# DETERMINING THE IMPORTANCE OF LEARNING THE UNDERLYING DYNAMICS OF SEA CLUTTER FOR RADAR TARGET DETECTION

*M. R. Cowper, C.P. Unsworth, and B. Mulgrew*

Signals and Systems Research Group,
Department of Electronics and Electrical Engineering,
Edinburgh University, King's Buildings, Mayfield Road,
Edinburgh, EH9 3JL, UK.
mrc@ee.ed.ac.uk

**ABSTRACT**

Existing evidence for and against sea clutter being chaotic and nonlinearly predictable is briefly discussed. Despite the uncertainty surrounding the chaotic nature of sea clutter, and its nonlinear predictability, the purpose of this paper is to examine what the best design criterion is for a nonlinear predictor which is to be used to detect targets against clutter which is known to be chaotic: mean square error performance or capturing the chaotic clutter's underlying dynamics. Single pulse detection analysis using a Swerling I target and chaotic *"clutter"* is carried out using predictor-based detectors in an attempt to determine which criterion is most suitable. The predictor detectors are compared with standard detection strategies.

## 1. INTRODUCTION

Researchers at McMaster University in Canada have claimed that sea clutter is a chaotic process [1] and that nonlinear predictor (NLP) networks can be used to improve the performance of maritime surveillance radars [2, 3]. However, criticisms of the measures used in [1] to categorise chaotic behaviour have been levelled [4, 5] which have thus called in to question the chaotic nature of sea clutter. Recent evidence [6] has challenged the view that NLP's can improve maritime surveillance radar performance.

Despite the uncertainty surrounding the chaotic nature of sea clutter, and its nonlinear predictability, the purpose of this paper is to examine what the best design criterion is for a NLP which is to be used to detect targets against clutter which is known to be chaotic. The reason for this investigation is that researchers at McMaster University [3] have advocated designing the NLP so that it captures the underlying dynamics [3, 7] of the chaotic clutter. However, recently [8] it has been found that capturing the underlying dynamics of a signal is not necessarily consistent with achieving the best mean square error (MSE) prediction performance. Essentially, the aim of this paper is to investigate which would work better, if clutter *were* found to be chaotic: a NLP detector (NLPD) [7] that consists of a NLP which had been trained to capture the chaotic clutter's underlying dynamics, or one which consisted of a NLP that had been trained to perform better, in terms of MSE prediction performance, than the NLP that had learnt the underlying dynamics.

This work was supported by EPSRC and BAE Systems

This investigation was carried out using Lorenz data [9] corrupted by white Gaussian noise as the *"clutter"* signal. Lorenz data is known to be chaotic, and therefore has known associated underlying dynamics. Noise was added as it was felt that this would more closely model a situation found in practice, plus it makes the problem of capturing the underlying dynamics of the Lorenz signal more difficult [8].

The paper is structured as follows. In section 2 noisy Lorenz data generation is described. In section 3 the detection strategies used are explained. In section 4 single pulse detection analysis using a Swerling I [10] target and the noisy Lorenz *"clutter"* is presented. Conclusions are presented in section 5.

## 2. GENERATING NOISY LORENZ DATA

The following coupled system of three nonlinear differential equations [11],

$$
\begin{aligned}
\frac{dx(t)}{dt} &= \sigma_L(y(t) - x(t)) \qquad (1)\\
\frac{dy(t)}{dt} &= r_L x(t) - y(t) - x(t)z(t)\\
\frac{dz(t)}{dt} &= x(t)y(t) - b_L z(t)
\end{aligned}
$$

where $\sigma_L$, $r_L$, and $b_L$ are dimensionless, describe the dynamics of the Lorenz attractor. For $\sigma_L$=10, $b_L = \frac{8}{3}$, the system behaves chaotically, whenever the Rayleigh number $r_L$ exceeds a critical value, which is approximately 24.74

[9]. Equation (1) can be solved for $x(t)$ using a $4^{th}$ order Runge-Kutta [12] technique with a suitable discrete step-size to produce a discrete Lorenz time series. White Gaussian noise [12] was added to the Lorenz data, the signal to noise ratio (SNR),

$$\mathrm{SNR} = 10\log_{10}\left[\frac{\sigma^2_{\mathrm{signal}}}{\sigma^2_{\mathrm{noise}}}\right] \tag{2}$$

where $\sigma^2_{\mathrm{signal}}$ is the variance of the signal of interest (in this case the Lorenz data), and $\sigma^2_{\mathrm{noise}}$ is the variance of the noise, was 25dB.

## 3. DETECTION STRATEGIES

A block diagram of a predictor detector is given in Figure 1. A K-step ahead predictor (linear, Volterra series filter

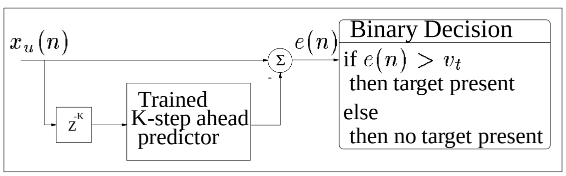



**Fig. 1**. Predictor-detector.

(VSF) [6,7,13], or normalised radial basis function network (NRBFN) [8,14]) was trained (for details on the training see [8]), then clutter samples which were not used to train the predictor $\{x_u(n)\}$ were presented to the predictor through a K-step delay. The clutter data $\{x_u(n)\}$ could consist of clutter alone, or clutter plus target. To determine if a target was present or not, the predictor error $e(n)$ was compared with a threshold level $v_t$. If the predictor error was greater than the threshold a target was declared to be present, and if it was less than the threshold level a target was not declared to be present.

Two different types of standard detectors [10] were used to compare with the predictor detectors. The two techniques used were a fixed threshold detector, and a cell-averaging constant false alarm rate detector (CA-CFAR). In particular, for the purposes of this paper, a distinction will be made between a CA-CFAR with a sliding window before the test cell, and a forward-backward CA-CFAR (FB-CA-CFAR), which has a sliding window either side of the test cell.

## 4. SINGLE PULSE DETECTION AGAINST A CHAOTIC CLUTTER BACKGROUND

The NRBFN predictors (NRBFNP's) used in [8] for the prediction analysis of the noisy Lorenz data were used for the detection analysis involving the noisy Lorenz *"clutter"*. In [8] it was shown that capturing a chaotic signal's underlying dynamics was not necessarily consistent with achieving the best MSE prediction performance. Specifically, it was shown that a NRBFNP with an embedding dimension (*i.e.* number of input nodes) of 7, an embedding delay (*i.e.* delay between each input node) of 3 samples, 400 kernels, and a training length of 2000 samples could be used to successfully capture the underlying dynamics of the noisy Lorenz data, whereas a NRBFNP with an embedding dimension of 7, an embedding delay of 1 sample, 400 kernels, and a training length of 2000 samples was not able to successfully capture the underlying dynamics, but did achieve a better MSE performance.

Detection analysis was carried out using the noisy Lorenz clutter, and a Swerling I target, for a signal to clutter ratio (SCR) of -6.99dB, where the target is the signal, and the noisy Lorenz data is the clutter. The following detectors were used. In addition to the NRBFN predictor detectors (NRBFNPD's), a linear predictor detector (LPD) with 30 taps, a VSF predictor detector (VSFPD) with an embedding dimension of 10 and embedding delay of 1 sample, a CA-CFAR with a window of size 1, a FB-CA-CFAR with a window of size 1 in the forward section and 1 in the backward section, and a fixed threshold detector were all used. Note that the sizes of the CFAR windows were found to be optimal in terms of detection performance. A prediction step of 1 sample (*i.e.* K=1) was used by each predictor. The following detection simulation parameters were common to all the detectors.

- Training data set length: 2000 samples.
- Non-training data set length: 35,000 samples.
- Number of target samples: 35,000.

The receiver operating curve (ROC) detection results for the above simulations are plotted in Figure 2. The prediction performance of the predictor detectors and CFAR detectors[1] are given in Table 1, in terms of normalised mean square error (NMSE),

$$\mathrm{NMSE} = 10\log_{10}\left(\frac{1}{\sigma_x^2 Y}\sum_{n=1}^{Y}(x(n+\mathrm{K}) - \hat{x}(n+\mathrm{K}))^2\right) \tag{3}$$

where $\sigma_x^2$ is the variance of $x$ over $Y$ samples, K is the prediction step, which was set equal to 1 sample for this detection analysis, $x(n+\mathrm{K})$ is the actual sample to be predicted at time step $n$ and $\hat{x}(n+\mathrm{K}))^2$ is the predictor's estimate of the actual sample to be predicted at time step $n$. From the

[1] A CFAR can be considered as a crude predictor, and consequently a NMSE figure can be worked out for it, using the CFAR error values, on the training and non-training data sets

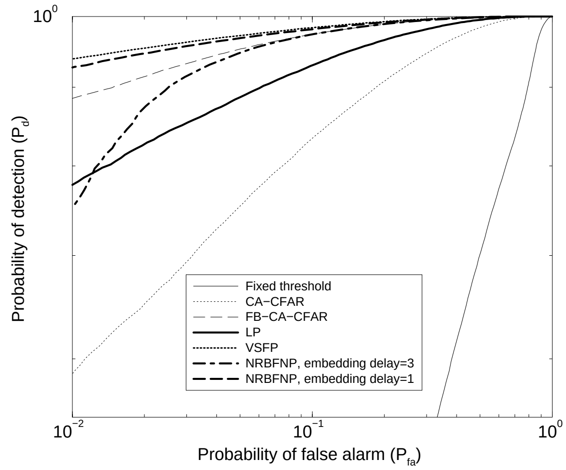


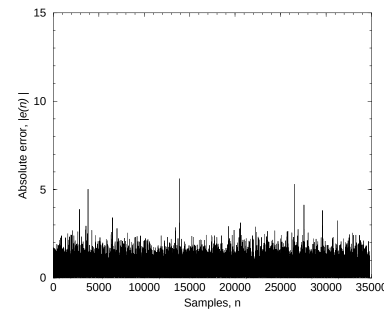

**Fig. 2**. Detection analysis using noisy Lorenz data for a SCR of -6.99dB.

| Predictor | Training NMSE [dB] | Non-training NMSE [dB] |
|---|---|---|
| NRBFNP, $\tau$=1 | -23.63 | -21.19 |
| NRBFNP, $\tau$=3 | -21.59 | -18.70 |
| VSFP | -23.38 | -21.79 |
| LP | -15.77 | -15.52 |
| FB-CA-CFAR | -19.90 | -19.92 |
| CA-CFAR | -11.02 | -11.08 |

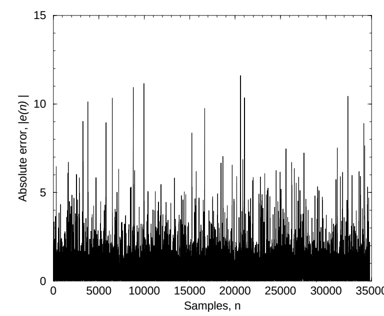
**Table 1**. Prediction NMSE performance for predictors on noisy Lorenz data, where $\tau$ denotes embedding delay.

results in Figure 2, it can be seen that the NRBFNPD with an embedding delay of 1 sample performed as well as, or better than, the NRBFNPD with an embedding delay of 3 samples. Taking into account the prediction results in Table 1, the NRBFNPD with the better NMSE performed better than the NRBFNPD which was able to capture the underlying dynamics of the noisy Lorenz data. From this evidence, given the choice between NMSE and capturing a signal's underlying dynamics, the predictor with the better NMSE should be incorporated into a predictor detector in favour of the predictor which captured the dynamics.

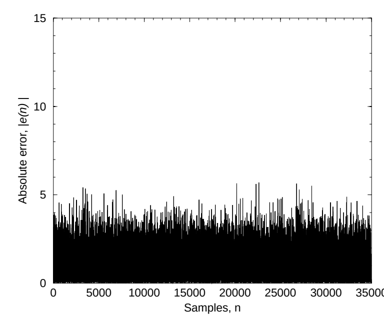
The rule that a better (*i.e.* more negative) NMSE value makes for a better predictor detector, or CFAR detector, seems to apply in Figure 2, except for the case of the NRBFNPD with an embedding delay of 3 samples (*i.e.* the network that managed to capture the underlying dynamics of the noisy Lorenz data). This NRBFNPD managed to achieve a better non-training data set NMSE value than the LPD, but for low $P_{fa}$ values the NRBFNPD with an embedding delay of 3 samples performed more poorly than the LPD. The reason for the poorer performance of the NRBFNPD, with an embedding delay of 3 samples, can be seen by considering the predictor errors plotted in Figure 3. The reason




(a)




(b)




(c)

**Fig. 3**. Predictor errors from the target detection analysis against a noisy Lorenz background: predictor errors for (a) a NRBFNPD with an embedding dimension of 7 and an embedding delay of 1 sample, (b) a NRBFNPD with an embedding dimension of 7 and an embedding delay of 3 samples, and (c) a LPD with 30 taps.

why the NRBFNPD with an embedding delay of 1 sample performed better than the NRBFNPD with an embedding delay of 3 samples, and the LPD, is because it produced smaller errors than the other predictor detectors, which allowed a better distinction to be made between error plus target samples and error only samples. The LPD was able to perform better than the NRBFNPD with an embedding delay of 3 samples, due to the same reason. Although the NRBFNPD with an embedding delay of 3 samples had a better *overall* NMSE value, it contained many error samples that were larger than those of the LPD, which resulted in poorer performance than the LPD, at low $P_{fa}$ values.

## 5. CONCLUSIONS

To summarise the results in section 4, for the case of training a NLP for use in a NLPD, it would appear that the smallest NMSE criterion would be preferred to the criterion of training a NLP to capture a signal's underlying dynamics, given that the network which had learnt the underlying dynamics had a poorer NMSE than the network which had not learnt the underlying dynamics. However, as in the case of the NRBFNPD with an embedding delay of 3 samples, using NMSE alone as a guide to the performance of a predictor detector can be deceiving.

In terms of processing chaotic signals, increasing the embedding delay has the effect of *"opening out"* the attractor in state space, which reduces the likelihood that noise will cause any vector to erroneously evolve (or jump) to the wrong part of the attractor. Avoiding such erroneous evolution eventualities results in correctly capturing the underlying dynamics of the chaotic signal in question. It would appear that increasing the embedding delay of a NLP has a negative impact on the performance of the corresponding NLPD. Therefore, capturing the underlying dynamics of chaotic clutter, and thus being able to reconstruct its attractor in state space do not appear to be consistent with designing the best NLPD from the evidence presented in this paper.

## 6. REFERENCES


[1] S. Haykin and S. Puthusserypady, "Chaotic dynamics of sea clutter," *Chaos*, vol. 7, no. 4, pp. 777–802, 1997.

[2] H. Leung and T. Lo, "Chaotic radar signal processing over the sea," *IEEE Journal Of Oceanic Engineering*, vol. 18, no. 3, pp. 287–295, July 1993.

[3] S. Haykin and X. B. Li, "Detection of signals in chaos," *Proceedings of the IEEE*, vol. 83, no. 1, pp. 95–122, January 1995.

[4] M. Davies, "Looking for non-linearities in sea clutter," in *IEE Radar and Sonar Signal Processing*, Peebles, Scotland, UK, July 1998, pp. 12/1–12/2.

[5] J. L. Noga, *Bayesian state-space modelling of spatio-temporal non-Gaussian radar returns*, Ph.D. thesis, Cambridge University, UK, 1998.

[6] M. R. Cowper, B. Mulgrew, and C. P. Unsworth, "An investigation into the use of nonlinear predictor networks to improve the performance of maritime surveillance radar target detectors," *IEE Proceedings on Radar and Sonar Navigation*, submitted July 2000.

[7] M. R. Cowper, *Nonlinear processing of non-Gaussian stochastic and chaotic deterministic time series*, Ph.D. thesis, Edinburgh University, March 2000.

[8] M. R. Cowper, B. Mulgrew, and C. P. Unsworth, "Nonlinear processing of chaotic signals using a normalised radial basis function network," *Signal Processing*, submitted October 2000.

[9] K. T. Alligood, T. D. Sauer, and J. A. York, *Chaos: An introduction to dynamical systems*, Springer-Verlag, 1997.

[10] M. I. Skolnik, *Introduction to radar systems*, McGraw-Hill, 2nd edition, 1981.

[11] S. Haykin and J. Principe, "Making sense of a complex world," *IEEE Signal Processing Magazine*, vol. 15, no. 3, pp. 66–81, May 1998.

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The art of scientific computing*, Cambridge University Press, 2nd edition, 1992.

[13] V. J Mathews, "Adaptive polynomial filters," *IEEE Signal Processing Magazine*, pp. 10–26, July 1991.

[14] J. Moody and C. J Darken, "Fast learning in networks of locally-tuned processing units," *Neural Computation*, vol. 1, no. 2, pp. 281–294, 1989.