SEA CLUTTER & CHAOS : IMPROVED SURROGATE-DATA TESTS

B.Mulgrew, C.P.Unsworth, M.R.Cowper, S.McLaughlin
The Signals & Systems Group,*
Department of Electronics & Electrical Engineering,
The University of Edinburgh, The Kings Buildings,
Edinburgh, Scotland, UK. EH9 3JL
Tel: +44 131 650 5580; fax: +44 131 650 6554
e-mail: Bernie.Mulgrew@ee.ed.ac.uk

ABSTRACT

Currently there is contention as to the nature of sea clutter for high resolution radar. Conventionally, sea clutter has been modelled as a compound stochastic k-distribution, originally suggested by Ward et al. However, recent work by Haykin et al. has suggested that the clutter can be modelled as a nonlinear deterministic process, otherwise referred to as a chaotic process. The paper presented here uses a new surrogate test which is designed specifically for this problem. The test is designed with a null hypothesis (H_0= The data can be approximated to by a compound stochastic k-distribution). Therefore, acceptance of such a test will accept the conventional k-distribution as a viable model sea clutter. In addition, a new surrogate statistic is introduced which is used to reject/accept the null hypothesis. This statistic is the normalised mean square error(NMSE) from a predictor and is a statistic which can be applied to any type of time-series. An overview of the method is presented together with results for a number a sea clutter data sets.

1 INTRODUCTION

The understanding and characterisation of the nature of sea clutter are critical in order to improve the performance of high resolution sea radar. Several models have been postulated in the radar literature to characterise sea clutter. The vast majority of these are stochastic in nature. Perhaps the most popular is the compound k-model[1] which has been used as the basis for the design of many operational radar systems. Part of the success of the model is its intuitive appeal combining a speckle component associated with short term motion of the sea with an underlying texture component associated with the wave motion. In contrast, Haykin et al.[2] claimed that high resolution sea clutter could be modelled quite differently as a nonlinear deterministic, or chaotic process, requiring only 5-6 coupled differential equations to describe the dynamics. Thus, presently a dilemma exists as to the true nature of seaclutter.

The statistical method of surrogate data testing was first introduced into nonlinear time-series analysis[3]. This was due to the difficulty encountered in the accurate determination of the chaotic invariants for real data. Haykin et al.[2] uses the same surrogate test, which makes use of the Fourier transform(FT), in his analysis of sea clutter. However, many problems have been identified with the FT surrogate[4-5] and recently in the invariants used as the surrogate statistic[6].

Work by Tough and Ward[7] has made it possible to synthesize compound stochastic k-distributed noise for a given shape and scale parameter. A more recent contribution by them[8] has extended this model to have the added benefit of retaining both the probability density function(pdf) and the power spectral density(PSD) of the texture component which until now was not possible.

In the paper presented here, an improved surrogate test designed specifically to re-address this issue will be introduced. This surrogate test uses the most recent model of Tough and Ward[8] to generate surrogate data. In addition, a new surrogate statistic is introduced which can be applied to any type of time-series. The result from this surrogate test will, for the first time, tell us if there is something that has been left un-modelled in the current compound stochastic k-distribution model, suggested by Tough and Ward.

2 SURROGATE DATA & HYPOTHESIS TESTING

Hypothesis testing[9] is a method used in statistical analysis. Essentially, one makes an assumption about the data, known as the null hypothesis (H_0) which one is hoping in general to statistically contradict.

One way a hypothesis can tested is by the method of surrogate data[3] which will now be described. The process involves measuring a statistic (Q_0) and then generating a set of synthetic time-series, known as surrogates, based around this statistic. There are many methods to generate the surrogate data. Regardless of the method employed, each method aims to achieve the same two

* This work was supported by BAE Systems, DERA Malvern, EPSRC and the Royal Society. Sea clutter data was provided courtesy of S.Lycett, DERA, Malvern and W.Dawber, J.Branson, DERA, Portsmouth.

effects.

(i)to retain the dynamics of the observed data;

(ii)to impart the null hypothesis (H_0) onto the surrogate.

Hence, the surrogate that is generated is the same as the original, in the sense that it retains the dynamics, but is different since additional features have been written into it from the null hypothesis. Thus, N surrogates of the observed data set are generated. Next, the same statistic is measured for all the N surrogates which results in a set of surrogate statistics $(Q_1, Q_2, Q_3, Q_4, ..., Q_N)$. The surrogate statistics together with the observed statistic (Q_0) are then ranked in order of numerical size. What is important now is the position of the observed statistic (Q_0) in this ranking structure. If H_0 is accepted Q_0 will appear in the main body of the ranked statistics. For H_0 to be rejected the observed statistic (Q_0) will fall at either end of the ranked statistics. Hence, a 'two tailed' test will reject H_0 at a particular significance level given by :

$$Pr(H_0 rejected) = \frac{Rank Q_0}{N+1} \quad (1)$$

2.1 Surrogate Methods

Various methods of surrogate data generation are widely employed in nonlinear time series analysis. The method employed in [2]to generate surrogate data makes use of the Fourier transform (FT)[3]. Various authors[4], have demonstrated that this can lead to artifacts in the surrogates which result in the spurious detection of nonlinearity. In addition, it has been shown that FT surrogates break down for systems with strong periodic components[5].

The main problem with the FT surrogate is that it retains only the PSD of the original data in the generated surrogate. The FT method does not retain the pdf of the original data in the surrogate. Instead the pdf is approximately Gaussian.

In this paper, the model used retains both the the PSD and pdf of the original data in the generated surrogate data. In addition, it generates compound stochastic k-distributed data for a specified set of parameters. Hence, this model as well as retaining both the PSD and pdf of the original data additionally integrates the hypothesis of the radar community into the surrogate test.

2.2 Measured Surrogate Statistics

The measured surrogate statistic used in the nonlinear time series analysis of sea clutter in [2] was the noise robust invariant documented by Schouten[9] known as the maximum likelihood of the correlation dimension, or (DML) value. Recently, Unsworth, et al. [6] demonstrated that this statistic falsely predicts chaos for known white and correlated stochastic time series.

In this paper a new surrogate statistic is introduced. This statistic is the normalised mean square error(NMSE) from both a linear and nonlinear predictor. This measure is also known to work for both deterministic and stochastic signals and does not have the same problems that have been identified with chaotic invariants. Hence, it is a more global statistic in the sense that it can be applied to any type of time series or distribution.

3 METHODOLOGY TO GENERATE COMPOUND STOCHASTIC k-DISTRIBUTED SURROGATES

The method of surrogate generation is summarised in the following steps (A schematic is shown in Figure 1):

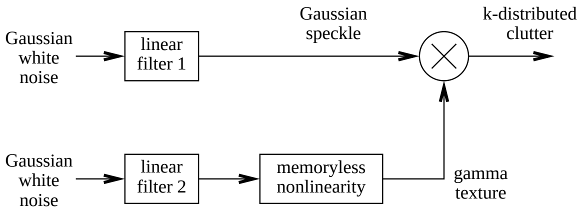

Figure 1: The compound k-dist. generator

(i) Data sets from radar sea trials and an experimental wave-tank radar systems are available which will be described later. For each data set an estimate of the observable features of the compound k-model is made. These include; the parameters of the gamma distribution; the autocorrelation sequence of the texture component; the autocorrelation sequence of the speckle component;

(ii) The parameters of the models are estimated. In particular the coefficients of filter 1 are obtained using spectral factorisation of the autocorrelation of the speckle sequence. The methods described in [8] are used to define the memoryless nonlinearity and the coefficients of filter 2.

(iii) Using these model parameters described in ii), surrogate data sets are generated. The quality of these data sets are checked by measuring the observable features from the surrogates and comparing these with those estimated from the original data.

Thus, the dynamics of the original data are retained in the surrogates since the parameters in (ii) are preserved and the null hypothesis is adhered to since the surrogates are generated to a known model which produces k-distributed noise.

3.1 Interpreting Surrogate Results

The correct interpretation of a surrogate test is very important, as described by Davies[10]. It is very easy to assume that if a null hypothesis is rejected that the hypothesis is wrong. However, if the 'method' that created the surrogate data is flawed, as in the FT surrogate[4], then rejection can also occur. However, this rejection is rather artifact resulting from a poor method of surrogate generation. This does not serve as a rejection of the hypothesis which could indeed be correct. Therefore, it is very important to be able to generate surrogate data by a robust method.

Thus regarding the work presented here, acceptance of H_0 will imply that the compound stochastic k-distribution model, suggested by Tough and Ward, is adequate to model the observed data. Rejection of H_0 will mean ONLY that the model does not describe the data well. Rejection does not imply the contrary, i.e. that the data is deterministic. It will simply state that this model is not appropriate.

4 SEA CLUTTER RESULTS

Two sea clutter data sets were used in this analysis. The first data set shall be referred to as the Dawber set and was taken from a stationary land-based radar pointed toward the sea. The second data set shall be referred to as the Wavetank set and was taken from a radar situated in a large wavetank laboratory. The radar parameters for both sets are shown below in Table 1.

Parameter	*Dawber*	*Wavetank*
Frequency	3GHz	15.75GHz
Pulse compression	not used	39ns Freq. chirp 500MHz BW
Resolution	150m	0.3m
Windspeeds	12.8m/s	12m/s
Beaufort scale	6(Strong Breeze)	6
Polarisation	VV	HH
PRF	20kHz	1kHz
Grazing angle	0.12^o	6^o
Beamwidth	6^o	5^o
No.datapoints	25,600	30,000

Table 1 : Radar parameters of sea clutter sets

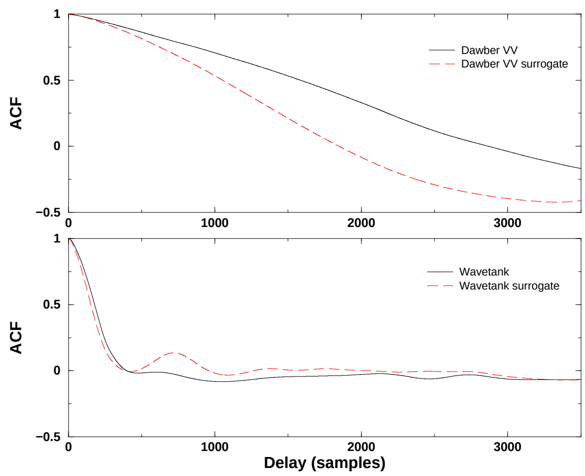

Figure 2: ACF of sea clutter set to its surrogate

For both sets the shape and scale parameters of the sea clutter were estimated as if the clutter were k-distributed. The Dawber gave shape=23 and scale=661. The Wavetank gave shape=0.1 and scale=0.035. For each data set the shape and scale parameters together with the autocorrelation function(ACF) and pdf of the set were passed to the Tough and Ward model[8] and surrogates data sets were produced. 50 surrogate data set were produced for each sea clutter set and each surrogate consisted of the same number of data points as its sea clutter counterpart.

The ACF for the surrogate and its sea clutter set are shown in Figure 2. It can be seen that the Wavetank ACF matches its surrogate to a delay of 400. A larger quantity of data would result in a better fit. The Dawber is even poorer than the Wavetank with a lag of 100. This is due to the high PRF with respect to the Wavetank PRF.

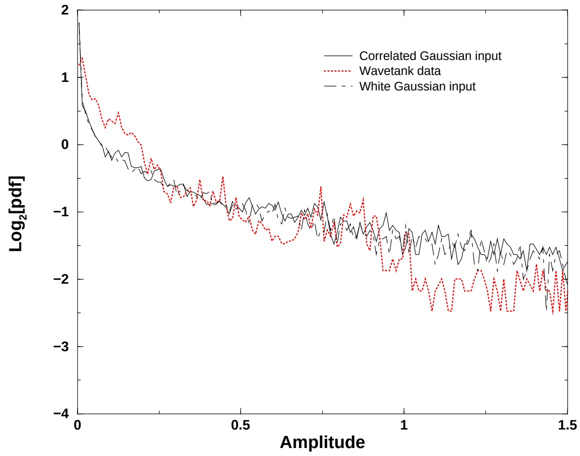

Figure 3: pdf comparison

Initially, moment estimation of the surrogate to the original was used to compare PDF. However, correlations effect moment estimation and due to a high degree of correlations in the data it becomes difficult to estimate these accurately. A difference in the gamma PDF, shown in Figure 3, of the surrogate and the actual clutter of the wavetank data also exists. However, the PDF's are affected to a lesser extent by the correlations than the moments are. What is needed is some parameter estimation technique which is independent of correlations in the data. Although it was not possible to estimate the parameters accurately it was still possible to demonstrate that the Tough and Ward model[8] was generating data correctly with the required pdf. This was achieved by simply using a white Gaussian input. As can be seen in Figure 3, the pdf for both a white Gaussian input and correlated Gaussian input are similar which implies that the model is generating data to the required pdf correctly.

4.1 Improved Surrogate Results

A 10 tap linear predictor(LP) and a cubic Volterra series filter nonlinear predictor(VSFP) with an embedding dimension of 10 were applied to each of the sea clutter and surrogate data sets. Both predictors were trained using the Householder transform. A training length of 10,000 and 8,000 was used for the wavetank and Dawber sets respectively. The validation results of the ranked NMSE for 50 surrogates together with the respective sea clutter set for both the LP and VSFP are shown if Figures 4 & 5 respectively. In [2], RBF and MLP predictors were used to perform the nonlinear prediction on sea clutter. Here we have elected to use the VSFP due to the work performed by Cowper[11] which showed that there was no difference between an RBF and VSFP on the clutter sets used here.

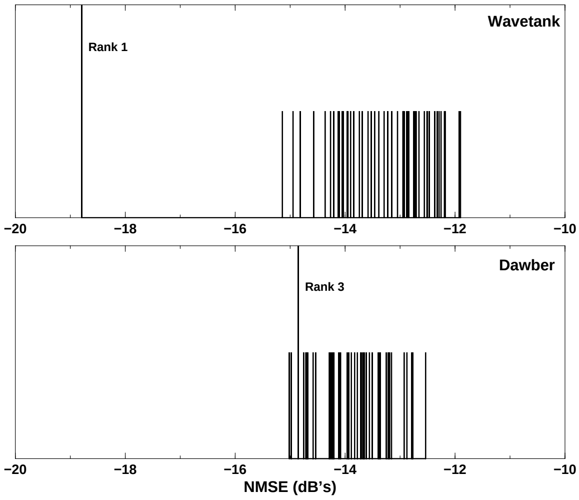

Figure 4: 10 tap linear predictor results

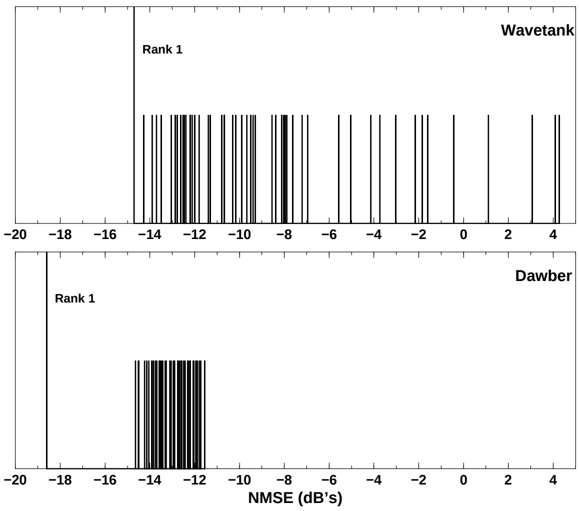

Figure 5: $d_E = 10$, cubic nonlinear VSFP results

For each of the figures, the surrogate statistics are represented by the short lines and the sea clutter statistic is represented by the tall line, together with its ranking in the distribution of the statistics. It is quite clear that the sea clutter statistic, in each case, is an outlier in the tail of the distribution. In the cases of LP wavetank, VSFP wavetank and VSFP Dawber, the sea clutter statistic is ranked 1st. This is evidence to suggest strong rejection of the the null hypothesis. For the LP Dawber the null hypothesis can be rejected at the 12% level. Hence, the results imply that the Tough and Ward model[8] does not sufficiently model the data.

5 Conclusion

In conclusion, a new method of surrogate generation has been introduced which uses Tough and Wards method[8] to generate surrogate data to a compound stochastic k-distribution model. This method has the benefit that it retains both pdf and PSD of the observed data in the surrogate which was not possible with the FT method used in [2]. Furthermore, this model allows the radar communities hypothesis to be integrated into the null hypothesis of the surrogate test. In addition, a new surrogate statistic has been introduced which is the NMSE from a predictor. This statistic is known to work for both deterministic and stochastic signals and does not have the same problems that have been identified with chaotic invariants. The new surrogate test has been applied to two sea clutter sets. The results show an overall rejection of the null hypothesis which implies that the Tough and Ward model does not sufficiently model the data presented here.

References

[1] K.D. Ward and C.J. Barker and S. Watts, "Maritime Surveillance Radar Part 2: Detection Performance Prediction in Sea Clutter," IEE Proceedings Part F, Vol.137, no.2, pp.63-72, 1990.

[2] S. Haykin and S.Puthusserypady, "Chaotic Dynamics of Sea Clutter," Chaos, Vol.7, no.4, pp.777-802, 1997.

[3] J.Theiler and S.Eubank, "Testing for nonlinearity in a time series: the method of surrogate data," Physica D, Vol.58, pp.77-94, 1992.

[4] A.M.Albano and P.E.Rapp, "Phase-randomised surrogates can produce spurious identification of non-random structure"

[5] C.J.Stam and J.P.M.Pijn and W.S.Pritchard, "Reliable detection of nonlinearity in experimental time series with strong periodic components", Physica D, Vol.112, pp.361-380, 1998.

[6] C.P.Unsworth and M.R.Cowper and S.McLaughlin and B.Mulgrew, "False detection of chaotic behaviour in the stochastic compound k-distributed model of radar sea clutter," Accepted to the 10th IEEE Workshop on statistical signal & array processing, Pennsylvania, August 14-16th 2000.

[7] K.D. Ward and R.J.A. Tough, "The Generation of Correlated k-distributed noise,", Tech.Rep.DRA/CIS/CBC3/WP9400/2.0/, DERA, Farnborough, England. GU14 6TD, 1994.

[8] R.J.A. Tough and K.D. Ward, "The Correlation Properties of Gamma & Other Non-Gaussian Processes Generated by Memoryless Nonlinear Transformations," submitted to the Journal of Physics D, 1999.

[9] J.C.Schouten and F.Takens and C.M.Van den Bleek, "Estimation of the dimension of a noisy attractor," Phys.Rev.E, Vol.50, No.3, 1994.

[10] M.Davies, "Looking for nonlinearities in seaclutter," IEE Radar & Sonar Signal Processing, pp.12/1-12/2, July 1998.

[11] M.R.Cowper, "Thesis-Nonlinear processing of non-Gaussian stochastic & chaotic deterministic time series," University of Edinburgh, Submitted Dec 1999.