Monograph:
Pan, Y. and Billings, S.A. (2006) Model validation of spatiotemporal systems using correlation function tests. Research Report. ACSE Research Report no. 915 . Automatic Control and Systems Engineering, University of Sheffield

Model Validation of Spatiotemporal Systems Using Correlation Function Tests

Y. Pan, S. A. Billings
Dept. of Automatic Control and Systems Engineering,
University of Sheffield
Sheffield, S1 3JD, UK
February, 2006

Abstract

Model validation is an important and essential final step in system identification. Although model validation for nonlinear temporal systems has been extensively studied, model validation for spatiotemporal systems is still an open question. In this paper, correlation based methods, which have been successfully applied in nonlinear temporal systems are extended and enhanced to validate models of spatiotemporal systems. Examples are included to demonstrate the application of the tests.

1. Introduction

Spatiotemporal dynamic systems have become an increasingly important research area for a large range of scientific subjects including chemistry, biology, ecology, meteorology and finance. Spatiotemporal systems have traditionally been described using nonlinear Partial Differential Equations (PDE) or in discrete time form as Lattice Dynamical Systems (LDS) or a subset of LDS called Coupled Map Lattices (CML). A CML model is defined over a *d*-dimensional lattice where each site evolves in time through a discrete map which describes the influence of past states and neighbouring sites. CML were initially introduced in the1980s by Kaneko (1985, 1986). The CML model is discrete in both the time and space domain but has a continuous state value. PDE's can be finitely approximated by CML, provided that certain conditions of spatial and temporal resolution have been met. Due to the computational efficiency and richness of dynamical behaviors, the analysis and identification of CML has been studied by several authors.

A fundamental feature of CML is that the local state-space variables associated to each lattice node are the same over the whole lattice. In other words, these variables represent the same set of physical quantities at each node of the given lattice. The CML model can be shown to be composed of two parts: a local term involving only the local input and output variables and a spatial coupling term which describes the interactions with the neighboring lattice sites.

The purpose of model validation is to validate the correctness of the model structure and the unbiasedness of the estimated model parameters and usually involves testing the identified model on another independent set of data. It is a final and essential stage in most system identification procedures. Model validation methods can also be used to check if an identified model is under or over fitted.

Model validation for linear temporal systems is now well established. If the model structure is correct and the estimated parameters are unbiased, the model residuals or the one-step-ahead prediction errors should be a random time sequence with zero mean and finite variance. The auto-correlation function (ACF) and the cross-correlation function (CCF) have been widely used in linear temporal model validation (Bohlin, 1971, 1978, Soderstrom and Stoica, 1990). It is well known that the ACF of the residuals and the CCF between the residuals and input should fall within preset confidence intervals if the identified model is correct and the residual sequence is white.

Simple auto and cross correlation tests for linear models cannot be applied directly to the model validation of nonlinear temporal systems since they cannot detect all possible missing nonlinear terms in the residuals (Billings and Voon, 1983). Model validation methods for nonlinear temporal systems based on higher order correlation tests between the input and the residuals were first introduced by Billings and Voon (1983, 1986) to detect the missing nonlinear terms in the residuals. In order to achieve more discriminatory power with less computational cost, improved correlation tests based on correlation functions between the input, output and residuals were introduced in later studies (Billings and Zhu 1994, 1995; Mao and Billings 2000).

But all these methods are for purely temporal systems and unfortunately model validation tests for spatiotemporal systems are more complex. Given a derived or identified model of a spatiotemporal system in the form of a PDE, CML or LDS, model validation tests are required to determine whether the model can adequately describe the underlying dynamics of the spatiotemporal system. The only model validation methods which are available for spatiotemporal systems are based on subjectively judging the quality of the one-step-ahead prediction errors or the model predicted output (Partilz and Merkwirth 2000, Coca and Billings 2001, 2003; Timer *et al.* 2000; Muller and Timmer 2002. An alternative method is to compare specific dynamical characteristics like the bifurcation diagram between the modeled system and the real system (Aguirre and Billings 1994, 1995a, 1995b, Guo and Billings 2004). But a disadvantage of the latter method is that a priori information about the dynamical characteristics of the spatiotemporal system under study must be available.

In this paper, model validation methods based on higher order correlation function tests are introduced for a wide class of spatiotemporal systems and examples are included to demonstrate the performance of the new methods. This paper is arranged as follows. Section 2 formulates the problem of model validation for spatiotemporal systems. Section 3 reviews the correlation test methods for nonlinear temporal systems, while section 4 introduces the new model validation methods for spatiotemporal systems based on a set of correlation test functions. Three numerical examples are included in section 4 to illustrate the application of the new model methods and to demonstrate how the new tests can be used to detect missing or over-fitted model terms.

2. Problem Statement

Consider the general form of the stochastic input-output CML model for spatially invariant lattice dynamical systems (Billings and Coca, 2002, Coca and Billings, 2003)

$$
\begin{aligned}
y_i(t) &= f(q^{n_y} y_i(t), q^{n_u} u_i(t), q^{n_\varepsilon} \varepsilon_i(t), q^{n_y} s^{m_y} y_i(t), q^{n_u} s^{m_u} u_i(t), q^{n_\varepsilon} s^{m_\varepsilon} \varepsilon_i(t)) + \varepsilon_i(t) \\
&= f_l(q^{n_y} y_i(t), q^{n_u} u_i(t), q^{n_\varepsilon} \varepsilon_i(t)) \\
&\quad + f_c(q^{n_y} y_i(t), q^{n_u} u_i(t), q^{n_\varepsilon} \varepsilon_i(t), q^{n_y} s^{m_y} y_i(t), q^{n_u} s^{m_u} u_i(t), q^{n_\varepsilon} s^{m_\varepsilon} \varepsilon_i(t)) + \varepsilon_i(t)
\end{aligned} \tag{2-1}
$$

where, spatial invariance means that the underlying dynamics in each lattice node are the same for all lattice nodes. Here, $i \in I^d$ is the spatial index of a d-dimensional space and t is the temporal index; $y_i(t)$ and $u_i(t)$ are the output and input variables respectively at lattice i and time t, and $\varepsilon_i(t)$ is an independent zero mean random sequence. q^n is a temporal backward shift operator

$$q^n = (q^{-1}, q^{-2}, ..., q^{-n}) \tag{2-2}$$

so that

$$
\begin{aligned}
q^{n_y} y_i(t) &= (y_i(t-1), y_i(t-2), ..., y_i(t-n_y)) \\
q^{n_u} u_i(t) &= (u_i(t-1), u_i(t-2), ..., u_i(t-n_u)) \\
q^{n_\varepsilon} \varepsilon_i(t) &= (\varepsilon_i(t-1), \varepsilon_i(t-2), ..., \varepsilon_i(t-n_\varepsilon))
\end{aligned} \tag{2-3}
$$

where n_y, n_u and n_ε denote the maximum temporal lags corresponding to output y, input u and the residual sequence ε.

In (2-1), s^m is a multi-valued spatial shift operator

$$s^m = (s^{p^1}, s^{p^2}, ..., s^{p^m}) \tag{2-4}$$

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

$$
\begin{aligned}
s^{m_y} y_i &= (y_{i-p^1}, y_{i-p^2}, ..., y_{i-p^{m_y}}) \\
s^{m_y} u_i &= (u_{i-p^1}, u_{i-p^2}, ..., u_{i-p^{m_u}}) \\
s^{m_\varepsilon} \varepsilon_i &= (\varepsilon_{i-p^1}, \varepsilon_{i-p^2}, ..., \varepsilon_{i-p^{m_\varepsilon}})
\end{aligned} \tag{2-5}
$$

The parameters m_y, m_u and m_ε denote the maximum temporal lags corresponding to the output y input u and the residual sequence ε. The model $f : Y \times U \to Y$ is composed of the local map $f_l(\cdot)$ and the coupling map $f_c(\cdot)$, which can both take the form of a very wide class of linear or nonlinear models, including polynomial or wavelet expansions.

The main object of model validation is to check the goodness of fit of any given model by using model validation tests. A model validation test can be formulated as a statistical hypothesis testing problem. For instance, the identified model f is set as the hypothesis H_o. Then, in the first step, a parameter-free statistic is formed, which is a statistical function of the available data. Therefore, the distribution of the statistic variable is known if the hypothesis H_o is true. In this paper, the residual sequence or the one-step-ahead prediction error $\varepsilon_i(t)$ associated with model (2-1) is used as a statistic variable. So if the hypothesis H_o for the identified model is true, the residual sequence $\varepsilon_i(t)$ at lattice i should be completely random and unpredicted from all past inputs and outputs at all other spatial sites, so that

$$\varepsilon_i(t) = e(t), \qquad i \in I^d \tag{2-6}$$

where $e(t)$ is an independent random sequence with zero mean and finite variance. In order to validate the accuracy of (2-6) using sample means, 95% confidence limits are often used.

Before the new correlation tests can be developed for spatiotemporal systems existing results for purely temporal nonlinear models will be reviewed in the next section.

3. Correlation Tests for Temporal Models

Consider the nonlinear but purely temporal model

$$y(t) = f(y^{t-1}, u^{t-1}, \varepsilon^{t-1}) + \varepsilon(t) \tag{3-1}$$

where $t(t = 1,2,...)$ is a discrete time index and

$$\begin{aligned} y^{t-1} &= (y(t-1),..., y(t-d)) \\ u^{t-1} &= (u(t-1),..., u(t-d)) \\ \varepsilon^{t-1} &= (\varepsilon(t-1),..., \varepsilon(t-d)) \end{aligned} \tag{3-2}$$

are the delayed output, input and residual vectors respectively with maximum time lag *d*. Higher order correlation tests for nonlinear systems involving the output, input and residual are defined as follows (Billings and Zhu, 1994).

$$\begin{aligned} \phi_{\alpha\varepsilon^2}(\tau) &= \frac{\sum_{t=1}^{N-\tau}(\alpha(t)-\overline{\alpha})(\varepsilon^2(t-\tau)-\overline{\varepsilon^2})}{[(\sum_{t=1}^{N}(\alpha(t)-\overline{\alpha})^2)\sum_{t=1}^{N}(\varepsilon^2(t)-\overline{\varepsilon^2})^2]^{1/2}} \\ \phi_{\alpha u^2}(\tau) &= \frac{\sum_{t=1}^{N-\tau}(\alpha(t)-\overline{\alpha})(u^2(t-\tau)-\overline{u^2})}{[(\sum_{t=1}^{N}(\alpha(t)-\overline{\alpha})^2)\sum_{t=1}^{N}(u^2(t)-\overline{u^2})^2]^{1/2}} \end{aligned} \tag{3-3}$$

where

$$\alpha(t) = y(t)\varepsilon(t) \tag{3-4}$$

and $\overline{\bullet}$ denotes the time average.

The output $y(t)$ in (3-1) can be represented by the one-step-ahead predicted output and the residual as

$$y(t) = f(y^{t-1}, u^{t-1}, \varepsilon^{t-1}) + \varepsilon(t) = \hat{y}(t) + \varepsilon(t) \tag{3-5}$$

Thus, Equation (3-3) can be written as

$$\begin{aligned} \phi_{\alpha\varepsilon^2}(\tau) &= \phi_{(y\varepsilon)\varepsilon^2}(\tau) = k_1\phi_{(\hat{y}\varepsilon)\varepsilon^2}(\tau) + k_2\phi_{\varepsilon^2\varepsilon^2}(\tau) \\ \phi_{\alpha u^2}(\tau) &= \phi_{(y\varepsilon)u^2}(\tau) = k_1\phi_{(\hat{y}\varepsilon)\varepsilon^2}(\tau) + k_2\phi_{\varepsilon^2 u^2}(\tau) \end{aligned} \tag{3-6}$$

where

$$k_1 = \frac{\left(\sum_{t=1}^{N}(\hat{y}(t)\varepsilon(t)-\overline{\hat{y}\varepsilon})^2\right)^{1/2}}{\left(\sum_{t=1}^{N}(y(t)\varepsilon(t)-\overline{y\varepsilon})^2\right)^{1/2}},\quad k_2 = \frac{\left(\sum_{t=1}^{N}(\varepsilon^2(t)-\overline{\varepsilon^2})^2\right)^{1/2}}{\left(\sum_{t=1}^{N}(y(t)\varepsilon(t)-\overline{y\varepsilon})^2\right)^{1/2}} \tag{3-7}$$

When the model structure is correct and the estimated parameters are unbiased, the residual sequence $\varepsilon(t)$ should be a totally random sequence with zero mean and finite variance. These conditions will hold when

$$\begin{aligned} &\phi_{(\hat{y}\varepsilon)\varepsilon^2}(\tau) = 0, \forall \tau > 0 \\ &\phi_{(\hat{y}\varepsilon)u^2}(\tau) = 0, \forall \tau \end{aligned} \tag{3-8}$$

and Equation (3-6) consequently becomes

$$\begin{aligned} &\phi_{\alpha\varepsilon^2}(\tau) = \phi_{(y\varepsilon)\varepsilon^2}(\tau) = k_2\phi_{\varepsilon^2\varepsilon^2}(\tau) \\ &\phi_{\alpha u^2}(\tau) = \phi_{(y\varepsilon)u^2}(\tau) = k_2\phi_{\varepsilon^2u^2}(\tau) \end{aligned} \tag{3-9}$$

According to the Central Limit Theorem, for sufficiently large N the correlation function estimates given in (3-3) are asymptotically normal with zero mean and finite variance, and the 95% confidence interval is approximately equal to $\pm 1.95/\sqrt{N}$, where N is the data length.

This set of higher order correlation test functions can detect almost all possible missing linear and nonlinear terms in the residuals, even if the variances of the input and residual are small. The discriminatory power of this method is greatly enhanced compared with the correlation tests only involving the residual and input (Billings and Zhu, 1994).

However, these correlation test functions may also have a disadvantage in some practical situations. For example in (3-6), $\phi_{\alpha\varepsilon^2}(\tau)$ is composed of two parts, $k_1\phi_{(\hat{y}\varepsilon)\varepsilon^2}(\tau)$ and $k_2\phi_{\varepsilon^2\varepsilon^2}(\tau)$ with k_1 , k_2 determined by (3-7). In an ideal situation, the residual will not correlate with the predicted output and input and Equation (3-6) can be converted into (3-8). But if the variances of the one-step-ahead predicted output and the residual are significantly different, k_1 and k_2 may take quite different values in those conditions. For example if k_1 is ten times larger than k_2 , $\phi_{\alpha\varepsilon^2}(\tau)/k_2$ may not be an approximate Dirac delta function even though the residual is a totally random sequence with zero mean value and finite variance.

4. Correlation Tests for Spatiotemporal Systems

It will be assumed throughout that the spatiotemporal systems under study are spatially invariant lattice dynamical systems. That is to say, the dynamics in each lattice can be described by the same parameter-invariant model, for example the model in (2-1). Another assumption is that all the signals from the stochastic spatiotemporal system under study are ergodic processes over both the time and space domains. Based on the first assumption, we do not need to study the dynamics of variables at every site of a spatiotemporal system. The overhead of computing correlation functions of the inputs, outputs and residuals from all lattice sites can be avoided by randomly selecting N sufficiently large data at different locations to calculate the correlation functions. From the latter assumption, it can easily be seen that this characteristic has two implications for the model residuals $\varepsilon_i(t)$, $\varepsilon_j(t)$ at the sites $i, j \in I^d$

$$\phi_{\varepsilon_i\varepsilon_i}(\tau) = E[\varepsilon_i(t)\varepsilon_i(t-\tau)] = \delta(\tau), \tau \in Z \tag{4-1a}$$

$$\phi_{\varepsilon_i\varepsilon_j}(\tau) = E\lfloor\varepsilon_i(t)\varepsilon_j(t-\tau)\rfloor = 0, \forall\tau \in Z, i \neq j \tag{4-1b}$$

where, $\delta(\tau)$ is a Dirac function and Z is the set of positive integers. From (4-1a), it can be seen that the residuals at a spatial location at different times are independent to each other while (4-1b) means that the residual variable at different sites are independent. These assumptions generally hold for a wide class of spatiotemporal systems.

New correlation tests can now be introduced for spatiotemporal systems based on cross correlation functions between the inputs, one-step-ahead predicted outputs and the residuals. From (2-1)

$$\begin{aligned} y_i(t) &= f(q^{n_y}y_i(t), q^{n_u}u_i(t), q^{n_\varepsilon}\varepsilon_i(t), q^{n_y}s^{m_y}y_i(t), q^{n_u}s^{m_u}u_i(t), q^{n_\varepsilon}s^{m_\varepsilon}\varepsilon_i(t)) + \varepsilon_i(t) \\ &= \hat{y}_i(t) + \varepsilon_i(t) \end{aligned} \tag{4-2}$$

where $\hat{y}_i(t)$ is the one-step-ahead predicted output and $\varepsilon_i(t)$ is the residual. The model predicted output of the CML model is defined as

$$y_i^{mpo}(t) = f(q^{n_y}y_i^{mpo}(t), q^{n_u}u_i(t), q^{n_y}s^{m_y}y_i^{mpo}(t), q^{n_u}s^{m_u}u_i(t)) \tag{4-3}$$

Two new tests $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$ are defined below, where N data samples of the input, one-step-ahead predicted output and residual sequences are randomly selected without repetitions from the space and time domains to compute the normalized correlation functions.

$$\begin{aligned} \phi_{\beta\varepsilon^2}(\tau) &= \frac{\sum_{(i,t)=S(0)}^{S(N-1)} \beta_i^0(t)\varepsilon_i^{2^0}(t-\tau)}{\left[\left(\sum_{(i,t)=S(0)}^{S(N-1)} \left(\beta_i^0(t)\right)^2\right)\left(\sum_{(i,t)=S(0)}^{S(N-1)} \left(\varepsilon_i^{2^0}(t)\right)^2\right)\right]^{1/2}} \\ \phi_{\beta u^2}(\tau) &= \frac{\sum_{(i,t)=S(0)}^{S(N-1)} \beta_i^0(t)u_i^{2^0}(t-\tau)}{\left[\left(\sum_{(i,t)=S(0)}^{S(N-1)} \left(\beta_i^0(t)\right)^2\right)\left(\sum_{(i,t)=S(0)}^{S(N-1)} \left(u_i^{2^0}(t)\right)^2\right)\right]^{1/2}} \end{aligned} \tag{4-4}$$

In (4-4) the vector S indicates the selection of the random locations (i_k, t_k) in both the time domain and space domains.

$$S = ((i_0, t_0), (i_1, t_1), ..., (i_{N-1}, t_{N-1})), i_k \in I^d, t_k \in T, k = 0, ..., N-1 \tag{4-5}$$

The normalized variables $\beta_i^0(t)$, $\varepsilon_i^{2^0}(t)$ and $u_i^{2^0}(t)$ in (4-4) are defined as follows

$$\beta_i^0(t) = \frac{\beta_i(t)}{\left[\frac{1}{N}\sum_{(i,t)=S(0)}^{S(N-1)} \beta_i^2(t)\right]^{1/2}}$$

$$\varepsilon_i^{2^0}(t) = \frac{(\varepsilon_i^2(t) - \overline{\varepsilon^2})}{\left[\frac{1}{N}\sum_{(i,t)=S(0)}^{S(N-1)} (\varepsilon_i^2(t) - \overline{\varepsilon^2})^2\right]^{1/2}} \tag{4-6}$$

$$u_i^{2^0}(t-\tau) = \frac{(u_i^2(t-\tau) - \overline{u_\tau^2})}{\left[\frac{1}{N}\sum_{(i,t)=S(0)}^{S(N-1)} (u_i^2(t-\tau) - \overline{u_\tau^2})^2\right]^{1/2}}$$

where, $\varepsilon_i(t)$ and $u_i(t)$ are the residual and input at lattice i and time t respectively. $\beta_i(t)$ is a normalized compound variable which is a function of the residual $\varepsilon_i(t)$ and one-step-ahead predicted output $\hat{y}_i(t)$.

$$\beta_i(t) = \frac{\hat{y}_i(t)\varepsilon_i(t) - \overline{\hat{y}\varepsilon}}{\left[\frac{1}{N}\sum_{(i,t)=S(0)}^{S(N-1)} (\hat{y}_i(t)\varepsilon_i(t) - \overline{\hat{y}\varepsilon})^2\right]^{1/2}} + \frac{\varepsilon_i(t)\varepsilon_i(t) - \overline{\varepsilon^2}}{\left[\frac{1}{N}\sum_{(i,t)=S(0)}^{S(N-1)} (\varepsilon_i(t)\varepsilon_i(t) - \overline{\varepsilon^2})^2\right]^{1/2}}$$

$$= \hat{y}\varepsilon_i^0(t) + \varepsilon\varepsilon_i^0(t) \tag{4-7}$$

In (4-6) and (4-7), $\overline{\bullet}$ denotes the time average over the specific domain defined by the vector *S*. For example, $\overline{\varepsilon^2}$ and $\overline{u_\tau^2}$ are defined as follows

$$\overline{\varepsilon^2} = \frac{1}{N}\sum_{(i,t)=S(0)}^{S(N-1)} \varepsilon_i^2(t) \tag{4-8}$$

$$\overline{u_\tau^2} = \frac{1}{N}\sum_{(i,t)=S(0)}^{S(N-1)} u_i^2(t-\tau) \tag{4-9}$$

Note that the mean value of the input variable $\overline{u_\tau^2}$ is defined as dependent on the value of τ. This is because in most practical spatiotemporal systems, the length of time *T* will not be large enough compared with the temporal system case. The value of τ in the correlation functions will therefore affect the mean values and variances of the selected data from variables of the spatiotemporal system. Actually, these statistical characteristics of the variables from spatiotemporal systems will have a significant difference in some situations. This will be illustrated in Example 3. Also in the proposed correlation tests (4-4), the compound variable $y_i(t)\varepsilon_i(t)$ used in the correlation method (3-3), is substituted by the combination of two normalized variables $\hat{y}\varepsilon_i^0(t)$ and $\varepsilon\varepsilon_i^0(t)$. An evident advantage for this improvement is that the new method can be practically feasible for spatiotemporal systems where the variances of the output and residual are quite different.

From the above definitions, (4-4) can be converted into

$$\phi_{\beta\varepsilon^2}(\tau) = k_1'\phi_{(\hat{y}\varepsilon)\varepsilon^2}(\tau) + k_2'\phi_{\varepsilon^2\varepsilon^2}(\tau)$$
$$\phi_{\beta u^2}(\tau) = k_1'\phi_{(\hat{y}\varepsilon)u^2}(\tau) + k_2'\phi_{\varepsilon^2 u^2}(\tau) \tag{4-10}$$

where

$$k_1' = \frac{\left(\sum_{(i,t)=S(0)}^{S(N-1)} (\hat{y}\varepsilon_i^0(t))^2\right)^{1/2}}{\left(\sum_{(i,t)=S(0)}^{S(N-1)} (\beta^0(t))^2\right)^{1/2}}, \quad k_2' = \frac{\left(\sum_{(i,t)=S(0)}^{S(N-1)} (\varepsilon\varepsilon_i^0(t))^2\right)^{1/2}}{\left(\sum_{(i,t)=S(0)}^{S(N-1)} (\beta^0(t))^2\right)^{1/2}} \tag{4-11}$$

From the definition in (4-7), it can be easily seen that k_1' is equal to k_2' which is close to the value of $1/\sqrt{2}$ when the model under study is correct. In the ideal situation, the residual should be unpredictable from all inputs and outputs, to give

$$\phi_{(\hat{y}\varepsilon)\varepsilon^2}(\tau) = 0, \forall \tau$$
$$\phi_{(\hat{y}\varepsilon)u^2}(\tau) = 0, \forall \tau > 0 \tag{4-12}$$

Equation (4-10) can now be written as

$$\phi_{\beta\varepsilon^2}(\tau) = \phi_{(y\varepsilon)\varepsilon^2}(\tau) = k_2'\phi_{\varepsilon^2\varepsilon^2}(\tau)$$
$$\phi_{\beta u^2}(\tau) = \phi_{(y\varepsilon)u^2}(\tau) = k_2'\phi_{\varepsilon^2 u^2}(\tau) \tag{4-13}$$

According to the Central Limit Theorem, for sufficiently large *N*, the estimates of the correlation function estimates given in (4-4) will be asymptotically normal with zero mean and finite variance, and the 95% confidence intervals, for $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$, will be approximately $\pm 1.95/\sqrt{N}$.

When these new correlation functions are applied to validate a spatiotemporal system, the inputs and outputs from neighbouring sites, for example the terms $s^{m_y} y_i(t), s^{m_u} u_i(t)$, should be treated as inputs in the correlation functions (4-4).

5. Numerical Examples

Three simulated spatiotemporal systems will be used to illustrate the new model validation method using the correlation tests. In the *Example 1*, a linear spatiotemporal system is studied and the new correlation method is illustrated by using the exact solution of the PDE. In *Example 2,* the model validation method is applied to a spatiotemporal system described by a CML model. Finally, an identified CML model of the Lokta-Volterra system is validated in *Example 3.*

5.1 *Example 1 - A Linear Spatiotemporal System*

The first example is based on the following linear diffusion equation

$$\frac{\partial^2 y(t,x)}{\partial t^2} - C\frac{\partial^2 y(t,x)}{\partial x^2} = u(t,x), \quad x \in [0,1], \tag{5-1}$$

where *x* is the spatial coordinate, with initial conditions

$$y(0,x) = 0, \qquad \frac{y(0,x)}{dt} = 4\exp(-x) + \exp(-0.5x), \tag{5-2}$$

and

$$u(t,x) = -13\exp(-x)\cos(1.5t) - 9.32\exp(-0.5x)\cos(2.1t). \quad (5\text{-}3)$$

For $C = 1.0$ the exact solution $y(t,x)$ of the above diffusion equation with the input as (5-3) is

$$y(t,x) = 4\exp(-x)\cos(1.5t) + 2\exp(-0.5x)\cos(2.1t) - 4\exp(-x-t) - 2\exp(-0.5x-0.5t) \quad (5\text{-}4)$$

In order to discretize the continuous system, the input and output were equally and spatially sampled on the spatial domain $\Omega = [0,1]$ at a grid size of 0.05, so that $x = (x_1, x_2, ..., x_{21}) = (0, 0.05, ..., 0.95, 1)$. In the time domain $(0, 10\pi)$, the input and output variables were evenly sampled at the rate $\Delta t = \pi/100$ so that $t = (t_1, t_2, ..., t_{1001}) = (0, \Delta t * 1, ..., 10\pi)$. The sampling functions at the location i and time t can be consequently written as

$$\begin{aligned} &u_i(k) = u(t_k, x_i) \\ &y_i(k) = y(t_k, x_i) + \varepsilon_i(k), k = 1,2,...,1001, i = 1,2,...,21 \end{aligned} \quad (5\text{-}5)$$

where $\varepsilon_i(k)$ is the residual in the corresponding location.

In this example, the data for the correlation tests comprised of 900 input and output data randomly selected from different locations in both the space and the time domains. The input and output data from the neighboring locations were treated as inputs in the correlation function. The spatially coupled terms were combined together as $u_{i+1}(t) + u_{i-1}(t)$ and $y_{i+1}(t) + y_{i-1}(t)$ due to the symmetry of the diffusive coefficients. Thus there are three inputs in the correlation functions, which are given in (5-6).

$$\begin{aligned} &u_{1,i}(k) = u_i(k) \\ &u_{2,i}(k) = u_{i+1}(k) + u_{i-1}(k) \\ &u_{3,i}(k) = y_{i+1}(k) + y_{i-1}(k) \end{aligned} \quad (5\text{-}6)$$

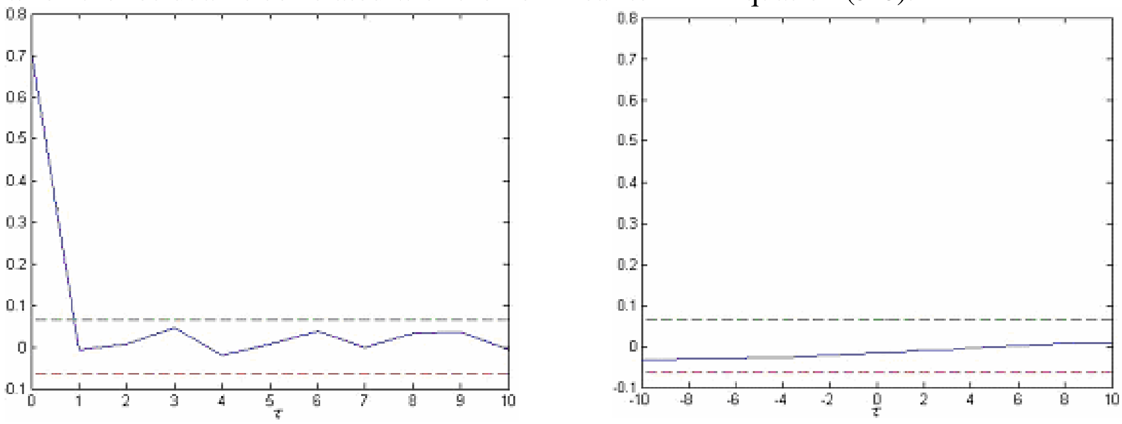
The residual $\varepsilon_i(k)$ was initially set as a purely random sequence $e_i(k)$ with the standard deviation was $\sigma = 0.3258$. The correlation functions $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$ given in (4-3) were then calculated and the corresponding results are showed in Figure (5-1), where the input in the correlation function $\phi_{\beta u^2}(\tau)$ represents the combination of the normalized inputs in (5-6), which is given as.

$$u_i(k) = \frac{u_{1,i}(k)}{\left(\frac{1}{N}\sum_{(i,k)=S(0)}^{S(N-1)} u_{1,i}{}^2(k)\right)^{1/2}} + \frac{u_{2,i}(k)}{\left(\frac{1}{N}\sum_{(i,k)=S(0)}^{S(N-1)} u_{2,i}{}^2(k)\right)^{1/2}} + \frac{u_{3,i}(k)}{\left(\frac{1}{N}\sum_{(i,k)=S(0)}^{S(N-1)} u_{3,i}{}^2(k)\right)^{1/2}} \quad (5\text{-}7)$$

In order to demonstrate the capability of detecting the wrong terms in the residual, the residual $\varepsilon_i(k)$ was deliberately set to be correlated with the input and the output of the neighboring site.

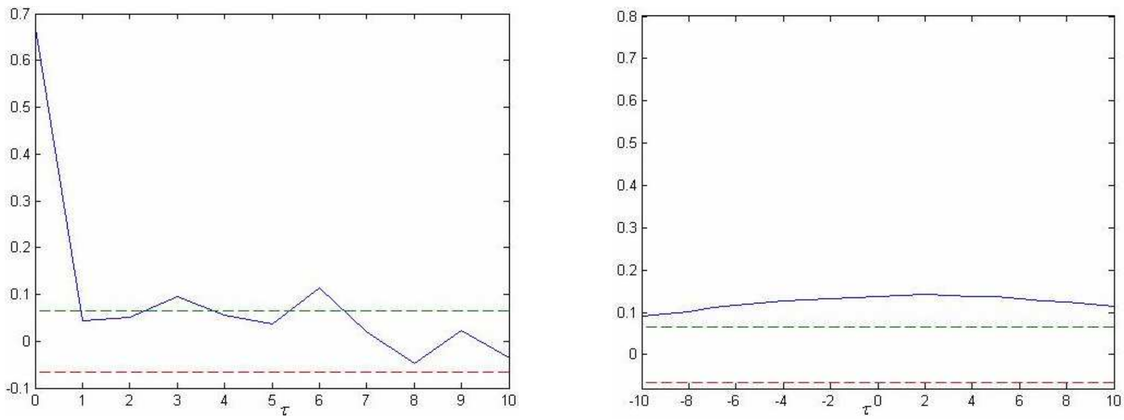
$$\varepsilon_i(k) = e_i(k) + 0.003 y_i(k-1) y_i(k-2) \quad (5\text{-}8)$$

From Figure (5-1) and Figure (5-2), it can be seen that the estimates of the correlation functions $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$ for this example are located within the 95% confidence intervals when the residual is random, but the estimates exceed the confidence intervals when the residual is correlated with the nonlinear term in Equation (5-8).



(a) (b)

Figure (5-1) Correlation tests for Example 1 with a random residual, (a) $\phi_{\beta\varepsilon^2}(\tau)$ test, (b) $\phi_{\beta u^2}(\tau)$ test.



(a) (b)

Figure (5-2) Correlation tests for Example 1 with the correlated residual defined in Equation (5-8), (a) $\phi_{\beta\varepsilon^2}(\tau)$ test, (b) $\phi_{\beta u^2}(\tau)$ test.

5.2 *Example 2 - A Nonlinear Spatiotemporal System Described by a CML Model*

Consider the following diffusively coupled map model in a 2-dimensional $L \times L$ lattice (Kaneko, 1989)

$$x_{i,j}(t) = (1-\theta) * f(x_{i,j}(t-1)) + \frac{\theta}{4} * (f(x_{i-1,j}(t-1)) + f(x_{i+1,j}(t-1)) + f(x_{i,j-1}(t-1)) + f(x_{i,j+1}(t-1))) \tag{5-9}$$

where $x_{i,j}(t), i, j = 1,..., L$ is the state at the discrete time t and the location of the lattice (i, j). Here L is chosen to be 50 and θ is the parameter defining the coupling length. The dynamics of the CML at the lattice sites are governed by θ and the local map f. In this example, the mapping function f is chosen as the logistic map

$$f(x) = 1 - ax^2 \tag{5-10}$$

This model has been extensively studied and it is known that a rich set of bifurcations will occur as the bifurcation parameter a is changed when $\theta > 0.3$ (Kaneko, 1989, Guo and Billings, 2004). In this example, the parameters in (5-9) were set as $a = 1.5$ and $\theta = 0.4$.

The CML model (5-9) was simulated with the parameters set above for 100 steps over the 50×50 lattice I^2 starting from a randomly generated initial population and periodic boundary conditions. Snapshots of the spatiotemporal patterns at different times are shown in Figure (5-3). Here, the measurement function at the location of the lattice (i, j) is given as

$$y_{i,j}(t) = x_{i,j}(t) + \varepsilon_{i,j}(t) \tag{5-11}$$

where the residual $\varepsilon_{i,j}(t)$ denotes the measurement noise at the specific location (i, j) and time t.

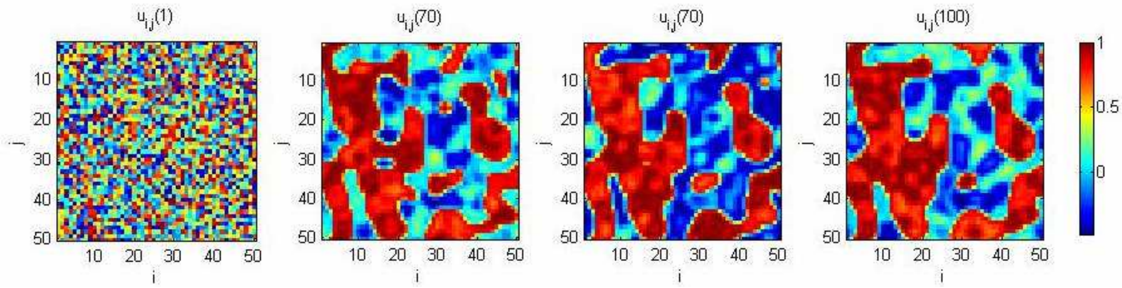


Figure (5-3) Snapshots of $x_{i,j}(t)$ taken at the times $t = 1, t = 40, t = 70, t = 100$

The new model validation methods were implemented with N is set to be 2500. The outputs in (5-9) from four neighbouring sites were treated as inputs, and the input in $\phi_{\beta u^2}(\tau)$ was selected to be a combination of four normalized inputs (similar to (5-7)).

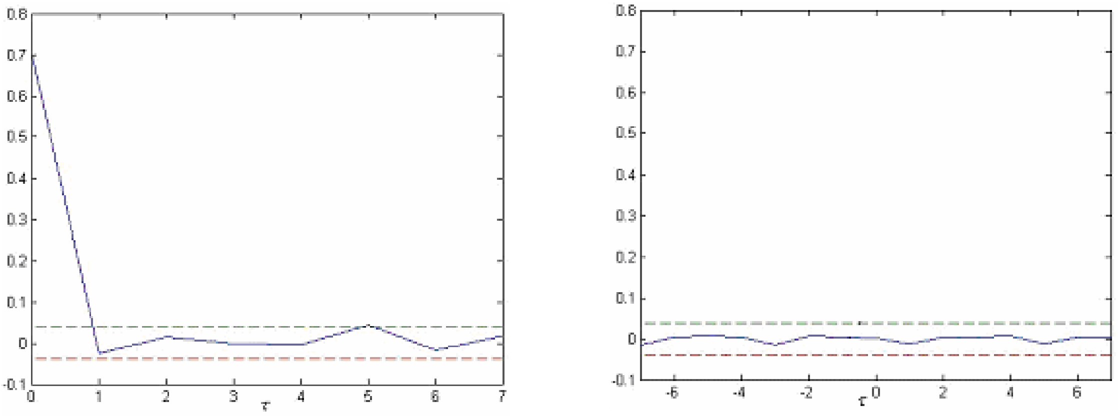
$$\begin{aligned} u_{i,j}^1(t) &= y_{i-1,j}(t) \\ u_{i,j}^2(t) &= y_{i+1,j}(t) \\ u_{i,j}^3(t) &= y_{i,j-1}(t) \\ u_{i,j}^4(t) &= y_{i,j+1}(t) \end{aligned} \tag{5-12}$$

Figure (5-3) shows the results of the correlation tests for the CML model (5-9) $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$ for the case where $\varepsilon_{i,j}(t)$ was a totally random spatiotemporal sequence. Figures (5-4), (5-5) show the correlation tests $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$ under the two conditions where the residual $\varepsilon_{i,j}(t)$ is random and correlated with the nonlinear terms defined in Equation (5-13) and Equation (5-14).

$$\varepsilon_{i,j}(t) = e_{i,j}(t) + 0.03 y_{i-1,j}(t-1) y_{i-1,j}(t-2) \tag{5-13}$$

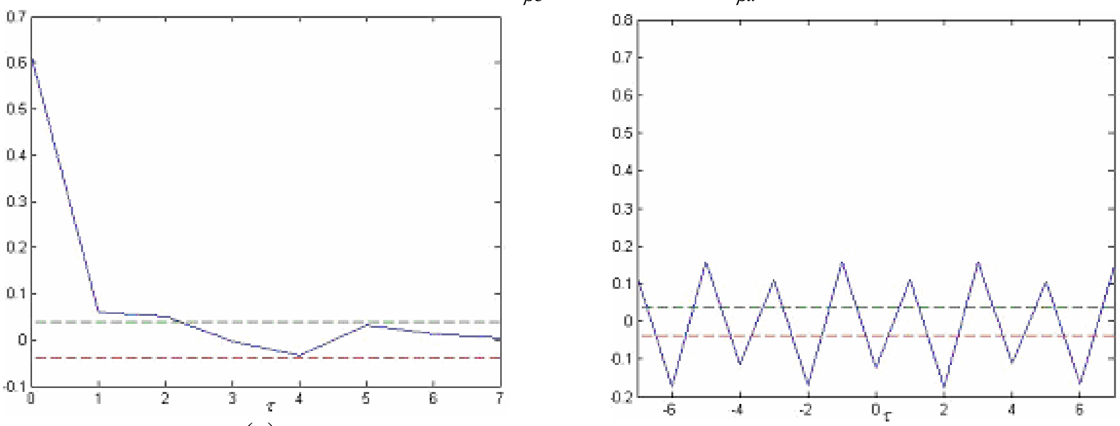
$$\varepsilon_{i,j}(t) = e_{i,j}(t) + 0.02 y_{i,j}(t-1) y_{i,j}(t-1) \tag{5-14}$$

It can be seen that the results of the correlation tests are within the 95% confidence limits for the random residuals case and are outside the confidence bounds and therefore correctly determine the model deficiency in Figure (5-5), (5-6).



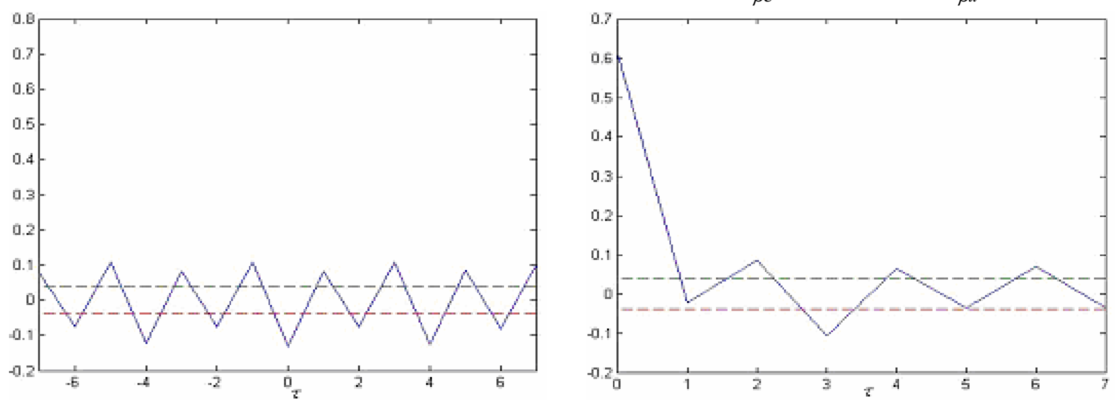
(a) (b)

Figure (5-4) Correlation tests for Example 2 with $\varepsilon_{i,j}(t) = e_{i,j}(t)$, where $e_{i,j}(t)$ is a random spatiotemporal sequence, (a) $\phi_{\beta\varepsilon^2}(\tau)$ test, (b) $\phi_{\beta u^2}(\tau)$ test.



(a) (b)

Figure (5-5) Correlation tests for Example 2 when the residuals are correlated with the nonlinear term defined in Equation (5-13), (a) $\phi_{\beta\varepsilon^2}(\tau)$ test, (b) $\phi_{\beta u^2}(\tau)$ test.



(a) (b)

Figure (5-6) Correlation tests for Example 2 when the residuals are correlated with the nonlinear term defined in Equation (5-14), (a) $\phi_{\beta\varepsilon^2}(\tau)$ test, (b) $\phi_{\beta u^2}(\tau)$ test.

5.3 *Example 3 – An Identified Nonlinear Spatiotemporal System*

In this example, the new correlation tests are used to validate an identified CML model of a nonlinear spatiotemporal system described by a partial differential equation.

Consider the Lotka-Volterra predator-prey model in two dimensions (Wilson *etc*, 1993) described by the following parabolic PDE as

$$\begin{aligned} \dot{u} &= c_{11}\nabla \cdot u + c_{12}\nabla \cdot u + a_1 u - r_1 uv \\ \dot{v} &= c_{21}\nabla \cdot v + c_{22}\nabla \cdot v - a_2 v + r_2 uv \end{aligned} \tag{5-15}$$

where $u = u(t,x,y)$ and $v = v(t,x,y)$ present the prey population density and the predator population density at time t and location (x,y) respectively. The corresponding coefficients in the above PDE were set as $a_1 = 0.47, r_1 = 0.024, a_2 = 0.76$ and $r_2 = 0.023$. The diffusive coefficients were set as $c_{11} = c_{12} = 0.1$ and $c_{21} = c_{22} = 0.01$ which signify that the prey diffuses faster than the predators through the space domain.

The Lotka-Volterra equation (5-15) was numerically simulated on the space domain $(0,1)\times(0,1)$ with the Neumann boundary conditions and initial conditions set as

$$\begin{aligned} u(0,x,y) &= \begin{cases} 1, (x-1/2)^2 + (y-1/2)^2 \le 1/16 \\ 0, otherwise \end{cases} \\ v(0,x,y) &= \begin{cases} 1, (x-1/2)^2 + (y-1/2)^2 \ge 1/16 \\ 0, otherwise \end{cases} \end{aligned} \tag{5-16}$$

The discrete observation for the identification are given by

$$\begin{aligned} u_{i,j}(k) &= u(k\Delta t, i\Delta x, j\Delta y) + \alpha_{i,j}(k) \\ v_{i,j}(k) &= v(k\Delta t, i\Delta x, j\Delta y) + \beta_{i,j}(k), i, j = 1,2,...,50; k = 1,2,...15 \end{aligned} \tag{5-17}$$

The numerical solution for (5-15) was sampled on the spatial grid $\Delta x = 0.02, \Delta y = 0.02$ with a time step $\Delta t = 0.06$.

In (5-17), $\alpha_{i,j}(k)$ and $\beta_{i,j}(k)$ are random sequences with standard deviations $\sigma_\alpha = 0.0058, \sigma_\beta = 0.01$ respectively. A CML model was identified by using an Orthogonal Forward Regression algorithm (Billings *etc* 1988, Guo and Billings 2004), and the results are shown in Table (5-1). In this example, for simplicity of illustration, only the model of subsystem *u* is investigated.

Table (5-1) Terms and parameters of the identified CML model for Example 3 and Subsystem *u*

Model Terms	Estimated Parameters
$u_{i,j}(k-1)$	0.3161
Constant	0.0316
$u^*_{i,j}(k-1)$ ¶	0.1681
$u_{i,j}(k-1)$^2	-0.3485

¶: $u^*_{i,j}(k-1) = u_{i-1,j}(k-1) + u_{i+1,j}(k-1) + u_{i,j-1}(k-1) + u_{i,j+1}(k-1)$

The identified CML model in Table (5-1) can be expressed as.

$$u_{i,j}(k) = 0.3161u_{i,j}(k-1) + 0.0316 + 0.1681u^{*}_{i,j}(k-1) - 0.3485u^{2}_{i,j}(k-1) + e_{i,j}(k) \tag{5-18}$$

where $e_{i,j}(k)$ is the residual sequence. Figure (5-7), (5-8) (5-9) show some snapshots of the measured output, model predicted output (4-3) and the residuals at different times. As noted in Section 4, it can be seen from Figure (5-7) that the variance of $u_{i,j}(k)$ changes with the time *k*.

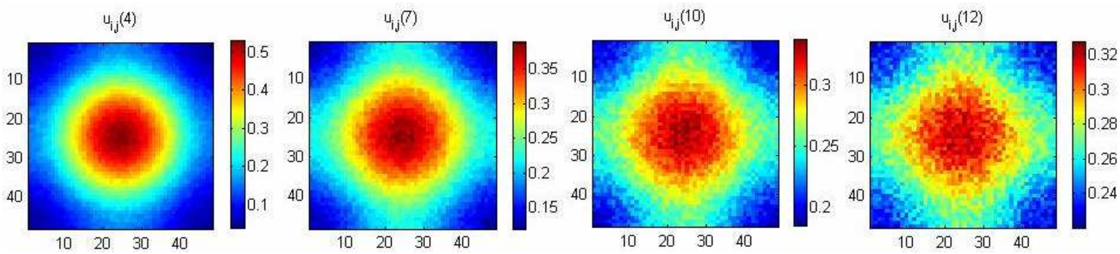


Figure (5-7) Snapshots of the measured output of the identified CML model Equation (5-18) taken at different times

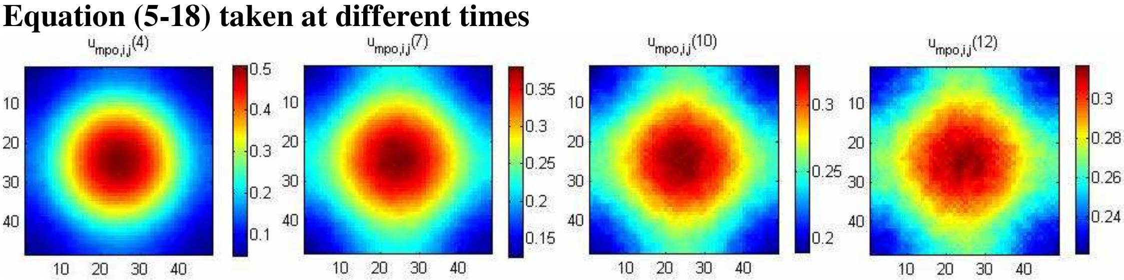


Figure (5-8) Snapshots of the model predicted output of the identified CML model Equation (5-18) taken at different times

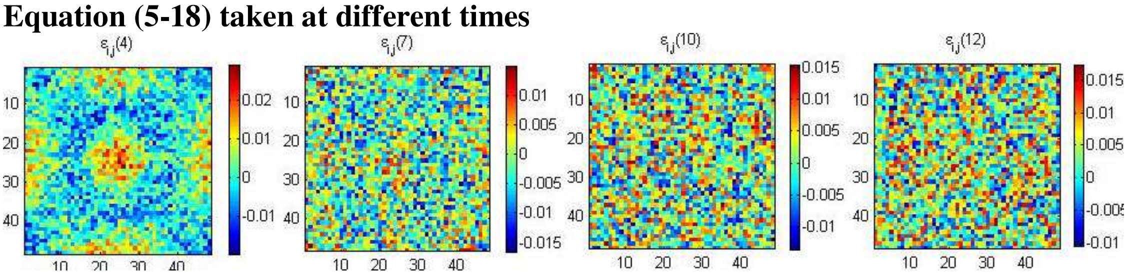
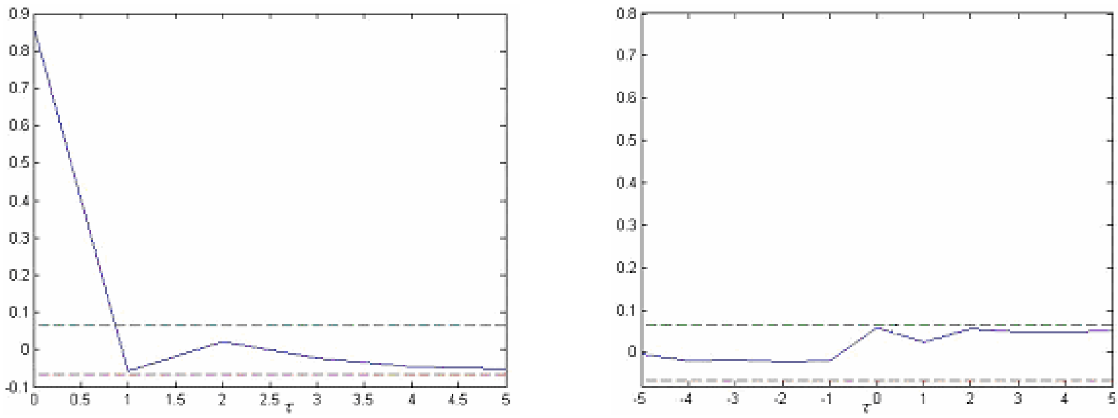


Figure (5-9) Snapshots of the residuals of the identified CML model Equation (5-18) taken at different times

The system output, one-step-ahead predicted output and the residuals generated from the CML model (5-18) were used to test the validity of the identified model. The correlation test results are given in the Figure (5-10). It can be seen that the estimates of correlation functions $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$ in Figure (5-10) are located within the 95% confidence bounds, indicating that the identified model is an adequate representation of the system. In order to test the identified model with missing or over-fitted terms, the estimated coefficient of the term $u^{2}_{i,j}(k-1)$ in (5-18) was assumed to be biased so that the corresponding residual $e'_{i,j}(k)$ would therefore be correlated with $u^{2}_{i,j}(k-1)$. The coefficient of $u_{i,j}^2(k-1)$ was therefore changed from the correct value of -0.3485 in Equation (5-18) to the incorrect or biased value of -0.3000 in Equation (5-19).

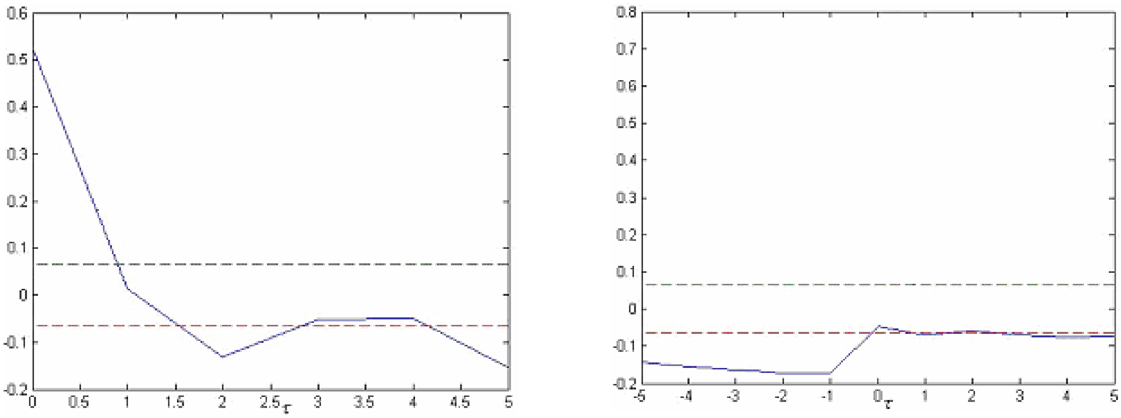
$$u_{i,j}(k) = 0.3161u_{i,j}(k-1) + 0.0316 + 0.1681u_{i,j}^*(k-1) - 0.3000u_{i,j}^2(k-1) + e'_{i,j}(k) \tag{5-19}$$

The correlation tests for (5-19) are given in Figure (5-11) and the incorrect estimate can clearly be detected in the model from the correlation functions $\phi_{\beta\varepsilon^2}(\tau)$ and $\phi_{\beta u^2}(\tau)$ which are now located outside the 95% confidence limit.



(a) (b)

Figure (5-10) Correlation tests for identified model (5-18) in Example 3, (a) $\phi_{\beta\varepsilon^2}(\tau)$ test, (b) $\phi_{\beta u^2}(\tau)$ test.



(a) (b)

Figure (5-11) Correlation tests for identified model (5-19) in Example 3 with a biased estimate, (a) $\phi_{\beta\varepsilon^2}(\tau)$ test, (b) $\phi_{\beta u^2}(\tau)$ test.

6 Conclusions

The problem of validating spatiotemporal systems has been investigated and new correlation-based tests have been proposed by extending the correlation test methods used for nonlinear temporal systems. New correlation functions have been constructed based on the inputs, one-step-ahead predicted outputs and the residuals. The overhead of computing all the states over the whole lattice has been avoided by randomly selecting the data from both the time and space domain. Replacing the compound variable $y_i(t)\varepsilon_i(t)$ by the combination of two normalized variables $\hat{y}\varepsilon_i^0(t)$ and $\varepsilon\varepsilon_i^0(t)$ was shown to make the new correlation tests more practically feasible and robust for the spatiotemporal systems case where variances of the output and residual can be quite different. The new model validation methods have been developed for SISO and SIMO spatiotemporal systems, but the application to MIMO spatiotemporal systems is straightforward by introducing the ideas from MIMO temporal systems (Billings and Zhu, 1995).

Reference:

1. Aguirre, L.A., Billings S.A. (1994). Validating identified nonlinear models with chaotic dynamics, *Int. J. Bifurcation and Chaos*, **4**, 109-125.
2. Aguirre, L.A., Billings S.A. (1995a). Retrieving Dynamical invariants from chaotic data using NARMAX models , *Int. J. Bifurcation and Chaos*, **5**, 449-474.
3. Aguirre, L.A., Billings S.A. (1995b). Dynamical effects of over parameterization in nonlinear models, Physica D, **80**, 26-40.
4. Billings, S.A. and Chen, S. and Kronenberg, M.J. (1988). Identification of MIMO non-linear systems using a forward regression orthogonal estimator. *Int. J. Control*, *49*, 2157-2189.
5. Billings, S.A. and Coca, D. (2002). Identification of coupled map lattice models of deterministic distributed parameter systems. *Int. J. Systems Science*, **33**, 623-634.
6. Billings,S.A. and Zhu,Q.M. (1994). Nonlinear model validation using correlation tests. *Int. J. Control*, **60**, 1107- 1120.
7. Billings,S.A. and Zhu,Q.M. (1995) Model validation tests for multivariable nonlinear models including neural networks. *Int. J. Control*, **62**, 749-766.
8. Bohlin,T. (1971) On the problem of ambiguities in maximum likelihood identification. *Automatica*, **7**, 199-210;
9. Bohlin,T. (1978), Maximum power validation of models without higher order fitting. *Automatica*, **14**, 137-146.
10. Coca, D. and Billings, S.A. (2003). Analysis and reconstruction of stochastic coupled map lattice models. *Phys. Lett. A*, **31**, 61-75.
11. Guo, L.Z. and Billings, S.A. (2004). Identification of coupled map lattice models of stochastic spatio-temporal dynamics using wavelets. *Dynamical System*, **19**, 265-278.
12. Kaneko, K. (1985). Spatiotemporal intermittency in coupled map lattices. *Prog. Theor. Phys*., **74**, 1033-1044.
13. Kaneko, K. (1986). Turbulence in coupled map lattices. *Physica D*, **18**, 475-476.
14. Kaneko, K. (1989). Spatiotemporal chaos in one-and two-dimensional coupled map lattices. *Physica D*, **37**, 60-82.
15. Mao, K.Z. and Billings, S.A. (2000). Multi-directional model validity tests for non-linear system identification. *Int. J. Control*, **73**, 132-143.
16. Ostavik, S. and Stark, J. (1998). Reconstruction and cross-prediction in coupled map lattices using spatio-temporal embedding techniques. *Phys. Lett. A*, **247**, 145-160.
17. Soderstrom, T. and Stoica, P. (1990). On covariance function tests used in system identification. *Automatica*, **26**, 125-133.

18. Wilson, W.G., Roos, A.M. and McCauley, E. (1993). Spatial instabilities within the diffusive Lotka-Volterra system: individual-based simulation results. *Theoretical Population Biology*, **43**, 91-127