*Kalman Filtering and Neural Networks*, Edited by Simon Haykin

ISBNs: 0-471-36998-5 (Hardback); 0-471-22154-6 (Electronic)

# 4

# CHAOTIC DYNAMICS

Gaurav S. Patel

*Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada*

Simon Haykin

*Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada*
(haykin@mcmaster.ca)

## 4.1 INTRODUCTION

In this chapter, we consider another application of the extended Kalman filter recurrent multilayer perceptron (EKF-RMLP) scheme: the modeling of a chaotic time series or one that could be potentially chaotic.

The generation of a *chaotic process* is governed by a coupled set of nonlinear differential or difference equations. The hallmark of a chaotic process is sensitivity to *initial conditions*, which means that if the starting point of motion is perturbed by a very small increment, the deviation in

**Table 4.1 Summary of data sets used in the study**

| | Network size | Training length | Testing length | Sampling frequency $f_s$(Hz) | Largest Lyapunov exponent $\lambda_{\max}$ (nats/sample) | Correlation dimension $D_{ML}$ |
|---|---|---|---|---|---|---|
| Logistic | 6-4R-2R-1 | 5,000 | 25,000 | 1 | 0.69 | 1.04 |
| Ikeda | 6-6R-5R-1 | 5,000 | 25,000 | 1 | 0.354 | 1.51 |
| Lorenz | 3-8R-7R-1 | 5,000 | 25,000 | 40 | 0.040 | 2.09 |
| $NH_3$ laser | 9-10R-8R-1 | 1,000 | 9,000 | 1[a] | 0.147 | 2.01 |
| Sea clutter | 6-8R-7R-1 | 40,000 | 10,000 | 1000 | 0.228 | 4.69 |

[a]The sampling frequency for the laser data was not known. It was assumed to be 1 Hz for the Lyapunov exponent calculations.

the resulting waveform, compared to the original waveform, increases exponentially with time. Consequently, unlike an ordinary deterministic process, a chaotic process is predictable only in the short term.

Specifically, we consider five data sets categorized as follows:

- The logistic map, Ikeda map, and Lorenz attractor, whose dynamics are governed by known equations; the corresponding time series can therefore be numerically generated by using the known equations of motion.
- Laser intensity pulsations and sea clutter (i.e., radar backscatter from an ocean surface) whose underlying equations of motion are unknown; in this second case, the data are obtained from real-life experiments.

Table 4.1 shows a summary of the data sets used for model validation. The table also shows the lengths of the data sets used, and their division into the training and test sets, respectively. Also shown is a partial summary of the dynamic invariants for each of the data sets used and the size of network used for modeling the dynamics for each set.

## 4.2 CHAOTIC (DYNAMIC) INVARIANTS

The *correlation dimension* is a measure of the complexity of a chaotic process [1]. This chaotic invariant is always a fractal number, which is one reason for referring to a chaotic process as a "strange" attractor. The other

chaotic invariants, the *Lyapunov exponents*, are, in part, responsible for sensitivity of the process to initial conditions, the occurrence of which requires having at least one positive Lyapunov exponent. The *horizon of predictability* (HOP) of the process is determined essentially by the largest positive Lyapunov exponent [1]. Another useful parameter of a chaotic process is the *Kaplan–York dimension* or *Lyapunov dimension*, which is defined in terms of a Lyapunov spectrum by

$$D_{KY} = K + \frac{\sum_{i=1}^{K} \lambda_i}{|\lambda_{K+1}|}, \tag{4.1}$$

where the $\lambda_i$ are the Lyapunov exponents arranged in decreasing order and $K$ is the largest integer for which the following inequalities hold

$$\sum_{i=1}^{K} \lambda_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{K+1} \lambda_i < 0.$$

Typically, the Kaplan–Yorke dimension is close in numerical value to the correlation dimension. Yet another byproduct of the Lyapunov spectrum is the *Kolmogorov entropy*, which provides a measure of information generated due to sensitivity to initial conditions. It can be calculated as the sum of all the positive Lyapunov exponents of the process. The chaotic invariants were estimated as follows:

1. The correlation dimension was estimated using an algorithm based on the method of maximum likelihood [2] – hence the notation $D_{ML}$ for the correlation dimension.
2. The Lyapunov exponents were estimated using an algorithm, involving the QR - decomposition applied to a Jacobian that pertains to the underlying dynamics of the time series.
3. The Kolmogorov entropy was estimated directly from the time series using an algorithm based on the method of maximum likelihood [2] – hence the notation $KE_{ML}$ for the Kolmogorov entropy so estimated. The indirect estimate of the Kolmogorov entropy from the Lyapunov spectrum is denoted by $KE_{LE}$.

## 4.3 DYNAMIC RECONSTRUCTION

The attractor of a dynamical system is constructed by plotting the evolution of the state vector in state space. This construction is possible when we have access to every state variable of the system. In practical situations dealing with dynamical systems of unknown state-space equations, however, all that we have available is a set of measurements taken from the system. Given such a situation, we may raise the following question: Is it possible to reconstruct the attractor of a system (with many state variables) using a single time series of measurements? The answer to this question is an emphatic yes; it was first illustrated by Packard et al. [3], and then given a firm mathematical foundation by Takens [4] and Mañé [5]. In essence, the celebrated *Takens embedding* theorem guarantees that by applying the *delay coordinate method* to the measurement time series, the original dynamics could be reconstructed, under certain assumptions. In the delay coordinate method (sometimes referred to as the method of delays), delay coordinate vectors are formed using time-delayed values of the measurements, as shown here:

$$\mathbf{s}(n) = [s(n), s(n-\tau), \ldots, s(n-(d_E-2)\tau), s(n-(d_E-1)\tau)]^T,$$

where $d_E$ is called the *embedding dimension* and $\tau$ is known as the *embedding delay*, taken to be some suitable multiple of the sampling time $\tau_s$. By means of such an embedding, it is possible to reconstruct the true dynamics using only one measurement. Takens' theorem assumes the existence of $d_E$ and $\tau$ such that mapping from $s(n)$ to $s(n+\tau)$ is possible. The concept of dynamic reconstruction using delay coordinate embedding is very elegant, because we can use it to build a model of a nonlinear dynamical system, given a set of measured data on the system. We can use it to "reverse-engineer" the dynamics, i.e., use the time series to deduce characteristics of the physical system that was responsible for its generation. Put it another way, the reconstruction of the dynamics from a time series is in reality an *ill-posed inverse problem*. The direct problem is: given the dynamics, describe the iterates; and the inverse problem is: given the iterates, describe the dynamics. The inverse problem is ill-posed because, depending on the quality of the data, a solution may not be stable, may not be unique, or may not even exist. One way to make the problem well-posed is to include prior knowledge about the input–output mapping. In effect, the use of delay coordinate embedding inserts some prior knowledge into the model, since the embedding parameters are determined from the data.

To estimate the embedding delay $\tau$, we used the method of mutual information proposed by Fraser [6]. According to this method, the embedding delay is determined by finding the particular delay for which the mutual information between the observable time series and its delayed version is minimized for the first time. Given such an embedding delay, we can construct a delay coordinate vector whose adjacent samples are as statistically independent as possible.

To estimate the embedding dimension $d_E$, we use the method of false nearest neighbors [1]; the embedding dimension is the smallest integer dimension that unfolds the attractor.

## 4.4 MODELING NUMERICALLY GENERATED CHAOTIC TIME SERIES

### 4.4.1 Logistic Map

In this experiment, the EKF-RMLP-based modeling scheme is applied to the logistic map (also known as the quadratic map), which was first used as a model of population growth. The logistic map is described by the difference equation:

$$x(k+1) = \alpha x(k)[1 - x(k)], \tag{4.2}$$

where the nonlinearity parameter $\alpha$ is chosen to be 4.0 so that the produced behavior is chaotic. The logistic map exhibits deterministic chaos in the interval $\alpha \in (3.5699, 4]$. An initial value of $x(0) = 0.5$ was used, and 35,000 points were generated, of which the first 5000 points were discarded, leaving a data set of 30,000 samples. A training set, consisting of the first 5000 samples, was used to train an RMLP on a one-step prediction task by means of the EKF method. An RMLP configuration of 6-4R-2R-1, which has a total of 61 weights including the bias terms, was selected for this modeling problem. The training converged after only 5 epochs and a sufficiently low MSE was achieved as shown in Figure 4.1.

***Open-Loop Evaluation*** A test set, consisting of the unexposed 25,000 samples, was used to evaluate the performance of the network at the task of one-step prediction as well as recursive prediction. Figure 4.2*a* shows the one-step prediction performance of the network on a short portion of the test data. It is visually observed that the two curves are

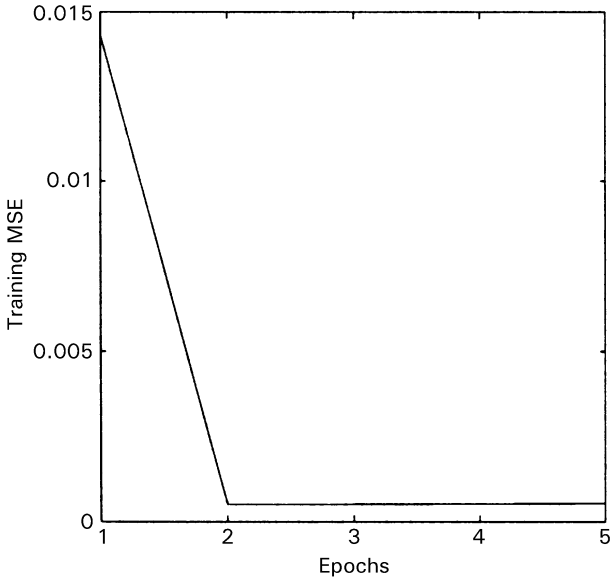



**Figure 4.1** Training MSE versus epochs for the logistic map.

almost identical. Also, for numerical one-step performance evaluation, signal-to-error ratio (SER) is used. This measure, expressed in decibels, is defined by

$$\text{SER} = 10 \log_{10} \frac{\text{MSS}}{\text{MSE}}, \tag{4.3}$$

where MSS is the mean-squared value of the actual test data and MSE is the mean-squared value of the prediction error at the output. MSS is found to be 0.374 for the 25,000-testing sequence and MSE is found to be $1.09 \times 10^{-5}$ for the trained RMLP network prediction error. This gives an SER of 45.36 dB, which is certainly impressive because it means that the power of the one-step prediction error over 25,000 test samples is many times smaller than the power of the signal.

***Closed-Loop Evaluation*** To evaluate the autonomous behavior of the network, its node outputs are first initialized to zero, it is then seeded with points selected from the test data, and then passed through a *priming phase* where it operates in the one-step mode for $p_l = 30$ steps. At the end of priming, the network's output is fed back to its input, and autonomous

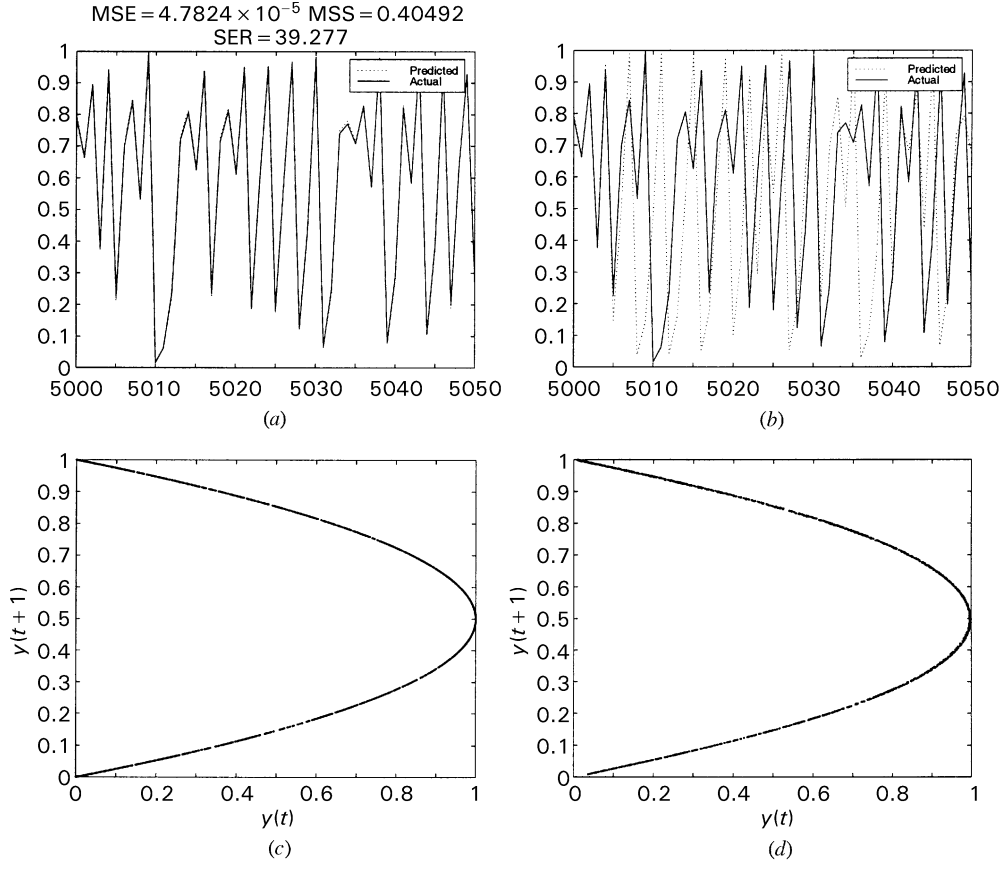



**Figure 4.2** Results for the dynamic reconstruction of the logistic map. (*a*) One-step prediction. (*b*) Iterated prediction. (*c*) Attractor of original signal. (*d*) Attractor of iteratively reconstructed signal.

operation begins. At this point, the network is operating on its own without further inputs, and the task that is asked of the network is indeed challenging. The autonomous behavior of the network, which begins after priming, is shown in Figure 4.2*b*, and it is observed that the predictions closely follow the actual data for about 5 steps on average [which is close to the theoretical horizon of predictability (HOP) of 5 calculated from the Lyapunov spectrum], after which they start to deviate significantly. Figure 4.3 plots the one-step prediction of the logistic map for three different starting points

The overall trajectory of the predicted signal, in the long term, has a structure that is very similar to the actual logistic map. The similarity is clearly seen by observing their attractors, which are shown in Figures 4.2*c* and 4.2*d*. For numerical autonomous performance evaluation, the dynamical invariants of both the actual data and the model-generated data are compared in Table 4.2. For the logistic map, $d_L = 1$; it therefore has only one Lyapunov exponent, which happens to be 0.69 nats/sample. This means that the sum of Lyapunov exponents is not negative, thus violating one of the conditions in the Kaplan–Yorke method, and it is for this reason that the Kaplan–Yorke dimension $D_{KY}$ could not be calculated. However, by comparing the other calculated invariants, it is seen that the Lyapunov exponent and the correlation dimension of the two signals are in close agreement with each other. In addition, the Kolmogorov entropy values for the two signals also match very closely. The theoretical horizons of predictability of the two signals are also in agreement with each other. These results demonstrate very convincingly that the original dynamics have been accurately modeled by the trained RMLP. Furthermore, the robustness of the model is tested by starting the predictions from various locations on the test data, corresponding to indices of $N_0 = 3060$, 5060, and 10,060. The results, shown in Figure 4.4, clearly indicate that the RMLP network is able to reconstruct the logistic series beginning from any location, chosen at random.

**Table 4.2 Comparison of chaotic invariants of logistic map**

| Time series | $d_E$ | $\tau$ | $d_L$ | $D_{ML}$ | $D_{KY}$ | $\lambda_1$ (nats/sample) | $KE_{LE}$ (nats/sample) | $KE_{ML}$ (nats/sample) | HOP (samples) |
|---|---|---|---|---|---|---|---|---|---|
| Actual logistic | 6 | 5 | 1 | 1.04 | —[a] | 0.69 | 0.69 | 0.64 | 5 |
| Reconstructed | 6 | 12 | 1 | 1.00 | —[a] | 0.61 | 0.61 | 0.65 | 6 |

[a]Since the sum of Lyapunov exponents is not negative, $D_{KY}$ could not be calculated.

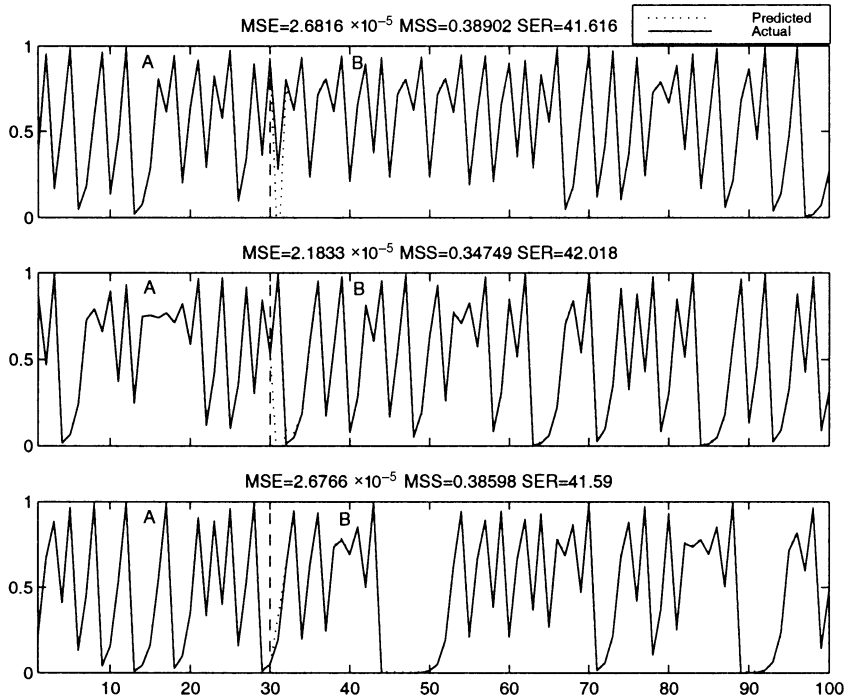



**Figure 4.3** One-step prediction of logistic map from different starting points. Note that A = initialization and B = one-step phase.

### 4.4.2 Ikeda Map

This second experiment uses the Ikeda map (which is substantially more complicated than the logistic map) to test the performance of the EKF-RMLP modeling scheme. The Ikeda map is a complex-valued map and is generated using the following difference equations:

$$m(k) = 0.4 - \frac{6.0}{1 + x_1^2(k) + x_2^2(k)}, \tag{4.4}$$

$$x_1(k+1) = 1.0 + \mu\{x_1(k)\cos[m(k)] - x_2(k)\sin[m(k)]\}, \tag{4.5}$$

$$x_2(k+1) = 1.0 + \mu\{x_1(k) + x_2(k)\cos[m(k)]\}, \tag{4.6}$$

where $x_1$ and $x_2$ are the real and imaginary components, respectively, of $x$ and the parameter $\mu$ is carefully chosen to be 0.7 so that the produced behavior is chaotic. The initial values of $x_1(0) = 0.5$ and $x_2(0) = 0.5$ were selected and, as pointed out earlier, a data set of 30,000 samples was

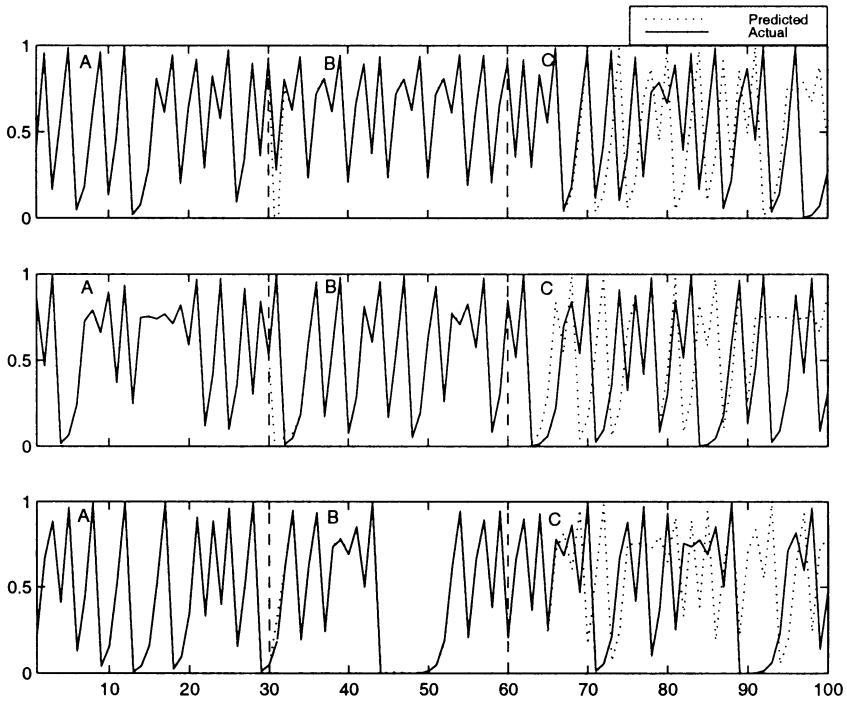



**Figure 4.4** Iterative prediction of logistic map from different starting points, corresponding to $N_0 = 3060$, 5060, and 10,060, respectively, with $p_l = 30$. Note that A = initialization, B = priming phase, and C = autonomous phase.

generated. In this experiment, only the $x_1$ component of the Ikeda map is used, for which the embedding parameters of $d_E = 6$ and $\tau = 10$ were determined. The first 5000 samples of this data set were used to train an RMLP with the EKF algorithm at one-step prediction. During training, a truncation depth $t_d = 10$ was used for the backpropagation through-time (BPTT) derivative calculations. The RMLP configuration of 6-6R-5R-1, which has a total of 144 weights including the bias terms, was chosen to model the Ikeda series. The training converged after only 15 epochs, and a sufficiently low incremental training mean-squared error was achieved, as shown in Figure 4.5.

***Open-Loop Evaluation*** The test set, consisting of the unexposed 25,000 samples of data, is used for performance evaluation, and Figure 4.6*a* shows one-step performance of the network on a short portion of the test data. It is indeed difficult to distinguish between the actual and predicted signals, thus visually verifying the goodness of the predictions.

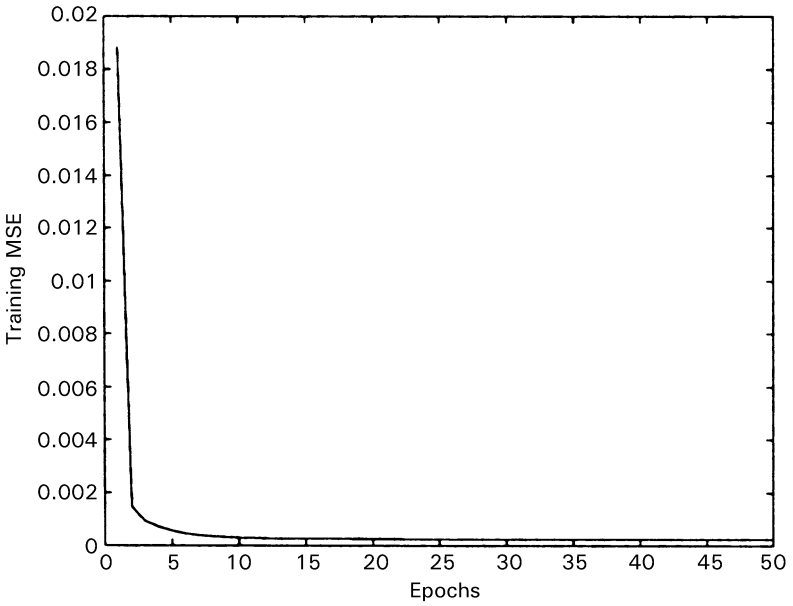



**Figure 4.5** Training MSE versus epochs for the Ikeda map.

For a numerical measure, the mean-squared value of the 25,000 sample Ikeda test series was calculated to be MSS $= 0.564$ and the mean-squared prediction error of MSE $= 1.4 \times 10^{-4}$ was produced by the trained RMLP network, thus giving an SER of 36.02 dB.

***Closed-Loop Evaluation*** To begin autonomous prediction, a delay vector consisting of 6 taps spaced by 10 samples apart is constructed as dictated by the embedding parameters $d_E$ and $\tau$. The RMLP is initialized with a delay vector, constructed from the test samples, and passed through a priming phase with $p_l = 60$, after which the network operates in closed-loop mode. The autonomous continuation from where the training data end is shown in Figure 4.6*b*. Note that the predictions follow closely for about 10 steps on average, which is in agreement with the theoretical horizon of predictability of 11 calculated by from the Lyapunov spectrum. A length of 25,000 autonomous samples were generated using the trained EKF-RMLP model, and the reconstructed attractor is plotted in Figure 4.6*d*. The reconstructed attractor has exactly the same form as the original attractor, which is plotted in Figure 4.6*c* using the actual Ikeda samples. These figures clearly demonstrate that the RMLP network has captured the underlying dynamics of the Ikeda map series. For numerical performance

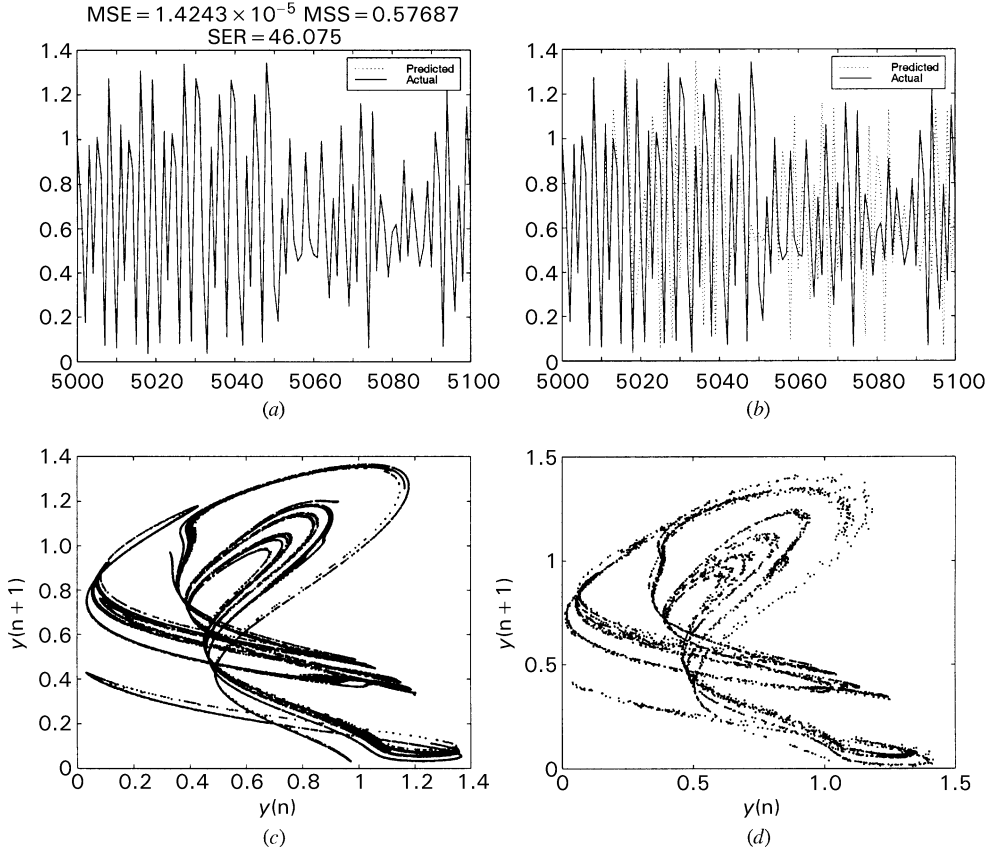



**Figure 4.6** Results for the dynamic reconstruction of the Ikeda map. (*a*) One-step prediction. (*b*) Iterated prediction. (*c*) Attractor of original signal. (*d*) Attractor of iteratively reconstructed signal.

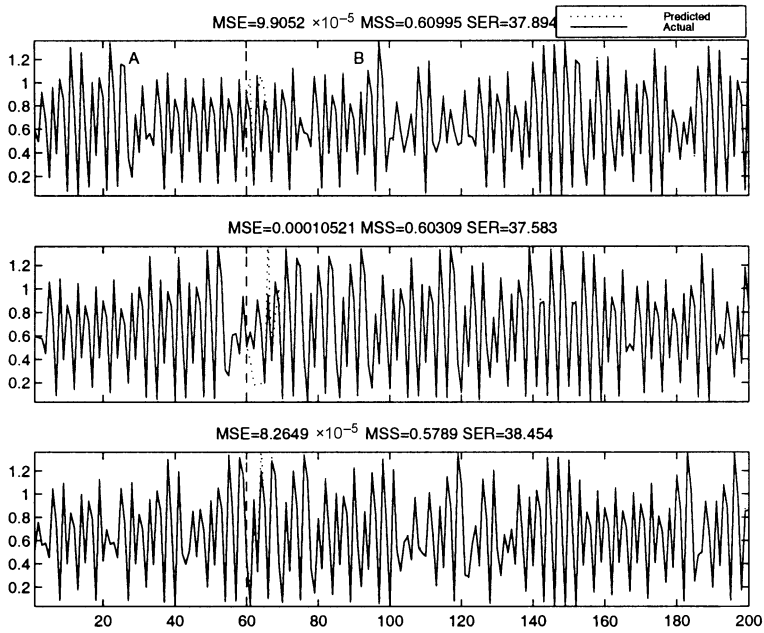



**Figure 4.7** One-step prediction of Ikeda series from different starting points. Note that A = initialization and B = one-step phase.

evaluation, the correlation dimension, Lyapunov exponents and Kolmogorov entropy of both the actual Ikeda series and the autonomously generated samples are calculated. Table 4.3, which summarizes the results, shows that the dynamic invariants of both the actual and reconstructed signals are in very close agreement with each other. This illustrates that the true dynamics of the data were captured by the trained network. Figure 4.7 plots the one-step prediction of the Ikeda map for three different starting points. The reconstruction produced here is robust and stable, regardless of the position of the initializing delay vector on the test data, as demonstrated in Figure 4.8, which shows autonomous operation starting at indices of $N_0 = 3120$, 10,120, and 17,120, respectively.

***Noisy Ikeda Series*** It was shown above that the noise-free Ikeda series can be modeled by the RMLP scheme. In a real environment, observables signals are usually corrupted by additive noise, which makes the problem more difficult. Thus, to make the modeling task more challenging than it already is, computer-generated noise is added to the Ikeda series such that the resulting signal-to-noise ratios (SNRs) of two sets of the noisy observables signals are 25 dB and 10 dB, respectively.

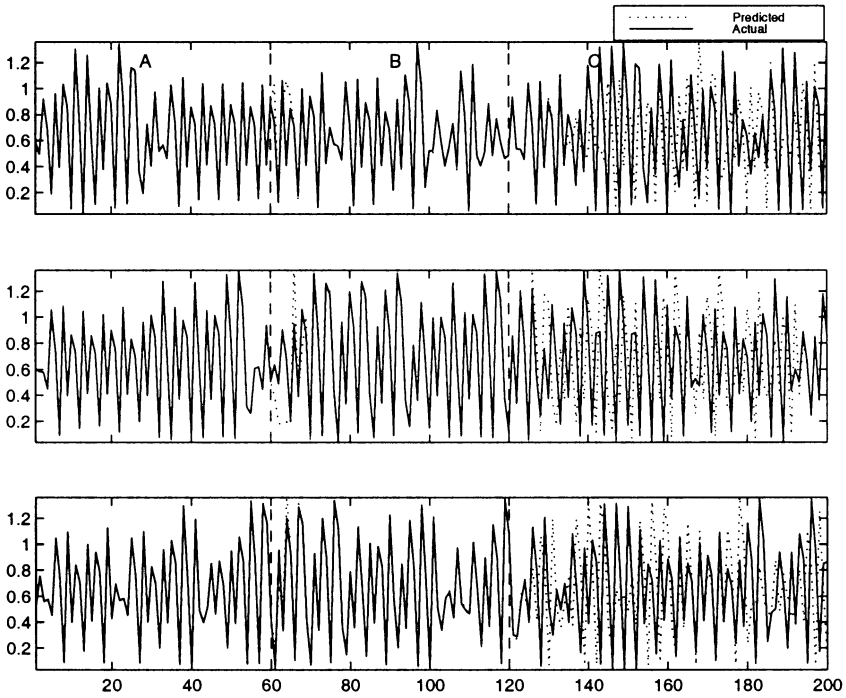



**Figure 4.8** Iterative prediction of Ikeda series from different starting points, corresponding to indices of $N_0 = 3120$, 10,120 and 17,120, respectively, with $p_l = 60$. Note that A = initialization, B = priming phase and C = autonomous phase.

The attractors of the noisy signals are shown in the left-hand parts of Figures 4.9*a* and 4.9*b*, respectively. The increase in noise level was more substantial for the 10 dB case, and this corrupted the signal very significantly. It is apparent from Figure 4.9*b* that the intricate details of the attractor trajectories are lost owing to the high level of noise.

The noisy signals were used to train two distinct 6-6R-5R-1 networks using the first 5000 samples in the same fashion as in the noise-free case. The right-hand plots of Figures 4.9*a* and 4.9*b* show the attractors of the autonomously generated Ikeda series produced by the two trained RMLP networks. Whereas the network trained with a 25 dB SNR was able to capture the Ikeda dynamics, the network trained with a 10 dB SNR was unable to do so. This shows that because of the substantial amount of noise in the 10 dB case, the network was unable to capture dynamics of the Ikeda series. However, for the 25 dB case, the network was not only able to capture the predictable part of the Ikeda series but also filtered the noise. Table 4.3 displays the chaotic invariants of the original and

**Table 4.3 Comparison of chaotic invariants of Ikeda map under noiseless and noisy conditions**

| Time series | $d_E$ | $\tau$ | $d_L$ | $D_{ML}$ | $D_{KY}$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $\lambda_4$ | $\lambda_5$ | $\lambda_6$ | $KE_{LE}$ | $KE_{ML}$ | HOP (samples) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | (nats/sample) | | | | | | (nats/sample) | | |
| Actual Ikeda | 6 | 10 | 2 | 1.51 | 1.44 | 0.354 | −0.802 | — | — | — | — | 0.354 | 0.316 | 11 |
| Reconstructed | 6 | 19 | 2 | 1.68 | 1.33 | 0.248 | −0.736 | — | — | — | — | 0.248 | 0.347 | 16 |
| 25 dB SNR Ikeda | 6 | 1 | 5 | 1.92 | 3.9 | 0.354 | 0.076 | −0.081 | −0.296 | −0.765 | — | 0.3548 | 0.314 | 14 |
| Reconstructed | 6 | 15 | 3 | 1.54 | 2.15 | 0.248 | −0.14 | −0.792 | — | — | — | 0.262 | 0.304 | 15 |
| 10 dB SNR Ikeda | 6 | 6 | 6 | 3.46 | —[a] | 1.145 | 0.948 | 0.778 | 0.543 | 0.215 | 0.495 | 0.363 | 0.662 | 3 |
| Reconstructed | 6 | 2 | 3 | 0.26 | —[a] | −3.47 | −6.01 | −3.47 | — | — | — | —[a] | —[a] | —[a] |

[a]Quantities that could not be calculated.

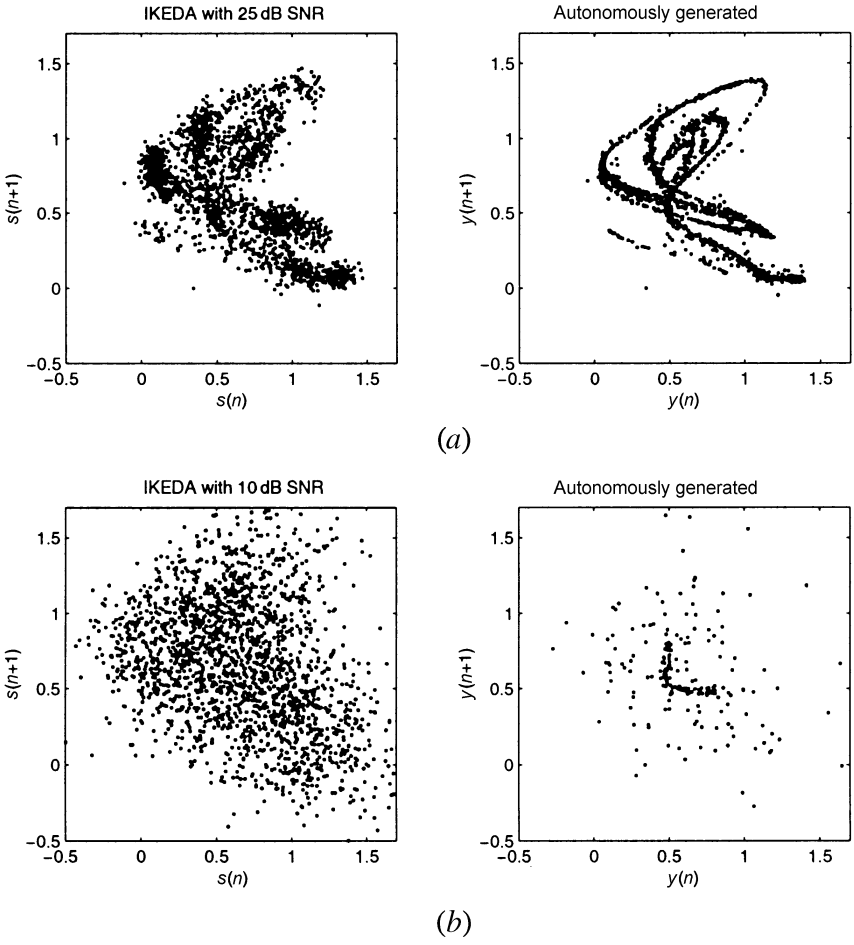



**Figure 4.9** Dynamic reconstruction of the noisy Ikeda map. (*a*) Autonomous performance for Ikeda map with 25 dB SNR. (*b*) Autonomous performance for Ikeda map with 10 dB SNR. Plots on the left to noisy original signals, and those on the right to reconstructed signals.

reconstructed signals for both levels of noise. The addition of noise has the effect of increasing the number of active degrees of freedom, and thus the number of Lyapunov exponents increases in a corresponding way. The invariants of the reconstructed signal corresponding to the 25 dB SNR case match more closely with the original noise-free Ikeda invariants as compared with the noisy invariants. However, for the failed reconstruction in the 10 dB case, there is a large disagreement between the reconstructed invariants and the actual invariants, which is to be expected. In fact, some of the invariants could not even be calculated.

**Table 4.4 SER results for several test cases for the Ikeda series**

| | Training noise | | |
|---|---|---|---|
| Testing noise | Clean | 25 dB SNR | 10 dB SNR |
| Clean | 36.0 | 27.1 | 16.3 |
| 25 dB SNR | 14.1 | 18.1 | 14.9 |
| 10 dB SNR | 1.7 | 5.8 | 6.6 |

Evaluation of one-step prediction for the more challenging noisy cases was also done. Table 4.4 summarizes the one-step SER results collected over a number of distinct test cases. For example, the first column of the table shows how a network trained with clean training data performs on test data with various levels of noise. It is important to note that it is difficult to achieve an SER larger than the SNR of the test data. The best overall generalization results are obtained when the network is trained with 25 dB SNR.

### 4.4.3 Lorenz Attractor

The Lorenz attractor is more challenging than the Ikeda or logistic map; it is described by a coupled set of three nonlinear differential equations:

$$\dot{x} = \sigma[y(n) - x(n)], \tag{4.7a}$$

$$\dot{y} = x(n)[r - z(n)] - y(n), \tag{4.7b}$$

$$\dot{z} = x(n)y(n) - bz(n), \tag{4.7c}$$

where the fixed parameters $r = 45.92$, $\sigma = 16$, and $b = 4$ are used, and $\dot{x}$ means the derivative of $x$ with respect to time $t$. As before, a data set of 30,000 samples was generated at 40 Hz, of which the first 5000 samples were used to train the RMLP model and the remaining 25,000 were used for testing. For the experimental Lorenz series, an embedding dimension of $d_E = 3$ and a delay of $\tau = 4$ were calculated. An RMLP network configuration of 3-8R-7R-1, consisting of 216 weights including the biases, was trained with the EKF algorithm, and the convergence of the training MSE is shown in Figure 4.10.

***Open-Loop Evaluation*** The results shown in Figure 4.11 were arrived at in only 10 epochs of working through the training set. The

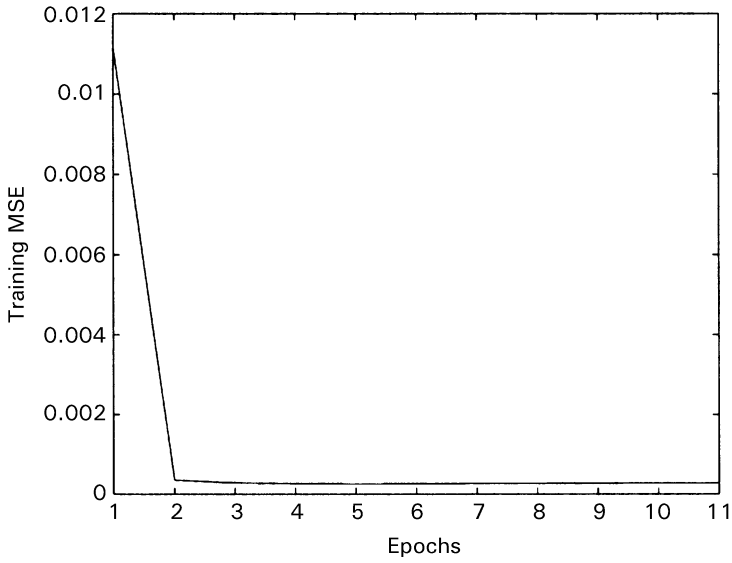



**Figure 4.10** Training MSE versus epochs for the Lorenz series.

plot of one-step prediction over a portion of the test data is depicted in Figure 4.11*a*. The one-step MSE over 25,000 samples of test data was calculated to be $1.57 \times 10^{-5}$, which corresponds to an SER value of 40.2 dB.

***Closed-Loop Evaluation*** The autonomous continuation, from where the test data end, is shown in Figure 4.11*b*, which demonstrates that the network has learned the dynamics of the Lorenz attractor very well. The iterated predictions follow the trajectory very closely for about 80 time steps on average, and then demonstrate chaotic divergence, as expected. This is in close agreement with the theoretical horizon of predictability of 97 calculated from the Lyapunov spectrum. A further testament to the success of the EKF-RMLP model is that the reconstructed attractor, shown in Figure 4.11*d*, is similar in shape to the attractor of the original series, which is shown in Figure 4.11*c*, demonstrating that the network has indeed captured the dynamics well. In addition, the dynamic invariants of the original and reconstructed series are compared in Table 4.5, which shows close agreement between their respective correlation dimension, Lyapunov spectrum, and Kolmogorov entropy, thus indicating the strong presence of the original dynamics in the reconstructed signal. Figure 4.12

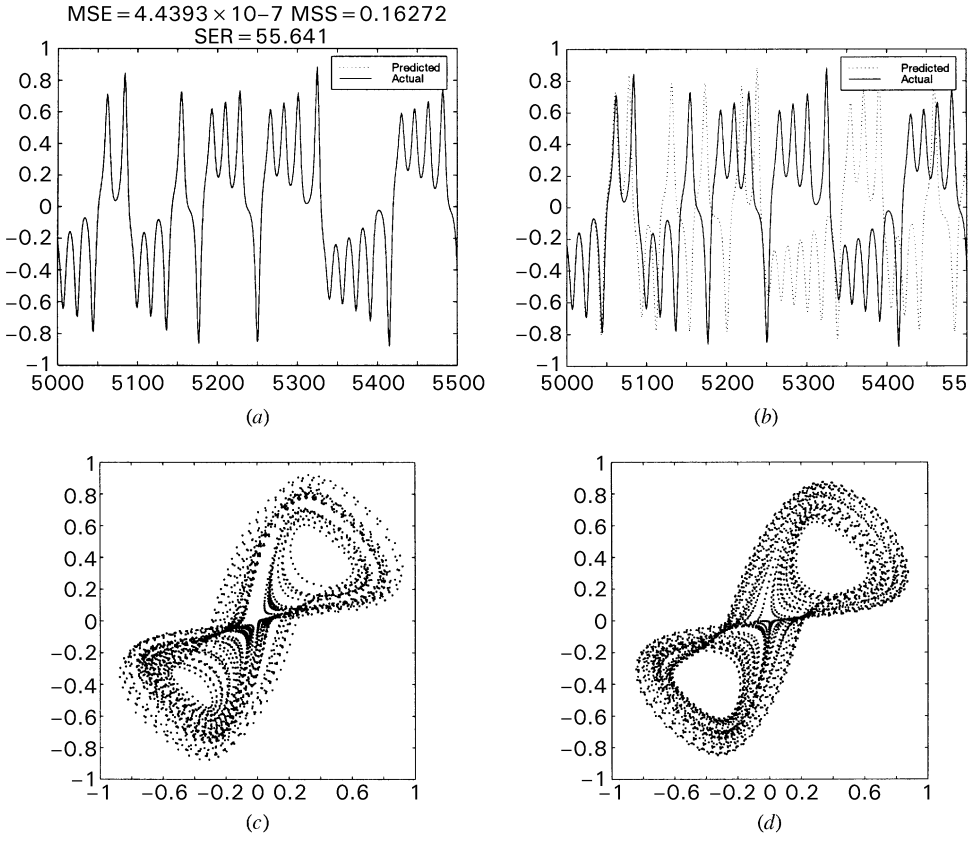



**Figure 4.11** Results for the dynamic reconstruction of the Lorenz series. (*a*) One-step prediction. (*b*) Iterated prediction. (*c*) Attractor of original signal. (*d*) Attractor of iteratively reconstructed signal.

**Table 4.5 Comparison of chaotic invariants of Lorenz series**

| Time series | $d_E$ | $\tau$ | $d_L$ | $D_{ML}$ | $D_{KY}$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $\lambda_4$ | $\lambda_5$ | $\lambda_6$ | $\lambda_7$ | $KE_{LE}$ | $KE_{ML}$ | HOP (samples) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | (nats/sample) | | | | | | | (nats/sample) | | |
| Actual Lorenz | 3 | 4 | 3 | 2.09 | 2.03 | 0.040 | −0.0005 | −1.286 | — | — | — | — | 0.040 | 0.042 | 97 |
| Reconstructed | 3 | 4 | 3 | 1.99 | 2.26 | 0.050 | −0.0009 | −0.1828 | — | — | — | — | 0.050 | 0.037 | 78 |
| 25 dB SNR Lorenz | 5 | 4 | 4 | 2.84 | 3.75 | 0.489 | 0.249 | −0.092 | −0.850 | — | — | — | 0.738 | 0.151 | 8 |
| Reconstructed | 3 | 4 | 3 | 2.00 | 2.11 | 0.039 | −0.006 | −0.303 | — | — | — | — | 0.039 | 0.022 | 99 |
| 10 dB SNR Lorenz | 7 | 14 | 7 | 7.18 | 6.71 | 1.554 | 0.419 | 0.269 | 0.109 | −0.10 | −0.442 | −1.128 | 1.35 | 1.06 | 7 |
| Reconstructed | 3 | 5 | 3 | 1.84 | 2.18 | 0.062 | 0.003 | −0.360 | — | — | — | — | 0.065 | 0.024 | 63 |

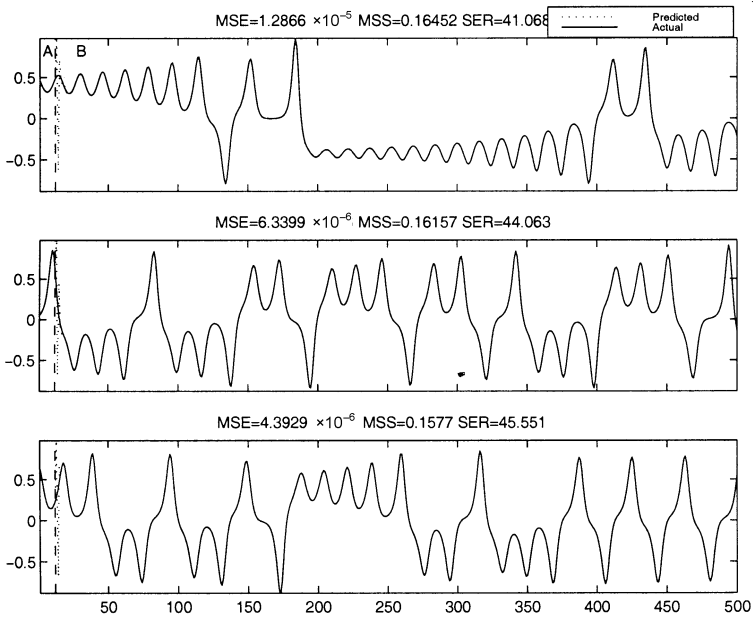



**Figure 4.12** One-step prediction of Lorenz series from different starting points. Note that A = initialization and B = one-step phase.

plots one-step predictions of the Lorenz time series for three different starting points. The reconstruction of the Lorenz series was indeed very robust and stable, regardless of the starting position on the test series, as demonstrated in Figure 4.13, which shows autonomous operation starting at indices of $N_0 = 3042$, 10,042, and 17,042, respectively.

***Noisy Lorenz Series*** To make the problem even more challenging than it already is, a significant level of noise was added to the clean Lorenz series just as was done for the Ikeda map. An RMLP network architecture was selected similar to the noise-free case, and two distinct networks were trained using the noisy Lorenz signals with 25 dB SNR and 10 dB SNR, respectively. The networks were trained with a learning rate of $p_r = 0.001$ for 15 epochs through the first 5000 samples, as was done for the noise-free case. Then, both one-step prediction and autonomous prediction results were obtained with the remaining unexposed 25,000 samples. Figures 4.14*a* and 4.15*b* show that excellent dynamic reconstruction of the Lorenz series was possible even in the presence of 25 dB and 10 dB SNR,

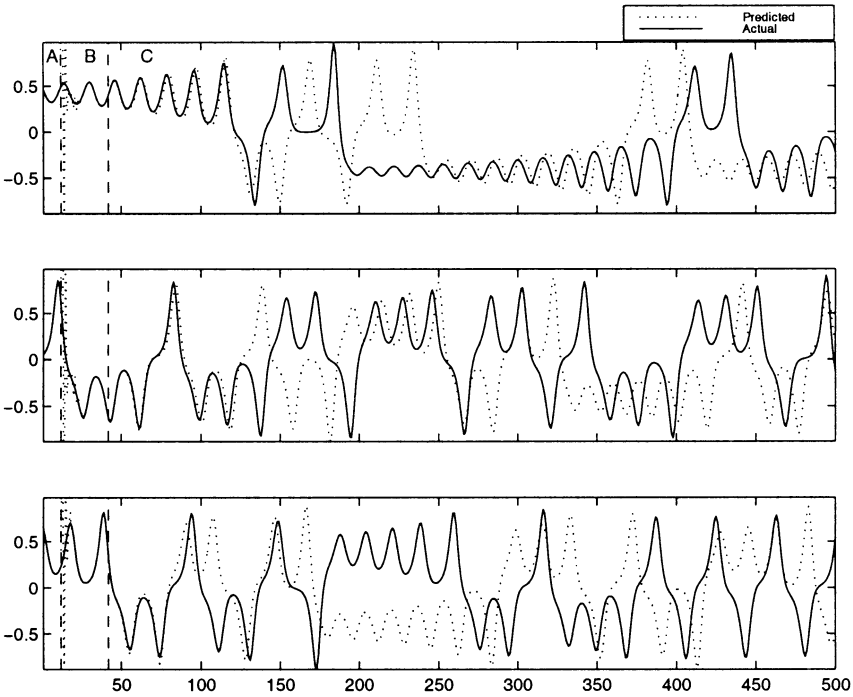



**Figure 4.13** Iterative prediction of Lorenz series from different starting points, corresponding to indices of $N_0 = 3042$, 10,042, and 17,042, respectively with $p_l = 30$. Note that A = initialization, B = priming phase, and C = autonomous phase.

respectively, despite the fact that the additive noise corrupts the attractors considerably. Moreover, comparison of the dynamic invariants of the noisy and reconstructed signals in Table 4.5 reveals that the reconstructed signals and their invariants are reasonably close to the noise-free signal and the iterated predictions are smoother in comparison to the noisy signals, as shown in Figs. 4.15*a* and 4.15*b*.

It was shown earlier that dynamic reconstruction of the Ikeda map was not possible for the case of 10 dB SNR. In order to explain this, the normalized frequency spectra of both the Ikeda map and Lorenz series are plotted in Figure 4.16. This figure illustrates how much more significantly the Ikeda spectrum is corrupted by 10 dB noise as compared with the Lorenz spectrum. Whereas the Ikeda spectrum is completely drowned at 10 dB SNR, the Lorenz spectrum is still discernible, and as such its dynamic reconstruction was possible. Another way of explaining the

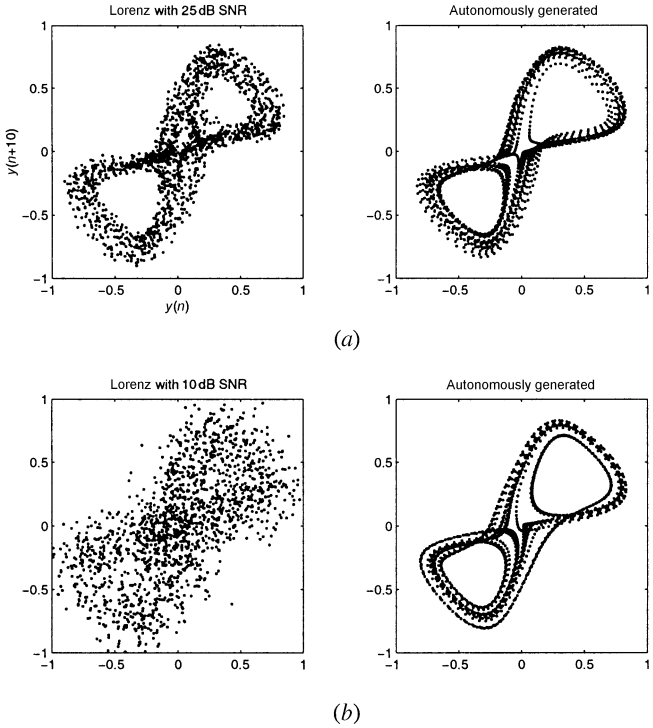
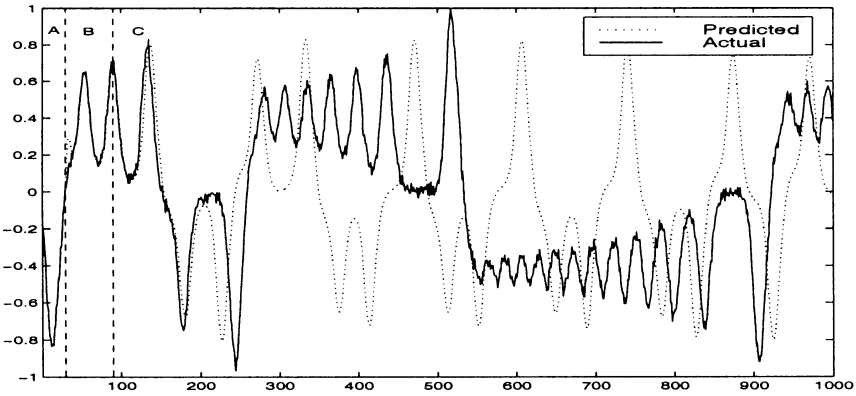


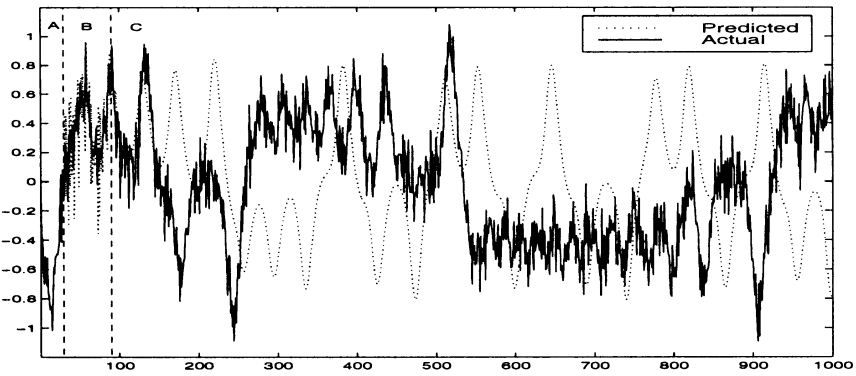

**Figure 4.14** Dynamic reconstruction of the noisy Lorenz series. (*a*) Autonomous performance for Lorenz series with 25 dB SNR. (*b*) Autonomous performance for Lorenz series with 10 dB SNR. Plots on the left pertain to the original data, and those on the right to reconstructed data.

reason for the robust behavior of the Lorenz attractor to the presence of additive noise is to recall that the correlation dimension of the Lorenz attractor is greater than that of the Ikeda map; see Table 4.1. Accordingly, the Lorenz attractor is more complex than the Ikeda map – hence the greater robustness of the Lorenz attractor to noise.

Evaluation of one-step performance for the more challenging noisy cases was also done. Table 4.6 summarizes the one-step SER results collected over a number of distinct test cases. For example, the first column of the table shows how a network trained with clean training data performs on test data with various levels of noise. The best overall generalization results are obtained when the network is trained with 25 dB SNR.




(*a*)




(*b*)

**Figure 4.15** Iterative prediction of noisy Lorenz series with (*a*) 25 dB and (*b*) 10 dB SNR. Note that A = initialization, B = priming phase and C = autonomous phase.

## 4.5 NONLINEAR DYNAMIC MODELING OF REAL-WORLD TIME SERIES

### 4.5.1 Laser Intensity Pulsations

In the first experiment with real-world data, a series recorded from a far-infrared laser in a chaotic state is used. This data set was provided as part of a data package used in the Santa Fe Institute Time Series Prediction and Analysis Competition [7]. During the time of the competition, only the first 1000 samples were provided, and the goal was to predict the following

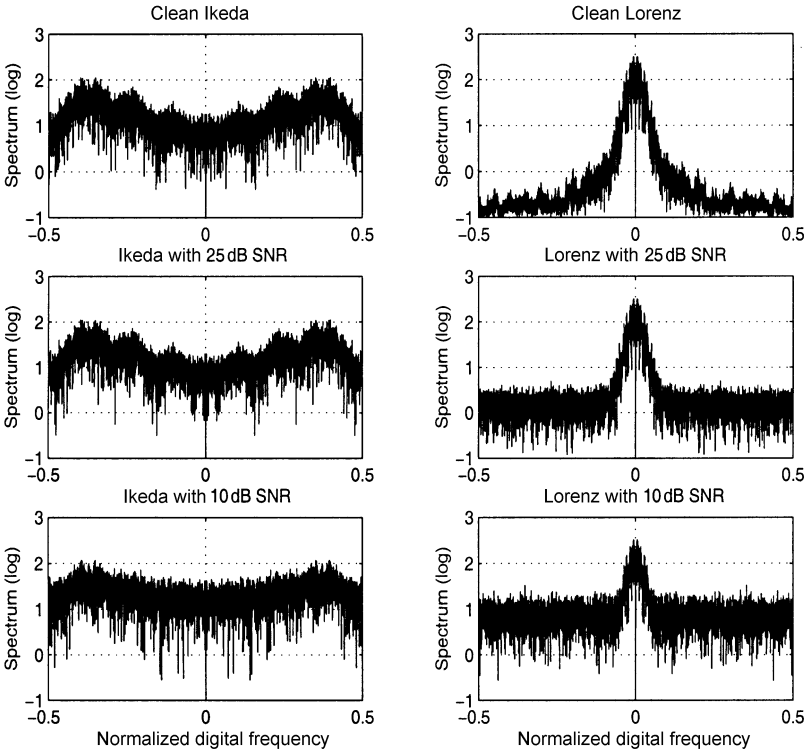



**Figure 4.16** Lorenz and Ikeda series: normalized frequency spectra.

**Table 4.6 SER results for several test cases for the Lorenz series**

| | Training noise | | |
|---|---|---|---|
| Testing noise | Clean | 25 dB SNR | 10 dB SNR |
| Clean | 40.2 | 30.5 | 36.8 |
| 25 dB SNR | 17.8 | 23.1 | 17.2 |
| 10 dB SNR | 3.5 | 9.3 | 3.8 |

100 samples as accurately as possible. Now, however, the full data set consisting of 10,000 samples is available. The first 1100 samples are plotted in Figure 4.17. The data consist of high-frequency pulsations, which gradually rise in amplitude and suddenly collapse from a high to low amplitude. The rapid decays of oscillations in this data set occur with no periodicity, and are a challenge to model. During these collapses, the

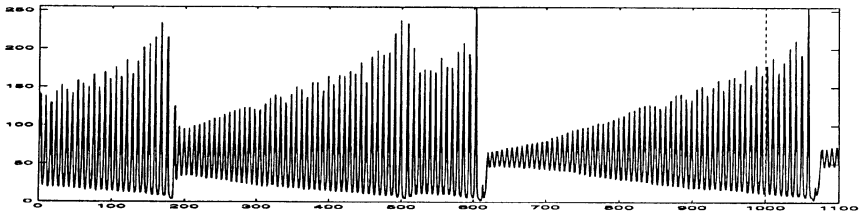


**Figure 4.17** First 1100 samples of original laser series.

signal abruptly jumps from one region of the attractor to another distinct region, and these transitions are the most difficult parts of this series to model.

To remain consistent with what was done in the SFI competition, only the first 1000 samples were presented to the network, while the remaining 9000 samples of data were reserved for testing purposes and not shown to the network. The optimal embedding delay was found to be $\tau = 2$ and the embedding dimension to be $d_E = 9$. A network architecture of 9-10R-8R-1, consisting of 361 weights including the biases, was used to model this data. The network was trained through the first 1000 points for 50 epochs.

***Open-Loop Evaluation*** The network makes excellent one-step predictions as shown in Figure 4.18*a*, with very low prediction errors in most regions. The accuracy of prediction drops slightly during the sudden signal collapses, but then increases quickly as the network adapts. The SER value for the one-step prediction performance on the unexposed 9000 samples of the test set was calculated to be 26.05 dB.

***Closed-Loop Evaluation*** The iterated predictions of the network are also very remarkable. The autonomous continuation from sample 1000, where the training data end, is shown in Figure 4.18*b*. This was achieved by first initializing the network with a delay vector, consisting of 9 taps, each spaced 2 samples apart, from the beginning of the training set and then priming for $p_l = 982$ steps, after which the network operates in an autonomous mode. Since the network was initialized with the first samples of the actual data, autonomous operation does not begin until the point $(9 \times 2) + 982 = 1000$, as shown in Figure 4.19. The autonomous output of the network closely follows the original series for about 60 steps (which is higher than the average horizon of predictability of 26 calculated from the Lyapunov spectrum) after which it begins to diverge. Figure 4.20 shows the one-step prediction of the laser time series with three different starting points.

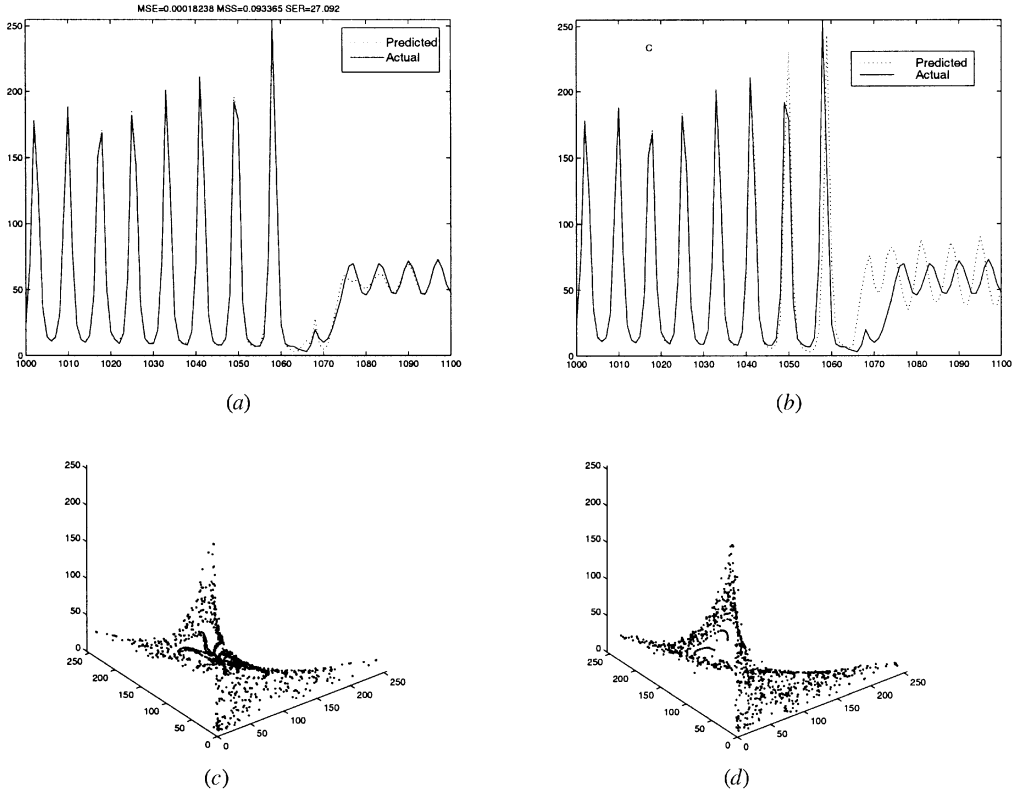



**Figure 4.18** Results of the dynamic reconstruction of the laser series. (*a*) One-step prediction. (*b*) Iterated prediction. (*c*) Attractor of original signal. (*d*) Attractor of iteratively reconstructed signal.

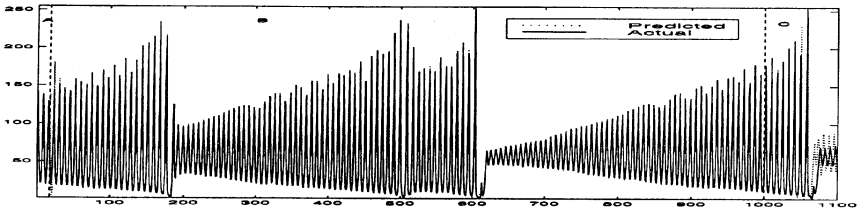



**Figure 4.19** Iterated prediction of the laser data beginning at sample 1000, using $p_l = 982$. Note that A = initialization, B = priming phase, and C = autonomous phase. Phase C is enlarged in Figure 4.18*b*.

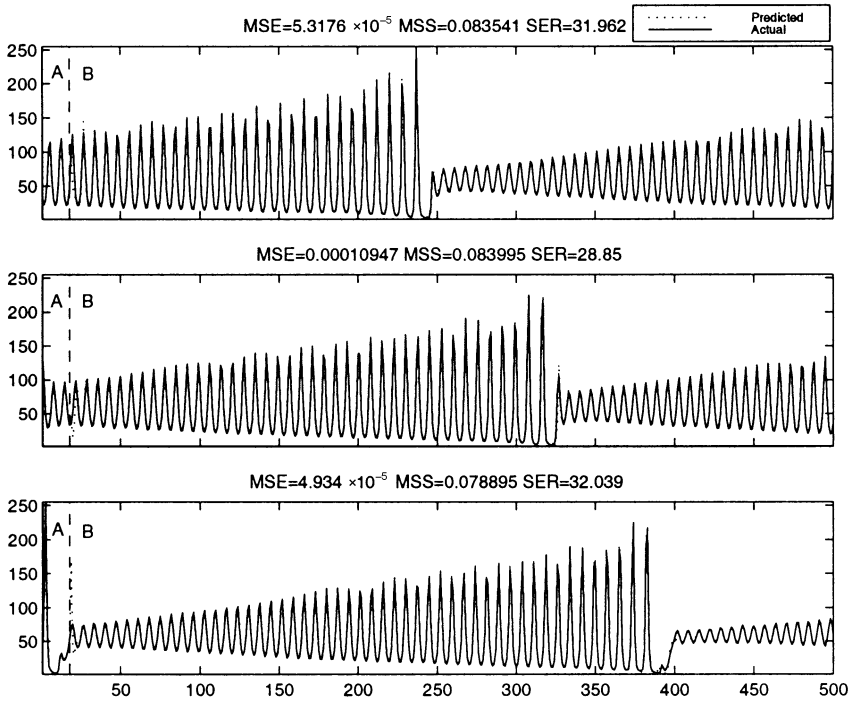



**Figure 4.20** One-step prediction of laser series with different starting points. Note that A = initialization and B = one-step phase.

Figure 4.21 shows that the network is capable of stable and robust reconstruction even when it is initialized from different starting points on the test set, corresponding to indices of $N_0 = 48$, 1048, and 4048, respectively. Note that the network has not seen this data set before, and is still able to predict the signal collapses up to an impressive degree of accuracy. Figure 4.22, which shows the 9000-sample iterated output of the network, demonstrates that the chaotic pulsations have been learned by the

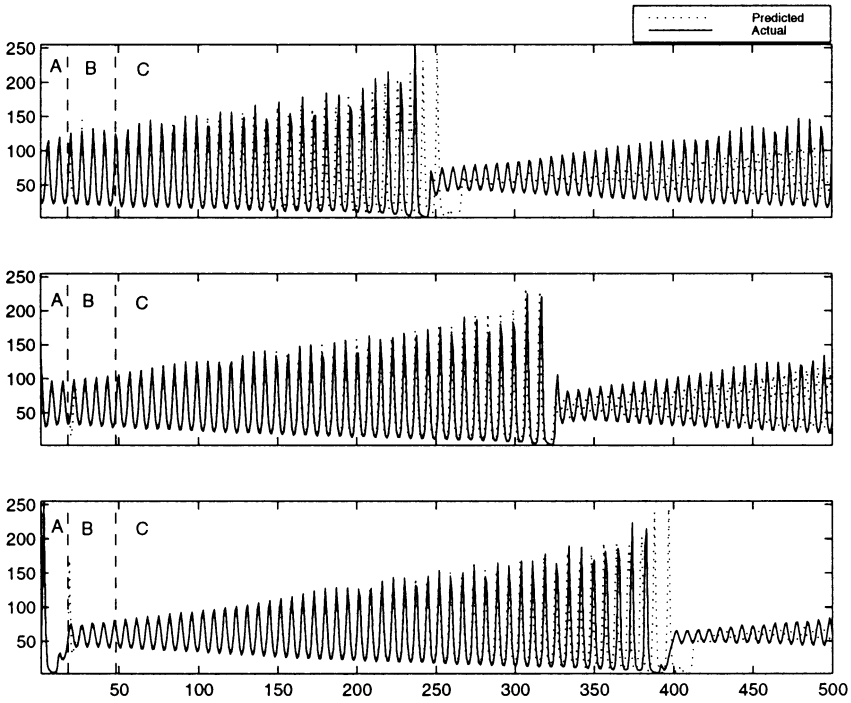



**Figure 4.21** Iterative prediction of laser series from different starting points, corresponding to indices of $N_0 = 48$, 1048, and 4048, respectively, with $p_l = 30$. Note that A = initialization, B = printing phase, and C = autonomous phase.

network from just the first 1000 samples. The attractor of the original series is plotted in Figure 4.18*c* using the actual laser data. The attractor in Figure 4.18*d* is plotted using the reconstructed points as produced by the trained RMLP network. It is clear from the attractor of the reconstructed signal that the dynamics of the laser data has been faithfully captured by the EKF trained RMLP network. Table 4.7 shows that the correlation dimension, Lyapunov exponents, and Kolmogorov entropy of the reconstructed signal are in close agreement with those obtained from the original data, thus indicating a faithful reconstruction of the original dynamics. The correlation dimension of $NH_3$ laser pulsations calculated here is consistent with what is reported by Hubner et al. [8].

***Longer Training Length*** In another experiment, the same-sized network was trained on the first 5000 points, as compared with the first 1000 used above. The iterated output produced by the network is shown in Figure 4.23. Since the network was exposed to more of the signal

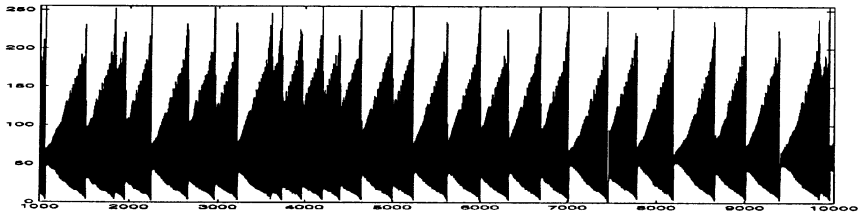


**Figure 4.22** 9000 samples of iterated output of the RMLP trained on the first 1000 samples.

**Table 4.7 Comparison of chaotic invariants of laser series**

| | | | | | | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $KE_{LE}$ | $KE_{ML}$ | HOP |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Time series | $d_E$ | $\tau$ | $d_L$ | $D_{ML}$ | $D_{KY}$ | (nats/sample) | | | (nats/sample) | | (samples) |
| Actual laser | 9 | 2 | 3 | 2.01 | 2.19 | 0.147 | −0.039 | −0.547 | 0.147 | 0.195 | 26 |
| Reconstructed | 9 | 2 | 3 | 1.90 | 2.13 | 0.115 | −0.037 | −0.602 | 0.115 | 0.133 | 34 |

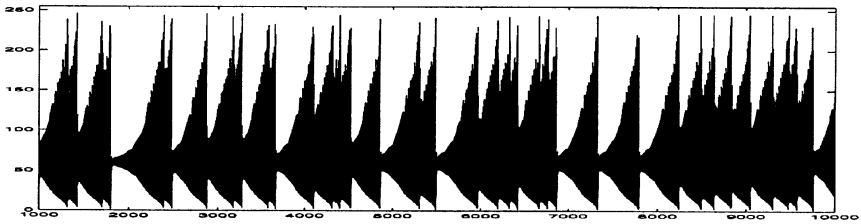


**Figure 4.23** 9000 samples of iterated output of the RMLP trained on the first 5000 samples.

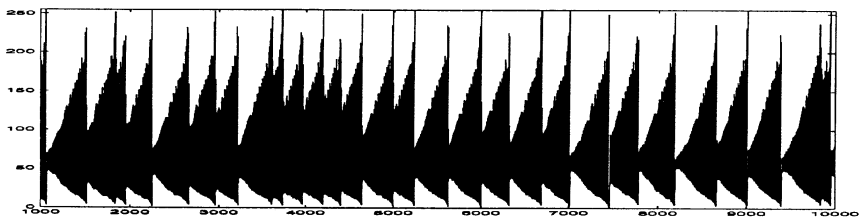


**Figure 4.24** Last 9000 samples of the actual laser series.

dynamics, in this case, the iterated output of the network demonstrates more frequent chaotic collapses and is closer in appearance to the actual signal shown in Figure 4.24.

### 4.5.2 Sea Clutter Data

*Sea clutter*, or *sea echo*, refers to the radar backscatter from an ocean surface. Radars operating in a maritime environment have a serious limitation imposed on their performance by unwanted sea clutter. Therefore, a problem of fundamental interest in the radar literature is the modeling of sea clutter to design optimum target detectors.

For many years, clutter echoes in radar systems were modeled as a stochastic process. However, in recent years, some researchers have concluded that the generation of sea clutter is governed by deterministic chaos [9 – 13]. These claims, in varying degrees, have been largely based on the estimation of chaotic invariants of sea clutter, namely, the correlation dimension, Lyapunov exponents, Kaplan-Yorke dimension, and the Kolmogorov entropy, the latter two are derived from the Lyapunov exponents. But, recently, serious concerns have been raised on the *discriminative power* of the state-of-the-art algorithms currently available for distinguishing between sea clutter and different forms of surrogate data derived from the original data; the surrogate data are known to be stochastic by purposely introducing randomization into their generation. By estimating the correlation dimension and Lyapunov exponents, it is found that there is little difference between the chaotic invariants of sea clutter and those of their surrogate counterparts [14, 15], which casts doubt on deterministic chaos being the nonlinear dynamical model foe sea clutter.

In [14], it is pointed out physical realities dictate the presence of noise not just in the measurement equation (owing to instrument errors) but also in the process equation (owing to the rate of variability of the forces affecting the ocean dynamics). According to Sugihara [16], there is unavoidable practical difficulty in disentangling the process noise from the measurement noise when we try to reconstruct an invariant measure such as the Lyapunov exponents. This may explain the reason for the inability of chaotic-invariant estimation algorithms to distinguish between sea clutter and its surrogates.

Our primary interest in this chapter is in dynamic reconstruction. In this context, we may raise the following question: Can the EKF-RMLP be configured to perform dynamic reconstruction on sea clutter in a robust manner? In what follows, an attempt to answer this question is presented;

**Table 4.8 Comparison of chaotic invariants of sea clutter**

| Time series | $d_E$ | $\tau$ | $d_L$ | $D_{ML}$ | $D_{KY}$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $\lambda_4$ | $\lambda_5$ | $KE_{LE}$ | $KE_{ML}$ | HOP (samples) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | (nats/sample) | | | | | (nats/sample) | | |
| Actual sea clutter | 6 | 10 | 5 | 4.69 | 4.42 | 0.228 | 0.102 | 0.005 | −0.118 | −0.448 | 0.335 | 0.369 | 17 |
| Reconstructed from $N_2 = 0$ | 6 | 9 | 5 | 3.12 | 4.51 | 0.214 | 0.078 | 0.010 | −0.083 | −0.427 | 0.302 | 0.292 | 18 |
| Reconstructed from $N_0 = 5160$ | 7 | 6 | 5 | 3.58 | 4.75 | 0.216 | 0.129 | 0.049 | −0.112 | −0.373 | 0.394 | 0.303 | 18 |
| Reconstructed from $N_0 = 6160$ | 6 | 8 | 5 | 4.43 | 4.62 | 0.232 | 0.136 | 0.023 | −0.113 | −0.434 | 0.391 | 0.317 | 17 |
| Reconstructed from $N_0 = 7160$ | 6 | 6 | 5 | 3.67 | 4.23 | 0.191 | 0.088 | 0.029 | −0.142 | −0.713 | 0.308 | 0.220 | 20 |

the answer has bearing on whether sea clutter is chaotic or not. Figure 4.25 shows the in-phase component of typical sea clutter, recorded by an instrument-quality radar on the East Coast of Canada, with the radar operating in a dwelling mode at a low grazing angle. The data consist of 50,000 samples, of which the first 40,000 are used for network training, while the remaining 10,000 samples are reserved for testing. An embedding dimension of $d_E = 6$ and embedding delay of $\tau = 10$ were calculated for sea clutter.

***Open-Loop Evaluation*** The results for the modeling of sea clutter by the EKF-RMLP method are shown in Figure 4.26. A 6-8R-7R-1 network trained with $p_r = 0.5$ was used. The SER over the unexposed 10,000 test samples was found to be 36.65 dB, which is certainly impressive, since in the EKF-RMLP based method presented here, the synaptic weights remain fixed after the training was completed.

***Closed-Loop Evaluation*** The iterated continuation from where the training data end, corresponding to an index of $N_0 = 0$, is shown in Figure

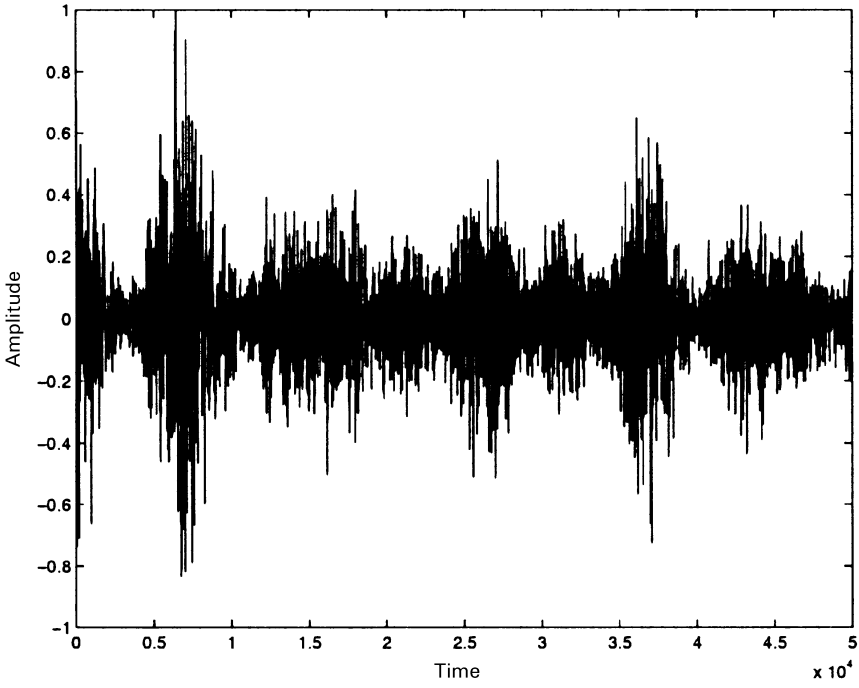



**Figure 4.25** In-phase component of typical sea-clutter data.

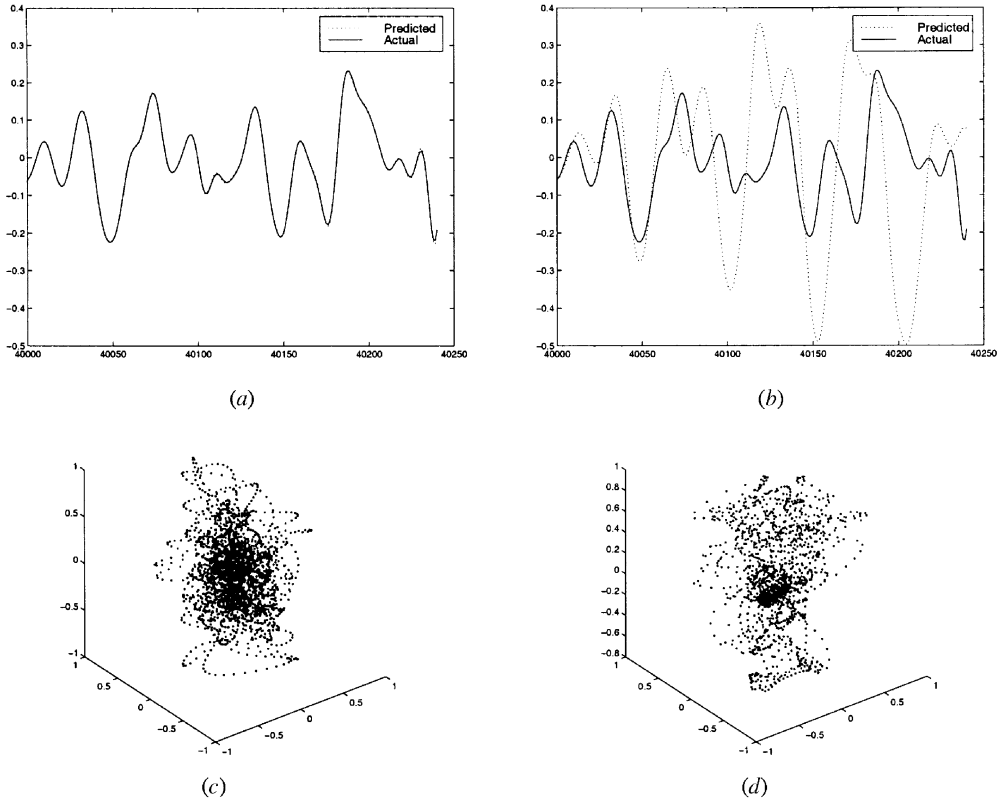



**Figure 4.26** Results for the dynamic reconstruction of sea clutter. (*a*) One-step prediction. (*b*) Iterated prediction. (*c*) Attractor of original signal. (*d*) Attractor of iteratively reconstructed signal.

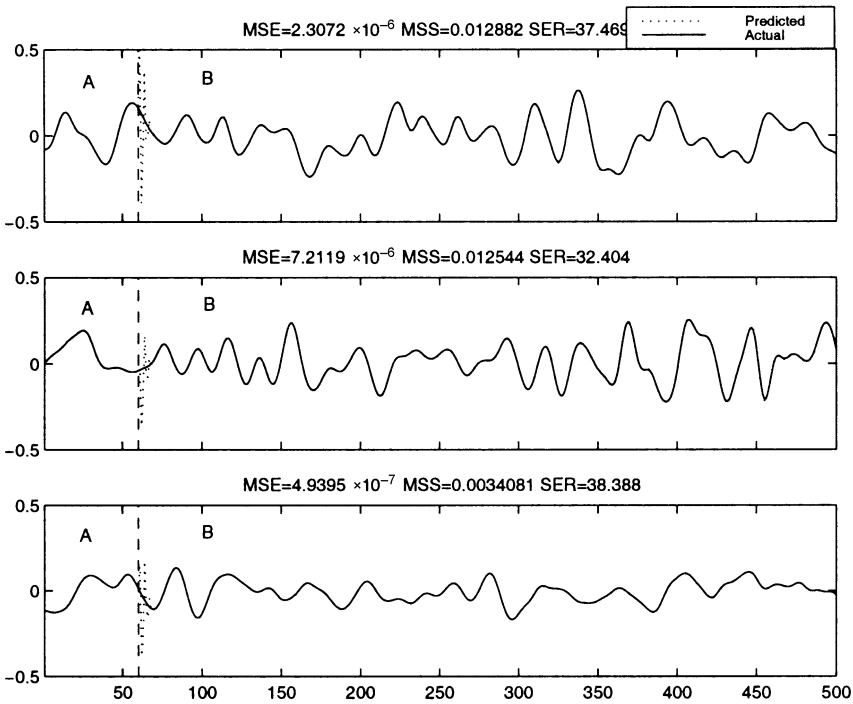



**Figure 4.27** One-step prediction of sea clutter from different starting points. Note that A = initialization and B = one-step mode of operation.

4.26*b*. Observe that the prediction is close to the actual data for roughly less than 10 steps (which is significantly less than the average theoretical horizon of predictability of 17 calculated from the Lyapunov spectrum), after which it diverges into a distinct trajectory of its own. The network, when iterated in this fashion for several thousand steps, produces an output whose attractor is plotted in Figure 4.26*d*. It is difficult to compare the reconstructed attractor with the original attractor in Figure 4.26*c* because we are looking at a five-dimensional attractor in three dimensions.

Since it is impossible to plot attractors in five dimensions, we may determine the success of dynamic reconstruction by a numerical comparison of the dynamic invariants of the reconstructed signal with those of the actual signal. As shown in Table 4.8, the Lyapunov exponents of the reconstructed signal are in close agreement with those obtained from the actual data. In addition, the correlation dimension and Kolmogorov entropy of both signals are also close, except for the slightly lower than expected value of $D_{ML}$ for the reconstructed signal. Furthermore, $D_{KY}$ and

$KE_{LE}$ are close to $D_{ML}$ and $KE_{ML}$, respectively. Figure 4.27 plots one-step predictions of sea clutter for three different starting points.

***Stability and Robustness*** Furthermore, the more difficult case of autonomous prediction from different starting points chosen from the test data is shown in Figure 4.28. This shows that stable reconstruction is also possible from indices of $N_0 = 5160$, 6160, and 7160 on the test data. The dynamic invariants calculated from signals reconstructed at these indices are in close agreement with the actual Table 4.8. Again, there seems to be a slight disagreement in the values of $D_{ML}$.

This shows that the EKF-RMLP scheme was able to reconstruct sea-clutter dynamics on several occasions in a reasonable manner, depending on the data used. However, it is important to point out that *robust reconstruction was not possible* from all initialization points – for this and other types of networks used. The different types of reconstruction

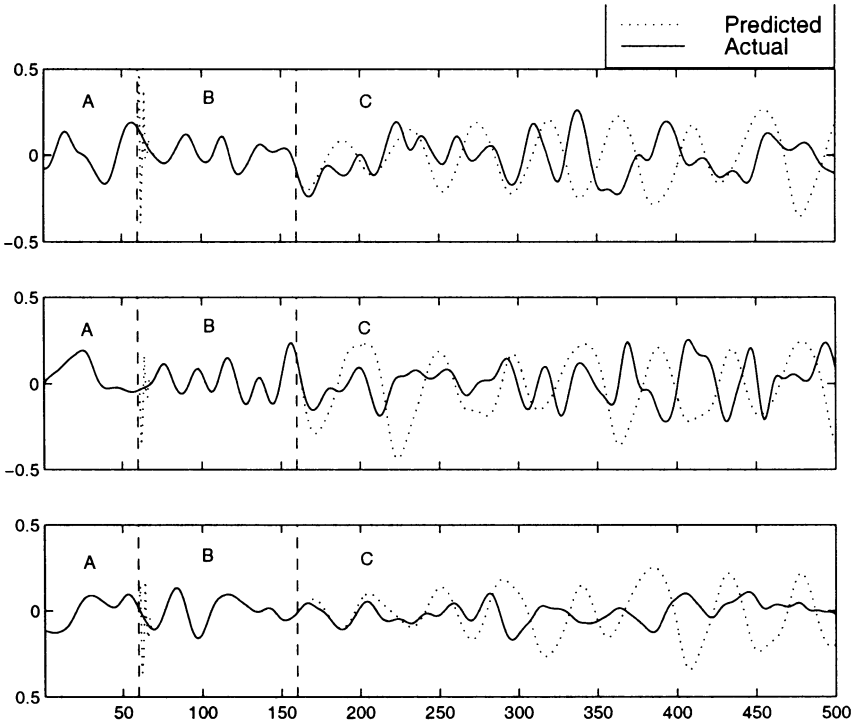



**Figure 4.28** Iterative prediction of sea clutter from different starting points, corresponding to indices of $N_0 = 5160$, 6160, and 7160, respectively, with $p_l = 100$. Note that A = initialization, B = priming phase, and C = autonomous phase.

failures encountered are illustrated in Figure 4.29. In the first type of reconstruction failure, the output becomes constant after a few steps. In other words, the output behaves as a fixed-point attractor. In the second type of failure, the output becomes periodic after a few time steps. In the third type of failure, the output breaks down after a few steps of good prediction. How do we explain the reconstruction failures of Figure 4.29? Dynamic reconstruction, in accordance with Taken's embedding theorem tolerates a moderate level of measurement noise, but the process equation has to be noise-free. The presence of dynamical noise in the process equation is therefore in violation of Taken's theorem — hence the difficulty in devising a robust dynamic reconstruction system for sea clutter.

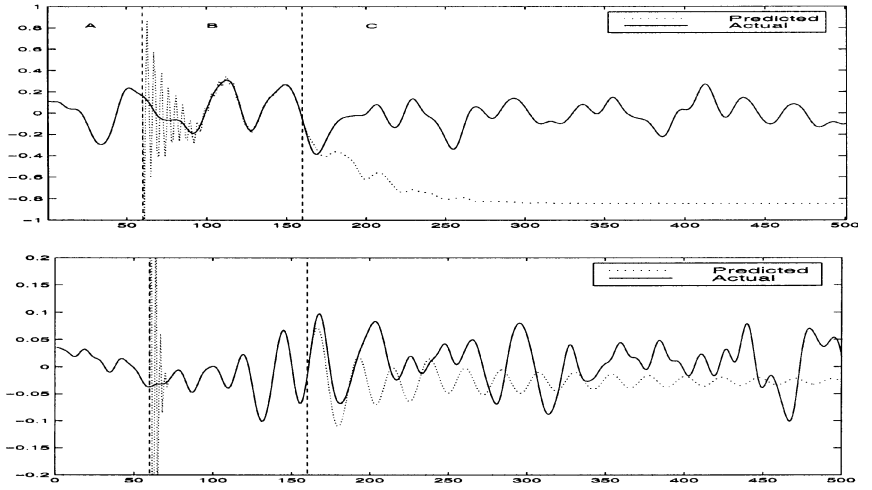
## 4.6 DISCUSSION

Takens' theorem provides the theoretical basis for an experimental approach for the autoregressive modeling of a nonlinear dynamical system. Specifically, the state evolution of the system with a $d$-dimensional state space can be reconstructed by observing $2d + 1$ time lags of a single output of the system for an arbitrary delay. The theorem requires that the observable time series be noiseless and of infinite length. Unfortunately, neither of these two conditions can be satisfied in practice. To mitigate these difficulties, the recommended procedure is to do the following:

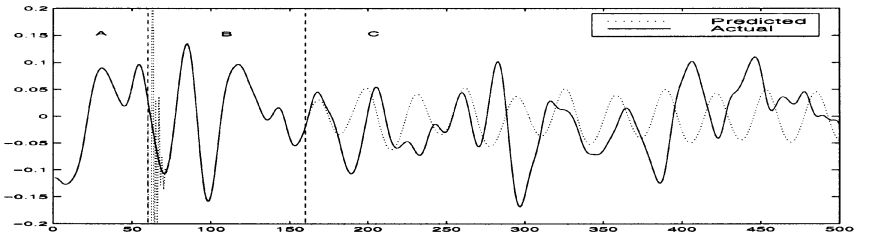
- Use an "optimum" time delay for the delay coordinate method, for which the adjacent samples are as statistically independent as practically possible.

For the experimental study presented here, we minimized the mutual information between the observable time series and its delayed version as the basis for estimating the optimum time delay.

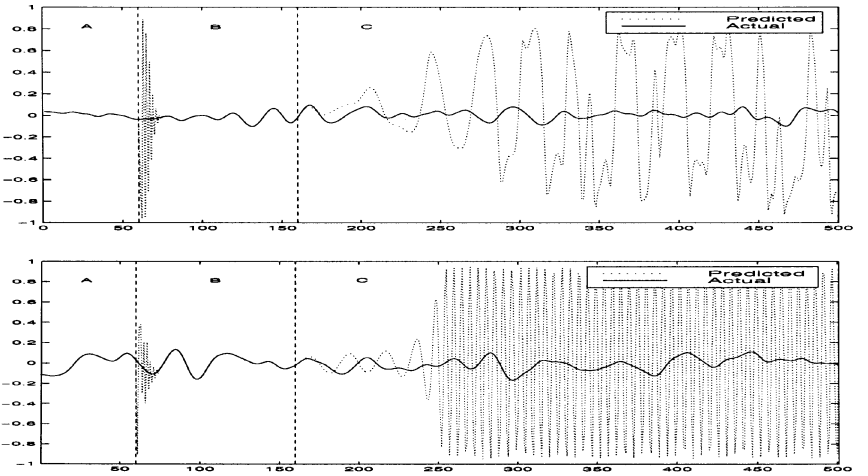
In this chapter, we have presented experimental results demonstrating the ability of the EKF-RMLP scheme to reconstruct the dynamics of chaotic processes in accordance with Takens' embedding theorem. The results presented for (1) the numerically generated logistic map, Ikeda map, and Lorenz attractor, and (2) the experimentally generated laser pulsations are indeed remarkable in that the reconstructed data closely match all the important chaotic characteristics of the original time series. Moreover, the dynamic reconstruction performed here is not only stable


A
B
C
Predicted
Actual

(*a*)


A
B
C
Predicted
Actual

(*b*)


A
B
C
Predicted
Actual

(*c*)

**Figure 4.29** Illustrative failure modes in iterative prediction of sea clutter by various networks. (*a*) Output converges to a constant (fixed point attractor); $N_0 = 4160$ and 160, respectively. (*b*) Output becomes periodic; $N_0 = 7160$. (*c*) Output breaks down after a few time steps; $N_0 = 160$ and 7160, respectively. Note that A = initialization, B = priming phase, and C = autonomous phase.

but also robust with respect to the choice of initial conditions and the presence of additive noise.

The results of this study also show that dynamic reconstruction of sea clutter is possible using the EKF-RMLP-based method depending on the data set used, but unfortunately as the results obtained with the EKF-MLP method were prone to numerous failures and therefore not as robust as one would expect from a good model. The results of this study clearly demonstrate that the dynamics of sea clutter is difficult to reconstruct. We attribute the difficulty to the unavoidable presence of dynamical noise in the process equation.

## REFERENCES

[1] H.D.I. Abarbanel, *Analysis of Observed Chaotic Data*. New York: Springer-Verlag, 1995.

[2] J.C. Schouten, F. Takens, and C.M. Van den Bleek, "Estimation of the Dimension of a Noisy Attractor," *Physical Review A*, **33**, 1134–1140 (1994).

[3] N.H. Packard, J.P. Crutchfield, J.D. Farmer, and R.S. Shaw, "Geometry from a time series," *Physical Review Letters*, **45**, 712–716 (1980).

[4] F. Takens, "Detecting Strange Actuators in Turbulence," in D.A. Rand and L.S. Young, Eds. *Dynamical Systems and Turbulence, Warwick, 1980*, Lecture Notes in Mathematics, Vol. 898.

[5] R. Mañé, "On the dimension of compact invariant sets of certain nonlinear maps," in D.A. Rand and L.S. Young, Eds. *Dynamical Systems and Turbulence, Warwick 1980*, *Lecture Notes in Mathematics* Vol. 898. 1981, p. 230. Berlin: Springer-Verlag.

[6] A.M. Fraser, "Information and entropy in strange attractors," *IEEE Transactions on Information Theory* **35**, 245–262 (1989).

[7] A.S. Weigend, "Paradigm change in prediction," *Philosophical Transactions of the Royal Society of London, Ser. A*, 405–420 (1994).

[8] U. Hubner, W. Klische, and C.O. Weiss, "Generalized dimensions of laser attractors," *Physical Review A*, **45**, 2128–2130 (1992).

[9] S. Haykin and S. Puthusserypady, "Chaotic dynamics of sea clutter" *Chaos*, **7**, 777 – 802 (1997).

[10] S. Haykin and Xiao bo Li, "Detection of signals in chaos," *Proceedings of the IEEE*, **83** (1995).

[11] H. Leung and S. Haykin, "Is there a Radar Clutter Attractor?" *Appl. Phys. Letters*, **56**, 393 – 395 (1990).

[12] H. Leung and T. Lo, "Chaotic Radar Signal-Processing Over the Sea," *IEEE J. Oceanic Engineering*, **18**, 287 – 295 (1993).

[13] A.J. Palmer, R.A. Kropfli, and C.W. Fairall, "Signature of Deterministic Chaos in Radar Sea Clutter and Ocean Surface Waves," *Chaos*, **6**, 613 – 616 (1995).

[14] S. Haykin, R. Bakker, and B. Currie, "Uncovering Nonlinear Dynamics: The Case Study of Sea Clutter," *Proc. IEEE*, under review.

[15] C.P. Unsworth, M.R. Cowper, B. Mulgrew, and S. McLaughlin, "False Detection of Chaotic Behaviour in Stochastic Computed K-distribution Model of Radar Sea Clutter," *IEEE Workshop on Statistical Signal and Array Processing*, pp. 296 – 300 (2000).

[16] G. Sugihara, "Nonlinear Forecasting for the Classification of Natural Time Series," *Phil. Trans. R. Soc. London*, A348, pp. 477 – 495 (1994).