Neighborhood Detection for the Identification of Spatiotemporal Systems

Y. Pan and S. A. Billings

***Abstract*—Neighborhood detection and local state vector construction for the identification of spatiotemporal systems is considered in this paper. Determining the neighborhood size both in the space and time domain can considerably reduce the complexity of the set of candidate model terms for the identification of coupled map lattice models. The computation requirements of the model identification algorithm can also be greatly reduced instead of the more direct identification approach of searching over the entire spatiotemporal neighborhood in the original space. In this paper, a new neighborhood detection method is introduced based on embedding theory for nonlinear dynamical systems to produce an initial spatiotemporal neighborhood for the identification of spatiotemporal systems. Numerical examples are provided to demonstrate the feasibility and applicability of the new neighborhood detection method.**

***Index Terms*—Neighborhood detection, spatiotemporal systems, system identification.**

I. Introduction

RECENTLY, various methods have been proposed for the identification of coupled map lattice (CML) models from spatiotemporal observations [1]–[4]. As a preliminary step in all of these identification methods, neighborhood detection, whose objective is to find the significant neighbors in both space and time domain, is a task frequently encountered during the modeling and prediction of spatiotemporal systems because of the lack of *a priori* knowledge about the observed system. Neighborhood detection involves the reconstruction of the local state vectors at some specified sites from the measured data, i.e., determining the spatiotemporal neighboring region or sites which influence the dynamics of that site. The determination of this spatiotemporal region plays a critical role in finding an appropriate representation of the underlying spatiotemporal dynamics. By determining the neighborhood size for the identification of spatiotemporal systems, dimensions of systematic variables in the spatiotemporal model can be obtained. Then, the candidate model terms of the CML model can be properly prespecified, and unnecessary computational costs can also be avoided.

It is well known that the neighborhood detection problem is closely related to the embedding problem. There is a large body of work devoted to finding the minimum embedding dimensions and the reconstruction of state-space vectors of time-series data [5]–[13]. However, the reconstruction problem of local state vectors for the identification of continuous spatiotemporal systems has received little attention. In [1], [14], and [15], the spatiotemporal neighborhood of the CML model was roughly prespecified with some finite radius in the space and time domains based on empirical information. The accurate neighborhood was finally searched from a wide spatial and temporal extent using a model selection method. Similarly, in [2] and [4], the state vector was reconstructed from a set of neighboring values in a rectangular or triangular spatiotemporal region for some simple classes of spatiotemporal systems. The reconstruction problem of spatiotemporal state space was first addressed in [3], which showed that the model prediction capability could be greatly improved by properly choosing the embedding dimensions. In [3], the determination of the embedding dimensions was indirectly implemented through the simulation of the cross-prediction of the fitted model with different embedding dimensions. Recently, the neighborhood detection problem for cellular automata modeling in a binary space has been studied in [16] and [17]. However, the application of these neighborhood-selection methods to spatiotemporal data with continuous-state value is not so promising. For example, in [17], the neighborhood was selected on the basis of the estimation of the mutual information associated with different reconstructed state-space vectors. But, for the continuous state space, to obtain an accurate estimation of the mutual information is always challenging and often involves subjective judgements [18]. Guo *et al.* [19] proposed a neighborhood detection method for continuous spatiotemporal systems by converting the real-valued pattern to a binary pattern via thresholding and then applying cellular automata neighborhood detection methods to produce an initial neighborhood to prime the CML modeling. Apparently, such neighborhood-selection results could be very coarse, since some of the dynamical information might be lost during the thresholding of the original real-valued pattern.

As a basic problem of modeling spatiotemporal systems, the important problem of reconstruction of the local state-space vectors is investigated in this paper. The reconstruction techniques for determining the embedding dimensions of the time-series data are first introduced. Then, a new local state vector-reconstruction approach is proposed to find the minimum necessary spatiotemporal dimensions of the reconstruction vector from the observed spatiotemporal data. At the end of this paper, numerical examples are provided to illustrate the application of the new method and demonstrate its effectiveness.

Manuscript received August 9, 2007; revised January 11, 2008. This paper was recommended by Associate Editor B. J. Oommen.

Y. Pan was with the Department of Automatic Control and Systems Engineering, Sheffield University, S1 3JD Sheffield, U.K. He is now with the Department of Psychology, Sheffield University, S1 3JD Sheffield, U.K.

S. A. Billings is with the Department of Automatic Control and Systems Engineering, Sheffield University, S1 3JD Sheffield, U.K.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2008.918571

II. Determination of Embedding Dimensions

Time-series embedding theory for a d-dimensional dynamical system has been established by Packard *et al.* [20], Takens [21], and Sauer *et al.* [22]. These results theoretically justify the existence of unfolding the attractor of a system from observations of a single dynamical variable. These theoretical results were initially dedicated to time-series embedding and the prediction of a d-dimensional autonomous dynamical system. In [22], it was shown that the delay coordinate embedding theorem for single variable measures could be extended to the case of multivariate measures.

The embedding theorem states that the dynamics of physical systems can be represented by geometric objects in a d-dimensional physical phase space. For example, suppose that the dynamical system under consideration evolves in a d-dimensional phase space, i.e.,

$$\mathbf{x}(k) = [x_1(k), x_2(k), \ldots, x_d(k)] \tag{1}$$

$$\mathbf{x}(k+1) = \mathbf{F}[\mathbf{x}(k)] \tag{2}$$

where $y(k)$ is an observed variable of a scalar component of $\mathbf{x}(k)$ or a scalar projection of the $\mathbf{x}$ space: $y(k) = g(\mathbf{x}(k))$. If the embedding time lag is correctly determined, a d_E-dimensional space which inherits many of the properties of the original d-dimensional $\mathbf{x}$ space can be reconstructed even though the $\mathbf{x}$ space is unknown. The following reconstructed state vector can be used to represent the original dynamical system:

$$\mathbf{v}_{d_E}(k) = [y(k), y(k-\tau), \ldots, y(k-(d_E-1)\tau)]. \tag{3}$$

In $\mathbf{v}_{d_E}(k)$, τ is the reconstructed time delay, which is an integer multiple of the sampling time, and d_E is the embedding dimension. The sufficient condition for a correct embedding is that d_E should be large enough ($d_E > 2d_A$), where d_A is the box-counting dimension of the orbit in the original space. For a correct embedding, the points in the reconstructed-phase space are mapped to the original space through a diffeomorphism, namely, a smooth, differentiable, and invertible transformation.

It is of great importance to choose an adequate dimension value d_E of the reconstructed state-space vectors for the modeling and prediction of experimental dynamical systems. When the dimension d_E is increased during system identification, the size of the set of candidate model terms can increase dramatically. Thus, it is often impractical to vary d_E over a large range of values. For example, when the reconstructed state-space vectors are employed for the identification of nonlinear dynamical systems, using an unnecessarily high embedding dimension has a significant penalty on the computational cost, and the number of free parameters can become excessive. Therefore, finding the minimum necessary embedding dimension is an important task in state-space reconstruction for nonlinear dynamical systems.

The usual method of choosing the minimum embedding dimension d_E is to compute some invariants on the attractor. By increasing the embedding dimension used for computation, the value of the invariant when it stops changing can be detected. The main problem with this class of approaches is that they are often very data intensive and are certainly subjective. Apart from these approaches, two kinds of methods, which are both based on the properties of the diffeomorphical map between the original, physical state space, and the time-series reconstructions, have been widely used to find the minimum embedding dimensions from observations of dynamical systems. One method to determine the dimensionality of the system is the false nearest neighbor method developed by Kennel *et al.* [6]. Pecora *et al.* [8] developed another statistical method to select the embedding dimension for a given level of confidence.

A. False Nearest Neighbor Method

The main idea behind the false nearest neighbor method is that, for deterministic systems, points which are close in the reconstructed state space stay close under forward iteration. This statement means that points with state-space neighbors are a result of the dynamics rather than being projected from near one another as an artifact of using too low an embedding dimension. Choosing an inappropriately low-dimension d_E will cause points which are widely separately distributed in the original state space to overlap spuriously in the reconstructed space. Some points will actually be far from each other in the original state space and appear as neighbors in the reconstructed state space, because the geometric structure of the attractor has been projected down onto a smaller space. These points are referred to as false neighbors. Otherwise, when the embedding dimension is chosen to be large enough, all neighbors of every point in the multivariate state space will be true neighbors in the reconstructed space.

Suppose that $\mathbf{v}_d(k), d \in \mathbf{Z}^+$ are embedding vectors which should be far apart in the original state space but are close in the reconstructed space because d is too low. Denote the rth nearest neighbor of $\mathbf{v}_d(k)$ by $\mathbf{v}_d(k^{(r)})$, then from (3), the square of the Euclidian distance between the point $\mathbf{v}_d(k)$ and the neighbor is

$$\begin{aligned} R_d^2(k,r) &= \left\| \mathbf{v}_d(k) - \mathbf{v}_d\left(k^{(r)}\right) \right\|^2 \\ &= \sum_{p=1}^{d-1} \left[y(k-p\tau) - y(k^{(r)}-p\tau) \right]^2 \end{aligned} \tag{4}$$

where $k^{(r)}$ is the index of the nearest neighbor $\mathbf{v}_d(k^{(r)})$ and τ is the time delay. When the embedding dimension is increased from d to $d+1$, the first criterion to determine a false nearest neighbor is based on the following inequality:

$$\frac{\left\| y(k-d\tau) - y\left(k^{(r)}-d\tau\right) \right\|}{\left\| \mathbf{v}_d(k) - \mathbf{v}_d\left(k^{(r)}\right) \right\|} > R_{\mathrm{tol}}. \tag{5}$$

In (5), $\|y(k-d\tau) - y(k^{(r)}-d\tau)\|$ is obtained from (4) as $R_{d+1}^2(k,r) - R_d^2(k,r)$. R_{tol} is an arbitrary positive threshold which is empirically fixed, $R_{\mathrm{tol}} \geq 10$ or $R_{\mathrm{tol}} \approx 15$ [6]. It was shown in [9] that the minimum reasonable R_{tol} should be given by the maximum of the local deterministic expansion rate, which can be much larger than $e^{\lambda_{\max}\tau}$, where $\lambda_{\max}$ is the maximal Lyapunov exponent.

However, in some situations when the time series are very noisy, the criterion (5) may fail to detect the false nearest neighbors. Consider an extreme case where white noise is considered, this criterion will report erroneously that the white

noise could be embedded in a quite small dimensional space. The second complementary criterion is defined as

$$\frac{R_{d+1}(k,r)}{R_A} > A_{\text{tol}} \tag{6}$$

where R_A^2 is the sample variance of the time series and $A_{\text{tol}} \approx 2$ is an arbitrary threshold. This pair of criteria works jointly to declare a nearest neighbor as false if either test fails. Then, the embedding dimension can be selected when the percentage of false nearest neighbors stops decreasing with the increasing of the embedding dimension.

As can be seen from (5) and (6), to implement the nearest neighbor method requires choosing two arbitrary thresholds R_{tol} and A_{tol}. Cao [23] has suggested a scale-free false nearest neighbor method, which can be defined as follows:

$$a_d(k) \triangleq \frac{\left\|\mathbf{v}_{d+1}(k) - \mathbf{v}_{d+1}\left(k^{(r)}\right)\right\|}{\left\|\mathbf{v}_d(k) - \mathbf{v}_d\left(k^{(r)}\right)\right\|} \tag{7}$$

where $k \in [1, N - d\tau]$. The mean of $a_d(k)$ is defined as

$$\overline{a}_d \triangleq \frac{1}{N - d\tau} \sum_{k=1}^{N-d\tau} a_d(k). \tag{8}$$

Define

$$E_d \triangleq \frac{\overline{a}_{d+1}}{\overline{a}_d}. \tag{9}$$

E_d stops changing when $d \geq d_E$. However, using Cao's method to determine the embedding dimension still lacks robustness against noise [13].

B. Continuity Test

As has been proved by the embedding theorem, for a correct embedding, the original state space can be mapped to the reconstructed space through a diffeomorphical mapping, which is a smooth, differentiable, and invertible transformation. The continuity test [8] was developed on the basis of certain analysis properties of the diffeomorphical mapping between the original state space and the reconstructed space. The minimum embedding dimension of the reconstructed state space is then determined using statistical tools to test for the null hypothesis.

Consider a multivariate time-series data set $\{y_j(k), j = 1, \ldots, M\}$ measured at an equal time interval τ. Suppose that $\mathbf{v}_d(k) = [y_{j_1}(k - \tau_1), y_{j_2}(k - \tau_2), \ldots, y_{j_d}(k - \tau_d)]$ is a d-dimensional embedding vector chosen from the multivariate data, $j_p \in \{1, 2, \ldots, M\}$, $p = 1, 2, \ldots, d$. Generally, each τ_p might be different from each component. For the reconstruction of the time-series data, the determination of the embedding dimension can be converted into the mathematical testing of the existence of the following:

$$y_{j_{d+1}} = f\left(y_{j_1}(k - \tau_1), y_{j_2}(k - \tau_2), \ldots, y_{j_d}(k - \tau_d)\right) \tag{10}$$

for any function $f : \mathbf{R}^d \to \mathbf{R}^1$. If there does not exist any function f, such that $y_{j_{d+1}}(k)$ can be mathematically represented by the embedding vector $\mathbf{v}_d(k)$, the component $y_{j_{d+1}}$ should be added to the previous embedding vector $\mathbf{v}_d(k)$, i.e., increase the dimension of the embedding space from d to $d + 1$. Using this approach, the new components are continuously added to the embedding vector $\mathbf{v}(k)$ until all possible candidate components from all time series and for all τ value are functions of previous components.

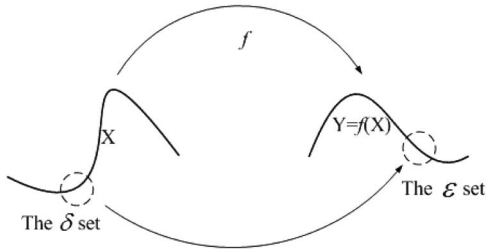

Fig. 1. Geometry of the mapping between the δ set and ϵ set.

The continuity statistic was developed by testing the continuity property of the function f. Consider a mapping f from a metric space $\mathbf{X}$ to another space $\mathbf{Y}$. Assume there is a metric $\|\cdot\|$ on each space. The function f is continuous at point $\mathbf{x}_o \in \mathbf{X}$ if $\forall \varepsilon > 0$, $\exists \delta > 0$, such that if $\|\mathbf{x} - \mathbf{x}_0\| < \delta$, then $\|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \varepsilon$. The geometric meaning of this statement is shown in Fig. 1.

The objective of the continuity test is to determine the embedding dimension of the reconstructed space by capturing the mathematical property of continuity on a given confidence level. The continuity statistic can be formulated in the following three steps.

First, a certain data number of data points are selected in δ and ε sets, which are consistent to the definition of continuity. Suppose that the vector $\mathbf{v}_d(k)$ has d coordinates. Equation (10) is a test for evidence of a continuous mapping $f : \mathbf{R}^d \to \mathbf{R}^1$ between $\mathbf{v}_d(k)$ and the $d + 1$th component $y_{j_{d+1}}(k)$. Following the definition of continuity, a positive ε is initially selected around a fiducial point $y_{j_{d+1}}(k_0)$ in $\mathbf{R}^1$. Accordingly, a positive δ is chosen around the fiducial point $\mathbf{v}_d(k_0)$ in $\mathbf{R}^d$ with respect to $y_{j_{d+1}}(k_0)$. Suppose that there are k points in the δ set. l points of these land in the ε set.

The second stage is to introduce the null hypothesis $\mathbf{H}_0$, which assumes that l out of k data points from the given δ set map to points in the ε set by a coin flip which comply with the binomial distribution [13], [24], $k \sim B(k, p)$. Namely, any data point in the δ set lands in the ε set by chance with probability $p = 0.5$ regardless of the size of ε. The probability density function of the binomial distribution is given by

$$F(l; k, p) = \frac{l!}{(l-k)!k!} p^k (1-p)^{l-k}. \tag{11}$$

To reject the null hypothesis, a given, usually 95%, confidence interval is required that the points in the δ set would not map to the ε set by a guess. The probability of rejecting the null hypothesis lies in the tail of the binomial distribution. Table I shows the number of points in the δ set that must be in the ε set to reject the null hypothesis at the 95% confidence level ($\alpha = 0.05$). For example, according to Table I, the probability of finding at least nine data points out of ten in the δ set map into the ε set is less than $\alpha = 0.05$, so the null hypothesis is rejected.

The final step of the continuity test is formulated on the basis of the minimum ε that can be used to reject the null hypothesis

TABLE I
NUMBER OF POINTS FROM THE δ SET MUST MAP INTO THE ε SET TO REJECT THE NULL HYPOTHESIS $\mathbf{H}_0$ AT THE 95% CONFIDENCE INTERVAL

Number of δ points	Number of points in the ε set	Number of points not in the ε set
7	7	0
8	8	1
9	8	1
10	9	1
11	9	2
12	9	3
13	10	3

$\mathbf{H}_0$ at each point. For each test point, δ and ε are assigned with an initial value (usually, $\varepsilon = \delta = 3\sigma$, where σ is the standard deviation of the state-space vector [24]). The number of points in the δ set around the fiducial point $\mathbf{v}_d(k_i)$ is n, and there are m number of points mapped to the ε set around the corresponding image point $y_{j_{d+1}}(k_i)$. Then, compute the binomial distribution with parameters $(n, 0.5)$ to find whether the cumulative probability of m or more points in the ε set is less than 0.05. If yes, the null hypothesis is rejected for this point, and the value of ε is recorded. If the null hypothesis is not rejected, ε is decreased until the null hypothesis is rejected. Denote this value as ε^*. To give a more robust estimate of ε^*, the average value, denoted as $\overline{\varepsilon^*}$, is usually computed from N test points $\mathbf{v}_d(k_i), i = 1, 2, \ldots, N$, which are selected at random times.

In the reconstruction process of the state-space vector, the value of $\overline{\varepsilon^*}$ is used as a criterion to find the minimum embedding dimension. If $\overline{\varepsilon^*}$ is reduced by adding a new component to the reconstruction vector, that component should be included in the reconstruction vector. If $\overline{\varepsilon^*}$ can no longer be reduced, the reconstruction vector is the final result.

Unlike the false nearest neighbor method, the continuity test is a data-adaptive method, and no arbitrary threshold needs to be prespecified. The minimum embedding dimension is faithfully determined on a given confidence level. In the meantime, this null hypothesis allows some points from the δ set to map outside the ε set but requires 95% confidence in the case of noise.

III. NEIGHBORHOOD DETECTION FOR SPATIOTEMPORAL SYSTEMS

The time-series embedding theoretical problem for dynamical systems with external inputs has also been investigated in [25] and [26]. Recently, this embedding result is extended to the case of finite-lattice LDSs with external inputs in a rigorous mathematical manner [27]. It has been shown that, under similar conditions, the state space of any finite-lattice dynamical system can be embedded into a reconstructed space as long as the dimension of the reconstructed space is larger than twice the size of the lattice. It has also been demonstrated in [27] that, for many practical modeling problems of LDSs, the original system dynamics can be embedded in the reconstruction vectors with relatively low dimensions. However, there is still no useful method proposed for finding the dimension of the reconstruction vector.

Compared with the false nearest neighbor method, the continuity test provides a powerful tool for finding the minimum embedding dimension on a given confidence level without any subjective decisions. The main idea behind this method is to test the functional independence between two vectors by the analytical property of continuity. In this section, this idea is employed to develop a new neighborhood detection method for the identification of CML models for continuous spatiotemporal systems whose cell entries have continuous value.

Consider the deterministic input–output CML model defined over a g-dimensional lattice I^g which consists of the set of all integer coordinate vectors $(i_1, \ldots, i_g) \in I^g$.

$$y_i(k) = f\left(\mathbf{q}^{n_y} y_i(k), \mathbf{q}^{n_u} u_i(k), \mathbf{q}^{n_y}\mathbf{s}^{m_y} y_i(k), \mathbf{q}^{n_u}\mathbf{s}^{m_u} u_i(k)\right) \quad (12)$$

where $i \in I^g$ is the spatial index of a g-dimensional space and $k \in T$ is the temporal index; $y_i(k)$ and $u_i(k)$ are the output and input variables at lattice i and time k, respectively. $\mathbf{q}^n(k)$ is a temporal backward shift operator, and s^m is a multivalued spatial shift operator, which are defined as follows:

$$\mathbf{q}^n = (q^{-1}, q^{-2}, \ldots, q^{-n}) \quad (13)$$

so that

$$\mathbf{q}^{n_y} y_i(t) = (y_i(t-1), y_i(t-2), \ldots, y_i(t-n_y))$$
$$\mathbf{q}^{n_u} u_i(t) = (u_i(t-1), u_i(t-2), \ldots, u_i(t-n_u)) \quad (14)$$
$$\mathbf{s}^m = (s^{p^1}, s^{p^2}, \ldots, s^{p^m}) \quad (15)$$

where $p^j \in I^d$ is the spatial translation multiindex, such that

$$\mathbf{s}^{m_y} y_i = (y_{i-p^1}, y_{i-p^2}, \ldots, y_{i-p^{m_y}})$$
$$\mathbf{s}^{m_u} u_i = (u_{i-p^1}, u_{i-p^2}, \ldots, u_{i-p^{m_u}}). \quad (16)$$

The identification task of the spatiotemporal system is to construct the CML model (12) that establishes a mathematical link between the past input, output variables, and the current system output so that the model-predicted output can adequately approximate the observed spatiotemporal dynamics using the identified model. Here, the model-predicted output y_i^{mpo}, which is defined in (17), is a tough model prediction capability test

$$y_i^{\mathrm{mpo}}(k) = f\left(\mathbf{q}^{n_y} y_i^{\mathrm{mpo}}(k), \mathbf{q}^{n_u} u_i(k)\ , \mathbf{q}^{n_y}\mathbf{s}^{m_y} y_i^{\mathrm{mpo}}(k), \mathbf{q}^{n_u}\mathbf{s}^{m_u} u_i(k)\right). \quad (17)$$

The main aim of neighborhood detection for spatiotemporal systems is to determine the corresponding minimum dimensions of m_y, m_u, n_y, and n_u associated with corresponding systematic variables. It is supposed here that function $f : \mathbf{Y} \times \mathbf{U} \to \mathbf{Y}$ in (12) is a continuous map, which represents a wide range of lattice dynamical systems. The continuity test, which was introduced in Section II-B, is employed to find the minimum dimension of the reconstructed state-space vector. The new neighborhood detection method based on the continuity test is implemented by testing the existence of the functional relationship (12) for any function $f : \mathbf{R}^d \to \mathbf{R}^1$, where $d = m_y + m_u + n_y + n_u$. Denote the d-dimensional vector $\Phi_d^i(k) = [\phi_1^i(k-\tau_1), \phi_2^i(k-\tau_2), \ldots, \phi_d^i(k-\tau_d)]$, $i \in I^g, k \in T$, which is composed of components from the

candidate variable set, $\phi_j^i(k-\tau_j) \in \{\mathbf{q}^{n_y} y_i(k), \mathbf{q}^{n_u} u_i(k), \mathbf{q}^{n_y}\mathbf{s}^{m_y} y_i(k), \mathbf{q}^{n_u}\mathbf{s}^{m_u} u_i(k)\}$, $j = 1, 2, \ldots, d$, as a reconstruction vector for the output $y_i(k)$ in (12).

To decide whether the neighborhood dimension should be increased from d to $d+1$ by adding a new component $\phi_{d+1}^i(k-\tau_{d+1})$ to the reconstruction vector $\Phi_d^i(k)$, the associated average minimal value of $\overline{\varepsilon^*}_{d+1}$ is compared with the previous $\overline{\varepsilon^*}_d$, where $\overline{\varepsilon^*}_j$, $j = 1, 2, \ldots$, is obtained by just rejecting the null hypothesis $\mathbf{H}_0$ on a given confidence level. Under the assumption that the model f in (12) between the image $y_i(k)$ and the reconstructed vector $\Phi_d^i(k)$ is continuous, if the reconstructed state-space vector is well done, the value of $\overline{\varepsilon^*}_j$ will no longer decrease by adding new components, i.e., the average minimal $\overline{\varepsilon^*}_d$ will no longer be larger than $\overline{\varepsilon^*}_{d+1}$. Usually, the variance of the sampled spatiotemporal data is normalized first to provide more faithful results of the value of $\overline{\varepsilon^*}$. Using this approach, the spatiotemporal neighborhood of the output $y_i(k)$ can be determined step by step.

In summary, the new neighborhood detection method for spatiotemporal systems which is based on the continuity test can be formulated as follows.

1) Normalize all the spatiotemporal data first to reduce any possible impact of different variances associated with each measured variable to the estimation of the minimum ε. In the first step, assign an initial reconstruction vector $\Phi_{d_0}(k)$ which is usually composed of some basic system variables; for instance, $\Phi_{d_0}(k) = \{y_i(k-1), u_i(k-1), y_{i-1}(k-1), y_{i+1}(k-1)\}$.
2) Randomly select N data points from the space and time domain. Sequentially choose one of the nearest neighboring components around the reconstruction vector $\Phi_d(k)$ to be added. Find each minimal value of ε_d^* associated with the new reconstruction vector when the null hypothesis $\mathbf{H}_0$ can just be rejected on a given confidence level (usually 95%). Calculate the average minimum value $\overline{\varepsilon^*}$ among N random data points associated with each potential component.
3) If the associated average minimal value $\overline{\varepsilon^*}_d \leq \overline{\varepsilon^*}_{d+1}$, the neighborhood size can be determined to be d based on the continuity test, and accordingly, the reconstruction vector is also determined. If not, repeat step 2) by adding a new component to the reconstruction vector.

IV. Numerical Applications

In this section, two numerical examples are included to illustrate the application of the new developed method for the neighborhood detection of spatiotemporal systems. In these two examples, both the noise-free and noisy cases are considered to test the capability of the new neighborhood detection method under random perturbations. Finally, the model-predicted outputs of the identified models are presented as tests of the model prediction capabilities and the neighborhood detection results.

Example 1—Linear Spatiotemporal System: Consider the following linear-diffusion equation:

$$\frac{\partial^2 y(t,x)}{\partial t^2} - C\frac{\partial^2 y(t,x)}{\partial x^2} = u(t,x), x \in [0,1] \qquad (18)$$

TABLE II
Neighborhood-Detection Results for Example 1 Without Noise

Selection order	Components	Average minimum of ε ($\overline{\varepsilon^*}$)
0	$\{y_i(k-1), y_{i-1}(k-1), u_i(k-1)\}$	0.1401
1	$y_i(k-2)$	0.0666
2	$y_i(k-3)$	0.0600
3	$y_{i+1}(k-1)$	0.0534
4	$y_{i+2}(k-1)$	0.0527
5	$y_{i+2}(k-2)$	0.0533

where x is spatial coordinate, with initial conditions

$$v(0,x) = 0 \qquad (19a)$$

$$\frac{dy(0,x)}{dt} = 4\exp(-x) + \exp(-0.5x) \qquad (19b)$$

where

$$u(t,x) = -13\exp(-x)\cos(1.5t) - 9.32\exp(-0.5x)\cos(2.1t). \qquad (20)$$

For $C = 1.0$, the exact solution of the earlier diffusion equation is

$$y(t,x) = 4\exp(-x)\cos(1.5t) + 2\exp(-0.5x)\cos(2.1t) - 4\exp(-x-t) - 2\exp(-0.5x-0.5t). \qquad (21)$$

The discrete data of output $y_i(k)$ and input $u_i(t)$ were evenly sampled on the space $\Omega = [0,1]$ and time domain $t = (0, 10\pi)$ at the rate of $\Delta x = 0.05$ and $\Delta t = \pi/50$, respectively.

In this example, to identify a CML model (12) from the simulated data, the initial reconstructed vector associated with lattice i at time k was initially composed of the following three components: $y_i(k-1)$, $u_i(k-1)$, and $y_{i-1}(k-1)$. To detect other neighborhoods that critically influence the dynamics of lattice i at time k, the proposed neighborhood detection procedure was applied. The input and output data were normalized first to reduce any possible bias of the estimates of ε; the standard deviations of the normalized input and output data were accordingly set to be one. Then, $N = 30$ points of spatiotemporal data were randomly chosen from the space and time domain to give a faithful estimate of ε by averaging the minimum value. The neighborhood detection results are then obtained by comparing the corresponding $\overline{\varepsilon^*}$ associated with the new reconstruction vector, which is sequentially added by the new component. Table II and Fig. 2 show the neighborhood detection results for this example. It is shown in Fig. 2 that the value of $\overline{\varepsilon^*}$ initially decreases, as each new component is added to the reconstruction vector but finally stops decreasing after the inclusion of the irrelevant components.

To make a comparative study, the system output $y_i(k)$ is assumed to be contaminated with a random noise sequence, which is evenly distributed between -0.03 and 0.03. In this situation, the neighborhood test results are shown in Table III and Fig. 3, where similar results have been obtained under two different conditions. Therefore, from these neighborhood test results, the final reconstruction vector for the system output $y_i(t)$ in this example can be believed to be composed of

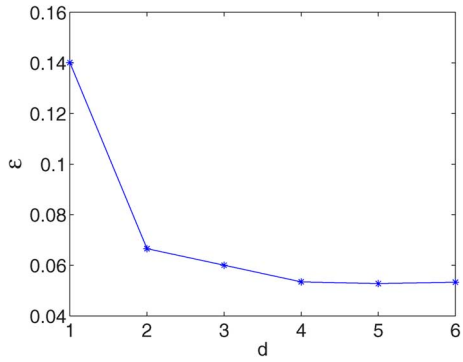

Fig. 2. Neighborhood detection test for Example 1 in a noise-free situation.

TABLE III
NEIGHBORHOOD-DETECTION RESULTS FOR EXAMPLE 1 WITH NOISE

Selection order	Components	Average minimum of ε $(\overline{\varepsilon^*})$
0	$\{y_i(k-1), y_{i-1}(k-1), u_i(k-1)\}$	0.1397
1	$y_i(k-2)$	0.0690
2	$y_i(k-3)$	0.0643
3	$y_{i+1}(k-1)$	0.0567
4	$y_{i+2}(k-1)$	0.0579

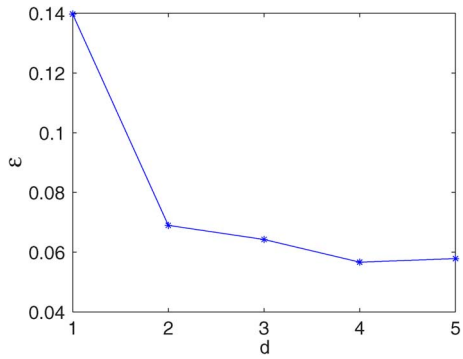

Fig. 3. Neighborhood detection test for Example 1 in a noisy situation.

these components: $u_i(k-1)$, $y_i(k-1)$, $y_i(k-2)$, $y_i(k-3)$, $y_{i-1}(k-1)$, and $y_{i+1}(k-1)$. Note that in the first neighborhood test, the value of $\overline{\varepsilon^*}$ associated with the addition of new component $y_{i+2}(k-1)$ is 0.0527, which is smaller than the previous value of 0.0533, but such a tiny difference can be ignored in practical applications. It can be said that the addition of this component makes no difference to the value of $\overline{\varepsilon^*}$, and it is not necessary to include this component for the sake of simplicity.

In the model identification procedure, the outputs of neighboring sites $y_{i-1}(t)$ and $y_{i+1}(t)$ were treated as a combined variable $y_i^*(t) = y_{i-1}(t) + y_{i+1}(t)$ to ensure the symmetry of topology. Based on the neighborhood detection results in Tables II and III, the maximum time lag of the candidate model terms for the identification was chosen to be three, and the nonlinearity degree was set to be one. Then, the identification procedure based on the l_1 regularized orthogonal forward regression algorithm, which was proposed in [28], was applied in this example. The finally identified CML model for the linear diffusive system is shown in Table IV. To test the model prediction accuracy of the identified model, the system-reference output and model-predicted output are shown in Fig. 4. It can be seen that the model-predicted output shows a good approximation of the dynamics of the linear-diffusive system, and the neighborhood detection results have been verified by the corresponding identified CML model.

TABLE IV
TERMS AND PARAMETERS OF THE IDENTIFIED CML MODEL FOR THE LINEAR DIFFUSIVE SYSTEM IN EXAMPLE 1

Model terms	Coefficients
$y_i(k-1)$	0.3790
$y_i(k-3)$	-0.2981
$y_i^*(k-1)$	0.4487
$u_i(k-3)$	-0.4536
$u_i(k-1)$	-0.5596
$u_i(k-2)$	1.0059

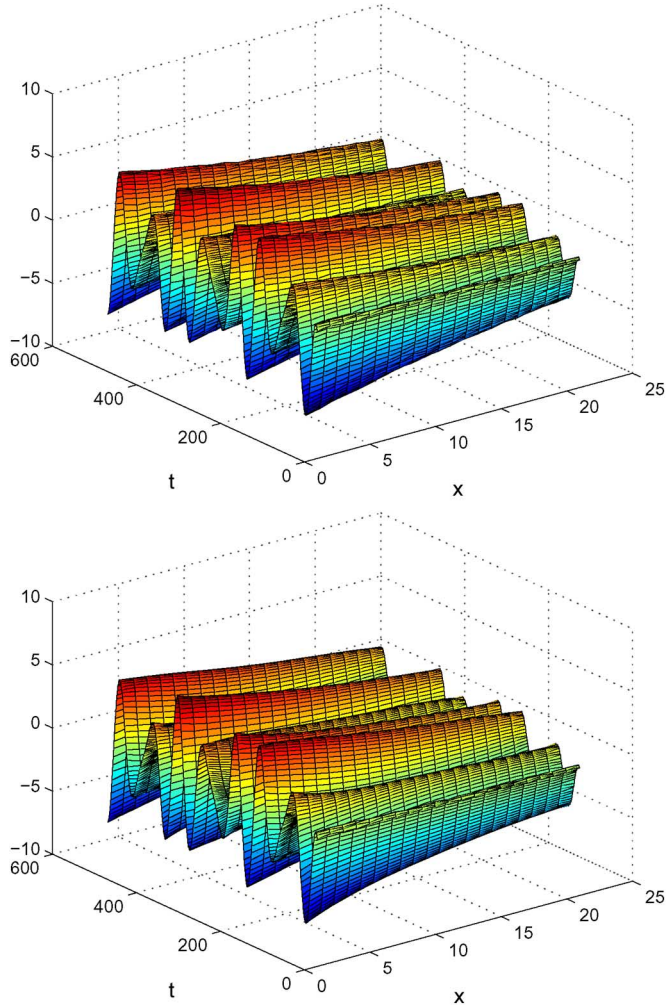

Fig. 4. (Top) System measurements and (bottom) model-predicted output error for the identified model in Example 1.

Example 2—Multivariate Spatiotemporal System: Consider the following multivariate spatiotemporal system described by the Fitz–Hugh–Nagumo reaction-diffusion equation:

$$\frac{\partial u}{\partial t} = d_1 \frac{\partial^2 u}{\partial x^2} + g(u) - v \tag{22a}$$

$$\frac{\partial v}{\partial t} = d_2 \frac{\partial^2 v}{\partial x^2} + \delta(u) - \gamma v \tag{22b}$$

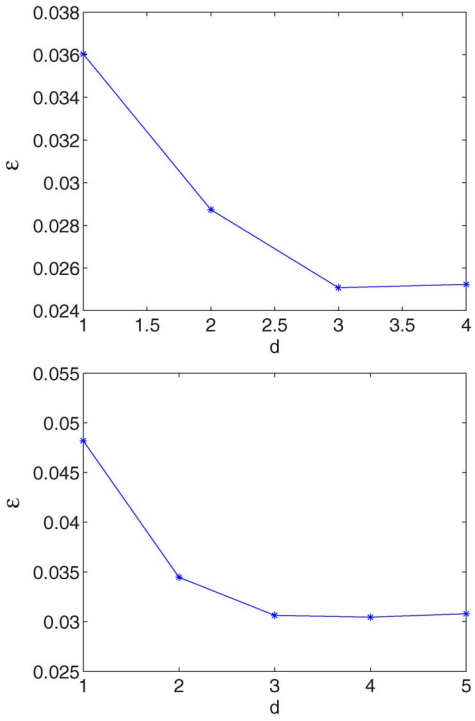

Fig. 5. Neighborhood detection test for u and v in Example 2 without noise, (top) u subsystem and (bottom) v subsystem.

with $g(u) = -u(u-\alpha)(u-1)$, $x \in \Omega = [0,1]$, and the Dirichlet boundary conditions were set as follows:

$$u(t,0) = u(t,1) = v(t,0) = v(t,1) = 0. \tag{23}$$

The initial conditions was set as

$$u(0,x) = v(0,x) = \sin\left(\frac{\pi x}{2}\right). \tag{24}$$

In this example, the parameters of the model (22) were set as $d_1 = d_2 = 6.188e-4$, $\delta = 40$, and $\gamma = -0.2$. The sampling parameters were chosen as $i\Delta x$, $\Delta x = 0.02$ in the spatial coordinate, and $k\Delta t$, $\Delta t = 0.01$ in the time coordinate. The neighborhoods of subsystems u and v are detected, respectively, using the proposed method with $N = 30$ samples of data randomly selected from the time and space domain. Fig. 5 and Table V show the neighborhood detection results for this example. Based on these results, the maximum time and spatial lag for this spatiotemporal system is two and one, respectively. Correspondingly, the reconstruction vectors for subsystems u and v can be described by $\{u_i(k-1), u_i(k-2), u_{i-1}(k-1), u_{i+1}(k-1), v_i(k-1)\}$ and $\{v_i(k-1), v_i(k-2), v_{i-1}(k-1), v_{i+1}(k-1), u_i(k-1)\}$, respectively.

To test the robustness of the new neighborhood detection method against noise, the system outputs u and v were deliberately corrupted with the random noise, which is evenly distributed between the range $[-0.05, 0.05]$. Table VI and Fig. 6 show the neighborhood detection results for this example in the noisy situation, which are consistent with the previous results. Therefore, according to the neighborhood detection results, the maximum time lag and spatial lag in this example can be determined to be two and one, respectively.

TABLE V
RESULTS OF NEIGHBORHOOD DETECTION FOR SUBSYSTEMS u AND v IN EXAMPLE 2 WITHOUT NOISE

Selection order	Components	Average minimum of ε $(\overline{\epsilon^*})$
0	$\{u_i(k-1), u_{i-1}(k-1), v_i(k-1)\}$	0.0360
1	$u_i(k-2)$	0.0287
2	$u_{i+1}(k-1)$	0.0251
3	$u_{i+1}(k-2)$	0.0252
0	$\{v_i(k-1), v_{i-1}(k-1), u_i(k-1)\}$	0.0485
1	$v_i(k-2)$	0.0345
2	$v_{i+1}(k-1)$	0.0306
3	$v_{i+1}(k-2)$	0.0304
4	$v_{i+2}(k-1)$	0.0308

TABLE VI
RESULTS OF NEIGHBORHOOD DETECTION FOR SUBSYSTEMS u AND v IN EXAMPLE 2 UNDER RANDOM NOISE PERTURBATIONS

Selection order	Components	Average minimum of ε $(\overline{\epsilon^*})$
0	$\{u_i(k-1), u_{i-1}(k-1), v_i(k-1)\}$	0.0488
1	$u_i(k-2)$	0.0338
2	$u_{i+1}(k-1)$	0.0303
3	$u_{i+1}(k-2)$	0.0299
4	$u_{i+2}(k-1)$	0.0309
0	$\{v_i(k-1), v_{i-1}(k-1), u_i(k-1)\}$	0.0462
1	$v_i(k-2)$	0.0374
2	$v_{i+1}(k-1)$	0.0334
3	$v_{i+1}(k-2)$	0.0340

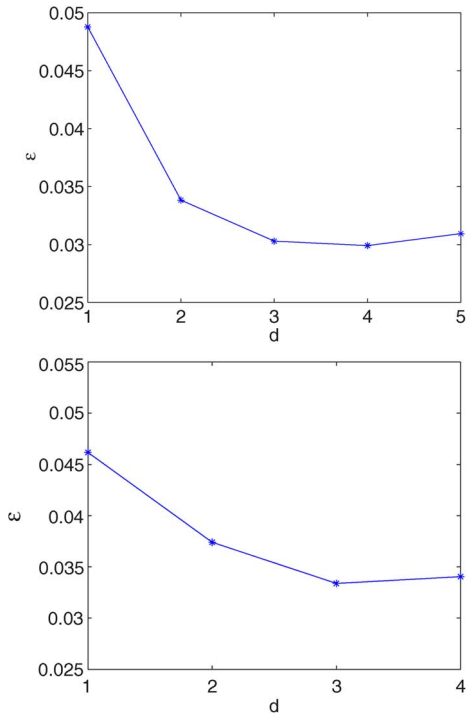

Fig. 6. Neighborhood detection test for subsystems u and v in Example 2 under random noise perturbations. (Top) u subsystem and (bottom) v subsystem.

TABLE VII
TERMS AND PARAMETERS OF THE IDENTIFIED CML MODEL FOR THE FITZ–HUGH–NAGUMO SYSTEM IN EXAMPLE 2

Subsystem	Model terms	Coefficients
$u_i(k)$	$u_i(k-1)$	0.3041
	$v_i(k-1)$	-0.0245
	$u_i(k-2)$	0.2047
	$u_{i-1}(k-1)$	0.2429
	$u_{i+1}(k-1)$	0.2406
	$u_i^2(k-1)$	0.0223
	$u_i^3(k-2)$	-0.0196
$v_i(k)$	$v_i(k-1)$	0.8230
	$u_i(k-1)$	0.9065
	$v_i(k-2)$	0.1426
	$v_{i-1}(k-1)$	0.0316

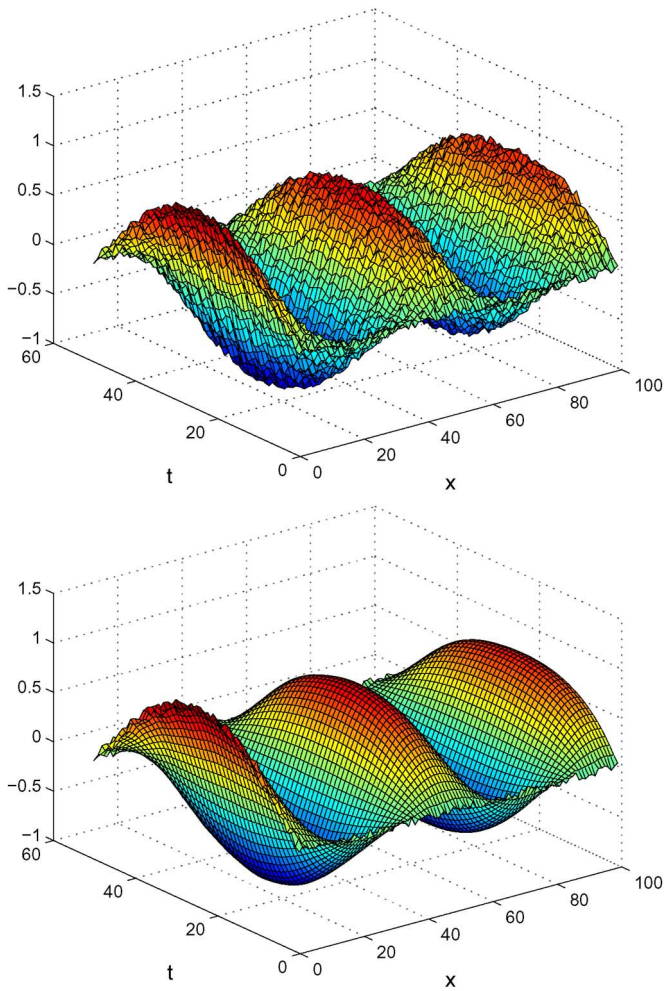

Fig. 7. (Top) System measurements and (bottom) model-predicted output for subsystem u in Example 2.

Based on the neighborhood detection results, a CML model was identified for this multivariate spatiotemporal system, which is given in Table VII.

To give a clearer demonstration of the identified results, the model-predicted output and error were shown in Figs. 7 and 8. It can be seen from the simulation results that the model-predicted output can closely approximate the output of the Fitz–Hugh–Nagumo system.

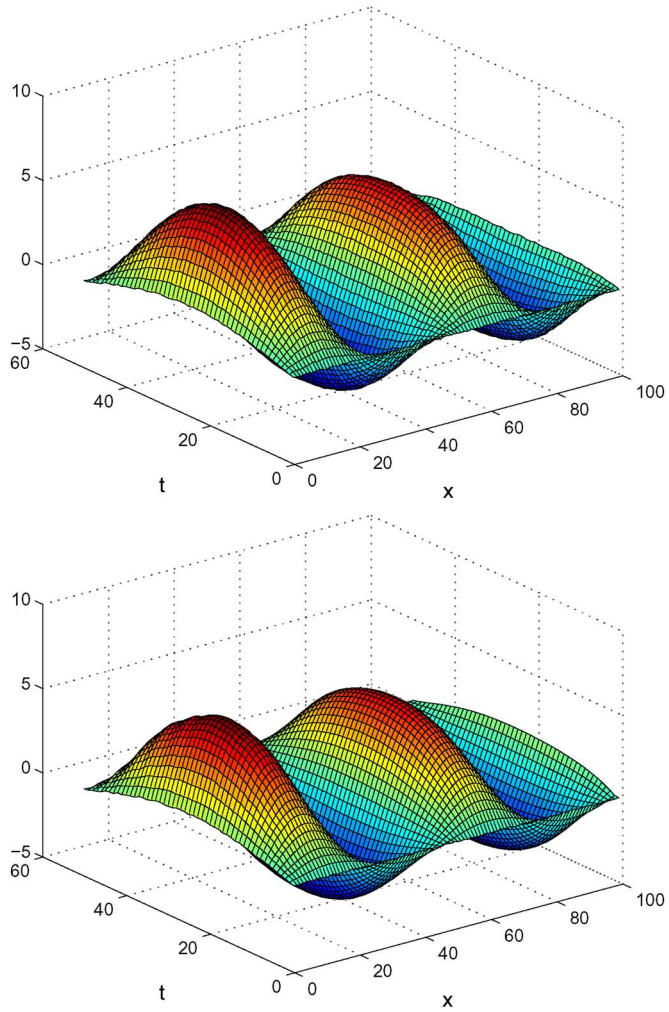

Fig. 8. (Top) System measurements and (bottom) model-predicted output for subsystem v in Example 2.

V. CONCLUSION

The neighborhood detection problem for the identification of spatiotemporal systems has been investigated in this paper. It is well known that the neighborhood detection task is closely related to the state-space reconstruction problem. Based on embedding theory for nonlinear dynamical systems, a new neighborhood detection approach has been developed in this paper for modeling spatiotemporal systems using CML models. The main idea behind the proposed method is to determine the spatiotemporal neighborhood size through testing of the continuity property of function f, which is independent of any model structure. The only assumption required for this method is that the function f in (12) is continuous, which can be satisfied for a wide range of spatiotemporal systems. If the system output $y_i(k)$ can be mathematically represented by a continuous model of the reconstructed vector of d-dimension, following the definition of the continuity, the value of ε reaches its minimum. Then, the minimal neighborhood size can be determined as d, and the selected spatiotemporal neighboring variables can form the reconstructed vector. Although the new neighborhood detection shares a similar theoretical background with the continuity test for time-series data, it provides a quite different procedure to determine the spatiotemporal

neighborhood from the method of finding minimum embedding dimensions for time-series data.

ACKNOWLEDGMENT

The authors would like to thank the Engineering and Physical Science Research Council (EPSRC) for their support to this paper.

REFERENCES

[1] D. Coca and S. A. Billings, "Identification of coupled map lattice models of complex spatio-temporal patterns," *Phys. Lett. A*, vol. 287, no. 1, pp. 65–73, Aug. 2001.

[2] S. Mandelj, I. Grabec, and E. Govekar, "Statistical approach to modelling of spatiotemporal dynamics," *Int. J. Bifur. Chaos*, vol. 17, pp. 2731–2738, 2001.

[3] S. Ostavik and J. Stark, "Reconstruction and cross-prediction in coupled map lattices using spatio-temporal embedding techniques," *Phys. Lett. A*, vol. 247, no. 1, pp. 145–160, Oct. 1998.

[4] U. Parlitz and C. Merkwirth, "Prediction of spatiotemporal time series based on reconstructed local states," *Phys. Rev. Lett.*, vol. 84, no. 9, pp. 1890–1893, Sep. 2000.

[5] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A*, vol. 33, no. 2, pp. 1134–1140, Feb. 1986.

[6] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phys. Rev. A*, vol. 45, no. 6, pp. 3403–3411, Mar. 1992.

[7] M. B. Kennel and H. D. I. Abarbanel, "False neighbors and false strands: A reliable minimum embedding dimension algorithm," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 66, no. 2, pp. 026 209-1–026 209-18, Aug. 2002.

[8] L. M. Pecora, T. L. Carrol, and J. F. Heagy, "Statistics for mathematical properties of maps between time series embeddings," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 52, no. 4, pp. 3420–3439, Oct. 1995.

[9] R. Hegger and H. Kantz, "Improved false nearest neighbor method to detect determinism in time series data," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 60, no. 4, pp. 4970–4973, Oct. 1999.

[10] L. T. Por and S. Puthusserypady, "Postprocessing methods for finding the embedding dimension of chaotic time series," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 72, no. 2, pp. 027 204-1–027 204-4, Aug. 2005.

[11] A. W. Jayawardena, W. K. Li, and P. Xu, "Neighbourhood selection for local modelling and prediction of hydrological time series," *J. Hydrol.*, vol. 258, no. 1–4, pp. 40–57, Feb. 2002.

[12] S. Boccaletti, D. L. Valladares, L. M. Pecora, H. P. Geffert, and T. Carroll, "Reconstructing embedding spaces of coupled dynamical systems from multivariate data," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 3, pp. 035 204-1–035 204-4, Mar. 2002.

[13] L. M. Pecora, L. Moniz, J. Nichols, and T. L. Carrol, "A unified approach to attractor reconstruction," *Chaos*, vol. 17, no. 1, pp. 013 110-1–013 110-9, Mar. 2002.

[14] D. Coca and S. A. Billings, "Analysis and reconstruction of stochastic coupled map lattice models," *Phys. Lett. A*, vol. 315, no. 1/2, pp. 61–75, Aug. 2003.

[15] L. Z. Guo and S. A. Billings, "Identification of coupled map lattice models of stochastic spatio-temporal dynamics using wavelets," *Dyn. Syst.*, vol. 19, no. 3, pp. 265–278, Sep. 2004.

[16] S. S. Mei, S. A. Billings, and L. Z. Guo, "A neighbourhood selection method for cellular automata models," *Int. J. Bifurc. Chaos*, vol. 15, no. 2, pp. 383–393, 2005.

[17] Y. Zhao and S. A. Billings, "Neighborhood detection using mutual information for the identification of cellular automata," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 36, no. 2, pp. 473–479, Apr. 2007.

[18] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 6, pp. 066 138-1–066 138-16, Jun. 2004.

[19] L. Z. Guo, S. S. Mei, and S. A. Billings, "Neighbourhood detection and identification of spatio-temporal dynamical systems using a coarse-to-fine approach," *Int. J. Syst. Sci.*, vol. 38, no. 1, pp. 1–15, 2007.

[20] N. Packard, J. Crutchfield, D. Farmer, and R. S. Shaw, "Geometry from a time series," *Phys. Rev. Lett.*, vol. 45, no. 9, pp. 712–715, Sep. 1980.

[21] F. Takens, *Detecting Strange Attractor in Turbulence*, vol. 898. New York: Springer-Verlag, 1981, pp. 366–381.

[22] T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," *J. Stat. Phys.*, vol. 65, no. 3/4, pp. 579–616, Nov. 1991.

[23] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, no. 1/2, pp. 43–50, Feb. 1997.

[24] L. Moniz, L. Pecora, J. Nichols, M. Todd, and J. R. Wait, "Dynamical assessment of structural damage using the continuity statistic," *Stuctural Health Monitoring*, vol. 3, no. 3, pp. 199–212, 2004.

[25] M. Casdagli, "A dynamical systems approach to modelling input–output systems," in *Nonlinear Modelling and Forecasting*, M. Casdagli and S. Eubank, Eds. Reading, MA: Addison-Wesley.

[26] A. Poncet, J. L. Poncet, and G. S. Moschytz, "On the input–output proximation of nonlinear systems," in *Proc. IEEE Symp. Circuits Syst.*, 1995, vol. 2, pp. 1500–1503.

[27] L. Z. Guo and S. A. Billings, "State-space reconstruction and spatio-temporal prediction of lattice dynamical systems," *IEEE Trans. Autom. Control*, vol. 52, no. 4, pp. 622–632, Apr. 2007.

[28] Y. Pan and S. A. Billings, "The identification of complex spatio-temporal patterns using coupled map lattice model," *Int. J. Bifurc. Chaos*, vol. 18, no. 4, 2008.

Y. Pan received the B.Sc. and M. Sc. degrees in electrical engineering from Harbin Institute of Technology, Harbin, China, in 2000 and 2002, respectively, and the Ph.D. degree from the University of the Sheffield, Sheffield, U.K., in 2007.

He is currently a Research Associate with the Department of Psychology, University of Sheffield. His research interests include statistics and the identification and model validation of complex nonlinear dynamic systems.

S. A. Billings received the B.Eng. degree (with first-class honors) in electrical engineering from the University of Liverpool, Liverpool, U.K., in 1972, the Ph.D. degree in control systems engineering from the University of Sheffield, Sheffield, U.K., in 1976, and the D.Eng. degree from the University of Liverpool in 1990.

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

Dr. Billings is a Chartered Engineer, a Chartered Mathematician, a Chartered Scientist, a Fellow of the Institution of Electrical Engineers (U.K.), and a Fellow of the Institute of Mathematics and Its Applications.