Physics Letters A 378 (2014) 869–873

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Accurate state and parameter estimation in nonlinear systems with sparse observations

Daniel Rey [a], Michael Eldridge [a], Mark Kostuk [a], Henry D.I. Abarbanel [a,b], Jan Schumann-Bischoff [c,d], Ulrich Parlitz [c,d,*]

[a] Department of Physics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0374, United States
[b] Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0374, United States
[c] Max Planck Institute for Dynamics and Self-Organization, Am Faßberg 17, 37077 Göttingen, Germany
[d] Institute for Nonlinear Dynamics, Georg-August-Universität Göttingen, Am Faßberg 17, 37077 Göttingen, Germany

ARTICLE INFO

Article history:
Received 20 December 2013
Accepted 14 January 2014
Available online 24 January 2014
Communicated by C.R. Doering

Keywords:
Data assimilation
Synchronization
Time series analysis
System identification

ABSTRACT

Transferring information from observations to models of complex systems may meet impediments when the number of observations at any observation time is not sufficient. This is especially so when chaotic behavior is expressed. We show how to use time-delay embedding, familiar from nonlinear dynamics, to provide the information required to obtain accurate state and parameter estimates. Good estimates of parameters and unobserved states are necessary for good predictions of the future state of a model system. This method may be critical in allowing the understanding of prediction in complex systems as varied as nervous systems and weather prediction where insufficient measurements are typical.

Testing the consistency of models of nonlinear, complex systems with observations, then using those models to predict future events appears in the analysis of a broad spectrum of physical and biological systems. This includes numerical weather prediction [1], systems biology [2,3], biomedical engineering [4], chemical engineering [5], biochemistry [6], coastal and estuarine modeling [7,8], cardiac dynamics [9], and nervous system networks [10,11], among many others.

For this task one seeks to use information from observed data to inform the model about its unknown parameters and unobserved states, enabling one to make testable predictions which can be used to validate the model. Accurate predictions are the metric of quality for the physical model. This process of testing and validating nonlinear models may encounter an impediment when the systems express chaotic trajectories because sensitivity to initial conditions may cause the search space to become irregular [12].

To assess how well a model output tracks observations one may try to achieve synchronization of the data and the model output [13,14]. Searches for unknown parameters and unobserved states may be accomplished by perturbing a quality metric toward a synchronized setting. If the synchronization manifold (SM) – where states of the model coincide with states of the system and model output and observations are expected to match – is *unstable*, this search encounters numerous local minima yielding bad estimates and inaccurate predictions [12]. For the search to succeed, the unstable directions on the SM must be stabilized by a sufficient number of measurements L_s [15–17]. If one has fewer than L_s measurements, the instability of the SM impedes the ability to make accurate estimates.

Typically, the total number of available measurements L, is sparse compared with the number of degrees of freedom D of the model dynamics. For example, in the analysis of a shallow water model of geophysical flow [18] it was shown that $L_s \approx 70\%$ of the dynamical variables in the model, while in operational weather prediction systems such as the one at the European Centre for Medium-Range Weather Forecasts only about 10^7 measurements for models with 10^8 or 10^9 degrees of freedom are typical [19]. This suggests that to achieve $L_s > L$ observations, additional means are required to ensure that sufficient information passes from observations to model dynamics.

The idea we explore here is that when one can only make $L < L_s$ measurements, one can use time-delays of the available measurements to provide this required additional information. We report on a technique which uses an observed state variable at a time t and the time-delays of those measurements as a control to drive the model to synchronize with the data.

Using time-delay coordinates in the description of nonlinear and chaotic systems is a well established method for reconstructing a proxy state space from limited measurements, thereby

* Corresponding author.
E-mail addresses: habarbanel@ucsd.edu (H.D.I. Abarbanel), ulrich.parlitz@ds.mpg.de (U. Parlitz).

0375-9601/$ – see front matter
http://dx.doi.org/10.1016/j.physleta.2014.01.027

providing a coordinate system to analyze nonlinear aspects of the system [20–25]. While our idea of using information in time-delayed measurements is similar to standard time-delay phase space reconstruction, its role here is quite distinct. Our results suggest that time-delay coordinates are equivalent to the additional measurements that are required to stabilize the SM. This permits accurate state and parameter estimation, and from that, accurate prediction.

We work with a set of L-dimensional measurements $y_\ell(t_n)$, $\ell = 1, 2, \ldots, L$, that are made at each observation time $\{t_0, t_1, \ldots, t_n, \ldots, t_N = T\}$ within an observation window $[t_0, T]$. The physical model developed to describe this system has D state variables $x_a(t)$, $a = 1, 2, \ldots, D$, which satisfy the deterministic ordinary differential equations

$$\frac{dx_a(t)}{dt} = F_a\big(\mathbf{x}(t), \mathbf{p}\big), \tag{1}$$

with fixed parameters $\mathbf{p}$. The index a collects the vector label of state variables and any discretized spatial coordinates for underlying dynamics of fields satisfying physical partial differential equations. In general there are model errors that may be represented as stochastic contributions to these deterministic dynamics.

The output of the model $\mathbf{x}(t)$ is related to the observations $y_\ell(t)$ by L observation functions $h_\ell(\mathbf{x}(t))$, and if synchronization is achieved then $y_\ell(t) \approx h_\ell(\mathbf{x}(t))$ and a high quality estimation of parameters and unobserved states is possible, and prediction will likely be accurate.

We show that this goal is achieved in two examples of chaotic dynamical systems with sparse observations: (i) the Lorenz '96 model [26] with $D = 20$ and the same forcing in each dynamical degree-of-freedom, and (ii) the Lorenz '96 model with $D = 10$ and different forcing presented to each dynamical variable. This model is chosen because it has been shown that L_s is proportional to D, which may be chosen freely [15]. This makes Lorenz '96 an excellent testing ground for investigating the effects of insufficient measurements. In both examples we view the fixed parameters as extended state variables with trivial dynamics $d\mathbf{p}/dt = 0$, and we presume that among the many state variables, only one is observed; we call this measured variable $y_1(t)$ and associate it with $x_1(t) = h(\mathbf{x}(t))$, so $L = 1 < L_s$. Using this observed quantity along with its time-delays, this technique allows us to accurately estimate the unobserved state variables as well as all of the unknown parameters.

In a subsequent paper, we will report on the successful application of this method to a ring of three coupled classical Lorenz '63 models [27] and to the four-dimensional Rössler model of 'hyperchaos' [28]. These investigations will include the effects of additive noise in the measurement on the viability of the method. In this paper however, we focus solely on the Lorenz '96 results.

We generate the data from a known model, so we are performing 'twin experiments', yet we proceed as if the only information available to us is the time series of measurements $y_1(t)$. We establish synchronization on the SM by comparing the D_M-dimensional time-delay data vector $Y_k(t) = \{y_1(t + (k-1)\tau)\}$, $k = 1, 2, \ldots, D_M$ with the D_M-dimensional time-delay model vector $S_k(t) = \{x_1(t + (k-1)\tau)\}$, $k = 1, 2, \ldots, D_M$, through monitoring the synchronization error

$$\mathrm{SE}_s^2(t) = \frac{1}{D_M}\sum_{k=1}^{D_M}\big(y_1\big(t + (k-1)\tau\big) - x_1\big(t + (k-1)\tau\big)\big)^2. \tag{2}$$

The quality of the estimation is then evaluated by the prediction of time series measurements for times greater than the observation period $T = t_N$. Although in a twin experiment we may examine the quality of the method for estimating unobserved states and unknown parameters, here we restrict our evaluation of estimates and predictions to observed quantities alone – in this case $x_1(t)$ as that mimics experimental settings. The luxury of comparing the unobserved state values is not available in actual experiments.

From the model output time series $x_1(t)$ we augment the dimension from one to $D_M > 1$ by considering the D_M-dimensional time-delay measurement vector constructed from the physical state variables $\mathbf{x}$ as

$$\mathbf{S}\big(\mathbf{x}(t)\big) = \big\{x_1(t), x_1(t+\tau), \ldots, x_1\big(t + (D_M - 1)\tau\big)\big\}, \tag{3}$$

where τ is a suitably chosen time-delay, and $\mathbf{S}(\mathbf{x})$ denotes a map from the physical space to a delay embedding space which is in general different from delay embedding space used for state space reconstruction. We are using the phrase "time-delay" in a general sense here, but use a time advanced embedding for computational ease when computing the Jacobian of the delay embedding map (7). The data $y_1(t)$ are used to create a D_M-dimensional data vector

$$\mathbf{Y}(t) = \big\{y_1(t), y_1(t+\tau), \ldots, y_1\big(t + (D_M - 1)\tau\big)\big\}.$$

In this way, the model output is compared to the observed data by asking when $\mathbf{Y}(t) \approx \mathbf{S}(t)$ through the evaluation of the synchronization error (2).

One way to proceed is to form the map from physical space $\mathbf{x}$ to time-delay space $\mathbf{S}$ via the construction (3). By casting the overall dynamics (1) into $\mathbf{S}$ space, one can use the full statistical physics path integral formulation of the problem in $\mathbf{S}$ space [17].

We focus here on a different approach, in which we use the equations of motion to advance the state $\mathbf{x}$ in physical space while incorporating the information from the observations via $\mathbf{Y}$ appropriately coupled into the $\mathbf{x}$ dynamics. Although the forward map $\mathbf{S}(\cdot)$ is explicit in Eq. (3), the global inverse map $\mathbf{S}^{-1}(\cdot)$ from D_M-dimensional space to D-dimensional physical space is not explicitly given in general. However, we can use the *local* version of the forward map which involves the $D_M \times D$ Jacobian matrix $\frac{\partial S}{\partial x}(t)$ and its pseudo-inverse $\frac{\partial x}{\partial S}(t)$ as a local inverse map.

To achieve synchronization, we proceed through the observation window using the equations of motion in physical space $\mathbf{x}$ to advance the state between observation times t_n. At each of the t_n we introduce a local control term in time-delay space of the form $g_n\delta(t - t_n)(\mathbf{Y}(t) - \mathbf{S}(t))$, where g_n denotes a coupling constant at each t_n, and $\delta(\cdot)$ is a delta function switching the coupling on at measurement times. Using the inverse of the delay embedding map (3) we express this equation of motion in physical space as follows:

$$\frac{dx_a(t)}{dt} = F_a\big(\mathbf{x}(t)\big) + \sum_{n=0}^{N} g_n\delta(t - t_n)\mathbf{S}_a^{-1}\big(\mathbf{Y}(t) - \mathbf{S}(t)\big), \tag{4}$$

over the $N + 1$ observations in the data assimilation window $[t_0, t_N = T]$.

For small deviations a Taylor expansion of the control term results in the following set of ODEs

$$\frac{dx_a(t)}{dt} = F_a\big(\mathbf{x}(t)\big) + \sum_{n=0}^{N} g_n\delta(t - t_n)\sum_{k=1}^{D_M}\frac{\partial x_a}{\partial S_k}(t)\big(Y_k(t) - S_k(t)\big). \tag{5}$$

At times t_n, the Jacobian matrix $\frac{\partial x}{\partial S}(t) = \{\frac{\partial x_a}{\partial S_k}(t)\}$ of $\mathbf{S}^{-1}$ couples the information from the time-delay data into all dynamical variables of the model [29,30].

The coupling term in the differential equation drives the solution $\mathbf{x}(t)$, represented through $\mathbf{S}(\mathbf{x}(t))$, to the observations $\mathbf{Y}(t)$ in a stable manner when g_n is sufficiently large, and when D_M provides enough information about the data waveform to the model.