Detecting nonlinear dynamics in spatio-temporal systems, examples from ecological models

Sarah Little a,*, Stephen Ellner b, Mercedes Pascual a, Michael Neubert a, Daniel Kaplan c, Timothy Sauer d, Hal Caswell a, Andy Solow a

a Woods Hole Oceanographic Institution, Wood Hole, MA 02543, USA
b North Carolina State University
c McGill University
d George Mason University

Abstract

Mathematical models of marine populations exhibit chaotic dynamics. However, we hypothesize that in moving water, Eulerian sampling of spatially heterogeneous populations may obscure any deterministic signal beyond the resolving capabilities of presently available nonlinear signal processing techniques. To examine this hypothesis we created two spatio-temporal models of population dynamics. To caricature actual ocean sampling limitations, we sampled the model output in two ways, random walks to simulate Eulerian sampling, and spatial averages to simulate population measurements from finite volumes. Results indicate that the ability to identify underlying nonlinear dynamics quickly degrades as the step size of a random walk sampling increases. On the other hand, the analysis techniques used are more robust in the face of spatial averaging.

Keywords: Nonlinear dynamics; Marine ecology; Spatio-temporal chaos; Sampling

1. Introduction

In this paper we examine how noise introduced by limitations in sampling of a chaotic spatio-temporal system affects the ability to decipher underlying nonlinear dynamics. This study arose out of work on sampling of marine ecosystems, but the results are extendable to spatio-temporal systems in general.

In the ocean, complex spatio-temporal interactions between predators, prey, and nutrients, coupled with ocean physics, govern the population dynamics of everything from phytoplankton to whales. Understanding the behavior of populations of oceanic organisms has great practical significance ranging from managing commercial fisheries such as cod and salmon, to controlling blooms of toxic plankton. However, the simple act of taking meaningful, long-term, spatio-temporal measurements of populations in a moving fluid is a formidable task itself. First, all point samples of a population must be spatial averages of some sort. Second, long-term ocean sampling is most easily accomplished in the Eulerian sense, that is, in a reference frame fixed with respect to the moving fluid, such as from a moored buoy or a pier. For an externally advected population, this effectively means...
that temporally successive samples are not taken from the same spatial location with respect to the spatial pattern of the population. The effects of these two sampling limitations on the ability to detect nonlinear deterministic behavior are addressed in this paper by applying several nonlinear time series analysis techniques to the output of two different spatio-temporal population models.

In theoretical ecology, there are many examples of temporal population models which exhibit chaos. The interaction of three variables in a predator–prey–nutrient system [1] is a well studied chaotic system, as is the simple logistic map [2]. Examples of spatio-temporal models are less common, but include population models in discrete time and space [3,4], and discrete time and continuous space [5].

The tests conducted in this paper use a one-dimensional continuous time and space model of a predator–prey–resource interaction [6], and a one-dimensional discrete time and space model of a single species; both models exhibit chaotic dynamics. Time series representing spatial averaging are generated by taking spatial averages of constant length and location from the output of the models. Time series which caricature Eulerian sampling are generated by sampling at points determined by a random walk in one-dimensional space. These series are compared to a time series of samples from a single spatial location.

The tests used in this paper utilize several recently developed methods [7–9] which are intended for situations where the data are scarce, the functional form of the underlying process is unknown, and noise may be present. The methods assay different aspects of nonlinear dynamics – predictability, continuity, and sensitive dependence on initial conditions. The purpose of this paper is not to compare methods, but to determine the relative effects of Eulerian sampling vs. spatial averaging on our ability to detect that the underlying dynamics are deterministic spatio-temporal chaos. Because the methods are new and their performance on spatio-temporal data has not been examined, we chose several methods so that consistent answers from all methods will give us confidence in the results.

These methods, and most other nonlinear dynamics signal processing ones, are based on phase space reconstruction using delay coordinate embedding

\[ V(t_i) = (x_i, x_{i+r}, x_{i+2r}, \ldots, x_{i+(d-1)r}), \]

where \( V \) is the embedded vector, \( x(t) \) the measured variable, \( d \) the embedding dimension, \( \tau \) the time lag, and \( i \) the index for time \( t \).

Here the methods diverge, [8,7] tests approach the problem of detecting nonlinear determinism with a two step process. First, they develop a measure of the data which will be sensitive to determinism. Their techniques capitalize on the geometric structure of the attractor in phase space which is the hallmark of deterministic systems. On the one hand, the complexity of this geometric structure is determined by the dynamics, and on the other, noise will tend to destroy the smoothness of the geometric structures. The techniques for measuring determinism here work by looking at the smoothness of local neighborhoods in phase space of the reconstructed attractor, the idea being that the smoother the structure, the more likely that it arose from a deterministic process and not a stochastic one.

Secondly, they test whether the determinism arises linearly or nonlinearly. They use the determinism measure to compare the data with linear surrogates which have the same power spectrum as the data but have a random phase spectrum [10]. The surrogates are representative of a linear random process with the same autocorrelation as the data. If a difference is found between the data and their linear surrogates, then the conclusion is drawn that the data do not arise from a linear stochastic system. Importantly though, the test does not rule out the situation where a deterministic process underlies the data but the data have a large noise component, nor does it rule out that the data arose from a nonlinear stochastic process.

The Nychka, Ellner, Gallant, McCaffrey test ([9,11]) (hereafter NEGM) approaches the question of detecting nonlinear determinism by measuring the Lyapunov exponent of a model fitted to the data. Because ecological time series typically have fewer than 1000 points, it is impossible to reliably estimate Lyapunov exponents from data directly. The approach generally taken, then, is to fit a nonlinear model to
the data and calculate the exponents directly from the model. The main pitfall of this approach is that the results can depend entirely on the choice of model. The issues of fitting an appropriate model to a data set are not trivial, and have been addressed in [9].

Analyses of population dynamics data have often been based on fitting the data with simple models having a few adjustable parameters (e.g. stock-recruitment models of fish population). The methods used in this paper are all “model-free” (i.e. “nonparametric”) in the sense of not presuming to know the underlying mechanisms or the form of the underlying equations. The former can be criticized because the choice of a specific model may strongly bias the outcome. Model-free or nonparametric methods have fewer assumptions, and their assumptions are usually qualitative rather than quantitative. Implemented carefully, such methods are flexible enough to avoid biases without requiring excessive amounts of data. However, model-free and nonparametric methods are nonmechanistic, so it is difficult to incorporate any information apart from the time-series population sizes. Deriving hybrid semi-mechanistic models, which use a mechanistic “skeleton” insofar as possible, incorporating what is known about the biology generating the dynamics, and using lags and nonparametrics to deal with areas of ignorance, is an area of current research.

2. Spatio-temporal population models

2.1. Continuous, spatially heterogeneous model

Pascual [6] studied the dynamics of a diffusive predator-prey model in one-dimensional space. Spatial heterogeneity is introduced by letting one of the parameters, the prey’s growth rate, vary linearly with space. In the absence of diffusion, the model is a standard predator-prey system, which exhibits limit cycles or equilibria [12]. The fixed (in time) spatial heterogeneity in prey resource (such as light or nutrients) induces chaos in this model.

In the model, both species diffuse at the same constant rate $D$. At any point $x$ and time $t$, the resource gradient $r$, and the population dynamics of the prey $p$, and predator $h$, are given by the following dimensionless reaction-diffusion model with logistic growth of the prey and a type II [13,14] functional response of the predator:

$$\frac{\partial p}{\partial t} = r(x)p(1 - p) - \frac{aph}{1 + bp} + D \frac{\partial^2 p}{\partial x^2}, \tag{2}$$

$$\frac{\partial h}{\partial t} = \frac{aph}{1 + bp} - mh + D \frac{\partial^2 h}{\partial x^2}, \tag{3}$$

$$r(x) = 2 - 1.4x, \tag{4}$$

where the boundary conditions are zero flux at $x = 0$ and $x = 1$. The model is started with predator and prey uniformly distributed along the spatial domain. The coefficients are $a = 5$, $b = 5$, $m = 0.6$, and $D = 10^{-4}$. With these parameter values, the resulting dynamics are chaotic in time [6]. The weak coupling of oscillators with different amplitudes and frequencies along the gradient leads to chaos in this model. The spatial heterogeneity of prey resource induces an additional effect on the dynamics: the temporal history of the model at the high end of the gradient is more regular than at the low end. This can be seen in Fig. 1 where the time axis shows regular cycles near $x = 0$ (high resource), and irregular oscillations near $x = 1$ (low resource).
The model is run for 1000 time units (after transients have died out) with the numerical scheme used in [6]. This scheme combines a fully implicit method for the diffusion terms, with a fourth-order Runge–Kutta method for the predator–prey interaction terms. The domain is on a 100 point spatial grid (from $x = 0$ to $x = 1$), and the model generates an output data matrix of 100 spatial points by 1000 time steps (Fig. 1). A control time series is generated by extracting all the time points from one spatial point in the matrix (a vector 1000 points long). Test time series are generated by point sampling in time the spatial axis of this matrix with random walks, and by taking time series of spatial averages at fixed locations. Each resulting test time series is a vector 1000 steps long.

The coherence length of the spatial pattern is calculated from the distance at which the autocorrelation of the spatial pattern first crosses zero. The mean coherence length, an average of the coherence length of the spatial pattern at each time step, is 20 points long. The random walk’s step sizes are all less than one quarter of this length, while the spatial averages start at approximately one quarter and increase to five times this length.

Five time series of randomly sampled locations are generated with random walk step sizes of 1, 2, 3, 4, and 5 spatial points. The random walk sequence is generated with a given step size, and this time series of spatial locations is used to pick points out of the model output matrix. Note that even for the smallest random walk step size of 1 point, the entire 100 point domain may be sampled by the time 1000 time steps have been completed. Seven time series of spatial averages are generated by averaging lengths of 6, 10, 20, 30, 50, 75 and the entire domain of 100 points. Selected examples of the time series generated by this model are shown in Fig. 2.

2.2. Discrete, spatially homogeneous model

The second model is a one-dimensional coupled map lattice which describes the growth and dispersal of a single population over a finite number of habitat patches (see e.g. [15,16]). In each patch the population behaves as a map with density dependent growth, that is, the growth rate is dependent on the population size. At each time step, the population at each point is allowed to disperse over a finite spatial kernel which
The dispersal kernel \( K(s - j) \) (graphed in Fig. 3) gives the fraction of the population in patch \( j \) which disperses to patch \( s \):

\[
K(s - j) = 0.2236 \left( \frac{2}{3} \right)^{|s-j|}, \quad |s - j| \leq 5,
\]

therefore \( \sum_s K(s) = 1. \) (6)

The density dependent growth in each patch is given by the Ricker curve \cite{17} \( f(n) \):

\[
f(n) = ne^{4.1(1 - n)}. \]  

This model exhibits chaos. Unlike the first model, this model is spatially homogeneous; the parameters are independent of position.

As for the first model, this model is run for 1000 time steps (after transients have died out), but with a spatial grid of 512 points, this generates a matrix of 512 spatial points by 1000 time steps. A control time series is generated by sampling at one spatial point in the matrix. Test data are generated, as above, by sampling this matrix with random walks, and with spatial averages.

The mean coherence length of the spatial pattern generated by this model is about 10 points long. The random walks’ step sizes are all less than this length, while the spatial averages start at approximately one quarter and increase to fifty times this length.

Four time series of randomly sampled locations are generated with random walk step sizes of 1, 2, 4 and 8 spatial points. The random walk sequence is generated...
3. Nonlinear analysis methods

3.1. A measure of determinism

To determine whether the spatial sampling schemes obscured the underlying determinism in our models, we used several techniques, three based on the concept of prediction, and one on the concept of continuity.

3.1.1. Predictability

Predictability is a comparison between the output of a prediction algorithm, and “out of sample” data points, the data not used by the algorithm. The general prescription for the prediction algorithms utilizes local neighborhoods of points in phase space and a variety of methods based on this have been published e.g. [18–21,7]. The time evolution of points on the attractor will be smooth in local neighborhoods if no noise is present in the system. Close trajectories will stay close in local regions. This property is used for prediction in the following ways.

For every point \( V_i \) in phase space, find its \( k \)-nearest neighbors, where \( k \) is a tunable parameter which usually is about 1–10% of the length of the data set. These neighbors are used in various ways to generate a prediction of one time step ahead.

One of the simplest prediction methods is to find the center of mass of the neighbors after one time step and compare it to the point \( V_i \) after one time step, \( v_{i+1} \), as described in Fig. 5. In other words, for each time \( i \) in the embedded time series \( V_i \), we calculate a predicted value at time \( i + 1 \) (or, more generally, \( i + h \), where \( h \) is a “prediction horizon”) in the following way. First, find the \( k \)-nearest neighbors to the embedded point \( V \) at time \( i \). Move each neighbor ahead one time step and find the center of mass of these images of the neighbors. This gives us a “predicted value” called \( y_{i+1} \).
Fig. 5. Center of mass prediction method used for Kaplan’s predictability in two dimensions. A discretely sampled time series is embedded using delay coordinates. For each point \( V(i) \), a prediction is made using the center of mass of the images of its nearest neighbors one time step ahead. The difference between this predicted point, \( y(i + 1) \), and \( V(i + 1) \), is the prediction error for that point.

Fig. 6. Local linear map prediction method on which Sauer’s method is based. A linear map is calculated, using least squares approximation, from the nearest neighbors to point \( V(i) \), to their images one time step ahead. This map is then applied to the point \( V(i) \) to get a predicted point \( y(i + 1) \). The prediction error for that point is the difference between \( y(i + 1) \) and \( V(i + 1) \).

which we compare to the actual value, \( V_{i+1} \). If the geometry is smooth, the distance between \( V_{i+1} \) and \( y_{i+1} \) will be small, and the prediction will be good. If the geometry is noisy, or highly chaotic, then the prediction will be worse. We evaluate the overall quality of prediction of the whole data set and get a measure of prediction error by looking at the root mean square (RMS) value of \( V_{i+1} - y_{i+1} \) for all \( i \). This analysis was performed using an embedding dimension of 3, and time lag of 1, using 5 nearest neighbors. This method is referred to as Kaplan predictability and the results are presented as the log of the RMS value of the prediction error.

In a slightly more complicated version of this [7], a linear mapping is fit in a least squares sense from the neighbors to their future values (typically one time step ahead) (see Fig. 6). This mapping then is applied to the point \( V_i \), and the resulting prediction \( y_{i+1} \) is compared to the actual \( V_{i+1} \). This method uses further refinements for noise reduction by performing local singular value decomposition (SVD) on the neighbors before calculating the mapping. In this process, the \( d \)-dimensional neighbors are projected onto a smaller \( (m + 1) \)-dimensional subdomain of the embedded space. The purpose of this is to reduce the unwanted effects of noise by keeping only the \( (m + 1) \) highest singular values. By taking more modes, or principal components, the method allows for greater flexibility in reducing noise and keeping signal. For very noisy data, few modes usually produce better predictions, while for clean data, retaining more modes produces better predictions. In this paper, all modes were calculated, but the results did not differ significantly so only \( m = 3 \) is presented. This analysis is
performed using embedding dimension 4, and time lag
of 1, and 10–35 nearest neighbors (depending on the
data set).

A further variation is used, NEGM, where a
weighted average is calculated of the neighbors one
time step ahead, using weights proportional to the
distances of the original neighbors to the point to be
predicted.

\[ y_{i+1} = \sum_{j} w_{j,i} V_{j+1}, \]

(8)

where \( y_{i+1} \) is the predicted value, \( w_{j,i} \) are the weights
calculated from distances between neighbors \( V_j \) to the
point \( V_i \), and \( V_{j+1} \) are the neighbors at one time step
ahead.

The advantage of this method is that nearby neigh-
bors, which should be better predictors, are weighted
more heavily. This method is called a kernel regres-
sion model [22]. Rather than comparing the prediction
to linear surrogates, however, NEGM uses ordinary
cross-validation [22] to find the optimum prediction
while varying the parameters of embedding dimen-
sion, time delay, and the weighting function. NEGM
uses this method to arrive at an optimum embedding
dimension which is then used in a program to esti-
mate the maximum Lyapunov exponent of a time se-
ries [9]. This analysis was performed with embedding
dimension 1–9, depending on the data set, and time
lag of 1.

### 3.1.2. Continuity

Continuity of trajectories in phase space can be used
as a test for smoothness. The idea behind this measure
is that in a deterministic system, the smooth geom-
etry is created by trajectories which stay close to each
other for (at least) short periods of time. In a noisy
system, the trajectories will diverge more rapidly. So
smoothness is quantified by looking at how rapidly
nearby points in phase space diverge. Continuity is
measured as follows [8]: For a deterministic system,
if two points, at time \( i \) and \( j \), are close together in
the embedding space, then their future values one time
step ahead, \( V_{i+1} \) and \( V_{j+1} \) should also be close to-
gether. By examining a plot of \( |V_{i+1} - V_{j+1}| \) (called
\( \varepsilon_{i,j} \)) vs. the distance between \( V_i \) and \( V_j \) in the embed-
ding space (called \( \delta_{i,j} \)), we can assess the amount of
determinism. On a plot of \( \varepsilon_{i,j} \) vs. \( \delta_{i,j} \), determinism
appears when \( \varepsilon_{i,j} \) tends to zero for those \( i, j \) where
\( \delta_{i,j} \) is small.

### 3.2. Linear surrogates

Surrogate data are realizations of a random process
with the same histogram and similar autocorrelation
(and power spectrum) as the original data. Theiler et
al. [10] describe Fourier transform methods for gen-
erating surrogate data. We used "amplitude adjusted"
surrogates [10, algorithm II] which tests the hypoth-
esis that the data are a nonlinear transformation of a
Gaussian linear stochastic process. We constructed ten
realization of the surrogates. If the prediction error of
the data is at least two standard deviations below the
mean prediction error of the surrogates, we conclude
that a nonlinear component is present in the underly-
ing dynamics.

### 3.3. Lyapunov exponent estimation

A nonmechanistic model for calculating the largest
Lyapunov exponent is used in this paper:

\[ x_{i+h} = f(x_i, x_{i-\tau}, x_{i-2\tau}, \ldots, x_{i-(d-1)\tau}) + \epsilon_i, \]

(9)

where \( h \) is the prediction horizon, \( \epsilon_i \) is a noise term,
and \( f \) is a statistical regression model to be fit to the
data \( x \) [9]. The analyses here used the feedforward
neural network model for \( f \) [23].

This method is often quite successful for model-
ling nonlinear and chaotic time series data e.g. [24].
In particular, the feedforward neural network is robust
against errors in choosing the embedding dimension
[23]. The model used explicitly allows for a random
component in the data. In such systems with exoge-
nous noise, Lyapunov exponents are calculated from
the derivatives of the estimated map \( f \), essentially by
proceeding as if \( f \) were the true map but evaluating
the derivatives along the observed trajectory (note that
this is not the same as computing the Lyapunov ex-
ponent of the model (Eq. (9))).
4. Model results

The different methods give extremely consistent results when applied to time series from the spatio-temporal models. In the controls, they all find high predictability and detect nonlinear deterministic behavior. In the random walk sampled data, they all find that the predictability and the ability to detect underlying nonlinear determinism falls rapidly as step size increases. In the spatially averaged data they all find that predictability and the ability to detect nonlinear determinism remains high even as averaging length increases beyond the spatial coherence length.

To compare consistency among techniques, the application of all the techniques to a subset of time series from only the continuous spatio-temporal model is qualitatively summarized in Table 1. The first three rows simply report the relative level of predictability measured. For each of three methods, the controls and spatial averages are highly predictable, while the random walks degrade rapidly. The next three rows report the comparison of predictability between the time series and their linear random surrogates. Again, for all methods the controls and the spatial averages are distinguishable from their linear surrogates, meaning that the underlying nonlinear deterministic model has been detected. The random walks series, however, are indistinguishable from their linear surrogates, except for Kaplan’s predictability method which is able to detect nonlinear deterministic behavior in the smaller of the two random walk step sizes. The last row, the results of NEGM Lyapunov calculations, confirm the previous results, and show positive Lyapunov exponents for the control and spatial averages, but not for the random walk samples.

Secondly, to compare consistency between models, the application of the simple center of mass prediction technique to all the time series from both models is presented in Figs. 7 and 8. On both models, the control time series from a single spatial location is easily distinguishable from its linear surrogates, as evidenced by the large separation between predictability estimates for controls vs. surrogates. On both models, the random walk sequences tend towards lower predictability and smaller separation between the predictability estimates for the times series and their surrogates. The details of this trend are different between models, with separation remaining out to the largest random walk step sizes for the discrete model but not for the continuous one. On both models the prediction estimates from spatial averages are distinguishable from their linear surrogates, except for the spatial average of the entire domain for the discrete model.

Table 1
Summary of results of methods applied continuous model

<table>
<thead>
<tr>
<th>Predictability, Kaplan</th>
<th>Predictability, Sauer</th>
<th>Predictability, Ellner</th>
<th>Linear surrogates, Kaplan continuity</th>
<th>Linear surrogates, Sauer predictability</th>
<th>Linear surrogates, Kaplan predictability</th>
<th>Lyapunov exponent, NEGM</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
<td>NL/DE</td>
<td>NL/DE</td>
<td>NL/DE</td>
<td>POS</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Indist.</td>
<td>NL/DE</td>
<td>POS</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>POS</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>POS</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>POS</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>POS</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>POS</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>POS</td>
</tr>
</tbody>
</table>
Fig. 7. Results of the application of Kaplan's center of mass prediction method to the sampled output of the continuous population model. The x-axis is a measure of distance: for the random walk it is step size, for the spatial averages it is averaging length. The prediction error for the control sample from a single point location is statistically distinguishable from that of its linear surrogates. For the random walk samples, the difference between data and their surrogates drops as step size increases, although not monotonically. Since the random walks are random samples, some may be less scrambled than others, and it is not surprising to see this kind of nonmonotonic behavior. For the spatially averaged samples, the overall prediction error is much lower for both data and surrogates, however, the difference between them remains statistically significant even out to a spatial average of the entire domain of 100 points.

It is clear that random walks through these spatio-temporal systems generate time series that quickly approach stochastic noise, while spatial averages preserve the character of the dynamics.

Since the methods used in this paper are searching for smoothness in phase space, it is instructive to look at phase portraits of the data. Fig. 9 presents phase portraits for the sampled model output. It is quite clear from these simple plots that (1) structure is present for spatial averages, and (2) structure is quickly lost for random walks.

5. Discussion

It is clear from the results of the model test that random walk sampling degrades the ability to detect nonlinear determinism much more rapidly than spatial averaging does. All the methods used in this paper agree with this finding.

The spatial average sample is a simple linear combination of spatial elements from the spatio-temporal system. This transformation should then, in theory, retain the dynamical information of the system. In some sense the spatial averaging is just another embedding, but in the spatial domain of a spatio-temporal system the effects of this embedding are not well understood. On the one hand, if there is noise present, this averaging will reduce noise levels and make detection of dynamics easier. On the other hand, spatial averaging also can reduce the signal level, as out-of-phase portions of space cancel each other (this can be seen in both Figs. 1 and 3, where in the noise-free case of spatial averages, the signal level drops as the spatial average length increases). In this study, an average of the entire spatial domain for the continuous model retained its underlying nonlinear deterministic structure, while an average of the entire spatial domain of the discrete model did not. The competing
Fig. 8. Results of the application of Kaplan's center of mass prediction method to the sampled output of the discrete population model. The x-axis is a measure of distance: for the random walk it is step size, for the spatial averages it is averaging length. The overall predictability is lower than for the continuous model. Nevertheless, the prediction error for the control sample from a single point location is statistically distinguishable from that of its linear surrogates. For the random walk samples, the difference between data and their surrogates drops as step size increases, although not monotonically, however, it is not surprising to see this kind of nonmonotonic behavior in random samples. For the spatially averaged samples, the overall prediction error is lower for data and slightly lower for the surrogates. The difference between data and surrogates remains statistically significant for spatial averages up to 17 points. When the entire spatial domain is averaged, the time series is indistinguishable from its linear stochastic surrogates.

The random walk sampling is quite different. In the cases studied here, the individual step size is typically small compared to the correlation length of the spatial pattern. At each step, the random walk sampling adds a random component to the data whose time correlation is dependent on the correlation of spatial patterns. The derivative of the random walk sampled time series is a combination of the local time derivative and spatial derivative. This added complexity quickly degrades the ability of the techniques used here to detect the underlying deterministic dynamics.

For the heterogeneous model, the random walk adds additional complexity, because over many time steps the spatial location of the sample can become quite far from the original sample point, far enough away to be in dynamic regimes that are apparently different. For the spatially heterogeneous case, although the form of the equations for the system is the same everywhere, certain locations are more regular than others (see for example Fig. 1, where the time history at spatial point 0 is nearly periodic, while that of spatial point 1 is quite aperiodic). The random walk sample goes between these, and over the long term the resultant signal is a scrambled combination of the different regimes. So the random walk sampling can add two levels of complexity to the data, local mixing of time and space derivatives, and slow drift into different dynamic regimes.

The ramifications of these results to uncovering nonlinear dynamics in the ocean are severe. They imply that the usual Eulerian sampling of the ocean may
be a very poor method to collect data for detecting underlying nonlinear deterministic population dynamics. A more general conclusion may also be drawn for experiments in other fields. When it is not possible to compensate for the movement of a spatio-temporal system in relation to the measurement point, it is better to take fixed spatial averages rather than moving individual point samples.

For oceanographic data, this amounts to obtaining spatial averaging of large sections of populations, for example using satellite images of plankton populations. Alternatively, drifting (Lagrangian)
measurements should be obtained, where the sample is taken from approximately the same point in the population each time step.

Acknowledgements

This work arose out of a workshop on Nonlinear Data Analysis in Marine Ecology held at the Woods Hole Oceanographic Institution in April 1994, sponsored by the Office of Naval Research. We thank the participants for discussions and ideas generated during the workshop. This work was funded in part by ONR grant number N00014-92-J-1527. WHOI contribution number 9095.

References