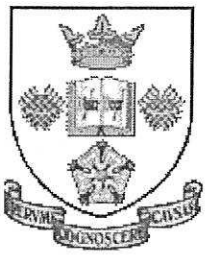
Monograph:
Guo, L.Z. and Billings, S.A. (2004) A Comparison of Polynomial and Wavelet Expansions for the Identification of Chaotic Coupled Map Lattices. Research Report. ACSE Research Report 860 . Department of Automatic Control and Systems Engineering, University of Sheffield.

A Comparison of Polynomial and Wavelet Expansions for the Identification of Chaotic Coupled Map Lattices

L. Z. Guo and S. A. Billings



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

Research Report No. 860
April 2004

A comparison of polynomial and wavelet expansions for the identification of chaotic coupled map lattices

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

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

Abstract

A comparison between polynomial and wavelet expansions for the identification of coupled map lattice (CML) models for deterministic spatio-temporal dynamical systems is presented in this paper. The pattern dynamics generated by smooth and non-smooth nonlinear maps in a well-known 2-dimensional CML structure are analysed. By using an orthogonal feedforward regression algorithm (OFR), polynomial and wavelet models are identified for the CML's in chaotic regimes. The quantitative dynamical invariants such as the largest Lyapunov exponents and correlation dimensions are estimated and used to evaluate the performance of the identified models.

1 Introduction

Complex spatio-temporal patterns have been widely observed and explored in recent years in many diverse fields including physical, chemical, biological, and ecological systems (Kaneko 1993, Sole, Valls and Bascompte 1992, Yanagita and Kaneko 1997, Tabuchi, Yakawa and Mallick et al. 2002, Kohler, Reinhard and Huth 2002, Bertram, Beta, Rotermund, and Ertl 2003, Goldman, et al. 2003, Adamatzky 2003). A large number of current studies of pattern formation phenomena involve observing what patterns are formed or changed under a variety of initial and boundary conditions. But an interesting and important question needs to be addressed: if an observed pattern formation follows some dynamical laws, then how can this dynamical behaviour be revealed effectively? In some instances, the dynamical origin of spatio-temporal pattern formation can be represented as a partial differential equation (PDE) or a coupled map lattice (CML). But in many other cases, such as for example in ecological systems, only a series of snapshots of the spatial pattern are available. At the same time, the study of the formation and evolution of spatio-temporal patterns normally requires a model with a specified accuracy. In both cases, however, obtaining or deriving such a dynamical model or PDE describing the pattern formation is by no means straightforward because either the interactions involved are too complex or there maybe no established laws on which to base the choice of the model. In this case, it would be advantageous if a model could be identified from the observed patterns. The identified model could then be used for the analysis of pattern formation or in control.

Various methods for the identification of local CML models from spatio-temporal observations have already been proposed (Coca and Billings 2001, Mandelj, Grabec and Govekar 2001, Marcos-Nikolaus, Martin-Gonzalez and Sole 2002, Grabec and Mandeji 1997, Parlitz and Merkwirth 2000, Billings, Wei, Mei, and Guo 2003, Billings, Guo, and Wei 2003), among which polynomial and wavelet methods have received more and more attention recently. In practice however, some of these approaches may fail to produce models that accurately describe the underlying spatio-temporal patterns either due to an inability to adapt the model structure to that of the unknown

system, or because the functions used to implement the model structure are not suitable for modelling the underlying dynamics. This is especially critical when an equivalent description of real-world systems is sought. In such cases the estimated model should provide very accurate information regarding the dynamical properties of the observed system. To evaluate the performance of different identification strategies and bring out the advantages and disadvantages of these, a comparison between polynomial and wavelet methods is conducted in this paper.

As a classical approach to the identification problem, polynomial expansions have been extensively studied and many good results have been obtained. Because of the arbitrary approximation properties to any sufficiently smooth function, polynomial methods have found wide applications in the field of smooth nonlinear approximation. On the other hand, recent theoretical studies have shown that the wavelet representation of any nonlinear function can be shown to be asymptotically near optimal in the sense that the convergence rates are equal to the best attainable using general nonlinear approximation schemes (DeVore, Jawerth, and Popov 1992). In addition wavelet approximations also provide similar rates of approximation for functions belonging to a wide variety of function spaces including functions with sparse singularities or functions that are not uniformly smooth or regular. All these properties suggest that wavelet multiresolution expansions should provide an excellent foundation for the development of identification algorithms for nonlinear CML models.

The paper is organised as follows. In section 2 a general input-output representation of CML models of spatio-temporal systems using polynomials and wavelets is derived. The identification algorithm is presented in section 3. In section 4 numerical simulation results and a detail comparison between the two methods is presented. Finally conclusions are drawn in section 5.

2 Parametric input-output representation of CML's using polynomials and wavelets

As a benchmark system, consider the following two-dimensional deterministic CML with symmetrical nearest neighbour coupling

$$x_{i,j}(t) = (1-\varepsilon) f(x_{i,j}(t-1)) + \frac{\varepsilon}{4}[f(x_{i,j-1}(t-1)) + f(x_{i,j+1}(t-1)) + f(x_{i-1,j}(t-1)) + f(x_{i+1,j}(t-1))] \quad (1)$$

where $x_{i,j}(t)$, $i, j = 1, 2, \cdots N$ is the state of the CML located at site (i, j) at discrete time instant t, ε is the coupling strength, N is the size of the lattice. The evolution of the CML on the lattice sites is governed by the local map f, which is generally a nonlinear function. The identification results using different local maps f will be generated and compared in later sections. Periodic boundary conditions are used throughout this study. Let $y_{i,j}$ is the observation variable of the CML at site (i, j). Then following the evolution of the CML (1), it is normally expected that the input-output behaviour of the CML (1) at the site (i, j) takes the following form

$$\begin{aligned} y_{i,j}(t) = g(&y_{i,j}(t-1), \cdots y_{i,j}(t-n_1), y_{i,j-1}(t-1), \cdots y_{i,j-1}(t-n_2), y_{i,j+1}(t-1), \cdots y_{i,j+1}(t-n_3), \\ &y_{i-1,j}(t-1), \cdots y_{i-1,j}(t-n_4), y_{i+1,j}(t-1), \cdots y_{i+1,j}(t-n_5)) \end{aligned} \quad (2)$$

where $y_{i,j-1}$, $y_{i,j+1}$, $y_{i-1,j}$, $y_{i+1,j}$ are the observation variables from the neighbouring sites, and n_1, n_2, n_3, n_4, and n_5 are the time lags for each observation variable.

Given observations $y_{i,j}(t-1),\cdots y_{i,j}(t-n_1), y_{i,j-1}(t-1),\cdots y_{i,j-1}(t-n_2), y_{i,j+1}(t-1),\cdots y_{i,j+1}(t-n_3)$, $y_{i-1,j}(t-1),\cdots y_{i-1,j}(t-n_4), y_{i+1,j}(t-1),\cdots y_{i+1,j}(t-n_5)$, the objective of CML identification is to approximate the input-output relationship function g from these observations. A practical solution is to approximate the unknown nonlinear function from the available observations using a known set of basis functions or regressors belonging to a given function class. Typical regressor classes include polynomials, spline functions, rational functions, radial basis functions, neural networks, and wavelets. In this paper, the algorithm and results for CML identification using polynomials and wavelets are presented and compared.

2.1 Approximation by polynomials

Let $\boldsymbol{\alpha} = (\alpha_1,\cdots,\alpha_n)$ be a multi-index, that is an n-tuple of nonnegative integers α_k, and denote by $\mathbf{x}^{\alpha}$ the monomial $x_1^{\alpha_1}\cdots x_n^{\alpha_n}$, which has degree $|\boldsymbol{\alpha}| = \sum_{k=1}^{n}\alpha_k$. Let s be a positive integer, and let $\Lambda=\{\alpha \mid |\alpha| \le s\}$ a set of multi-indices, then the set of polynomials of total order s is $\Sigma_s = \text{span}\{\mathbf{x}^{\alpha} \mid |\alpha| \le s\}$. Note that Σ_s is a L-dimensional space, where $L = 1+n+(n+1)n/2!+\cdots+(n+s-1)\cdots(n+1)n/s!$. For instance, there are 210 basis polynomial functions in the case that $n = 6$, $s = 4$. Approximating nonlinear function g in (2) using the polynomial approximation space Σ_s yields the following representation

$$y_{i,j}(t) = \theta_0 + \sum_{i_1=1}^{n}\theta_{i_1}x_{i_1}(t) + \sum_{i_1=1}^{n}\sum_{i_2=i_1}^{n}\theta_{i_1 i_2}x_{i_1}(t)x_{i_2}(t) + \cdots \sum_{i_1=1}^{n}\cdots\sum_{i_s=i_{s-1}}^{n}\theta_{i_1\cdots i_s}x_{i_1}(t)\cdots x_{i_s}(t) + e(t) \tag{3}$$

where $n = n_1+n_2+n_3+n_4+n_5$, all θ represent parameters and all $x(t)$ represent lagged terms in $y_{i,j}$, $y_{i,j-1}$, $y_{i,j+1}$, $y_{i-1,j}$, $y_{i+1,j}$, and $e(t)$ denotes the error of this approximate representation.

The approximation power of the class Σ_s can be described with the moduli of smoothness (Schumaker 1981). If an any given multivariate function $g \in L_p(\Omega)$, Ω is a convex set of $\mathbf{R}^n$, Λ is a complete set of multi-indices with boundary $\partial\Lambda$, and $1 \le q \le p \le \infty$, it is well known that for any $0 < \varepsilon < 1$, there exists a constant C such that

$$d(g,\Sigma_s)_q \le C\boldsymbol{\delta}^{1/q-1/p}\omega_{\partial\Lambda}(g;\boldsymbol{\delta})_p \tag{4}$$

provided p, q satisfy

$$\varepsilon \le \max\left\{\left\lfloor\frac{|\boldsymbol{\alpha}|}{n}\right\rfloor, \frac{1}{q}-\frac{1}{p}+\frac{|\boldsymbol{\alpha}|}{n}, \min(1-\frac{1}{p},\frac{1}{q})\right\}, \quad \boldsymbol{\alpha}\in\partial\Lambda \tag{5}$$

where $d(g,\Sigma_s)_q = \inf_{h\in\Sigma_s}\|g-h\|_{L_q(\Omega)}$ is the distance of g to Σ_s, $\boldsymbol{\delta} = (\delta_1,\cdots,\delta_n)$ with $\delta_i = \sup\{|y_i - x_i| \mid \mathbf{x}, \mathbf{y} \in \Omega\}$, and $\omega_{\partial\Lambda}(g;\boldsymbol{\delta})_p$ is the $\partial\Lambda$-modulus of smoothness in the p-norm of function $g\in L_p(\Omega)$.

2.2 Wavelet approximation

The wavelet decomposition of a multivariate function g defined on $\mathbf{R}^n$ can be described as follows. Let Φ be a bounded function defined on $\mathbf{R}^n$. For all $p \in \mathbf{Z}$ and $\mathbf{k} \in \mathbf{Z}^n$, a series of functions defined on $\mathbf{R}^n$ can be derived in terms of the translates and dyadic dilates of Φ: $\Phi_{p,\mathbf{k}}(\mathbf{x}) = \Phi(2^p\,\mathbf{x} - \mathbf{k})$. Then

if these functions $\Phi_{p,\mathbf{k}}$, $p \in \mathbf{Z}$ and $\mathbf{k} \in \mathbf{Z}^n$ form a Riesz basis, function g has a unique decomposition in terms of functions $\Phi_{p,\mathbf{k}}$

$$g(\mathbf{x}) = \sum_{p}\sum_{\mathbf{k}} \alpha_{p,\mathbf{k}} \Phi_{p,\mathbf{k}}(\mathbf{x}) \tag{6}$$

Such a Riesz basis in space $L^2(\mathbf{R}^n)$ can be constructed from the univariate scaling function φ and the associated wavelet function ψ in terms of the tensor product. The univariate scaling function considered in this paper is the m-th order cardinal B-spline function $\varphi(x) = \varphi^m(x) = \mathbf{B}^m(x)$ given by the recursive relation

$$\mathbf{B}^m(x) = \frac{x}{m-1}\mathbf{B}^{m-1}(x) + \frac{m-x}{m-1}\mathbf{B}^{m-1}(x-1) \tag{7}$$

where $\mathbf{B}^1(x)$ is the indicator function

$$\mathbf{B}^1(x) = \begin{cases} 1 & \text{if } x \in (0,1] \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The wavelet function is defined as a linear combination of scaling functions

$$\psi^m(x) = \sum_{l=0}^{3m-2} q_l \varphi^m(2x-l) \tag{9}$$

with the coefficients given by

$$q_l = \frac{(-1)^l}{2^{m-1}} \sum_{k=0}^{m} \binom{m}{k} \varphi^{2m}(l-k+1), l = 0,\cdots,3m-2 \tag{10}$$

If the nonlinear function g in eqn. (2) is in $L^2(\mathbf{R}^n)$. Then the B-spline wavelet representation of the input-output CML equation (2) can be described as follows

$$y_{i,j}(t) = \sum_{p}\sum_{\mathbf{k}}\sum_{l=1}^{2^n-1} \theta_{p,\mathbf{k},l} \Psi_{p,\mathbf{k}}^{(l)}(\mathbf{x}(t)) \tag{11}$$

where all θ represent parameters and $\mathbf{x}(t) = (x_1(t),\cdots,x_n(t))^T$ whose components represent lagged terms in $y_{i,j}$, $y_{i,j\text{-}1}$, $y_{i,j+1}$, $y_{i\text{-}1,j}$, $y_{i+1,j}$ as shown in (2), and $\Psi_{p,\mathbf{k}}^{(l)}(\mathbf{x})$ are the 2^n-1 n-dimensional wavelet functions produced by the tensor product of the univariate B-spline scaling and wavelet functions φ, ψ. According to the multiresolution analysis, eqn. (11) can equivalently be expressed as

$$y_{i,j}(t) = \sum_{\mathbf{k}} \theta_{p_0,\mathbf{k},0} \Phi_{p_0,\mathbf{k}}(\mathbf{x}(t)) + \sum_{p \geq p_0}\sum_{\mathbf{k}}\sum_{l=1}^{2^n-1} \theta_{p,\mathbf{k},l} \Psi_{p,\mathbf{k}}^{(l)}(\mathbf{x}(t)) \tag{12}$$

where p_0 is the starting resolution level.

The wavelet multiresolution approximation (12) is generally an infinite series expansion. In practice, however, it is not realistic to use all the terms in this infinite series expansion. Generally the objective of the identification algorithm is to obtain a truncated finite representation containing the terms up to some orders of scaling and dilation. Therefore the identified CML model will be an approximate representation of the underlying system, which can be equivalently described as an infinite wavelet series. Let s be a positive integer, the s-truncated space Σ_{s,p_0} with a starting resolution p_0 is the set of all functions

$$h(\mathbf{x}) = \sum_{\mathbf{k}} \theta_{p_0,\mathbf{k},0} \Phi_{p_0,\mathbf{k}}(\mathbf{x}) + \sum_{p_0 \le p \le s} \sum_{\mathbf{k}} \sum_{l=1}^{2^n-1} \theta_{p,\mathbf{k},l} \Psi^{(l)}_{p,\mathbf{k}}(\mathbf{x}) \tag{13}$$

Note that the series in space Σ_{s,p_0} are those up to dyadic level s, which may possibly be infinite because there is no limitation on the translation operation. In practice, the range of measured data is always finite so that there are only finite numbers of translation operations which produce non-empty intersections within the range of the data. Therefore, the identified wavelet series are always finite. Furthermore, in many applications, a 3-truncated space is often enough to obtain a good approximation result because wavelets with higher dyadic levels are most likely to have compact support which contains no data points. Using the approximation space Σ_{s,p_0} as a regressor class, a truncated approximation representation of (12) takes the form

$$y_{i,j}(t) = \sum_{\mathbf{k}} \theta_{p_0,\mathbf{k},0} \Phi_{p_0,\mathbf{k}}(\mathbf{x}(t)) + \sum_{p_0 \le p \le s} \sum_{\mathbf{k}} \sum_{l=1}^{2^n-1} \theta_{p,\mathbf{k},l} \Psi^{(l)}_{p,\mathbf{k}}(\mathbf{x}(t)) + e(t) \tag{14}$$

where $e(t)$ is the truncation error.

Note that the Fourier transform of the univariate B-spline scaling function $\varphi(x) = \varphi^m(x) = \mathbf{B}^m(x)$ is

$$\widetilde{B}^m(x) = \left(\frac{1 - e^{-i\omega}}{i\omega} \right)^m \tag{15}$$

by the Strang-Fix condition and Poisson's summation formula the polynomials 1, x, ..., x^{m-1} are linear combinations of the univariate translates $\varphi(x-k)$. It follows that the space Σ_{s,p_0} of multivariate wavelets contains the space Σ_{mn} of polynomials of total degree $< mn$, and this implies that

$$d(g, \Sigma_{s,p_0})_p = \inf_{h \in \Sigma_{s,p_0}} \|g - h\|_p \le d(g, \Sigma_{mn})_p = \inf_{h \in \Sigma_{mn}} \|g - h\|_p \tag{16}$$

More detailed discussions about wavelets and wavelet approximation can be found in (Chui 1992, Sweldens and Piessens 1994, and DeVore, Jawerth, and Popov 1992).

2.3 An alternative wavelet representation of the input-output relationship of CML's

For simplicity, let the nonlinear function to be identified be defined on the cube $[0, 1]^n$, and consider the number of wavelet terms in the basis in the space Σ_{s,p_0}. Let $\varphi(x)$ and $\psi(x)$ be the

univariate scaling and wavelet functions of B-spline functions of order m. Then the support of $\varphi(x)$ and its dilates and translates $\varphi_{p,k}(x) = 2^{p/2}\varphi(2^p x\text{-}k)$ are $[0, m]$ and $[2^{-p} k, 2^{-p} (m+k)]$, and the support of $\psi(x)$ and its dilates and translates $\psi_{p,k}(x) = 2^{p/2}\psi(2^p x\text{-}k)$ are $[0, 2m\text{-}1]$ and $[2^{-p} k, 2^{-p} (2m\text{-}1+k)]$. Assume that the domain of nonlinear functions to be identified in one component is $[0, 1]$ it is then sufficient that the translate parameter k for univariate scaling and wavelet functions falls into the intervals

$$\begin{aligned} \varphi: &\quad -m+1 \le k \le 2^p - 1 \\ \psi: &\quad -2(m-1) \le k \le 2^p - 1 \end{aligned} \tag{17}$$

It follows that the total number of terms in the basis of the space Σ_{s,p_0} for n-dimensional functions defined on the cube $[0, 1]^n$ is $\sum_{p=p_0}^{s} n_p^n, \quad n_p = 2^{p+1} + 3(m-1)$. For instance, if $n = 6$, $m = 4$, $p_0 = 0$, $s = 1$ which means the dimension of the approximated function is 6 with B-spline scaling and wavelet functions of order 4, starting scaling level 0 and the truncation 3, then the total number of the terms in the basis of the space Σ_{s,j_0} is 6,598,370, which is clearly a time-consuming number for any identification algorithm. To overcome this difficulty, in this paper the identified nonlinear function g is first decomposed into a number of functional components as follows

$$f(x_1, \cdots x_n) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1 \le i < j \le n} f_{ij}(x_i, x_j) + \sum_{1 \le i < j < k \le n} f_{ijk}(x_i, x_j, x_k) + \cdots f_{1\cdots n}(x_1, \cdots, x_n) \tag{18}$$

where f_0 is a constant. A truncated representation of (18) containing the functional components up to tri-variate terms is often sufficient to express a nonlinear function itself. Applying the above wavelet decomposition to each of the functional components significantly reduces the terms using for identification. Assume that a multivariate function is defined on the cube $[0, 1]^n$ again. Consider the l-variate functional components, the total number of significant terms (the intersection of its support and $[0, 1]^n$ is non-empty) can be calculated according to the following formula

$$\binom{n}{l} \sum_{p=p_0}^{s} n_p^l, \quad n_p = 2^{p+1} + 3(m-1) \tag{19}$$

Now consider the same example above with maximal functional components of the tri-variate. If $p_0 = 0$, and $s = 2$, $s = 1$ and $s = 0$ for uni, bi and tri-variate components, this yields a total of 31,145 terms and one constant term, 174 univariate terms, 4,350 bi-variate terms and 26,620 tri-varaite terms. This is a significant reduction compared to 6,598,370.

3 The parameter estimation algorithm

Given a set (candidate terms) of basis functions from a regressor class, either polynomials or wavelets, the objective of an identification algorithm is to select the significant terms from this set while estimating the corresponding parameters. In this paper, an Orthogonal Forward Regression algorithm (OFR) (Chen, Billings, and Luo 1989) is applied to a set of either polynomial or wavelet basis functions. The OFR algorithm involves a stepwise orthogonalisation of the regressors and a forward selection of the relevant terms based on the Error Reduction Ratio criterion (Billings, Chen, and Kronenberg 1988). The algorithm provides the optimal least-squares estimate of the parameters θ.

For a given candidate regressor set $G = \{\phi_i\}_{i=1}^{M}$, the OFR algorithm can be outlined as follows

- Step 1

$$I_1 = I_M = \{1, \cdots, M\}$$

$$w_i(t) = \phi_i(t), \qquad \hat{b}_i = \frac{w_i^T y}{w_i^T w_i} \tag{20}$$

$$l_1 = \arg\max_{i \in I_1} \left(\hat{b}_i \frac{w_i^T y}{w_i^T w_i} \right) = \arg\max_{i \in I_1} (err_i) \tag{21}$$

$$w_1^0 = w_{l_1}, \qquad c_1^0 = \frac{w_1^{0T} y}{w_1^{0T} w_1^0} \tag{22}$$

$$\alpha_{1,1} = 1 \tag{23}$$

- Step $j, j > 1$

$$I_j = I_{j-1} \setminus \{l_{j-1}\} \tag{24}$$

$$w_i(t) = \phi_i(t) - \sum_{k=1}^{j-1} \frac{w_k^{0T} y}{w_k^{0T} w_k^0} w_k^0, \qquad \hat{b}_i = \frac{w_i^T y}{w_i^T w_i} \tag{25}$$

$$l_j = \arg\max_{i \in I_j} \left(\hat{b}_i \frac{w_i^T y}{w_i^T w_i} \right) = \arg\max_{i \in I_j} (err_i) \tag{26}$$

$$w_j^0 = w_{l_j}, \qquad c_j^0 = \frac{w_j^{0T} y}{w_j^{0T} w_j^0} \tag{27}$$

$$\alpha_{k,j} = \frac{w_k^{0T} \phi_{l_j}}{w_k^{0T} w_k^0}, \quad k = 1, \cdots, j-1 \tag{28}$$

The procedure is terminated at the M_s step when the termination criterion

$$1 - \sum_{i=1}^{M_s} err_i < \rho \tag{29}$$

is met, where ρ is a designated error tolerance, or when a given number of terms in the finial model is reached.

The estimated parameters are calculated from the following equation

$$\begin{pmatrix} \theta_{l_1} \\ \theta_{l_2} \\ \vdots \\ \theta_{l_{M_s}} \end{pmatrix} = \begin{pmatrix} 1 & \alpha_{1,2} & \cdots & \alpha_{1,M_s} \\ 0 & 1 & \vdots & \alpha_{2,M_s} \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} c_1^0 \\ c_2^0 \\ \vdots \\ c_{M_s}^0 \end{pmatrix} \tag{30}$$

and the selected terms are $\phi_{l_1}, \phi_{l_2}, \cdots, \phi_{l_{M_s}}$.

4 Numerical results

4.1 Case 1: f is a smooth nonlinear map

First, consider the two-dimensional CML defined by (1) with the nonlinear function f chosen to be the logistic map

$$f(x) = 1 - ax^2 \tag{31}$$

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 1989).

In order to analyse and compare the capabilities of identification methods using polynomials and wavelets, the model (1) with (31) was simulated for a lattice of the size 50×50 with random initial conditions, periodic boundary conditions, and parameters ε = 0.4, a = 1.55. The observation variable was set to be $y_{i,j}(t) = x_{i,j}(t)$. All data were normalised to the interval [0, 1]. Some snapshot patterns are shown in Fig. 1. With these parameters, the system is actually in a chaotic regime with Lyapunov exponents λ_1 = 0.0648, λ_2 = 0.0622, λ_3 = 0.0158, λ_4 = -0.0014, λ_5 = -0.0106, λ_6 = -0.0275, λ_7 = -0.0478, λ_8 = -0.0811, λ_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, which is just the sum of all positive Lyapunov exponents. 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. These results are illustrated in Fig. 5.

In the identification, the same set of 100 observation pairs randomly selected among the data set were used for both the polynomial and wavelet methods. The neighbourhood was set to be the nearest four sites, that is, $(i, j\text{-}1)$, $(i, j\text{+}1)$, $(i\text{-}1, j)$, and $(i\text{+}1, j)$ and the time lag was set to be 1. For the polynomial identification, the maximal order of polynomial terms was set to be 3. This implies that for the polynomial identification, the set of the candidate regressors is $\Sigma_3 = \text{span}\{\mathbf{x}^\alpha \mid |\alpha| \le 3\}$ and the total number of candidate terms is 56. For the wavelet method, the time lag was set to be 1 and the initial wavelet model structure was chosen as

$$y_{i,j}(t) = f(x_1, x_2, \cdots, x_5) = f_0 + \sum_{i=1}^{5} f_i(x_i) + \sum_{1 \le i < j \le 5} f_{ij}(x_i, x_j) + \sum_{1 \le i < j < k \le 5} f_{ijk}(x_i, x_j, x_k) \tag{32}$$

where $x_1, \cdots, x_5$ represent $y_{ij}(t\text{-}1)$, $y_{ij\text{-}1}(t\text{-}1)$, $y_{ij\text{+}1}(t\text{-}1)$, $y_{i\text{-}1,j}(t\text{-}1)$, $y_{i\text{+}1,j}(t\text{-}1)$ and f_0 is a constant term. The starting resolution scale was set to be 0 for all three submodels and the maximal scales were set to be 2, 1, and 0 for uni, bi, and tri-variates, respectively. The univariate B-spline function of order

3 was used to generate all the higher-dimensional terms by tensor products. It follows that the total number of terms in the set of candidate model terms is 6871. For both methods, the maximal number of selected terms in the OFR selection algorithm was set to be 10 and the tolerance ρ was chosen as 10^{-3} that means if $1-\sum_{i=1}^{M_s} err_i < 10^{-3}$, the algorithm will terminate.

After applying the OFR algorithm, a 7-term polynomial model and a 10-term wavelet model were identified. These are listed in Tables 1 and 2, respectively, where ERR denotes the Error Reduction Ratio and STD denotes the standard deviations. The model predictive outputs from the two identified models are shown in Fig. 2 and Fig. 3. The model predictive errors at time instant 100 for the two models are shown in Fig. 4. By using Rosenstein's method to the data from site (25, 25), a positve value of $\lambda_1 = 0.0635$ and an estimated correlation dimension 0.494 for the polynomial model were found while $\lambda_1 = 0.0599$ and $C_m(r) = 0.465$ for the wavelet model. These quantities are listed in Table 3. To test the compression abilities of the two methods, the terms in the final models, whose coefficients (absolute values) are less than 0.01, were removed from the final models. The resulting largest Lyapunov exponents and correlation dimensions for the reduced models are also included in Table 3.

The identification results clearly show that both methods can provide satisfactory prediction performance for this specific smooth nonlinear CML. Both estimated largest Lyapunov exponents and correlation dimensions are quite close to the values calculated using the correct model. In this example, the polynomial model is slightly better than the wavelet model. Moreover, the absolute errors which are shown in Fig. 4 indicate that the polynomial model is more accurate than the wavelet model. However for the reduced models, these invariant quantities indicate the wavelet model is more robust than the polynomial model.

4.2 Case 2: f is a nonsmooth map

Now, consider the two-dimensional CML defined by (1) with the nonlinear function f chosen to be the following piecewise linear map (Miller and Huse 1993)

$$f(x) = \begin{cases} -3x-2, & -1 \le x \le -1/3 \\ 3x, & -1/3 \le x \le 1/3 \\ -3x+2, & 1/3 \le x \le 1 \end{cases} \tag{33}$$

According to Miller and Huse (1993), the CML dynamics are chaotic and ergodic for $\varepsilon = 0$. Miller and Huse found that the CML has a ferromagnetically ordered steady state for $0.8216 \le \varepsilon \le 0.96$. Moreover, this CML is chaotic for couplings in both the paramagnetic and ferromagnetic regimes. For the purpose of identification, the model (1) with (33) was simulated for a lattice of the size 32×32 with random initial conditions within [-1 1], periodic boundary conditions, and parameters ε = 0.8920. The observation variable was set to be $y_{i,j}(t) = x_{i,j}(t)$. All data were normalised to the interval [0, 1]. Some snapshot patterns are shown in Fig. 6. The largest Lyapunov exponent λ_1 = 0.0819 and an estimated correlation dimension 1.0 were obtained by using Rosenstein's method to the data from site (15, 15). Fig. 10 shows the estimated largest Lyapunov exponents and correlation dimensions with embedding dimensions 1 to 9.

Given the same settings as in case 1, after applying the OFR algorithm, a 10-term polynomial model and 10-term wavelet model were identified. These are listed in Tables 4 and 5, respectively.

The model predictive outputs from the two identified models are shown in Fig. 7 and Fig. 8. The model predictive errors at time instant 100 for the two models are shown in Fig. 9. By using Rosenstein's method to the data from site (15, 15), a positive value of $\lambda_1 = 0.0901$ and $C_m(r) = 1.0$ for the wavelet model were obtained. These quantities are listed in Table 6. Note that the final polynomial model cannot provide a prediction correctly so these two quantities cannot be calculated properly. In Fig. 7, the dark area indicates where no finite values can be calculated. The identification results show that in this example the polynomial method cannot provide a reasonable model for this non-smooth nonlinear CML while wavelet method does provide a satisfactory prediction performance. Note that both estimated largest Lyapunov exponents and correlation dimensions for this final wavelet model are quite close to that calculated by using the simulated data. This indicates that the wavelet method is more applicable than the polynomial method in this case.

5 Conclusions

A comparison between polynomial and wavelet identification methods for chaotic CML's has been conducted. The largest Lyapunov exponents and correlation dimensions have been estimated to validate the obtained models. The results show that for CML's with a smooth nonlinear map, both methods can provide satisfactory performance but the polynomial method is preferred because of the smaller number of candidate terms (56 v.s 6981 in the wavelet method using our settings) and good predictive ability. However, for CML's with a non-smooth nonlinear map, the wavelet method can still provide a good predictive performance but the polynomial method does not. This advantage of the wavelet method means that this approach is much more applicable in the case where the underlying dynamics are totally unknown.

Acknowledgement

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

References

Adamatzky, A., (2003) On patterns in affective media, *International Journal of Modern Physics* C, Vol. 14, No. 5.

Bertram, M., Beta, C., Pollmann, M., Mikhailov, A. S., Rotermund, H. H., and Ertl, G., (2003) Pattern formation on the edge of chaos: experiments with CO oxidation on a Pt(110) surface under global delayed feedback, *Phys. Rev.*, E67, No. 3, 036208.

Billings, S. A., Chen, S., and Kronenberg, M. J., (1988) Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator, *Int. J. Contr.*, Vol. 49, pp. 2157-2189.

Billings, S. A. and Coca, D., (2002) Identification of coupled map lattice models of deterministic distributed parameter systems, *Int. J. Systems Science*, Vol. 33, pp. 623-634.

Billings, S. A., Guo, L. Z., and Wei, H. L., (2003) Identification of coupled map lattice models for spatio-temporal patterns using wavelets, *Submitted for publication.*

Billings, S. A., Wei, H. L., Mei, S. S., and Guo, L. Z., (2003) Identification of spatio-temporal systems using multiresolution wavelet models, *Submitted for publication.*

Chen, S., Billings, S. A., and Luo, W., (1989) Orthogonal least squares methods and their application to non-linear system identification, *Int. J. Contr.*, Vol. 50, No. 5, pp. 1873-1896.

Chui, C. K., (1992) *An introduction to wavelets*, Academic Press, Inc., London.

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

DeVore, R. A., Jawerth, B., and Popov, V., (1992) Compression of wavelet decompositions, *American Journal of Mathematics*, Vol.114, pp.737-785.

Goldman, D. I., Shattuck, M. D., Sung, J. M., Swift, J. B., and Swinney, H. L., (2003) Lattice dynamics and melting of a nonequilibrium pattern, *Phys. Rev. Lett.*, Vol. 90, No. 10, 104302.

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.

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

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

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

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

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

Miller, J. and Huse, D. A., (1993) Macroscopic equilibrium from microscopic irreversibility in a chaotic coupled-map lattice, *Physical Review E*, Vol. 48, No. 4, pp. 2528-2535.

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

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.

Schumaker, L. L., (1981), *Spline functions: basic theory*, John Wiley & Sons, New York.

Sole, 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. 123-128.

Sweldens, W. and Piessens, R., (1994) Asymptotic error expansion of wavelet approximations of smooth function II, *Numerische Mathematik*, Vol. 68, No. 3, pp. 377-401.

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.

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

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 1 Case1: The terms and parameters for the estimated polynomial model

Terms	Estimates	ERR	STD
constant	9.5801e-01	8.1197e-01	2.9845e-01
$\varphi_{0,0}(y_{i,j}(t-1))$	-7.9769e-01	1.7574e-01	7.6291e-02
$\psi_{0,0}(y_{i,j}(t-1))$	1.3328e+00	6.7804e-03	5.1070e-02
$\varphi_{1,1}(y_{i,j}(t-1))$	1.0371e-01	1.7499e-03	4.2180e-02
$\varphi_{0,0}(y_{i,j-1}(t-1))\varphi_{0,0}(y_{i,j+1}(t-1))\varphi_{0,0}(y_{i-1,j}(t-1))$	-3.0985e-02	1.1610e-03	3.5059e-02
$\varphi_{1,0}(y_{i+1,j}(t-1))$	-1.8126e-01	6.0660e-04	3.0688e-02
$\psi_{2,3}(y_{i,j}(t-1))$	-2.5606e+00	5.8245e-04	2.5804e-02
$\psi_{1,0}(y_{i,j+1}(t-1))\psi_{1,1}(y_{i-1,j}(t-1))$	-8.6797e-02	1.7778e-04	2.4117e-02
$\varphi_{0,0}(y_{i,j-1}(t-1))$	-1.8009e-01	1.4308e-04	2.2668e-02
$\varphi_{0,0}(y_{i,j+1}(t-1))\varphi_{0,0}(y_{i-1,j}(t-1))$	-5.6931e-01	5.0346e-04	1.6593e-02

Table 2 Case1: The terms and parameters for the estimated wavelet model

Model	Total number of initial terms	Number of selected terms	Lyapunove Exponent		Correlation Dimension
			Jacobian	Rosenstein	
CML model (1)	×	×	0.0648	0.0644	0.495
Polynomials	56	7	×	0.0635	0.494
Wavelets	6871	10	×	0.0599	0.465
Reduced polynomials	×	5	×	0.0143	0.495
Reduced wavelets	×	8	×	0.0329	0.495

Table 3 Case1: A quantitative comparison of the polynomial and wavelet models

Terms	Estimates	ERR	STD
$y_{i-1,j}(t-1)$	4.1211e-01	9.8575e-01	5.6166e-02
$y_{i,j+1}(t-1)$	-8.6125e-02	5.5832e-03	4.3913e-02
constant	4.6978e-01	1.0987e-03	4.1075e-02
$y_{i,j+1}(t-1)^2 y_{i+1,j}(t-1)$	1.6728e+00	7.4366e-04	3.9006e-02
$y_{i,j}(t-1)^2 y_{i-1,j}(t-1)$	2.3264e+00	2.1194e-04	3.8396e-02
$y_{i-1,j}(t-1)^2 y_{i+1,j}(t-1)$	-3.1008e-01	8.5450e-05	3.8147e-02
$y_{i,j-1}(t-1)y_{i+1,j}(t-1)^2$	2.2322e+00	1.2704e-04	3.7774e-02
$y_{i,j}(t-1)y_{i,j-1}(t-1)y_{i+1,j}(t-1)$	-1.5469e+00	2.1955e-04	3.7121e-02
$y_{i+1,j}(t-1)$	-7.5232e-01	7.7395e-05	3.6888e-02
$y_{i,j}(t-1)y_{i,j-1}(t-1)y_{i,j+1}(t-1)$	-2.6006e+00	1.8818e-04	3.6315e-02

Table 4 Case2: The terms and parameters for the estimated polynomial model

Terms	Estimates	ERR	STD
constant	5.4427e-01	9.8461e-01	5.8550e-02
$\psi_{2,-1}(y_{i-1,j}(t-1))$	-1.2475e-01	8.5885e-03	3.8915e-02
$\psi_{2,-1}(y_{i+1,j}(t-1))$	-9.9230e-02	3.0578e-03	2.8863e-02
$\psi_{2,0}(y_{i,j+1}(t-1))$	-4.3860e-02	9.5855e-04	2.4890e-02
$\psi_{2,0}(y_{i,j-1}(t-1)))$	-3.5910e-02	5.6903e-04	2.2197e-02
$\varphi_{0,0}(y_{i,j+1}(t-1))\varphi_{0,0}(y_{i-1,j}(t-1))\varphi_{0,0}(y_{i+1,j}(t-1))$	-1.6474e+01	7.5748e-04	1.7998e-02
$\psi_{0,0}(y_{i-1,j}(t-1))$	-9.7201e-01	2.1340e-04	1.6624e-02
$\psi_{2,0}(y_{i,j}(t-1))$	-1.7811e-02	1.4834e-04	1.5599e-02
$\psi_{1,-1}(y_{i,j}(t-1))\psi_{1,-1}(y_{i,j-1}(t-1))$	-3.8368e+02	1.6433e-04	1.4377e-02

Table 5 Case2: The terms and parameters for the estimated wavelet model

Model	Total number of initial terms	Number of selected terms	Lyapunove Exponent		Correlation Dimension
			Jacobian	Rosenstein	
CML model (1)	×	×	×	0.0819	1.0
Polynomials	56	10	×	×	×
Wavelets	6871	10	×	0.0907	1.0

Table 6 Case2: A quantitative comparison of the polynomial and wavelet models

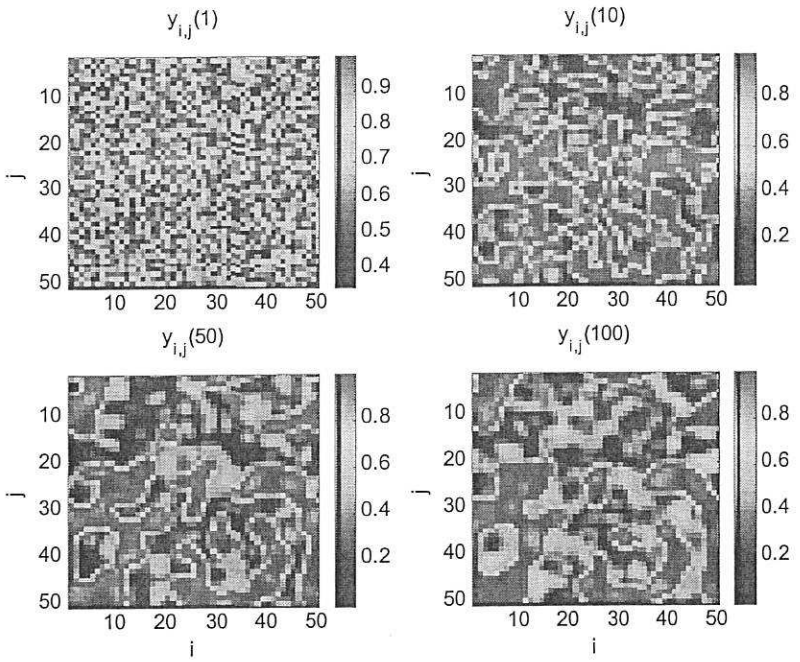


Figure 1 Case 1: Some snapshots (at t = 1, 10, 50, and 100) from simulated data

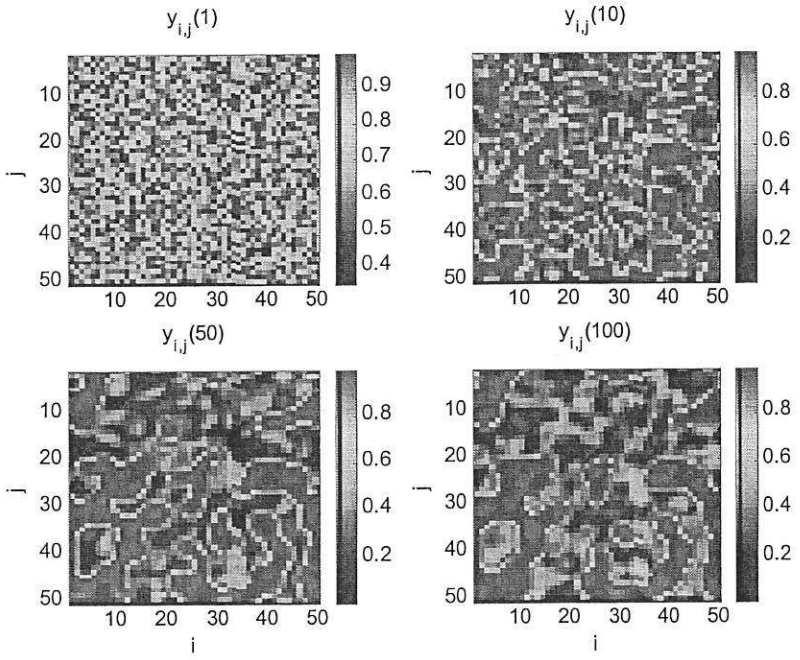


Figure 2 Case 1: Snapshots (at t = 1, 10, 50, and 100) from the estimated polynomial model predictive output

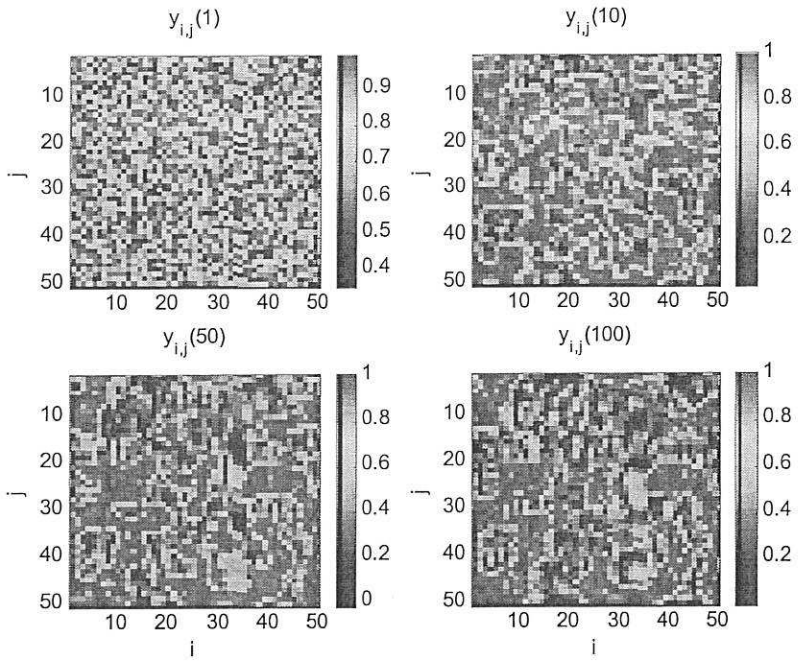


Figure 3 Case 1: Snapshots (at $t = 1$, 10, 50, and 100) from the estimated wavelet model predictive output

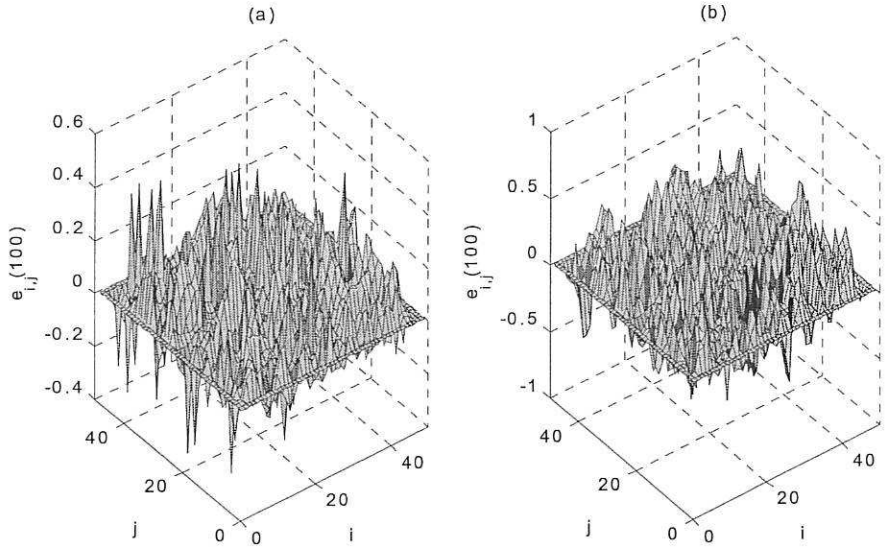


Figure 4 Case 1: Model predictive errors (at $t = 100$) for (a) polynomial and (b) wavelet models

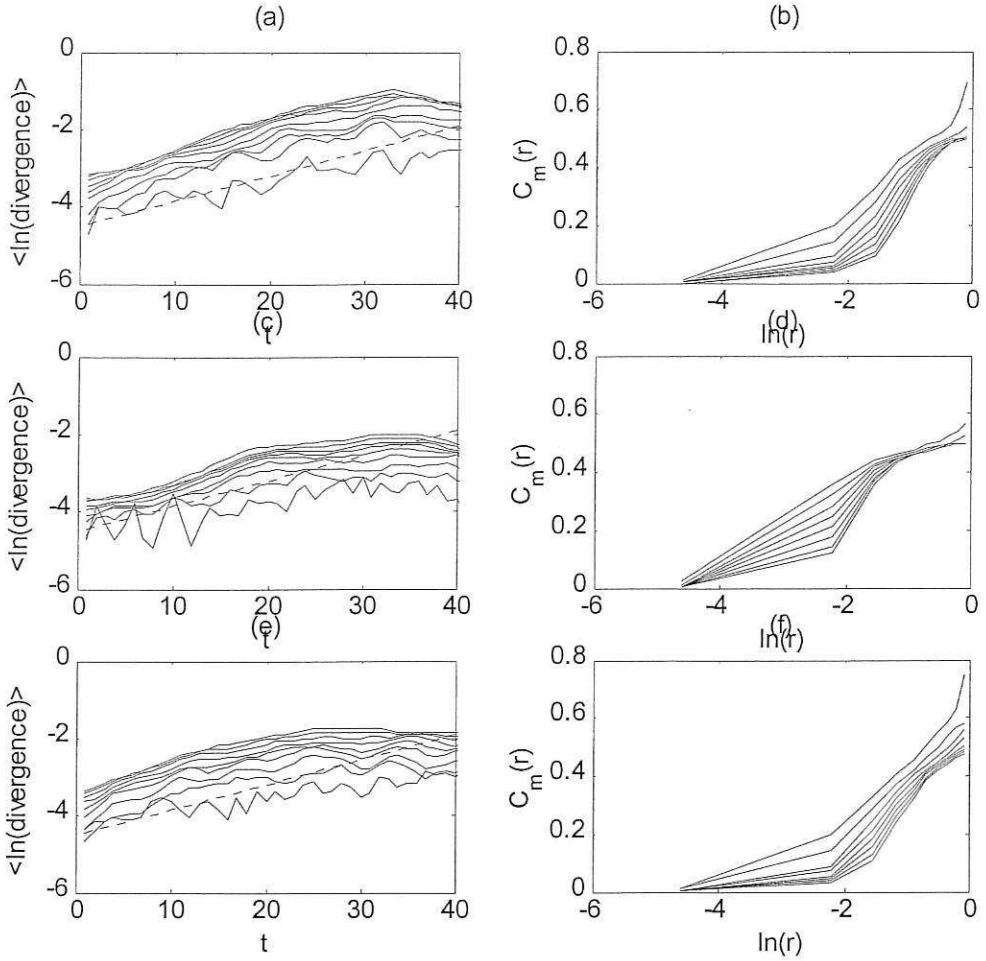


Figure 5 Case1: Lyapunov exponents and correlation dimensions (*a-b*) CML model, (*c-d*) polynomial and (*e-f*) wavelet models for embedding dimensions 1 to 9

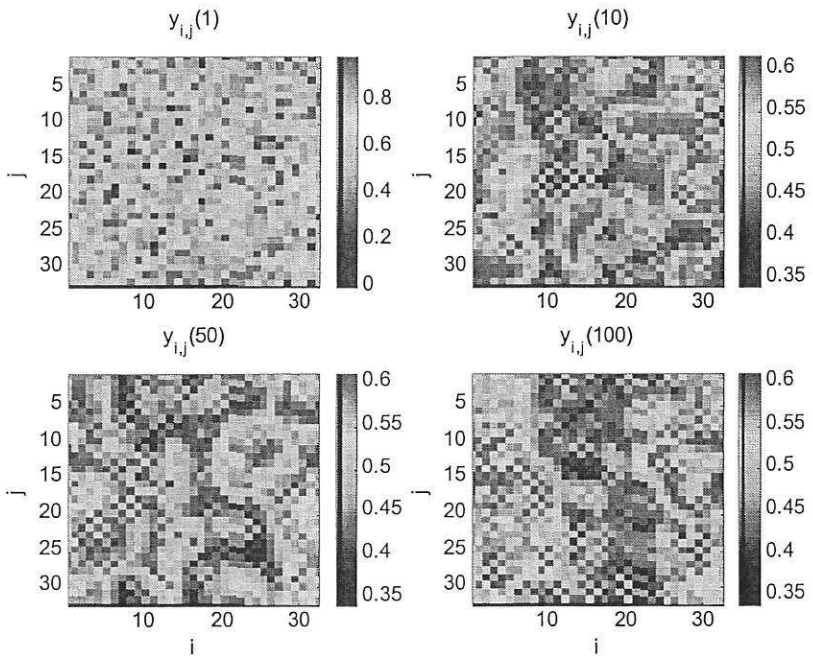


Figure 6 Case 2: Some snapshots (at t = 1, 10, 50, and 100) from simulated data

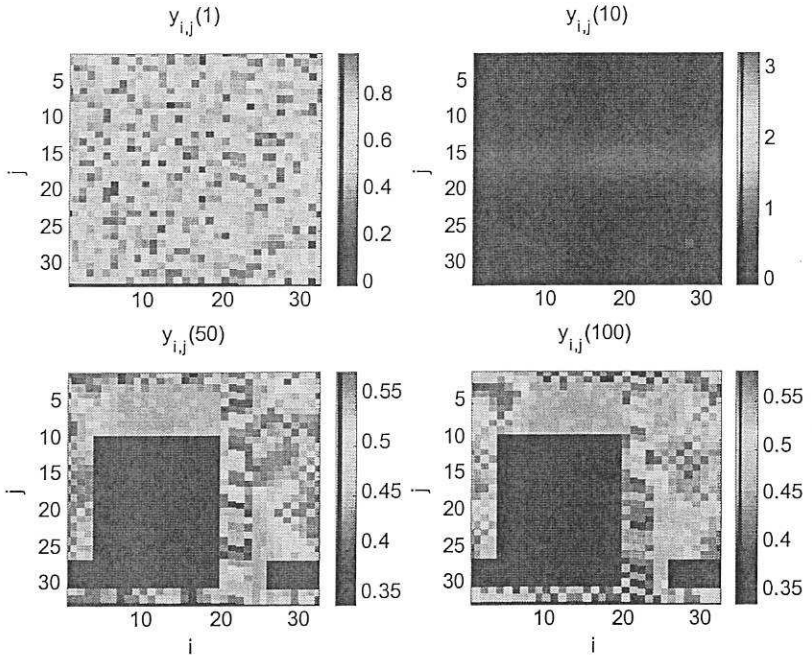


Figure 7 Case 2: Snapshots (at $t = 1$, 10, 50, and 100) from the estimated polynomial model predictive output

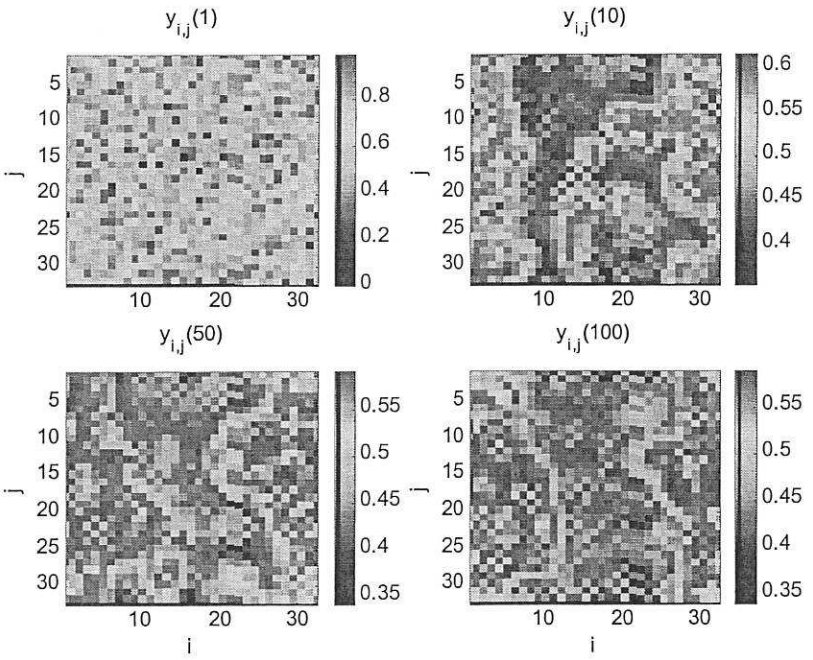


Figure 8 Case 2: Snapshots (at $t = 1$, 10, 50, and 100) from the estimated wavelet model predictive output

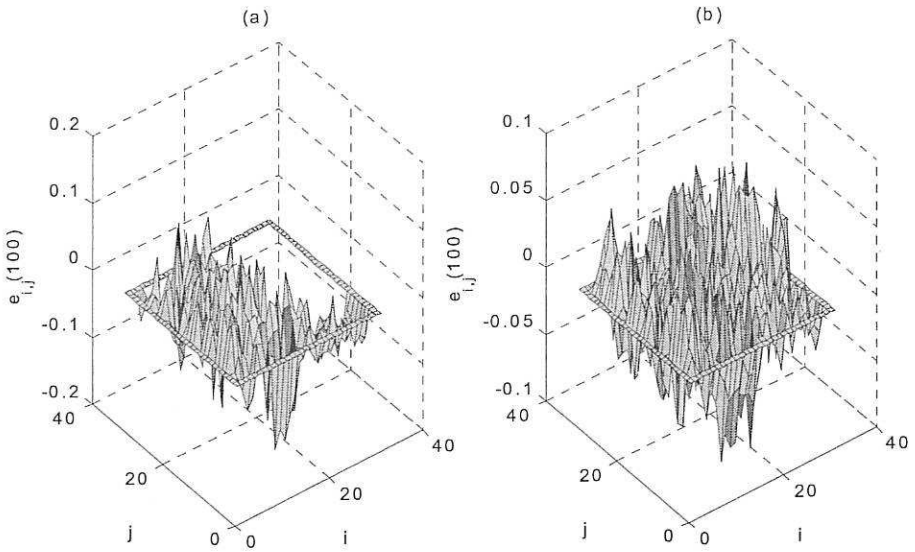


Figure 9 Case 2: Model predictive errors for (a) polynomial and (b) wavelets

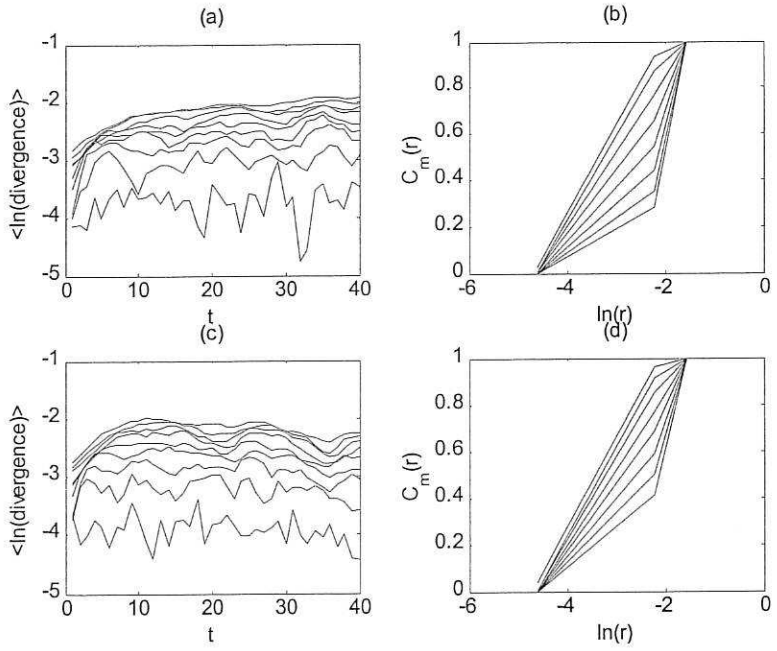


Figure 10 Case 2: Lyapunov exponents and correlation dimensions (*a-b*) CML model, and (*c-d*) wavelets for embedding dimensions 1 to 9