MODELLING SEA CLUTTER USING CONDITIONAL HETEROSCEDASTIC MODELS

Jacek L. Noga & William J. Fitzgerald *
Cambridge University Engineering Department
Trumpington Street, Cambridge CB2 1PZ, UK
Tel: +44 1223 332 767; fax: +44 1223 332 662
e-mail: jln@eng.cam.ac.uk

ABSTRACT

In this paper a class of conditional heteroscedastic models is introduced in the context of sea clutter modelling. In particular, an Auto-regressive (AR) process driven by conditional heteroscedastic (CH) errors (AR-CH model) is proposed as a model for the time evolution dynamics of the modulating component of sea clutter. The CH process parameters of the AR-CH model determine the weight of the tails of the marginal distribution, while the AR component largely determines the correlation structure. Different functional forms of conditional variance models are investigated using real sea clutter data.

1 INTRODUCTION

Sea clutter is widely agreed to be a compound process [8] where the fast oscillating speckle component is modulated by a slowly varying underlying process, σ, associated with the swell. It has been found empirically that the single point statistics of sea clutter is very well modelled by the K-distribution [8]. The K-distribution results from a Rayleigh distributed speckle, modulated by Gamma distributed process, σ. Sea clutter has also been modelled as a Spherically Invariant Random Process (SIRP) [2], of which the K-distribution is a member. SIRPs allow the modelling of correlations of the speckle component, but assume the underlying modulation process, σ, to be independent from one spherically invariant random vector (SIRV) to the next. However, modelling of the correlation structure of the modulating component becomes important, particularly in the context of CFAR detection.

*This work was supported in part by DRA Agreement Number CSM3202905

Using arguments based on birth-death migration processes, Jakeman [4] derives a Stochastic Differential Equation (SDE) model for the underlying modulating process of sea clutter. However, although the resulting process has a Gamma distributed single point statistics, such models are capable only of capturing a limited number of autocorrelation functions [7]. Furthermore, the practical utility of such models is limited in that when discretized, the resulting difference equations assign non-zero probability to negative values of σ.

In this paper models based on an auto-regressive process driven by conditionally heteroscedastic errors are proposed to model the time evolution dynamics of $\log \sigma$. Such models, just like Jakeman's SDE model, allow the noise variance to be state dependent (*i.e.* , depend on previous observations, prediction errors, or any other *past* information). Unlike the SDE based model, however, the auto-regressive process driven by conditional heteroscedastic noise provides a flexible framework for modelling processes with a wide range of marginal distributions and correlation functions.

In section 2 conditional heteroscedastic models studied in this paper are formally introduced. Results obtained using autoregressive models driven by conditional heteroscedastic noise applied to real sea clutter data are presented in section 3. Discussion of the results and implications to sea clutter modelling is presented in section 4, while the summary of the main points raised in the paper is presented in section 5.

2 THE MODEL

Consider a model for a process y_t, which in the present case will correspond to the logarithm of the underlying modulating component of sea clutter (*i.e.* , $y_t = \log \sigma_t$). Denote an auto-regressive model of order p driven by conditional heteroscedastic process of order q (with parameter vector $\beta = \{\beta_0, \beta_1, \ldots, \beta_q\}$) by AR(p)-CH(q). Such a process takes on the form

$$\begin{aligned} y_t &= \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i} + e_t \\ e_t &= h_t^{1/2} \epsilon_t \qquad \epsilon_t \sim N(0,1) \\ h_t &= f(\beta; e_{t-1}, \ldots, e_{t-q}; y_{t-1}, \ldots, y_{t-q}) \end{aligned}$$

The above model is conditionally Gaussian, but the process variance is not necessarily constant. Instead, the variance is allowed to depend on the previous observations, previous errors or any other *past* information [3].

The most widely studied conditional heteroscedastic model is the original ARCH model of Engle [3] denoted by ARCH(q). For such a model the current variance is a linear combination of previous residual squared errors, *i.e.* , $h_t = \beta_0 + \sum_{j=1}^{q} \beta_j e_{t-j}^2$. However, due to positivity constraints on the model parameters β, it proves to be more convenient to work with exponential conditional heteroscedastic models [1] of the form

$$\text{ECH(q)} \qquad \log h_t = \beta_0 + \sum_{j=1}^{q} \beta_j \log y_{t-j}^2$$

$$\text{EARCH(q)} \qquad \log h_t = \beta_0 + \sum_{j=1}^{q} \beta_j \log e_{t-j}^2$$

Such models have no positivity constraints on the parameters, and the model parameters can readily be estimated using simple scoring algorithm [3, 6].

A large body of literature exists on the moment structure [5], stationarity [3] and parameter estimation [6] of ARCH processes. Although conditionally Gaussian, the study of higher order moments of the marginal distribution indicates that the conditional heteroscedastic processes are leptokurtic (heavy tailed).

3 RESULTS

The results presented in this section are based on the data obtained using RSRE coherent I-band radar operating in polarimetric mode with recordings made in HH and VV polarisation at the effective sampling frequency of 500Hz per channel. The raw data has been averaged over blocks of 250 samples to remove the speckle. The histogram of 5000 samples of the averaged data corresponding to the modulating component of sea clutter is shown in figure 1 and 2, for HH and VV polarisation channels, respectively.

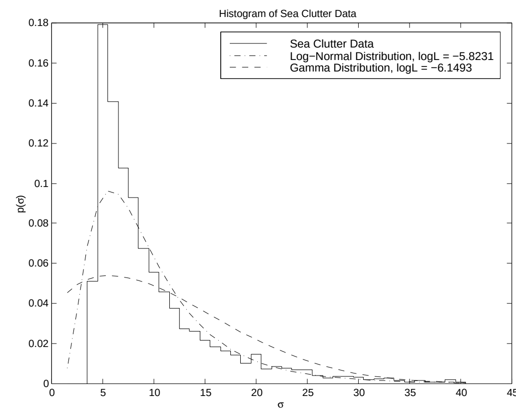

Figure 1: Histogram of HH data

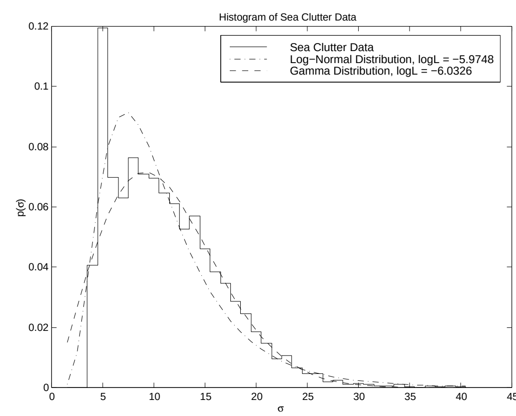

Figure 2: Histogram of VV data

The histogram in figure 1 shows the spiky nature of the horizontally polarised clutter. The log-Normal and Gamma distributions corresponding to maximum likelihood parameter estimates, shown in figure 1, demonstrate that the single point statistics of the modulating component of the horizontally polarised clutter is well modelled by the Log-Normal distribution. On the other hand, figure 2 indicates that the single point statistics of the modulating component of vertically polarised sea clutter is equally well modelled by log-Normal and Gamma distributions.

Although widely accepted to be Gamma distributed, results presented in figure 1 and 2 suggest that the marginal distribution of the modulating component varies widely with polarisation, sea state, viewing angle, *etc.* The good fit of the log-normal distribution to the marginal indicates that the modulating component of the sea clutter, σ, is best modelled in the logarithm domain.

AR(p)-EARCH(q) and AR(p)-ECH(q) models were fitted to 200 samples of the logarithm of averaged data (*i.e.* , $\log \sigma$). The parameters were estimated by maximum likelihood using the scoring algorithm, for different model orders p and q. It was found that no significant improvement in modelling ability is obtained for models with $q > 1$.

The averaged log likelihoods obtained for the 200 sample training set for models AR(p), AR(p)-EARCH(1) and AR(p)-ECH(1) are shown in figures 3 and 5 for horizontally and vertically polarised data, respectively. The corresponding averaged log likelihoods for a 200 sample test set (taken from the same clutter file) for the abovementioned models are shown in figure 4 and 6.

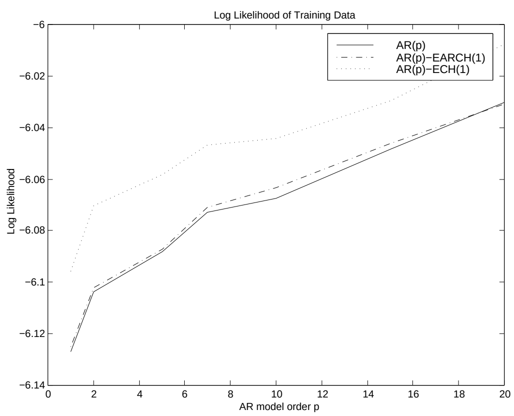

Figure 3: Log Likelihood of HH training data

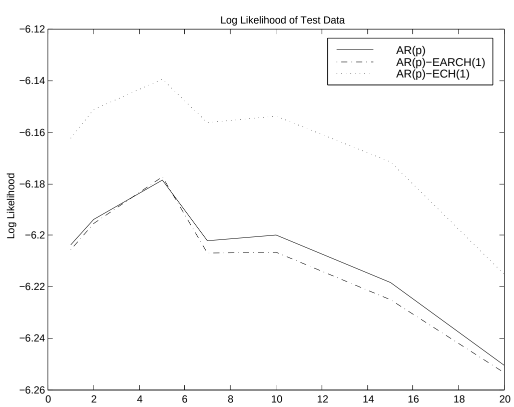

Figure 4: Log Likelihood of HH test data

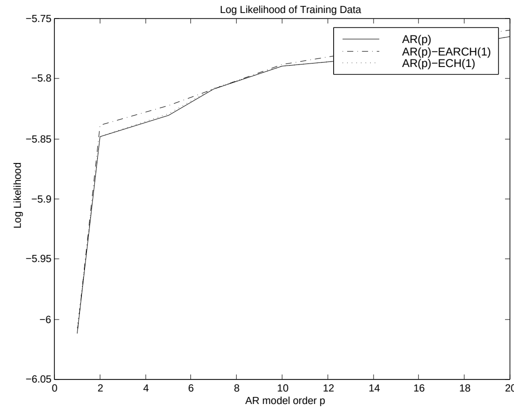

Figure 5: Log Likelihood of VV training data

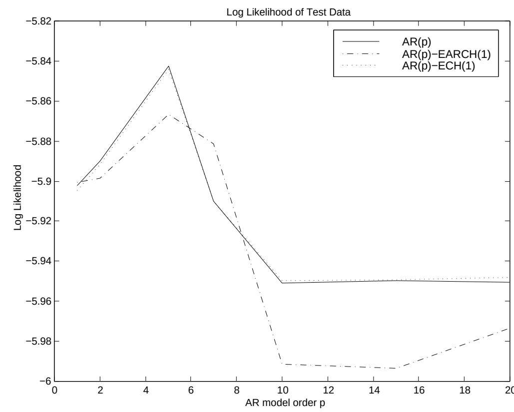

Figure 6: Log Likelihood of VV test data

4 DISCUSSION

The results obtained for vertically polarised clutter, shown in figure 5, indicate that the log likelihood for the training set is approximately the same for all three models (AR(p), AR(p)-EARCH(1) and AR(p)-ECH(1)). Although for the training data set, the AR(p)-EARCH(1) model outperforms the other two models, it can be seen from figure 6 that this model does not generalise well to the test data set. In fact, in the case of the vertically polarized data, the AR(p) and AR(p)-ECH(q) models are essentially indiscernible.

In the case of horizontally polarised clutter, it can be seen from figure 3 and 4 that the AR(p)-ACH(1) model outperforms the other two models in the average log likelihood sense on both the training and the test data sets. The improvement is very marked, both with respect to the pure autoregressive process AR(p), as well as the AR(p)-EARCH(1) model. This result indicates that the modulating component of sea clutter is not well modelled by log-Normal distribution, which would be consistent with a pure autoregressive process in the logarithmic domain. It does, however, indicate that there is a strong dependence of the process variance on the previous observations (*i.e.* , $h_t = f(\log y^2)$).

Such behaviour is not uncharacteristic of real world processes. In financial time series, where the ARCH models were first introduced, the process variance is a function of previous prediction errors. Such dependence stems from the fact that uncertainty in the past indicates uncertainty about the future.

In the case of sea clutter, the modulating component is associated with the swell structure of the sea. It is not unreasonable to expect that a stochastic model associated with this process has non-constant noise variance. Such state-dependent behaviour is clearly visible in the stochastic differential equation formulation based on birth-death migration processes used by Jakeman [4] to justify the Gamma marginal distribution for the modulating component of sea clutter.

Conditional heteroscedastic noise models, due to the inherent state dependence of the noise process variance, can also potentially explain the sea spike behaviour observed in sea clutter. Furthermore, it has been found from experiments on real sea clutter data that the auto-regressive models driven by conditional heteroscedastic noise can reproduce a wide range of marginal distributions and correlation functions observed in sea clutter.

5 SUMMARY

In this paper Auto-Regressive models with Conditional Heteroscedastic errors (AR-CH models) were introduced in the context of modelling of the modulating component of sea clutter.

Using empirical data, it was demonstrated that the logarithm of the modulating component is well modelled by the AR-ECH model, for which the log variance is a regressive function of the logarithm of previous observations. Very strong evidence for conditional heteroscedascity was found, particularly for spiky (*e.g.* , horizontally distributed) clutter.

Due to the conditional Gaussian form of the AR-CH models, the underlying modulating component of the clutter is conditionally log-Normal, although the corresponding marginal is not. In fact, depending on the model parameters, the marginal can be leptokurtic and a wide range of correlation functions can be entertained.

Though presented in the context of modelling the temporal dynamics of sea clutter, the models presented in this paper can readily be extended to model the spatio-temporal dynamics of the clutter. Such models form the basis for adaptive CFAR detectors that are currently being studied by the authors.

6 REFERENCES

[1] T Bollerslev, RT Chou, and KF Kroner. ARCH modelling in finance. *Journal of Econometrics*, 52:5–59, 1992.

[2] E Conte and M Longo. Characterisation of radar clutter as a Spherically Invariant Random Process. *IEE Proceedings F*, 134(2):191–97, April 1987.

[3] RF Engle. AutoRegressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, July 1982.

[4] E Jakeman and RJA Tough. Non-Gaussian models for the statistics of scattered waves. *Advances in Physics*, 37(5):471–529, 1988.

[5] Milhoj. The moment structure of ARCH processes. *Scandinavian Journal of Statistics*, 12:281–292, 1986.

[6] KW Newey and DG Steigerwald. Asymptotic bias for Quasi-Maximum-Likelihood estimators in conditional heteroskedasticity models. *Econometrica*, 65(3):587–599, May 1997.

[7] CJ Oliver and RJA Tough. On the simulation of correlated K-distributed random clutter. *Optica Acta*, 33(3):223–250, 1986.

[8] KD Ward, CJ Baker, and S Watts. Maritime survillance radar part 1: Radar scattering from the ocean surface. *IEE Proceedings F*, 137(2):51–62, April 1990.