department of Automatic Control and Systems Engineering, University of Sheffield
Mappin Street, Sheffield, S1 3JD, UK

A novel modelling structure for identifying spatio-temporal systems is proposed based on multiresolution wavelet decompositions. Two classes of spatio-temporal systems, coupled map lattices and cellular automata, are considered and the new modelling approach is applied to identify these systems. Several examples are provided to demonstrate the applicability and effectiveness of the new identification procedure.

1. Introduction

Spatio-temporal systems are complex systems where the system states evolve spatially as well as temporally. Unlike classical control systems where the current output is a function of previous inputs and outputs only in time, the output of a spatio-temporal system depends not only on the past values in time but also past values at different spatial locations. Spatio-temporal phenomena are widely found in biology, chemistry, ecology, geography, medicine, physics, and sociology [Kaneko 1985, 1986; Sole et al. 1992; Cross and Hohenberg 1993; Jahne 1993; Booth et al. 1995; Silva et al. 1997; Sole 1998; Czaran 1998].

In order to analyse, control or predict the dynamics of spatio-temporal systems, mathematical representations are usually required. A popular theoretical description of spatio-temporal systems is usually given in terms of partial differential equations (PDE's) or derivatives [Duarte and Solari 1997; Torres-Sorando and Rodriguez 1997; Tziperman et al. 1997]. While several results have been obtained based on this approach constructing partial differential equations is very difficult unless the form of the model is known. In addition, PDE's can only represent a limited class of spatio-temporal systems.

Ideally, a class of practical models are required which are relatively simple in structure but are still capable of representing spatio-temporal systems with a satisfied accuracy. Cellular automata (CA) and coupled map lattices (CML), which will be studied here, are among the most successful models which satisfy most of these constraints and which have been widely applied in practice. Other models including Boolean networks (BN) [Decarvalho et al. 1994; Li et al. 1996], cellular neural networks (CNN) [Chua and Shi 1995; Chua 1998] have also been widely used.

While the history of CA can be traced back to the early pioneering work of John von Neumann in the 1950s [von Nuemann 1951], the more recent incarnation as simple models of complexity in nature can arguably be traced to a single landmark review paper by Wolfram[1983]. This paper ignited the recent explosive development of CA's by physicists [Ilachinski 2000]. CA are characterized by states that exist only on lattices of discrete sites. Each site has a finite number of possible states and the site evolutions are determined by either a deterministic or a random rule, which involves some finite neighbourhood on the lattice and a finite number of previous time steps.

The coupled map lattice model was originally developed as a spatial counterpart of lumped discrete dynamical systems and can be used to represent spatio-temporal patterns [Kaneko 1984, 1985]. A CML is a dynamical system with discrete time and discrete space, and unlike a CA, has a continuous state. CML's are more complex than CA's and are capable of capturing local information with continuous state variables. Being discrete in time and space, CML's are computationally simpler than PDE's which generally require a substantial computational load to simulate. Similarities between the behaviour of CA's and CML's indicate the close connection between the two class of models [Crutchfield and Kaneko 1986; Kaneko 1989].

Whilst the forward problem in CA's and CML's has been extensively studied, the inverse problem, which is concerned with finding a set of rules or equations, and quantitatively reproducing a given set of prescribed spatio-temporal patterns, has received relatively little attention and relatively few results have been achieved. Identification plays an important role for solving the inverse problem in CA's and CML's where analytical models are not available, but few studies have been completed on CA and CML identification, see the recent work by Adamatzky [1994], Lopez et al. [2000], Parlizt and Merkwirth [2000], Yang [2000],Yang and Billings [2000], Coca and Billings [2001], Mandelj et al. [2001], Billings and Coca [2002], Marcos-Nikolaus et al. [2002], Billings and Yang [2003], and Sitz et al. [2003].

The objective of the present study is to introduce a new methodology to identify CA or CML models for nonlinear spatio-temporal systems based on only a finite set of observational data. In the new identification procedure here, a CA or a CML is initially described using an NARX model [Leontaritis and Billings 1985]. This is then expressed as a form of a quasi-ANOVA (analysis of variance) expansion, and each functional component of the quasi-ANOVA is approximated using multiresolution wavelet decompositions [Billings and Wei 2003; Wei and Billings 2003]. Finally a parsimonious wavelet model is obtained by means of the widely used orthogonal least squares (OLS) [Korenberg et al. 1988; Chen et al. 1989] or the recently proposed matching pursuit orthogonal least squares (MPOLS) algorithm [Billings and Wei 2003]. The paper is organized as follows. In Section 2 the general forms of CA and CML models are briefly summarized. In Section 3, the wavelet-NARX model for identifying CA and CML is introduced. This includes a brief review on multiresolution wavelet decompositions and the NARAMX representation, and a description of mutiresolution wavelet models. Some examples are given in Section 4, and conclusions are given in Section 5.

2. CA and CML Models

In this section, the general form of CA's and CML's are briefly reviewed. Only models up to two dimensional, which have obvious physical meanings and are widely studied in practice, will be considered. This can easily be extended to a higher dimensional case in a straightforward way.

2.1 CA models

Let $x(i,j;t)$ be the (i,j)th cell to be updated at time t. Define

$$z^t(i,j;p,q,d) = \{x(i+p, j+q; t-\tau) : 1 \le \tau \le d\},\ i,j \in Z^+, p,q \in Z, \tag{1}$$

and

$$N^d[x(i,j;t)] = \{z^t(i,j;p,q,d) : -r_1 \le p \le r_2, -s_1 \le q \le s_2; r_1, r_2, s_1, s_2 \in Z^+, p,q \in Z\} \tag{2}$$

A two-dimensional, dth-order time dependence CA rule, for which a particular site value at discrete time t depends on the rs-neighbourhood site values for the previous d iteration step can be expressed as

$$x(i,j;t) = f(N^d[x(i,j;t)]) \tag{3}$$

The two dimensional CA expressed by (3) can be defined on variety of different lattices and can use many different neighbourhood structures [Wolfram 1994; Ilachinski 2001]. The most commonly studied Euclidean lattice neighbourhoods are the *von Neumann* and the *Moore neighbourhoods* which can respectively be expressed as

$$x(i,j;t) = f(x(i-1,j;t-1), x(i,j-1;t-1), x(i,j;t-1), x(i,j+1;t-1), x(i+1,j;t-1)) \tag{4}$$

and

$$\begin{aligned} x(i,j;t) = f(&x(i-1,j-1;t-1), x(i-1,j;t-1), x(i-1,j+1;t-1), x(i,j-1;t-1), \\ &x(i,j;t-1), x(i+1,j-1;t-1), x(i+1,j;t-1), x(i+1,j+1;t-1)) \end{aligned} \tag{5}$$

The nonlinear mapping f in model (3) is generally represented by a set of Boolean functions of the values of the sites belonging to the neighbourhood.

2.2 CML models

CML's are a special class of spatial distributed interacting dynamical systems known as lattice dynamical systems (LDS's) and can be defined as a discrete-time, spatially invariant LDS's with symmetric coupling topology of finite radius and symmetric or anti-symmetric coupling functions[Billings and Coca 2002]. CML's can also be viewed as simple generalizations of generic CA's. As a special case of a large class of spatially extended dynamical systems, a generic CA system, where time, space and the local state space are all discretized, can be generalized by taking away the constraint on the local state variables $x(i,j;t)$ and allowing them to be real-valued. Such systems, where the values of the local state variables $x(i,j;t)$ are permitted to be any real number but time and space still remain discrete, are CML's [Ilachinski 2001]. Taking the two-dimensional case as an example, the CML model can be expressed as follows

$$x(i,j;t) = f(N^d[x(i,j;t)]) = f^{(1)}(C^d[x(i,j;t)]) + f^{(2)}(N^d[x(i,j;t)]) \tag{6}$$

where

$$C^d[x(i,j;t)] = \{z^t(i,j;p,q,d) : p = 0, q = 0\} \tag{7}$$

The function $f^{(1)}$ is a local map which involves only the local state space variable, and the function $f^{(2)}$ is a coupling map which describes the interactions with the neighbouring lattice sites.

A well known two-dimensional CML model involving nearest-neighbour coupling on a squared lattice with the von Neumann neighbourhood was proposed by Kaneko [1989]. This model is given by

$$\begin{aligned} x(i,j;t) &= f(x(i,j;t-1), x(i-1,j;t-1), x(i,j-1;t-1), x(i,j+1;t-1), x(i+1,j;t-1)) \\ &= (1-c)g(x(i,j;t-1) \\ &\quad + \frac{c}{4}[g(x(i-1,j;t-1) + g(x(i,j-1;t-1) + g(x(i,j+1;t-1) + g(x(i+1,j;t-1))] \end{aligned} \tag{8}$$

where the function g can be chosen to be the logistic map $g(x) = 1 - ax^2$. It was shown that this model can produce many spatio-temporal patterns. Other illustrative diffusive coupling lattices can also be found in the

work of Kaneko[1984,1985, 1986, 1989], Waller and Kapral [1984], Crutchfield and Kaneko [1986], and Sole et al. [1992].

Although CA and CML are assumed to live on infinitely large lattices, only adequate limit sets are considered and used in analysis. It is common to use *periodic boundary conditions* to update the values of the cells lying on the boundary of a sub-lattice for simulation and analysis, the null boundary condition is also used sometimes. Although the choice of boundary conditions seems to play a minor role providing that the automaton has many more sites than the number of sites in the local neighbourhood, it has been turned out that boundary conditions might affect the form of the resulting dynamics [Shih 2000; Ilachinski 2001].

3. Identification of CA and CML Systems Using Multiresolution Wavelet Models

In this section, multiresolution wavelet decompositions and the NARX representation are briefly reviewed and the wavelet-NARX model is then introduced. The intrinsic nonlinearity of CA and CML systems will be represented by means of the wavelet-NARX model.

3.1 Multiresolution wavelet decompositions

It is known that for identification problems it is useful to have a basis of orthogonal (semi-orthogonal or bi-orthogonal) functions whose support can be made as small as required and which provides a universal approximation to any function in $L^2(R)$ with arbitrary desired accuracy. Under some assumptions and considerations, an orthogonal wavelet system can be constructed using *multiresolution analysis* (MRA) [Mallat 1989; Chui 1992]. Assume that the wavelet φ and associated scaling function ϕ constitute an orthogonal wavelet system, then any function $f \in L^2(R)$ can be expressed as a *multiresolution wavelet decomposition*

$$f(x) = \sum_k a_{j_0,k}\phi_{j_0,k}(x) + \sum_{j\geq j_0}\sum_k d_{j,k}\varphi_{j,k}(x) \tag{9}$$

where $\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$ and $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$, $j,k \in Z$,and the wavelet approximation coefficient $a_{j_0,k}$ and the wavelet detail coefficient $d_{j,k}$ can be calculated in theory by the inner products:

$$a_{j_0,k} =< f, \phi_{j_0,k} >= \int f(x)\overline{\phi_{j_0,k}(x)}dx \tag{10}$$

$$d_{j,k} =< f, \varphi_{j,k} >= \int f(x)\overline{\varphi_{j,k}(x)}dx \tag{11}$$

The over-bar above the functions $\varphi(\cdot)$ and $\phi(\cdot)$ in (10) and (11) indicates complex conjugate, and j_0 is an arbitrary integer representing the lowest resolution or scaling level.

Using the concept of *tensor products*, the multiresolution decomposition (9) can be immediately generalised to the muti-dimensional case, where a multiresolution wavelet decomposition can be defined by taking the *tensor product* of the one-dimensional scaling and wavelet functions [Mallat,1989]. Let $f \in L^2(R^d)$, then $f(x)$ can be represented by the *multiresolution wavelet decomposition* as

$$f(x_1,\cdots,x_d) = \sum_k \alpha_{j_0,k}\Phi_{j_0,k}(x_1,\cdots,x_d) + \sum_{j\geq j_0}\sum_k\sum_{l=1}^{2^d-1}\beta_{j,k}^{(l)}\Psi_{j,k}^{(l)}(x_1,\cdots,x_d) \tag{12}$$

where $k=(k_1,k_2,\cdots,k_d)\in Z^d$ and

$$\Phi_{j_0,k}(x_1,\cdots,x_d)=2^{j_0 d/2}\prod_{i=1}^{d}\phi(2^{j_0}x_i-k_i) \tag{13}$$

$$\Psi^{(l)}_{j,k}(x_1,\cdots,x_d)\ =2^{jd/2}\prod_{i=1}^{d}\eta^{(i)}(2^{j}x_i-k_i) \tag{14}$$

with $\eta^{(i)}=\phi$ or φ (scalar scaling function or the mother wavelet) but at least one $\eta^{(i)}=\varphi$.

Notice that if j_0 is large enough, the approximation representation (9) can be expressed using only the scaling function ϕ, that is, there exists a sufficiently large integer *J*, such that

$$f(x_1,\cdots,x_d)=\sum_k \alpha_{J,k}\Phi_{J,k}(x_1,\cdots,x_d)=\sum_{k_1,k_2,\cdots,k_d}2^{Jd/2}\prod_{i=1}^{d}\phi(2^{J}x_i-k_i) \tag{15}$$

3.2 The NARMAX model

The NARMAX model representation, which was initially proposed by Leontaritis & Billings [1985], takes the form of the following nonlinear difference equation:

$$y(t)=f(y(t-1),\cdots,y(t-n_y),u(t-1),\cdots,u(t-n_u),e(t-1),\cdots,e(t-n_e))+e(t) \tag{16}$$

where f is an unknown nonlinear mapping, $u(t)$ and $y(t)$ are the sampled input and output sequences, n_u and n_y are the maximum input and output lags, respectively. The noise variable $e(t)$ with maximum lag n_e, is unobservable but is assumed to be bounded and uncorrelated with the inputs and the past outputs. The model (16) relates the inputs and outputs and takes into account the combined effects of measurement noise, modelling errors and unmeasured disturbances represented by the noise variable $e(t)$. As a general and natural representation for a wide class of linear and nonlinear systems, model (16) includes, as special cases, several model types, including the Volterra and Wiener representations, time-invariant and time-varying AR(X), NARX and ARMA(X) structures, output-affine and rational models, and the bilinear model [Pearson 1999].

One of the popular representations for the NARMAX model (16) is the polynomial representation which takes the function $f(\cdot)$ as a polynomial of degree ℓ and gives the form as

$$\begin{aligned}y(t)=\theta_0+&\sum_{i_1=1}^{n}f_{i_1}(x_{i_1}(t))+\sum_{i_1=1}^{n}\sum_{i_2=i_1}^{n}f_{i_1 i_2}(x_{i_1}(t),x_{i_2}(t))+\cdots\\&+\sum_{i_1=1}^{n}\cdots\sum_{i_\ell=i_{\ell-1}}^{n}f_{i_1 i_2\cdots i_\ell}(x_{i_1}(t),x_{i_2}(t),\cdots,x_{i_\ell}(t))+e(t)\end{aligned} \tag{17}$$

where $\theta_{i_1 i_2\cdots i_m}$ are parameters, $n=n_y+n_u+n_e$ and

$$f_{i_1 i_2\cdots i_m}(x_{i_1}(t),x_{i_2}(t),\cdots,x_{i_m}(t))=\theta_{i_1 i_2\cdots i_m}\prod_{k=1}^{m}x_{i_k}(t),\ 1\le m\le \ell,$$

$$x_k(t)=\begin{cases}y(t-k) & 1\le k\le n_y\\ u(t-(k-n_y)) & n_y+1\le k\le n_y+n_u\\ e(t-(k-n_y-n_u)) & n_y+n_u+1\le k\le n_y+n_u+n_e\end{cases} \tag{18}$$

The degree of a multivariate polynomial is defined as the highest order among the terms. For example, the degree of the polynomial $h(x_1,x_2,x_3)=a_1x_1^4+a_2x_2x_3+a_3x_1^2x_2x_3^2$ is $\ell=2+1+2=5$. Similarly, a NARMAX model with polynomial degree ℓ means that the order of each term in the model is not higher than ℓ.

The NARX model is a special case of the NARMAX model and takes the form

$$y(t)=f(y(t-1),\cdots,y(t-n_y),u(t-1),\cdots,u(t-n_u))+e(t) \tag{19}$$

Similar to (18), (19) can be described using a polynomial representation with

$$x_k(t)=\begin{cases}y(t-k), & 1\le k\le n_y\\ u(t-k+n_y), & n_y+1\le k\le n=n_y+n_u\end{cases} \tag{20}$$

3.3 The quasi-ANOVA expansions

Generally, a multivariate nonlinear function can often be decomposed into a superposition of a number of functional components similar to the well known functional analysis of variance (ANOVA) expansions as below

$$\begin{aligned}y(t)&=f(x_1(t),x_2(t),\cdots,x_n(t))\\&=f_0+\sum_{i=1}^{n}f_i(x_i(t))+\sum_{1\le i<j\le n}f_{ij}(x_i(t),x_j(t))+\sum_{1\le i<j<k\le n}f_{ijk}(x_i,x_j,x_k)+\cdots\\&\quad+\sum_{1\le i_1<\cdots<i_m\le n}f_{i_1i_2\cdots i_m}(x_{i_1}(t),x_{i_2}(t),\cdots,x_{i_m}(t))+\cdots+f_{12\cdots n}(x_1(t),x_2(t),\cdots,x_n(t))+e(t)\end{aligned} \tag{21}$$

where the first functional component f_0 is a constant to indicate the intrinsic varying trend; f_i, f_{ij},$\cdots$, are univariate, bivariate, etc., functional components. The univariate functional components $f_i(x_i)$ represent the independent contribution to the system output that arises from the action of the ith variable x_i alone; the bivariate functional components $f_{ij}(x_i,x_j)$ represent the interacting contribution to the system output from the input variables x_i and x_j, etc. Let $x_k(t)$ (k=1,2,...,n) be defined as (18) or (20), the quasi-ANOVA expansion (21) can then be viewed as a special form of the NARMAX or NARX models for dynamic input and output systems. In practice, the constant term f_0 can often be omitted since it can be combined into other functional components.

In practice, many types of functions, such as kernel functions, splines, polynomials and other basis functions can be chosen to express the functional components in the models (21) and (22). In the present study, however, mutiresolution wavelet decompositions will be chosen to describe the functional components. The functional components $f_{i_1i_2\cdots i_r}(x_{i_1}(t),x_{i_2}(t),\cdots,x_{i_r}(t))\,(1\le i_1<i_2<\cdots<i_r\le n)$ will be expressed using the multiresolution wavelet decompositions (12) and (15). For example, the univariate and bivariate functional component $f_p(x_p(t))$ (p=1,2,...,n) and $f_{pq}(x_p(t),x_q(t))\,(1\le p<q\le n)$ can be expressed using the multiresolution wavelet decompositions (12) and (15) as

$$f_p(x_p(t)) = \sum_k \alpha_{j_1,k}^{(p)} \phi_{j_1,k}(x_p(t)) + \sum_{j \geq j_1} \sum_k \beta_{j,k}^{(p)} \varphi_{j,k}(x_p(t)),\ p = 1,2,\cdots,n, \tag{22}$$

$$f_{pq}(x_p(t), x_q(t)) = \sum_{k_1} \sum_{k_2} \alpha_{J;k_1,k_2} \phi_{J,k_1}(x_p(t)) \phi_{J,k_2}(x_q(t)),\ 1 \leq p < q \leq n, \tag{23}$$

3.4 The wavelet-NARX model

Consider the NARX model (19) and assume that the nonlinear mapping f can be decomposed into a number of functional components up to the rth-variate functional terms using the quasi-ANOVA expansion (21), then the NARX model (19) can be expressed as

$$\begin{aligned} y(t) &= f(x_1(t), x_2(t), \cdots, x_n(t)) + e(t) \\ &= F_1(x(t)) + F_2(x(t)) + \cdots + F_r(x(t)) + e(t) \end{aligned} \tag{24}$$

where $x(t) = [x_1(t), x_2(t), \cdots, x_n(t)]^T$ and

$$x_k(t) = \begin{cases} y(t-k), & 1 \leq k \leq n_y \\ u(t-k+n_y), & n_y + 1 \leq k \leq n = n_y + n_u \end{cases} \tag{24a}$$

$$F_1(x(t)) = \sum_{i=1}^{n} f_i(x_i(t)) \tag{24c}$$

$$F_2(x(t)) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} f_{ij}(x_i(t), x_j(t)) \tag{24d}$$

$$F_r(x(t)) = \sum_{1 \leq i_1 < i_2 < \cdots < i_r \leq n} f_{i_1 i_2 \cdots i_r}(x_{i_1}(t), x_{i_2}(t), \cdots, x_{i_r}(t)),\quad 2 < r \leq n, \tag{24e}$$

The univariate and bivariate functional components $f_p(x_p(t))$ (p=1,2,...,n) and $f_{pq}(x_p(t), x_q(t))$ ($1 \leq p < q \leq n$) can be expressed using the multiresolution wavelet decompositions (22) and (23). And the higer-order functional components $f_{i_1 i_2 \cdots i_r}(x_{i_1}(t), x_{i_2}(t), \cdots, x_{i_r}(t))$ ($2 < r \leq n$) will be approximated using the multiresolution wavelet decompositions (15) in this study as

$$\begin{aligned} f_{i_1 i_2 \cdots i_r}(x_1(t), \cdots, x_r(t)) &= \sum_k \alpha_{J_r,k} \Phi_{J_r,k}(x_1(t), \cdots, x_r(t)) \\ &= \sum_{k_1, k_2, \cdots, k_r} 2^{rJ_r/2} \prod_{i=1}^{r} \phi(2^{J_r} x_i(t) - k_i) \end{aligned} \tag{25}$$

where J_r (r=3,4,...) are some appropriate integers for the approximation of the wavelet decompositions. The multiresolution wavelet model (24) was referred to as the wavelet-NARX model, or WANARX in [Wei and Billings, 2003], and will be used for identifying CA and CML models. Although many functions can be chosen as scaling and/or wavelet functions, most of these are not suitable in system identification applications, especially in the case of multidimensional and multiresolution expansions. An implementation, which has been tested with very good results, involves B-spline and B-wavelet functions in multiresolution wavelet decompositions [Billings and Coca 1999, Wei and Billings 2002]. B-spline wavelets were originally introduced by Chui and Wang [Chui 1992; Chui and Wang 1992] to define a class of semi-orthogonal wavelets. In the following, the B-spline wavelet and scaling functions will be considered and used in CA and CML identification. Some implementation issues on mutiresolution wavelet models such as data normalization, highest resolution

level determination, translation parameter selection and wavelet dictionary determination were discussed in detail in [Billings and Wei 2003; Wei and Billing, 2003].

Assume that M bases (scalar mother wavelet or scaling functions or multiplication of some scalar wavelet and scaling functions) are required to expand the NARX model (24), and for convenience of representation also assume that the M wavelet bases are ordered according to a single index m, that is, the wavelet dictionary $D = \{p_m\}_{m=1}^{M}$, then (24) can be expressed as a linear-in-the-parameters form as below:

$$y(t) = \sum_{m=1}^{M} \theta_m p_m(t) + e(t) \tag{26}$$

which can be solved using linear regression techniques. Note that for large n_y and n_u, the model (26) might involve a great number of model terms or regressors. Experience shows that very often many of the model terms are redundant and therefore are insignificant to the system output and can be removed from the model. An efficient orthogonal least squares (OLS) algorithm and an error deduction ratio (ERR) criterion [Korenberg et al. 1988; Chen et al. 1989] was developed to determine which terms should be included in the model. A new fast matching pursuit orthogonal least squares (MPOLS) algorithm [Billings and Wei 2003] can also be used to select significant model terms for low-dimensional CA models.

3.5 Identification of CA and CML models

From Section 2, both CA's and CML's can be expressed in the form

$$x(i,j;t) = f(N^d[x(i,j;t)]) \tag{27}$$

The objective of identification is to obtain a model which can approximate the unknown nonlinear mapping f providing that no explicit Boolean rules for CA model (3) are available, or no explicitly formulated functions $f^{(1)}$ and $f^{(2)}$ for CML model (6) are known beforehand but only some spatio-temporal patterns have been observed.

Although the nonlinear mapping f in CA models (3) is generally represented by a set of Boolean functions of the values of the sites belonging to the neighbourhood, it was shown by Billings and Yang [2003] that polynomial models, which can be identified from a limit set of observational records, can also be used to represent the nonlinear mapping f. This provides an alternative representation for CA's where no Boolean rules are available explicitly. It also follows that the CML model (6) can be approximated by the polynomial NARX model perfectly [Coca and Billings 2001; Billings and Coca 2002]. In the present study, however, the new modelling structure introduced in the previous section will be used to represent CA's and CML's.

Taking a simple two-dimensional case as an example by setting d=1, $r_1 = r_2 = s_1 = s_2 = 1$ in (2), where the *Moore* neighbourhood is involved during the system evolution, and this can be expressed as a form of (5). The values of the 9 cells in $N^1[x(i,j;t)]$, together with the value of the cell $x(i, j; t)$ to be updated at the present time step, will be referred to as *a group of observations*, or *an observational group* for identification. Denote the 9 values in a group by $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$, and x_9 in the sense that

$$x_1 \leftarrow x(i-1, j-1; t-1)$$
$$x_2 \leftarrow x(i-1, j; t-1)$$

$$x_3 \leftarrow x(i-1, j+1; t-1)$$
$$x_4 \leftarrow x(i, j-1; t-1)$$
$$x_5 \leftarrow x(i, j; t-1)$$
$$x_6 \leftarrow x(i, j+1; t-1)$$
$$x_7 \leftarrow x(i+1, j-1; t-1)$$
$$x_8 \leftarrow x(i+1, j; t-1)$$
$$x_9 \leftarrow x(i+1, j+1; t-1)$$

which can be viewed as the input of the group, and denote the value of the centre cell $x(i, j; t)$ by y, which can be viewed as the output of the group. Assuming that K observational groups for identification are chosen from the given spatio-temporal patterns, number these observational groups as $\{y(k), x_i(k)\}$, i=1,2, ..., 9; k=1, 2, ..., K. The output $y(k)$ and the input $\{x_i(k)\}$ are related by a nonlinear mapping f as

$$y(k) = f(x_1(k), x_2(k), \cdots, x_9(k)), \quad k = 1,2,\cdots,K, \tag{28}$$

Notice that the input-output data here are related to the observational groups $\{y(k), x_i(k)\}$ (k=1,2,...K) and are relevant to not only time but also space. The wavelet-NARX model (24) will be used to approximate the unknown nonlinear function f and this will be demonstrated by some examples proposed in the next section. Note that the definition of the variables x_i in (28) are different form those in (24). This two-dimensional case can easily be extended to a higher dimensional case with a higher order.

4. Simulation Examples

In this section, three examples, one 2D CML system, one 1D and one 2D CA's, are simulated and identified using the new proposed modelling approach. In the following sections, it will be assumed that the meaning of the symbol $x_{i,j}^t$ is identical to that of $x(i,j; t)$ and therefore these two symbols can be interchanged.

4.1 Example 1: A two-dimensional CML

Consider the 2D CML system defined by (8) with a=1.5, c=0.4. This system was simulated with random initial and periodic boundary conditions, where the initial values of the cells $x(i,j;0)$ were uniformly distributed in [0, 1]. Some patterns produced by this CML model are shown in figure 1. Altogether 500 observational groups were arbitrarily chosen from the initial (t=0), first (t=1) and second (t=2) step patterns of 100×100 site lattices and these observational groups were used for identification. For the purpose of identification, it was assumed that these patterns were generated from a two-dimensional CML of 1st-order, but the neighbourhood was unknown. Therefore all the 9 neighbouring cells in the Moore neighbourhood instead of the 5 cells in the von Neumann neighbourhood structure, were considered for identification. The initial wavelet model was chosen as

$$\begin{aligned} y(\tau) &= f(x_1(\tau), x_2(\tau), \cdots, x_9(\tau)) \\ &= f_0 + \sum_{i_1=1}^{9} f_{i_1}(x_{i_1}(\tau)) + \sum_{i_1=1}^{8}\sum_{i_2=i_1+1}^{9} f_{i_1 i_2}(x_{i_1}(\tau), x_{i_2}(\tau)) + \sum_{i_1=1}^{7}\sum_{i_1=i_2+1}^{8}\sum_{i_3=i_2+1}^{9} f_{i_1 i_2 i_3}(x_{i_1}(\tau), x_{i_2}(\tau), x_{i_3}(\tau)) \end{aligned} \tag{29}$$

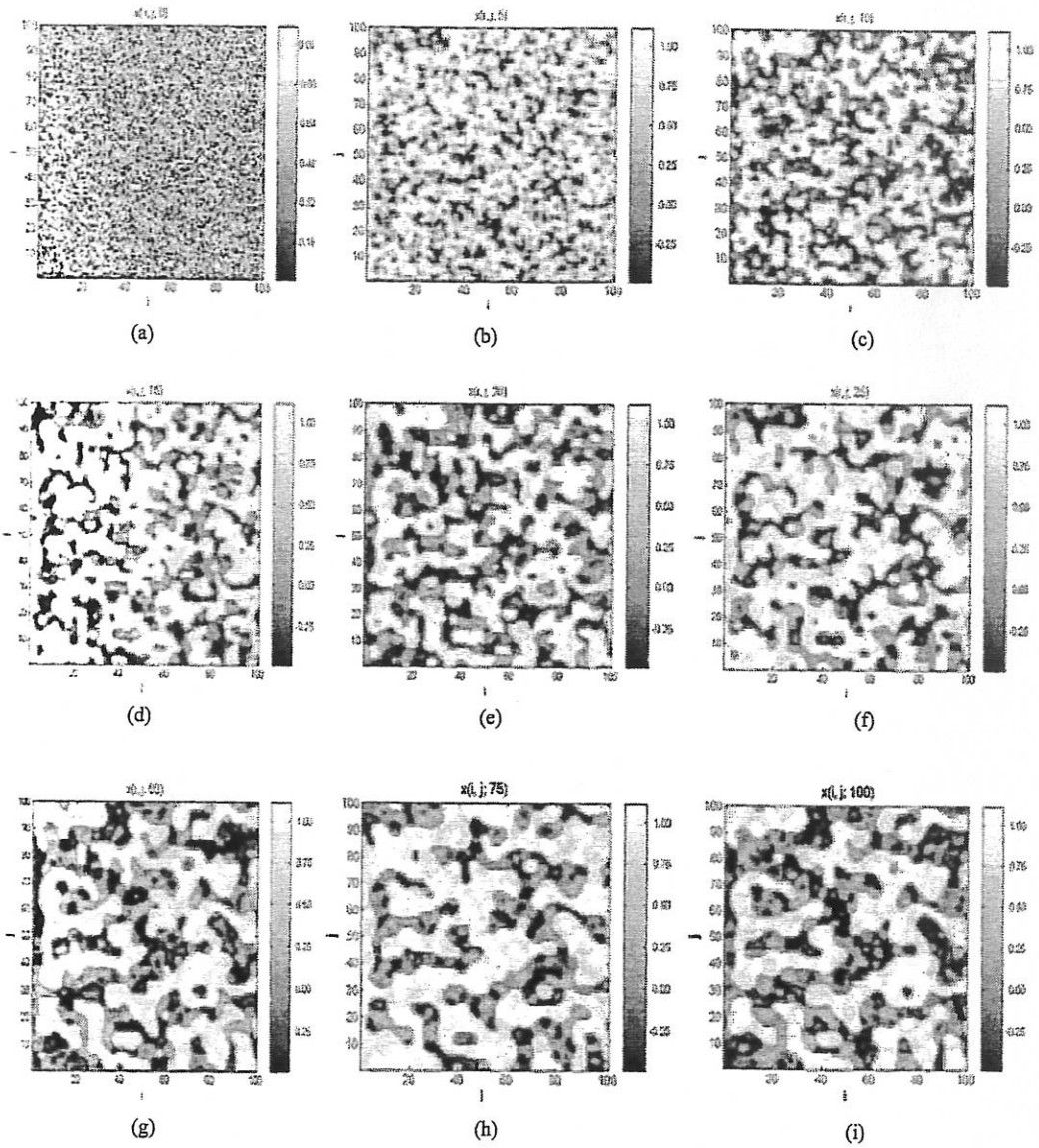


Fig. 1 Patterns generated from the 2D CML model (8) at different time steps. (a) for t=0; (b) for t=5; (c) for t=10; (d) for t=15; (e) for t=20; (f) for t=25; (g) for t=50; (h) for t=75 and (i) for t=100.

where $x_i(\tau)$ (i=1,2, …, 9)were defined as in Section 3.5, f_0 is a constant, the univariate functional components $f_{i_1}(x_{i_1}(\tau))$ (i_1=1,2, …, 9) were defined as (22) with starting resolution scale j_1=0 and the highest resolution scale J_1=4, the bivariate functional components $f_{i_1 i_2}(x_{i_1}(\tau), x_{i_2}(\tau))$ ($1 \le i_1 < i_2 \le 9$) were defined as (23) with a resolution scale J_2=2, and the tri-variate functional components $f_{i_1 i_2 i_3}(x_{i_1}(\tau), x_{i_2}(\tau), x_{i_3}(\tau))$ ($1 \le i_1 < i_2 < i_3 \le 9$)were defines as (25) with a resolution scale J_3=0. The model (29) involved 3690 regressors (model terms) when each of the functional components was decomposed into a wavelet model. Only 10 significant model terms were selected by performing the OLS based model term selection algorithm [Korenberg et al. 1988; Chen et al. 1989], and a parsimonious model was then obtained as below

$$y(\tau) = \sum_{k=1}^{10} \theta_k p_k(\tau) + e(\tau) \tag{30}$$

which is equivalent to

$$x(i,j;t) = \sum_{k=1}^{10} \theta_k p_k(i,j;t) + e(t) \tag{31}$$

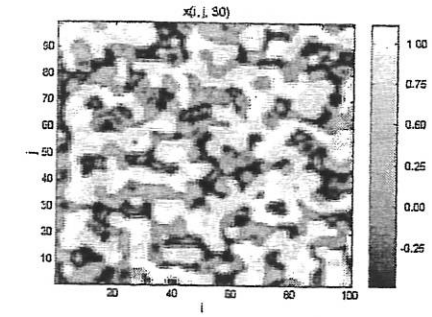
where the regressors $p_k(i,j;t)$, the parameters θ_k and the corresponding ERRs are listed in Table 1. The model (31) was used to reproduce spatio-temporal patterns with the same initial conditions as those in Figure 1, and the pattern reproduced by the model (31) at time step t=30 was compared with that produced by the original model (8), this is shown in Figure 2. In Figure 3, some patterns reproduced from the identified wavelet model (31) with different initial conditions are shown.

Table 1 Basis functions, parameters and the corresponding ERRs for the 2D CML in Example 1

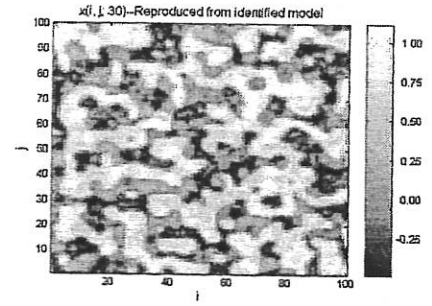
Index k	Terms $p_k(i,j;t)$	Parameters θ_k	$ERR_k \times 100\%$
1	$\phi_{0,-2}(x_{i,j-1}^{t-1})$	1.60242108e+000	9.53843943e+001
2	$\phi_{0,-1}(x_{i-1,j}^{t-1})\phi_{0,-1}(x_{i,j+1}^{t-1})\phi_{0,-1}(x_{i+1,j}^{t-1})$	-2.51739222e-001	2.75756235e+000
3	$\phi_{0,-1}(x_{i,j-1}^{t-1})$	-5.65822941e-001	8.74162094e-001
4	$\varphi_{0,-2}(x_{i,j}^{t-1})$	-1.53273442e-001	4.22358294e-001
5	$\phi_{0,-3}(x_{i,j}^{t-1})\phi_{0,-2}(x_{i,j+1}^{t-1})\phi_{0,-2}(x_{i+1,j}^{t-1})$	4.00966751e-001	8.28431865e-002
6	$\phi_{0,0}(x_{i-1,j}^{t-1})$	-7.85191475e-001	1.10418835e-001
7	$\phi_{0,-1}(x_{i,j-1}^{t-1})\phi_{0,-2}(x_{i,j+1}^{t-1})\phi_{0,-1}(x_{i+1,j}^{t-1})$	7.52240142e-001	1.05492604e-001
8	$\phi_{0,-3}(x_{i,j-1}^{t-1})\phi_{0,-1}(x_{i,j+1}^{t-1})\phi_{0,-1}(x_{i+1,j}^{t-1})$	-2.55627625e+000	7.61658071e-002
9	$\varphi_{0,0}(x_{i,j}^{t-1})$	5.63356952e+000	7.59801210e-002
10	$\phi_{2,0}(x_{i,j-1}^{t-1})\phi_{2,0}(x_{i+1,j}^{t-1})$	2.61603102e-002	2.81694124e-002

Note: 1) The threshold value $\rho = 10^{-3}$, $\sum_{k=1}^{10} ERR_k \ge 1-\rho$.

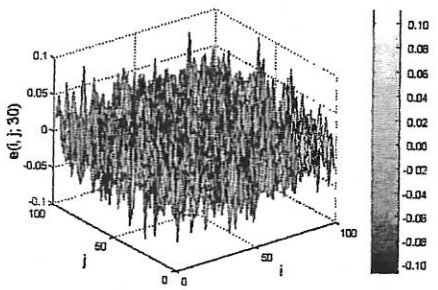
2) $\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$ and $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$ — the 4th-order B-splne wavelet and scaling functions.



(a)



(b)



(c)

Fig. 2 Comparison of the patterns generated from the original model (8) and the identified wavelet model (31) at time step t=30, where the same initial values were used for both models. (a) generated from the original model (8); (b) reproduced by the identified model (31) and (c) model predicted errors $e(i, j;30) = x(i, j;30) - \hat{x}(i, j;30)$, where $\hat{x}$ is generated from the identified model (31).

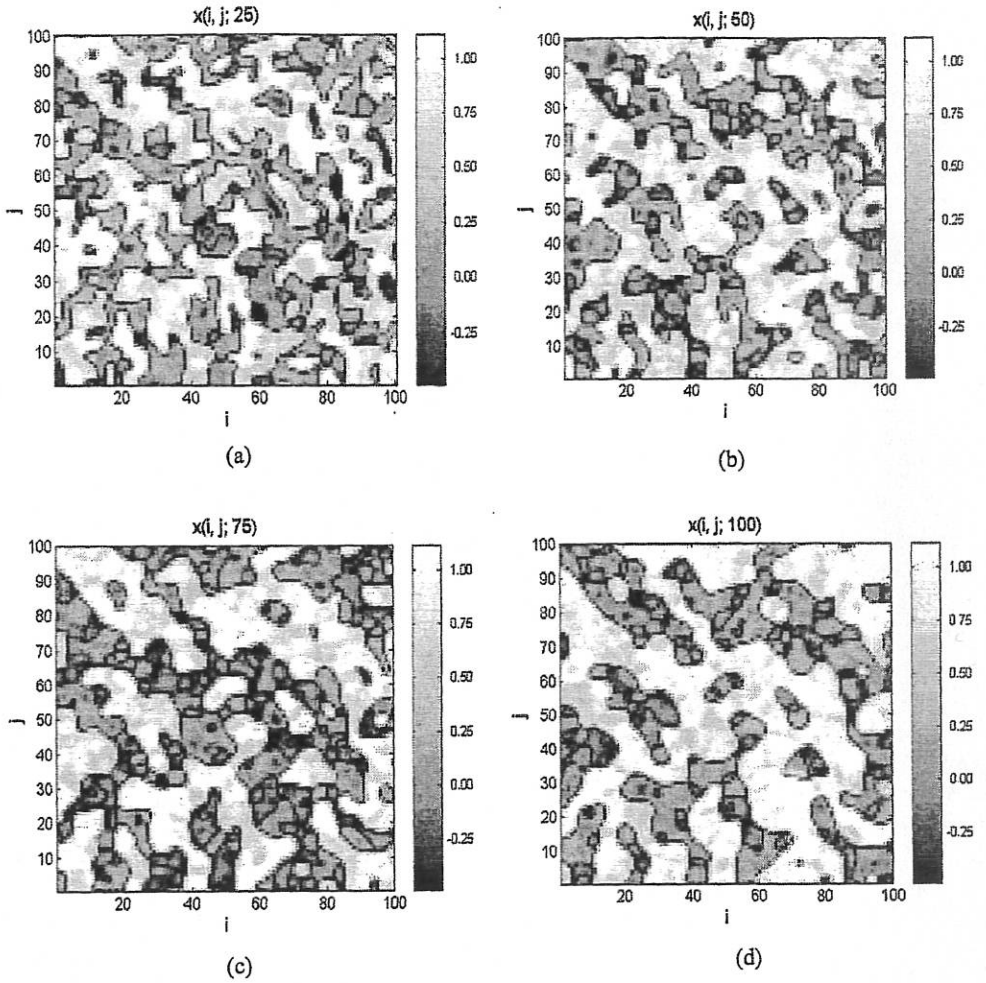


Fig. 3 Patterns reproduced by the identified wavelet model (31) with different initial values from those in Fig. 1. (a) for t=25; (b) for t=50; (c) for t=75; and (d) for t=100.

4.2 Example 2: A one-dimensional CA

One-dimensional CA's can be expressed by

$$s_j(t) = f(s_{j-p}(t-1), s_{j-p+1}(t-1), \cdots, s_{j-1}(t-1), s_j(t-1), s_{j+1}(t-1), \cdots, s_{j+q}(t-1)) \qquad (32)$$

where $s_j(t)$ indicates the jth cell at time t. Model (32) can be expressed using the multiresolution wavelet representation (24) and then the OLS [Korenberg et al. 1988; Chen et al. 1989] or MPOLS [Billings and Wei 2003] algorithms can be applied to obtain a parsimonious model, which can be used to reconstruct the original CA patterns. The property of the Haar wavelet and scaling functions makes the multiresolution Haar wavelet models particularly appropriate to represent binary cellular automata.

Taking Rule54 as an example, the rule for the 3-site von Neumann neighbourhood is shown in Table 2 and the pattern of a 200×200-site lattice generated by this rule with a set of random initial values is shown in Fig. 4, where periodic boundary conditions were considered so that the first site is a nearest neighbour of the n-th site in an n-site automata.

Table 2 The one dimensional three-site CA Rule 54--the local states and the updated central site values

$s_{j-1}(t-1)s_j(t-1)s_{j+1}(t-1)$	000	001	010	011	100	101	110	111
$s_j(t)$	0	1	1	0	1	1	0	0

To identify the rule, the multiresolution wavelet model was initially chosen as

$$\begin{aligned} s_j(t) &= f(s_{j-3}(t-1), s_{j-2}(t-1), s_{j-1}(t-1), s_j(t-1), s_{j+1}(t-1), s_{j+2}(t-1), s_{j+3}(t-1)) \\ &= f(x_1(t), x_2(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t)) \\ &= \sum_{i=1}^{7} f_i(x_i(t)) + \sum_{i_1=1}^{6}\sum_{i_2=i_1+1}^{7} f_{i_1 i_2}(x_{i_1}(t), x_{i_2}(t)) \\ &+ \sum_{i_1=1}^{5}\sum_{i_2=i_1+1}^{6}\sum_{i_3=i_2+1}^{7} f_{i_1 i_2 i_3}(x_{i_1}(t), x_{i_2}(t), x_{i_3}(t)) + \cdots \\ &+ \sum_{i_1=1}^{2}\cdots\sum_{i_6=i_5+1}^{7} f_{i_1 i_2 i_3 i_4 i_5 i_6}(x_{i_1}(t), x_{i_2}(t), \cdots, x_{i_6}(t)) \\ &+ f_{1234567}(x_1(t), x_2(t), \cdots, x_7(t)) \end{aligned} \tag{33}$$

where $x_k(t) = s_{8-k}(t-1)$. Each univariate function component was decomposed using the Haar wavelet and scaling (the first-order B-spline wavelet and scaling) functions as

$$f_i(x_i(t)) = a_{0,0}^{(i)}\phi_{0,0}(x_i(t)) + \sum_{j=0}^{5}\sum_{k\in K_j} d_{j,k}^{(i)}\varphi_{j,k}(x_i(t)) \tag{34}$$

where $K_j = \{0,1,\cdots,2^j-1\}$. Other function components were decomposed using the Haar scaling functions as

$$f_{i_1\cdots i_m}(x_{i_1}(t),\cdots,x_{i_m}(t)) = \sum_{k_1=0}^{2^{J_m}-1}\cdots\sum_{k_m=0}^{2^{J_m}-1} a_{J_m;k_1,\cdots,k_m}^{(i_1\cdots i_m)}\prod_{p=1}^{m}\phi_{J_m,k_m}(x_{i_p}(t)) \tag{35}$$

where J_2=3, J_3=2, J_4=J_5=J_6=J_7=1. $\phi(x)$ and $\varphi(x)$ are the Haar scaling and wavelet functions defined as

$$\phi(x) = \begin{cases} 1 & \text{for } 0\le x<1 \\ 0 & \text{otherwise} \end{cases} \tag{36}$$

$$\varphi(x) = \begin{cases} 1 & 0\le x<1/2 \\ -1 & 1/2\le x<1 \\ 0 & \text{otherwise} \end{cases} \tag{37}$$

The model (33) contains 5776 candidate basis functions (regressors) after decomposition into the multiresolution wavelet models. Both the OLS and MPOLS algorithms were used to select the significant model terms based on 200 observational groups arbitrarily selected from the given pattern. The results from the OLS and MPOLS were identical and these are listed in Table 3.

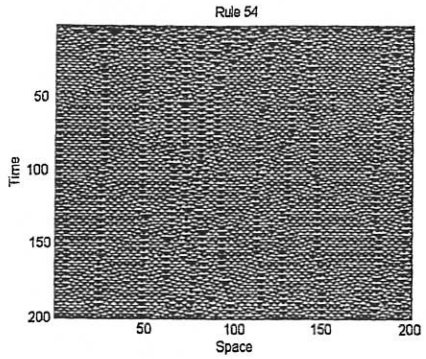


Fig. 4 A CA pattern generated from the Rule54.

Table 3 Basis functions, parameters and the corresponding ERRs for the1D CA in Example 2

Index k	Terms $p_k(j,t)$	Parameters θ_k	$ERR_k \times 100\%$
1	$\varphi_{4,0}(s(j,t-1))$	2.50000000e-001	5.16154273e+001
2	$\phi_{3,0}(s(j-1,t-1)) \times \phi_{3,0}(s(j+1,t-1))$	1.25000000e-001	2.31182796e+001
3	$\phi_{2,0}(s(j-1,t-1)) \times \phi_{2,0}(s(j,t-1))$ $\times \phi_{2,0}(s(j+1,t-1))$	-2.50000000e-001	2.52662931e+001
Note: 1) $ERR_1 + ERR_2 + ERR_3 = 1$. 2) The CPU time taken by OLS and MPOLS algorithms to select these model terms from all the candidate term set was 189.37s and 7.11s, respectively.			

The final identified wavelet model for Rule54 was found to be

$$\begin{aligned} s_j(t) &= 0.125\varphi_{4,0}(s_j(t-1)) \\ &\quad +0.125\phi_{3,0}(s_{j-1}(t-1))\phi_{3,0}(s_{j+1}(t-1)) \\ &\quad -0.25\phi_{2,0}(s_{j-1}(t-1))\phi_{2,0}(s_j(t-1))\phi_{2,0}(s_{j+1}(t-1)) \\ &= \varphi(16s_j(t-1)) + \phi(8s_{j-1}(t-1))\phi(8s_{j+1}(t-1)) - 2\phi(4s_{j-1}(t-1))\phi(4s_j(t-1))\phi(4s_{j+1}(t-1)) \end{aligned} \tag{38}$$

Noting that the values of $s_{j-1}(t-1)$, $s_j(t-1)$ and $s_{j+1}(t-1)$ are set to be 0 or 1 for all j and t, from the definition of the functions $\phi(x)$ and $\varphi(x)$, (38) can be simplified as

$$\begin{aligned} s_j(t) &= \varphi(s_j(t-1)) + \phi(s_{j-1}(t-1))\phi(s_{j+1}(t-1)) - 2\phi(s_{j-1}(t-1))\phi(s_j(t-1))\phi(s_{j+1}(t-1)) \\ &= [1-s_j(t-1)] + [1-s_{j-1}(t-1)][1-s_{j+1}(t-1)] \\ &\quad -2[1-s_{j-1}(t-1)][1-s_j(t-1)][1-s_{j+1}(t-1)] \end{aligned} \tag{39}$$

The model (39) is identical to the Boolean rule listed in Table 2 and the validity of this model can easily be verified. In this example, the multiresolution wavelet model (33) for a Boolean rule was finally simplified as a polynomial model. In fact by setting all the wavelet shift parameters (translations) to be zero in the initial wavelet model (33), this will result in a polynomial representation. This clearly indicates that the polynomial

representations for binary CA rules are a special case of the wavelet models where the basis functions are chosen to be the Haar scaling and wavelet functions.

Notice that only a small number of observational groups of the CA pattern were used for model estimation and these observational groups can be arbitrarily chosen from the pattern. No a priori knowledge of the system structure was assumed except that the neighbourhood radius was restricted to be not greater than 3 for this example.

4.3 Example 3: A two-dimensional CA

Consider a two-dimensional CA defined by the following outer totalistic rule

$$x(i,j,t)=\begin{cases}1 & X(i,j;t-1)=0 \text{ or } 1\\ 0 & X(i,j;t-1)=2,3 \text{ or } 4\end{cases} \tag{40}$$

where $X(i,j;t)=x(i-1,j;t)+x(i,j-1;t)+x(i,j+1;t)+x(i+1,j;t)$. This system was simulated with random initial and periodic boundary conditions, where the initial values of the cells $x(i,j;0)$ were chosen to be 0 with a probability $p_0=P\{x(i,j;0)=0\}$=0.475 and were chosen to be 1 with a probability $p_1=P\{x(i,j;0)=1\}$ =0.525. Some patterns produced by the rule (40) are shown in figure 5. Altogether 300 observational groups were arbitrarily chosen from the initial (t=0), first (t=1) and second (t=2) step patterns of 100×100 site lattices and these observational groups were used for identification. For the purpose of identification, it was assumed that these patterns were generated from a two-dimensional CML of 1st-order, but the neighbourhood was unknown. Therefore all the 9 neighbouring cells in the Moore neighbourhood instead of the 5 cells in the von Neumann neighbourhood structure, was considered for identification. The initial wavelet model was chosen as

$$\begin{aligned}y(\tau)&=f(x_1(\tau),x_2(\tau),\cdots,x_9(\tau))\\&=f_0+\sum_{i_1=1}^{9}f_{i_1}(x_{i_1}(\tau))+\sum_{i_1=1}^{8}\sum_{i_2=i_1+1}^{9}f_{i_1i_2}(x_{i_1}(\tau),x_{i_2}(\tau))\\&\quad+\sum_{i_1=1}^{7}\sum_{i_1=i_2+1}^{8}\sum_{i_3=i_2+1}^{9}f_{i_1i_2i_3}(x_{i_1}(\tau),x_{i_2}(\tau),x_{i_3}(\tau))+\cdots\\&\quad+\sum_{i_1=1}^{2}\cdots\sum_{i_8=i_7+1}^{9}f_{i_1i_2\cdots i_8}(x_{i_1}(\tau),x_{i_2}(\tau),\cdots,x_{i_8}(\tau))\\&\quad+f_{123456789}(x_1(\tau),x_2(\tau),\cdots,x_9(\tau))\end{aligned} \tag{41}$$

where $x_i(\tau)$ (i=1,2, ..., 9)were defined as in Section 3.5. Each univariate function component was decomposed using the Haar wavelet and scaling functions as (34) with starting resolution scale j_1=0 and the highest resolution scale J_1=5, and bivariate and other higher order functional components were decomposed using the Haar scaling functions as (35) with multiresolution scales J_2=2, J_3=J_4=1, and J_5=J_6=J_7=J_8=J_9=0, respectively. The initial model (41) contains 4096 candidate regressors (model terms) after decomposition into the multiresolution wavelet models. Both the OLS and MPOLS algorithms were used to select the significant model terms. The results from the OLS and MPOLS were identical and these are listed in Table 4.

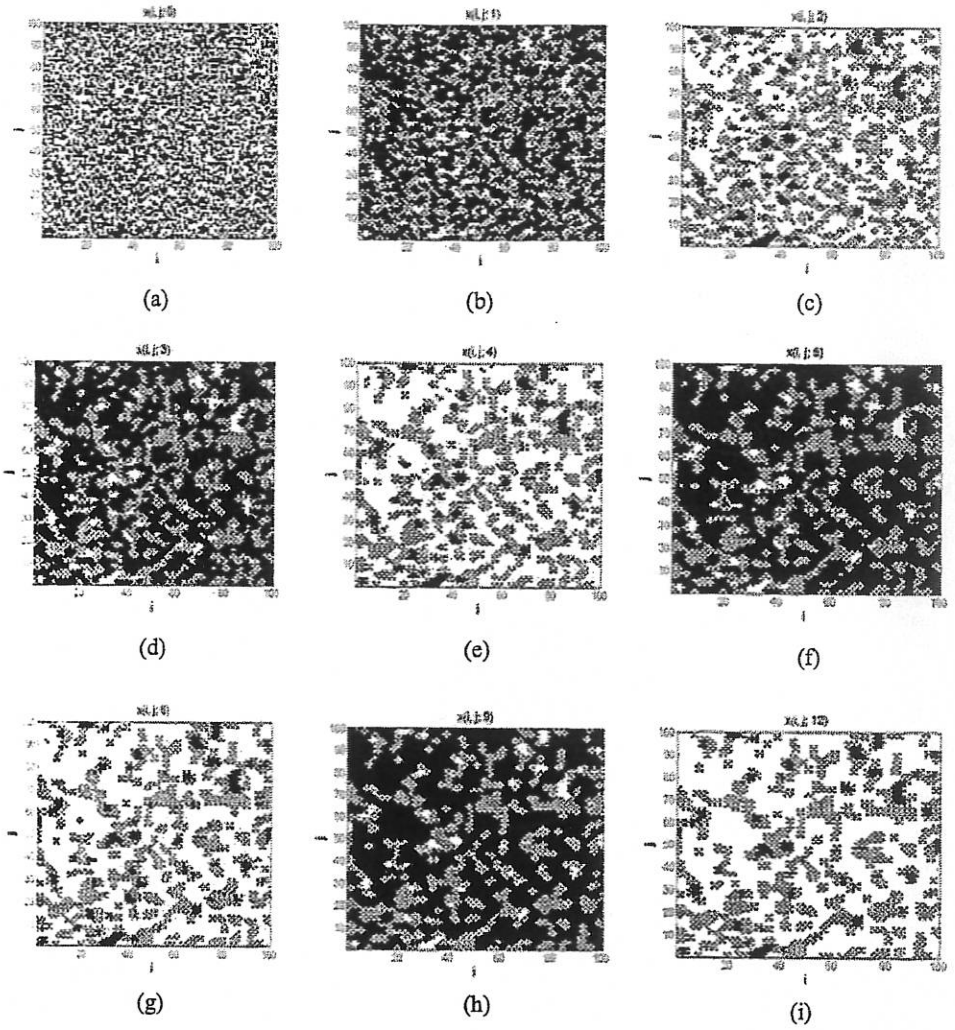


Fig. 5 Patterns generated from the 2D CA model (40) at different time steps. (a) for t=0; (b) for t=1; (c) for t=2; (d) for t=3; (e) for t=4; (f) for t=5; (g) for t=6; (h) for t=9 and (i) for t=12.

Table 4 Basis functions, parameters and the corresponding ERRs for the2D CA in Example 3

Index k	Terms $p_k(i,j;t)$	Parameters θ_k	$ERR_k \times 100\%$
1	$\phi_{1,0}(x_{i-1,j}^{t-1})\phi_{1,0}(x_{i,j-1}^{t-1})\phi_{1,0}(x_{i,j+1}^{t-1})$	3.53553391e-001	2.12765957e+001
2	$\phi_{1,0}(x_{i-1,j}^{t-1})\phi_{1,0}(x_{i,j-1}^{t-1})\phi_{1,0}(x_{i+1,j}^{t-1})$	3.53553391e-001	2.34042553e+001
3	$\phi_{1,0}(x_{i-1,j}^{t-1})\phi_{1,0}(x_{i,j+1}^{t-1})\phi_{1,0}(x_{i+1,j}^{t-1})$	3.53553391e-001	1.48936170e+001
4	$\phi_{1,0}(x_{i,j-1}^{t-1})\phi_{1,0}(x_{i,j+1}^{t-1})\phi_{1,0}(x_{i+1,j}^{t-1})$	3.53553391e-001	2.34042553e+001
5	$\phi_{0,0}(x_{i-1,j}^{t-1})\phi_{0,0}(x_{i,j-1}^{t-1})\phi_{0,0}(x_{i,j+1}^{t-1})\phi_{0,0}(x_{i+1,j}^{t-1})$	-3.00000000e+000	1.70212766e+001

Note: 1) $\sum_{k=1}^{5} ERR_k = 1$; 2) $\sqrt{2}/4$ =0.353553391

3) The CPU time taken by OLS and MPOLS algorithms to select these model terms from all the candidate term set was 217.86*s* and 15.42s, respectively.

The identified wavelet model for the totalistic rule described by (40) was found to be

$$\begin{aligned} x(i,j;t) =& \frac{\sqrt{2}}{4}\phi_{10}(x(i-1,j;t-1))\phi_{10}(x(i,j-1;t-1))\phi_{10}(x(i,j+1;t-1)) \\ &+\frac{\sqrt{2}}{4}\phi_{10}(x(i-1,j;t-1))\phi_{10}(x(i,j-1;t-1))\phi_{10}(x(i+1,j;t-1)) \\ &+\frac{\sqrt{2}}{4}\phi_{10}(x(i-1,j;t-1))\phi_{10}(x(i,j+1;t-1))\phi_{10}(x(i+1,j;t-1)) \\ &+\frac{\sqrt{2}}{4}\phi_{10}(x(i,j-1;t-1))\phi_{10}(x(i,j-1;t-1))\phi_{10}(x(i+1,j;t-1)) \\ &-3\phi_{00}(x(i-1,j;t-1))\phi_{00}(x(i,j-1;t-1))\phi_{00}(x(i,j+1;t-1))\phi_{00}(x(i+1,j;t-1)) \end{aligned} \quad (42)$$

Similar to (38), Eq. (42) can be simplified as

$$\begin{aligned} x(i,j;t) &= g_1 + g_2 + g_3 + g_4 + g_5 \\ &= \phi(x(i-1,j;t-1))\phi(x(i,j-1;t-1))\phi(x(i,j+1;t-1)) \\ &+\phi(x(i-1,j;t-1))\phi(x(i,j-1;t-1))\phi(x(i+1,j;t-1)) \\ &+\phi(x(i-1,j;t-1))\phi(x(i,j+1;t-1))\phi(x(i+1,j;t-1)) \\ &+\phi(x(i,j-1;t-1))\phi(x(i,j-1;t-1))\phi(x(i+1,j;t-1)) \\ &-3\phi(x(i-1,j;t-1))\phi(x(i,j-1;t-1))\phi(x(i,j+1;t-1))\phi(x(i+1,j;t-1)) \end{aligned} \quad (43)$$

The final identified model (43) was validated in Table 5, where $x_{i,j}^{t} = x(i,j;t)$ and $X = x_{i-1,j}^{t-1} + x_{i,j-1}^{t-1} + x_{i,j+1}^{t-1} + x_{i+1,j}^{t-1}$. Table 5 clearly shows that the identified model (43) is identical to the original totalistic rule (40).

Table 5 Model validation for the identified wavelet model (43) in Example 3

No.	$x_{i-1,j}^{t-1}$	$x_{i,j-1}^{t-1}$	$x_{i,j+1}^{t-1}$	$x_{i+1,j}^{t-1}$	g_1	g_2	g_3	g_4	g_5	X	$x_{i,j}^{t} = x(i,j;t)$	
											Rule (40)	Model (43)
1	0	0	0	0	1	1	1	1	-3	0	1	1
2	0	0	0	1	1	0	0	0	0	1	1	1
3	0	0	1	0	0	1	0	0	0	1	1	1
4	0	0	1	1	0	0	0	0	0	2	0	0
5	0	1	0	0	0	0	1	0	0	1	1	1
6	0	1	0	1	0	0	0	0	0	2	0	0
7	0	1	1	0	0	0	0	0	0	2	0	0
8	0	1	1	1	0	0	0	0	0	3	0	0
9	1	0	0	0	0	0	0	1	0	1	1	1
10	1	0	0	1	0	0	0	0	0	2	0	0
11	1	0	1	0	0	0	0	0	0	2	0	0
12	1	0	1	1	0	0	0	0	0	3	0	0
13	1	1	0	0	0	0	0	0	0	2	0	0
14	1	1	0	1	0	0	0	0	0	3	0	0
15	1	1	1	0	0	0	0	0	0	3	0	0
16	1	1	1	1	0	0	0	0	0	4	0	0

5. Conclusions

A novel effective approach for identifying spatio-temporal systems has been proposed based on multiresolution wavelet decompositions. The new proposed modelling framework has been successfully applied to identify a two-dimensional CML system and one and two dimensional CA rules.

The close relationship in structure and in the behaviour between CA and CML's make it possible to deal with these two classes of spatio-temporal systems in a uniform way as shown in the present study. Whilst the wavelet models based on higher order B-spline wavelet and scaling functions provide a good representation for CML's, the models with respect to the Haar wavelet and scaling functions (the first order B-spline wavelet and scaling function) provide accurate representations for CA rules. It was shown that the polynomial representation of Boolean rules is a special case of the Haar wavelet models, where all the wavelet shift parameters (translations) are set to be zero.

The new approach has been exemplified using noise-free simulated data from one and two dimensional deterministic spatio-temporal systems. However, it would be hoped that the approach is also applicable to noise corrupted or even stochastic spatio-temporal systems, and this will be illustrated in a separate study.

Acknowledgment

The authors gratefully acknowledge that part of this work was supported by EPSRC(UK).

References

Adamatzky, A., 1994, *Identification of Cellular Automata* (London: Taylors & Francis).

Billings, S.A., and Coca, D., 2002, Identification of coupled map lattice models of deterministic distributed parameter systems. *International Journal of Systems Science*, **33**, 623-634.

Billings, S.A., and Wei, H.L., 2003, The wavelet-NARMAX representation: a hybrid model structure combining polynomial models with multiresolution wavelet decompositions (submitted for publication).

Billings, S.A. and Yang, Y.X.,2003, Identification of the neighbourhood and CA rules from spatio-temporal CA Patterns. *IEEE Transactons on Systems, Man and Cybernetics*, **Part B 33(2)**, 332-339.

Booth, V., Erneux, T., and Laplante, J.P., 1995, Experimental and numerical study of weakly coupled bistale chemical reactors. *Journal of Physical Chemistry*, **98**, 6537-5640.

Chen,S., Billings, S.A., and Luo,W., 1989, Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, **50(5)**, 1873-1896.

Chua, L.O., and Shi, B., 1995, Autonomous cellular neural networks: a unified paradigm for pattern formation and active wave propagation, *IEEE Trans. Circuits and Systems I: Fundamental Theory and Applications*, **42**, 559-577.

Chua, L.O., 1998, *CNN: Paradigm for Complexity* (Singapore : Scientific World).

Chui,C. K.,1992, *An Introduction to Wavelets* (Boston : Academic Press).

Chui, C. K., and Wang, J. Z.,1992, On compactly supported spline wavelets and a duality principle. *Trans. of the American Mathematical Society.* **330(2)**, 903-915.

Coca, D., and Billings, S.A., 2001, Identification of coupled map lattice models of complex spatio-temporal patterns. *Physics Letters A*, **287**, 65-73.

Cross, M.C., and Hohenberg, P.C.,1993, Pattern-formation outside of equilibrium. *Review of Morden Physics*, **65**, 851-1112.

Crutchfield, J.P., and Kaneko, K., 1986, Phenomenology of spatio-temporal chaos. In B.H. Hao, (ed.), *Direction in Chaos* (Singapore : Scientific World).

Czaran, T., 1998, *Spatiotemporal Models of Population and Community Dynamics* (London: Chapman & Hall).

Decarvalho, A., Fairhurst, M.C., and Bisset, D.L., 1994, An integrated Boolean neural network for pattern classification, *Pattern Recognition*, **15**, 807-813.

Duarte, A.A., and Solari, H.G., 1997, Modelling the spatio-temporal dynamics of semiconductor lasers: the monochromatic solutions. *Optics Communications*, **144**, 99-108.

Ilachinski, A., 2001, *CellularAutomata: A Discrete Universe* (New Jersey : World Scientific).

Jahne, B., 1993, *Spatio-Temporal Image Processing : Theory and Scientific Applications* (New York; Berlin; London: Springer-Verlag).

Kaneko, K., 1984, Period-doubling of kink-antikink patterns, quasiperiodicity in antiferro-like structures and spatial intermittency in coupled logistic lattice: towards a prelude of a 'field theory of chaos'. *Progress of Theoretical Physics*, **72**, 480-486.

Kaneko, K., 1985, Spatiotemporal in intermittency in coupled map lattices. *Progress of Theoretical Physics*, 74, 1033-1044.

Kaneko, K., 1986, Turbulence in coupled map lattices. *Physics D*, **18**, 475-476.

Kaneko, K., 1989, Spatio-temporal chaos in one- and two-dimensional coupled map lattices. *Physica D*, **37**, 60-82.

Korenberg, M., Billings, S.A., Liu, Y. P., and McIlroy P.J.,1988, Orthogonal parameter estimation algorithm for non-linear stochastic systems. *International Journal of Control*, **48**, 193-210.

Leontaritis,I.J., and Billings, S.A.,1985, Input-output parametric models for non-linear systems, (part I: deterministic non-linear systems; part II: stochastic non-linear systems). *Int. Journal of Control*, **41**, 303-344.

Li, X.H., Bhide, S., and Kabuka, M.R., 1996, Labelling of MR brain images using Boolean neural network, *IEEE Trans. Medical Imaging*, **15**, 628-638.

Lopez, C., Alvarez, A., and Hernandez-Garcia, E.,2000, Forecasting confined spatiotemporal chaos with genetic algorithms. *Physical Review Letters*, **85**, 2300-2303.

Mallat,S.G., 1989, A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, **11(7)**, 674-693.

Mallat, S.G., and Zhang, Z., 1993, Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, **41(12)**, 3397-3415.

Mandelj, S., Grabec, I., and Govekar, E., 2001, Statistical approach to modelling of spatiotemporal dynamics. *International Journal of Bifurcation and Chaos*, **11**, 2731-2738.

Marcos-Nikolaus, P., Martin-Gonzalez, J.M., and Sole, R.V., 2002, Spatial forecasting : detecting determinism from single snapshots. *International Journal of Bifurcation and Chaos*, **12**, 369-376.

Parlizt, U., and Merkwirth, C., 2000, Prediction of spatiotemporal time series based on reconstructed local states. *Physical Review Letters*, **84**, 1890-1893.

Pearson, R.K., 1999, *Discrete-time Dynamic Models* (New York; Oxford: Oxford University Press).

Shih, C.W., 2000, Influence of boundary conditions on pattern formation and spatial chaos in lattice systems. *SIAM Journal on Applied Mathematics*, **61**, 335-368.

Silva, F.L., Principe, J.C., and Almeida, L.B., (ed.), 1997, *Spatiotemporal Models in biological and Artificial Systems* (Washington: IOS Press).

Sizt, A., Kurths, J., and Voss, H.U., 2003, Identification of nonlinear spatio-temporal systems via partitioned filtering. *Physical Review E*, **68**, 016202.

Sole, R.V., Valls, J., and Bascompte, J., 1992, Spiral waves, chaos and multiple attractors in lattice models of interacting populations, *Physics Letters A*, **166**, 123-128.

Sole, R.V., (ed.), 1998, *Modelling Spatiotemporal Dynamics in Ecology* (Berlin: Springer).

Torres-Sorando,L., and Rodriguez, D.J., 1997, Models of spatio-temporal dynamics in malaria. *Ecological Modelling*, **104**, 231-240.

Tziperman, E., Scher H, Zebiak., S.E., and Cane., M,A., 1997, Controlling spatio-temporal chaos in a realistic El Nino prediction model. *Physical Review Letters*, **79**, 1034-1037.

Von Neumann, J. 1951, The general logical theory of automata. In Jeffries, L.A.(ed.), *Cerebral Mechanisms in Behaviour – The Hixon Symposium* (New York: Wiley & Sons).

Waller, I., and Kapral, R., 1984, Spatial and temporal structure in systems of coupled, nonlinear oscillators. *Physicl Review A*, **30**, 2047-2055.

Wei, H.L., and Billings, S.A., 2002, Identification of time-varying systems using multi-resolution wavelet models, *International Journal of Systems Science*, **33**, 1217-1228.

Wei, H.L., and Billings, S.A., 2003, A unified wavelet-based modelling framework for nonlinear system identification: the WANARX model structure(submitted for publication).

Wolfram, S., 1983, Statistical mechanics of cellular automata. *Reviews of Modern Physics* **55(3)**, 601-644.

Wolfram, S., 1994, Cellular Automata and Complexity (New York: Addison-Wesley).

Yang, Y.X., 2000, Identification and analysis of a class of spatio-temporal systems. PhD Thesis, Department of Automatic Control and Systems Engineering, the University of Sheffield, UK.

Yang, Y.X., and Billings, S.A., 2000, Neighbourhood detection and rule selection from cellular automata patterns, *IEEE Transactons on Systems, Man and Cybernetics*, **Part A 30(6)**, 840-847.