We propose a GARCH model to represent the clutter in radar applications. We fit this model to real sea clutter data and we show that it represents adequately the statistics of the data. Then, we develop a detection test based on this model. Using synthetic and real radar data, we evaluate its performance and we show that the proposed detector offers higher probability of detection for a specified value of probability of false alarm than tests based on Gaussian and Weibull models, especially for low signal to clutter ratios.

© 2013 Published by Elsevier Inc.
detector obtained by assuming the data is Gaussian, and a third detector based on Weibull clutter envelope. For low values of false alarm probability, our method results in a higher detection probability than the others detectors, independently of the value of signal to clutter ratio (SCR). For high SCR and high false alarm probability, any detector achieves more or less the same detection probability.

2. Estimation problem and detection scheme

2.1. GARCH process

The class of stochastic processes that presents autoregressive conditional heteroscedasticity (ARCH) was proposed in [10]. This processes have been widely used in econometrics to model different financial variables. Their main feature is that they present a Gaussian conditional distribution whose variance changes over time, but their unconditional variance is constant. Later, an extension of the ARCH processes to GARCH processes was introduced [11], similar to the extension of the AR processes to ARMA processes in time series theory. An advantage of the last ones is that they allow a more flexible delays structure, especially with delays process in time series theory. An advantage of the last ones is that they allow a more flexible delays structure, especially with long memory structures.

A stochastic process is heteroscedastic if its variance changes over time. In the case of GARCH processes, the conditional heteroscedasticity (ARCH) was proposed in [10]. This justifies the name conditional variance used for its unconditional variance is finite and equal to

\[ \sigma^2 = \sum_{j=1}^{q} \alpha_j \sigma^2_{t-j} + \sum_{j=1}^{p} \beta_j |v_{t-j}|^2, \]

where \( C/N \) denotes circular normal distribution and \( k, \alpha_j \) and \( \beta_j \) are the process coefficients. The GARCH \((p, q)\) models the return as a Gaussian white noise process with nonconstant conditional variance, \( \sigma^2_t \) is a linear function of its past values, \( \sigma^2_{t-1}, \ldots, \sigma^2_{t-p} \), and a quadratic function of the past values of the return process envelope, \( |v_{t-1}|, \ldots, |v_{t-q}| \).

We must impose some constraints on the model coefficients to obtain desirable properties. The natural constraints are [11]

\[ k > 0, \quad \alpha_j \geq 0, \quad j = 1, \ldots, p, \quad \beta_j \geq 0, \quad j = 1, \ldots, q, \]

to ensure that the conditional variance is always positive.

Let \( \psi \) denote the set of all information up to time \( t \), i.e., \( \sigma^2_t \) and \( v_t \) for \( t \leq t - 1 \), then

\[ v_t/\psi \sim C/N(0, \sigma^2_t). \]

This justifies the name conditional variance used for \( \sigma^2_t \). Moreover it is possible to show that \( E[v_t] = 0 \), and if

\[ \sum_{j=1}^{p} \alpha_j + \sum_{j=1}^{q} \beta_j < 1, \]

its unconditional variance is finite and equal to [11]

\[ \text{Var}[v_t] = \frac{1}{1 - \sum_{j=1}^{p} \alpha_j - \sum_{j=1}^{q} \beta_j}. \]

In this case, because GARCH processes are serially uncorrelated, \( v_t \) is a wide sense stationary process.

It is not possible to find an explicit expression for the probability density function (pdf) of this kind of processes. However there are conditions such as (5) which ensure that the process moments of any order exist [10,11].

2.2. Quasi-maximum likelihood estimation

One of the aims of this work is to model the clutter as a GARCH process, that is fitting a GARCH model to signal samples from clutter measured by the radar. By fit we mean to estimate the coefficients of the process conditional variance. A possible way to perform this estimation would be using the maximum likelihood method, but there is no explicit expression for the pdf of a GARCH \((p, q)\) vector \( v = [v_1, \ldots, v_n]^T \), since the distribution of \( [\sigma_1, \ldots, \sigma_n]^T \) is not known. To overcome this difficulty, several authors consider instead the conditional likelihood function,

\[ \ell(\theta) = -n \ln(f(v/v_0, \ldots, v_{t-1})), \]

where \( \theta = [k \alpha_1 \cdots \alpha_p \beta_1 \cdots \beta_q]^T \) is the parameter vector that we want to estimate.

Then, the quasi-maximum likelihood estimator (QMLE), \( \hat{\theta} \), is the value of \( \theta \) that maximizes \( f(v/v_0, \ldots, v_{t-1}) \), or, equivalently, that minimizes \( \ell(\theta) \) [13]. This estimator must also satisfy the constraints for the coefficients (3) and (5). In summary, we may write

\[ \hat{\theta} = \arg\min_{\theta \in \Theta} \ell(\theta), \]

where \( \Theta \subset (0, \infty) \times \Omega \) with \( \Omega = ([\alpha_1 \cdots \alpha_p]^T \in [0,1]^p \land [\beta_1 \cdots \beta_q]^T \in [0,1]^q / \sum_{j=1}^{p} \alpha_j + \sum_{j=1}^{q} \beta_j < 1) \).

2.2.1. QMLE asymptotic properties

Let \( v_t \) be the stationary GARCH \((p, q)\) process defined by (1) and (2) with parameter vector \( \theta_0 = [k \alpha_1^0 \cdots \alpha_p^0 \beta_1^0 \cdots \beta_q^0]^T \in \Theta \). Suppose the polynomials \( \alpha^0(z) = \alpha_0^0 z + \cdots + \alpha_p^0 z^p \) and \( \beta^0(z) = 1 - \beta_1^0 z - \cdots - \beta_q^0 z^q \) do not have common zeros. Then the QMLE obtained from (8) is strongly consistent [13,14], i.e.,

\[ \hat{\theta} \to \theta_0 \quad \text{as} \quad n \to \infty. \]

Since \( z_1 \) is an iid normal sequence, and assuming that \( \theta_0 \) is in \( \Theta \), it is possible to show that the estimator \( \hat{\theta} \) is also asymptotically normal [13,14]. However, it is not possible to establish a number of samples needed to get a desired estimation error, because it is impossible to explicitly compute the covariance matrix of this asymptotic distribution.

2.3. Alternative method for coefficient estimation

As we shall see later on, to achieve an acceptable estimation error the method presented in the former subsection requires a number of samples that can be larger than the number available...
in realistic radar applications. To overcome this limitation, in this subsection we propose the use of several independent process realizations of short length instead of a single long one to estimate the values of the coefficients, and we derive the likelihood function using this approach.

Let \( v^{(1)}, v^{(2)}, \ldots, v^{(M)} \) be \( M \) independent process realizations, then the likelihood function, \( f(v^{(1)}, v^{(2)}, \ldots, v^{(M)}) \), may be written as the product of the \( M \) pdfs \( \{ f(v^{(1)}, v^{(2)}, \ldots, v^{(M)}) \}_{i=1}^{M} \), which have the form of \( f(v/v_0, \ldots, v_{1-r}) \).

The log-likelihood function is in this case
\[
\ell(\theta) = \sum_{i=1}^{M} \left[ n \ln(\pi) + \sum_{t=1}^{n} \ln(h_{t}^{(i)}) + \frac{|v_{t}^{(i)}|^2}{h_{t}^{(i)^2}} \right],
\]
where \( h_{t} = \sigma_{t}^2 \).

Then, we obtain the QMLE of \( \theta \) solving the optimization problem (8) with \( \ell(\theta) \) given by (9).

2.4. Detection

In a general setup, the radar transmits a pulse and the receiver samples the reflected signal at the output of the matched filter. If we denote the complex sample at the output of the quadrature demodulator stage as \( y_{t} \), then the detection procedure is given by the decision between the two hypotheses \( H_0 \) and \( H_1 \) after \( y_{t} \) has been received from the range cell under test
\[
\begin{align*}
H_0 &: y_{t} = v_{t}, \\
H_1 &: y_{t} = x_{t} + v_{t}. 
\end{align*}
\]

Under the null hypothesis \( H_0 \) it is assumed that the data consist only of clutter \( v_{t} \), which is modeled as a GARCH(p, q) process. We assume that the electronic noise is negligible or is part of the clutter model. Under the hypothesis \( H_1 \) it is assumed that the measurement is the combined result of clutter and echoes from a target, \( x_{t} \). Since \( y_{t} \) is the sample at the output of the matched filter we can think that \( x_{t} \) is a constant proportional to the energy of the signal. However, a perfect knowledge of the received signal implies knowing the range of the target very precisely. Because at microwave frequencies this precision is not feasible, it is more realistic to assume that it is known only to within a timing-error factor \( e^{\theta} \), where \( \theta \) is a random variable uniformly distributed over \( (0, 2\pi) \) [18]. Then the resulting signal model is \( x_{t} = E e^{\theta} \), where \( E \) is a deterministic constant.

Under the well-known Neyman–Pearson criterion, the decision is designed to maximize the probability of detection, \( P_d \), under the constraint that the probability of false alarm, \( P_{fa} \), does not exceed a given value. The solution of this optimization problem leads directly to the decision rule [18,19]
\[
\Lambda(y_{t}) = \frac{f_y(y_{t}/H_1)}{f_y(y_{t}/H_0)} \geq \eta,
\]
known as likelihood ratio test, where \( \Lambda(y_{t}) \) is called the likelihood ratio. \( f_y(y_{t}/H_1) \) and \( f_y(y_{t}/H_0) \) are the pdfs of \( y_{t} \) given that the target was present and given that the target was not present, respectively, and \( \eta \) is the decision threshold to be determined.

2.4.1. GARCH model of clutter and unknown parameters

Several issues make our problem depart from the classical hypothesis test above. First, we highlight that it is not possible to find an explicit expression of the pdf of a GARCH process. Thus we cannot find an expression for the likelihood ratio to obtain the decision rule. To overcome this difficulty, one can consider instead the conditional pdf given the process observations up to the range cell under test, \( y_{t-1}, \ldots, y_{1} \), in the same way we did in Section 2.2. This is a reasonable approach because, except for the first cell, the previous samples to the range cell under test are known. If we define \( \psi_{t} = [y_{t-1} \ldots y_{1}]^{T} \), then we use a conditional likelihood ratio, \( \Lambda(y_{t}/\psi_{t}) \), given by
\[
\Lambda(y_{t}/\psi_{t}) = \frac{f_y(y_{t}/\psi_{t}; H_1)}{f_y(y_{t}/\psi_{t}; H_0)},
\]
to make the decision.

On the other hand, we do not have a perfect knowledge of the conditional pdfs. Regarding the clutter model, we do not know the coefficients of the GARCH process and the phase of the signal model. This leads to a composite hypothesis testing problem.

We adopt a Bayesian approach for the signal phase [18] and a generalized likelihood ratio test (GLRT) [19] for the clutter coefficients.

If we denote the prior pdf of \( \theta \) as \( f_{\theta}(\theta) \), the conditional pdfs of the data are
\[
\begin{align*}
&f_{y}(y_{t}/\psi_{t}; H_0) = \int f_{y}(y_{t}/\psi_{t}; \theta) f_{\theta}(\theta) d\theta, \quad i = 0, 1.  \\
&f_{y}(y_{t}/\psi_{t}; H_1) = \int f_{y}(y_{t}/\psi_{t}; \theta) f_{\theta}(\theta) d\theta,
\end{align*}
\]
From (4) and (10), \( f_{y}(y_{t}/\psi_{t}; \theta; H_0) \) and \( f_{y}(y_{t}/\psi_{t}; \theta; H_1) \) are the pdfs of the distributions \( CN(0, \sigma_{t}^2) \) and \( CN(E e^{\theta}, \sigma_{t}^2) \), respectively, with
\[
\sigma_{t}^2 = \hat{k} + \hat{\alpha}_3 \sigma_{t-1}^2 + \cdots + \hat{\alpha}_p \sigma_{t-p}^2 + \hat{\beta}_1|v_{t-1}|^2 + \cdots + \hat{\beta}_q|v_{t-q}|^2,
\]
where \( \hat{k}, \hat{\alpha}_3, \ldots, \hat{\theta}_p, \hat{\beta}_1, \ldots, \hat{\beta}_q \) are the quasi-maximum likelihood estimates of the coefficients.

We see that \( f_{y}(y_{t}/\psi_{t}; H_0) = f_{y}(y_{t}/\psi_{t}; \theta; H_0) \) because it does not depend on \( \theta \). Furthermore, as in the Gaussian case [18], after integrating (13) over \( \theta \) for the hypothesis \( H_0 \) we get
\[
\begin{align*}
&f_{y}(y_{t}/\psi_{t}; H_1) = \frac{1}{\pi \sigma_{t}^2} e^{-(|y_{t}|^2 - \bar{E}^2)/\sigma_{t}^2} I_{0}(2 \bar{E}|y_{t}|/\sigma_{t}^2),
\end{align*}
\]
where \( I_{0}(\cdot) \) is the modified Bessel function of the first kind.

We obtain then the following decision rule
\[
\begin{align*}
&\left[ \ln \left( \frac{2 \bar{E}|y_{t}|}{\sigma_{t}^2} \right) \right] \overset{H_1}{\gtrless} \ln(\eta) + \frac{E^2}{\sigma_{t}^2} = \lambda,
\end{align*}
\]
Since the function \( \ln[I_{0}(\cdot)] \) is monotonically increasing, the same detection results can be obtained by simply comparing its argument to a modified threshold. The decision rule then becomes
\[
|y_{t}| \overset{H_1}{\gtrless} \lambda'.
\]
We obtained a decision criterion but the specific value of the threshold, \( \eta \) or \( \lambda' \), that will ensure that the false alarm probability does not exceed a given value has to be found.

The classic way to evaluate the performance of a detector is by means of its detection and false alarm probabilities. Once an expression for the \( P_{fa} \) in terms of \( \lambda' \) is obtained, it is inverted to obtain the threshold setting in terms of the \( P_{fa} \). Thus the distribution of \( u_{t} = |y_{t}| \) under each of the two hypotheses is needed. Again, we do not have explicit expressions of the pdfs of \( v_{t} \) or \( y_{t} \). Thus, to determine a threshold we resort to the conditional false alarm probability given \( y_{t-1}, \ldots, y_{1} \), denoted \( P_{fa}/\psi_{t} \).

Under the hypothesis \( H_0 \) the conditional pdf of \( u_{t} \) given \( \psi_{t} \), \( f_{u}(u/\psi_{t}; H_0) \), is Rayleigh with parameter \( \sigma_{t}^2/2 \). Then the \( P_{fa}/\psi_{t} \) is given by
\[
P_{fa}/\psi_{t} = \int_{0}^{\infty} f_{u}(u/\psi_{t}; H_0) du = e^{-\lambda'/|\sigma_{t}|^2}.
\]
Finally, we can solve (17) to obtain the threshold $\lambda'$, which results

$$\lambda' = \sqrt{-\sigma^2 \ln(P_{FA}/\psi)}.$$  (18)

If we set the $P_{FA}/\psi$, then from (18) we obtain a value for the threshold in each cell. With this procedure we get an adaptive threshold $\lambda'$ which depends on the previous clutter samples for each cell under test. A problem arises when the decision has to be made in a cell after a detection, because when the signal received by the radar corresponds to signal plus clutter, it is not possible to separate the contribution of the clutter. In this work we consider isolated point targets, so this is not a problem. The decision algorithm can be modified for extended targets by changing the way in the threshold is set after a presumed detection. An alternative would be to restart the detection algorithm after the statistic exceeds the threshold using the unconditional standard deviation, or to set the threshold after a presumed detection based on the last sample for which no detection occurred.

3. Simulations

3.1. Estimation error versus number of samples

In order to have a guideline to the estimation quality when real data are used, we performed an analysis of the estimation error as a function of the number of samples of the observation vector. This analysis was done through numerical simulations which consisted in the generation of independent process realizations of different length, $N$, and the estimation of the model coefficients for each realization. The estimation technique, described in [8], was solved using the Matlab Active Set Algorithm, implemented through the method SQP (Sequential Quadratic Programming) [20]. This procedure was repeated 500 times for each case, i.e., we obtained 500 values of the coefficient estimates for each value of $N$. We only include results corresponding to the GARCH(0, 1) process, but similar results were obtained for different process orders, $p$, $q$. The realizations were generated with coefficients $k^0 = 20$ and $\beta^0 = 0.8$. Table 1 shows mean and standard deviation values of the estimates as a function of $N$. It can be seen that the results verify the mentioned asymptotic behavior, with a rate of convergence close to $\sqrt{N}$. However, note that to achieve a reasonable estimation error, the number of required spatial samples may be larger than the number available in radar realistic situations.

Table 1: Coefficient estimates of a GARCH(0, 1) process.

<table>
<thead>
<tr>
<th>$N$</th>
<th>Coef.</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>$k$</td>
<td>20.2676</td>
<td>3.2321</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>0.7874</td>
<td>0.1045</td>
</tr>
<tr>
<td>1000</td>
<td>$k$</td>
<td>20.2198</td>
<td>1.3927</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>0.7904</td>
<td>0.0485</td>
</tr>
<tr>
<td>2500</td>
<td>$k$</td>
<td>20.0780</td>
<td>0.9642</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>0.7959</td>
<td>0.0339</td>
</tr>
<tr>
<td>7000</td>
<td>$k$</td>
<td>20.0364</td>
<td>0.5613</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>0.7986</td>
<td>0.0198</td>
</tr>
<tr>
<td>10000</td>
<td>$k$</td>
<td>20.0121</td>
<td>0.4891</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>0.7995</td>
<td>0.0159</td>
</tr>
</tbody>
</table>

To compare the original estimation method with the alternative method proposed in Section 2.3, numeric simulations were performed to compute mean, standard deviation and RMS error of the estimates. This simulations consisted in the generation of $M$ independent realizations of length $N$ of a GARCH(0, 1) process and the estimation of its coefficients solving the optimization problem (8) with $\ell(\theta)$ given by (9). The values of the coefficients with which the process realizations were generated are $k^0 = 20$ and $\beta^0 = 0.8$. This procedure was repeated 500 times again. The results are summarized in Table 2. This table also includes the same parameters computed for the estimator determined by the original method with a realization of 10,000 samples. Note that the estimators obtained with this procedure show a behavior similar to those obtained by the original method when the total number of samples is the same. It can be seen that the estimators are unbiased and that the orders of the magnitude of the error are equal. This gives an idea of the number of independent realizations of fixed length needed to achieve a given estimation error.

3.2. Detection and false alarm probabilities

In order to evaluate the performance of the detector presented in Section 2.4, which we call GARCH detector, Monte Carlo simulations were performed to estimate the $P_D$, for different values of SCR, and the $P_{FA}$. They consisted of generating realizations of a GARCH(0, 1) process of $N$ samples, where each sample corresponds to a range cell. The coefficients were set equal to those of Section 3.1 with an unconditional variance equal to 100. In a range cell, chosen at random, we added a sample signal modeled as described in Section 2.4. The threshold obtained in (18) was evaluated, where we fixed $P_{FA}/\psi = 10^{-3}$. Note that for these process orders the threshold only depends on the previous sample of the range cell under test, i.e., $\psi = y_{t-1}$. Thus, we used $y_{t-1}$ instead of $\psi$ instead, regardless if the statistic exceeds the detection threshold in the previous range cell or not. To calculate the first threshold we used the standard deviation calculated from (6). Then, the samples were compared with the threshold and we estimated the $P_D$ and the $P_{FA}$ by their relative frequencies.

We set a number of independents process realizations large enough to ensure that the detection and false alarm probabilities are accurately estimated [19], each one of length $N = 200$ samples. We repeat the simulations to estimate the $P_D$ for different values of SCR, keeping the noise power constant and varying the value of the signal energy, $\varepsilon$. For comparison purposes, we also consider a detector resulting from assuming that the data have normal distribution, which we call Gaussian detector. Assuming a normal distribution with zero mean and variance $\sigma^2$ for the clutter and the signal model described in Section 2.4, we get for the Gaussian detector a decision rule equal to (16) but with a threshold given by

$$\lambda_g' = \sqrt{-\sigma^2 \ln(P_{FA})},$$  (19)

note that in this case the resulting threshold is constant (the derivation is analogous to the GARCH detector, except that pdfs of the data model are used instead of the conditional pdfs [18]).

The simulations to determine the $P_D$ and the $P_{FA}$ were repeated with the Gaussian detector. However, for this detector simulations were made in two different ways. In the first we used the Gaussian detector with clutter generated from a normal distribution and in the second this detector was used with clutter generated as a GARCH(0, 1) process. In both cases the variance was set equal to 100.

Table 2: Comparison between estimates.

<table>
<thead>
<tr>
<th>$M$</th>
<th>$N$</th>
<th>Coef.</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>RMS error</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>250</td>
<td>$k$</td>
<td>20.0211</td>
<td>0.4738</td>
<td>0.4738</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta_1$</td>
<td>0.7990</td>
<td>0.0153</td>
<td>0.0153</td>
</tr>
<tr>
<td>1</td>
<td>10,000</td>
<td>$k$</td>
<td>20.0121</td>
<td>0.4891</td>
<td>0.4887</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta_1$</td>
<td>0.7995</td>
<td>0.0159</td>
<td>0.0159</td>
</tr>
</tbody>
</table>
Fig. 1. Comparison of detectors performance: probability of detection.

Table 3
Comparison of detectors performance: probability of false alarm.

<table>
<thead>
<tr>
<th>Detector – data</th>
<th>P_{FA}</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH detector – GARCH data</td>
<td>1.10 × 10^{-3}</td>
</tr>
<tr>
<td>Gaussian detector – GARCH data</td>
<td>16.70 × 10^{-3}</td>
</tr>
<tr>
<td>Gaussian detector – Gaussian data</td>
<td>0.99 × 10^{-3}</td>
</tr>
</tbody>
</table>

Fig. 1 shows the estimated values of the PD obtained by means of numerical simulations for both detectors. We observe that the PD of the GARCH detector is greater than that of the Gaussian detector for low SCR. Table 3 shows the estimated values of the P_{FA}, where it can be seen that for GARCH clutter and Gaussian detector the P_{FA} is higher than expected. Note that since the Gaussian is a particular case of a GARCH process (GARCH(0,0)), the GARCH detector for Gaussian data will behave as a Gaussian detector. In fact, it can be seen that if the variance is constant in time, the GARCH detector threshold (18) is the same as the Gaussian detector threshold (19).

In a similar way, we construct the receiver operating characteristic (ROC) curve using GARCH data for both detectors. A pair (PD, P_{FA}) corresponds to a point on the curve. By changing the threshold, any point of the ROC is computed. This curves are shown in Fig. 2. We can see the performance degradation that suffers the Gaussian detector compared to the GARCH detector for low values of the P_{FA}.

3.3. Sensibility analysis

A desired characteristic for a radar detector is that the variance of any unknown parameter does not affect significantly the detection and the false alarm probabilities. Thus, in the presence of an estimation error or small changes in the parameters the performance of the detector should not vary. For the GARCH detector, a lower sensibility would allow us to estimate the coefficients with a smaller number of samples or to maintain the coefficients for a longer period of time before updating them.

We performed a simulation where we computed the PD and the P_{FA}, assuming the coefficients were corrupted due to estimation errors. Since we cannot obtain the true distribution of the coefficient estimators, we assume for them a normal distribution. The parameters of these distributions are those obtained in the asymptotic analysis performed in Section 3.1, in particular we use the values obtained for N = 7000 samples, i.e., k ~ N(20, (0.5613)^2) and \( \beta_1 \sim N(0.8, (0.0198)^2) \). For the PD the simulation was repeated 10 times, i.e., for ten different values of the coefficients. Fig. 3 shows the relative error of the PD, \( (P_D - \hat{P}_D)/P_D \), obtained for different values of SCR. We observe that it has low PD error for a wide range of SCR values. In the case of the P_{FA} the simulation was repeated 100 times. The mean and standard deviation of the P_{FA} estimates are 1.10 × 10^{-3} and 1.34 × 10^{-4} respectively. This shows that the detector is robust to small changes in the parameters.

4. Real sea data

4.1. Clutter modeling

Below we present the results of the fit of a GARCH process to real clutter data. By fit we mean the procedure to choose the process orders and estimate the coefficients of the process conditional variance using real sea clutter measurements. We use data of the McMaster University IPIX radar, collected at the Osborne Head Gunnery Range (OHGR), Dartmouth, Nova Scotia, Canada [21]. Specifically, we use the data recorded on November 11, 1993, at 4:11 p.m. that correspond to the data set stare0. The IPIX radar has polarimetric information, the shown results correspond to the vertical polarization (VV). Data corresponds to echoes from a small
target in inhomogeneous sea clutter. The height of the sea waves was of approximately 0.7 m. The target consists of a beach ball wrapped with aluminum foil. Its nominal location is 2691 m. The fast time or range dimension consists of 54 samples, the sampling interval is 15 m and the radar range resolution is 30 m. The number of transmitted pulses, i.e., the number of samples of the slow time dimension, is 2048 with a pulse repetition frequency of 200 Hz. In our case the process realizations are taken in range, i.e., we have a realization for each transmitted pulse. However, because we are only interested in fitting the GARCH model to the clutter data, we consider that the realizations extend up to cell 45, as to exclude the cells that contain the target. We use the first 1024 realizations for the estimation procedure. The rest of the data set will be used for validation purposes to evaluate the performance of the detection algorithm.

In principle there is no reason to believe that the realizations are independent, an hypotheses needed to perform the estimation by the alternative method proposed in Section 2.3. However, we can think that after a time the environment changes and this consideration is approximately valid. In order to have a quantitative idea, we estimate the data autocorrelation function by means of the average of their temporal autocorrelation functions. Fig. 4 shows the estimated autocorrelation function results for the data of the channel VV. As we can see after 30 ms, or equivalently after six pulses, the autocorrelation function is approximately zero; in other words samples from two pulses separated a time interval longer than 30 ms are not correlated. While this does not necessarily imply independence, it sets a lower bound to the time between pulses so that independence may exist. We assume that realizations separated by six pulses or more will produce estimation results similar to independent realizations.

The coefficients estimation was performed using the alternative method proposed in Section 2.3. We chose $M = 156$ realizations, separated each six pulses, which as mentioned consists of $N = 45$ samples. A priori we do not know the order of the process that we want to fit to the data, for this reason the estimate was made for different orders starting with $p = 0$ and $q = 1$. Initially, we evaluated the possibility of using some model selection criteria, such as BIC (Bayesian Information Criterion), to determine the order. However, this idea was rejected for two reasons: i) because the deduction of the BIC assumes that the gradient of the likelihood function is zero for $\theta = \hat{\theta}$ [22], which in this case is not necessarily true since $\theta$ is a constrained maximum of this function; ii) when we perform the fit for higher model orders the estimators corresponding to the coefficients $k$ and $\beta_1$ take the same values as in the case of the GARCH(0,1) process fit and the estimates of the new coefficients do not appear to be statistically significant. This behavior was common not only for this data set, but also when we performed the fit to other available measurements not included in this paper. The values of the estimates $k$ and $\beta_1$ resulting from the fit of the GARCH(0,1) process to the real data are $5.3675 \times 10^{-3}$ and 0.7463, respectively.

GARCH processes tend to be impulsive (volatile in the economic literature). A small value of $k$ indicates that the conditional variance of the noise floor will be small and the value of $\beta_1$ responsible for the frequency and magnitude with which the impulses occur, for larger values of $\beta_1$ impulses are more likely to occur. Since the process orders are $(0, 1)$ the conditional variance depends only on the last value of the innovations and most impulses will be of short duration.

Fig. 5 shows the magnitude histograms of the real data, of a GARCH(0,1) process, of a circularly symmetric Gaussian process and of a process whose envelope distribution is Weibull. The phase was uniformly distributed in the first three cases. The samples of the GARCH(0,1) process were generated with the value of the estimated coefficients, the samples of the gaussian process were generated from a normal distribution with zero mean and variance equal to the data sample variance and the parameters of the Weibull density function were estimated by the maximum likelihood estimation method [23]. We can see the similarity between the histograms of the actual data and the histograms of the GARCH process. To measure the goodness of fit, we evaluated the root mean-squared error (RMSE) defined as [24]

$$\text{RMSE} = \sqrt{\frac{1}{L} \sum_{j=1}^{L} \left| f(j) - h(j) \right|^2};$$

where $f(\cdot)$ is the generic pdf whose parameters are estimated from the real data, $h(\cdot)$ is the real data histogram and $j$ is the generic point of the amplitude in which both the pdf and the histogram are evaluated. The pdf that gives the smallest RMSE value is considered to be the best fit to the data. In our case the GARCH model gives the smallest RMSE for all the data sets considered. It is also important to note that the data generated by the normal distribution have a low probability of taking values above three times the standard deviation, while actual clutter measurements
are more likely to take these values, as we discussed in the previous paragraph. Thus, from the results obtained in Section 3.2 it is expected that the GARCH detector will present a better performance than the Gaussian detector when used with real data. The Weibull distribution shows a good fit on the tail, but with less flexibility at low magnitudes.

### 4.2. Results of the detection algorithm

To evaluate the performance of the detectors in a real application situation, we use the same data set stare0 of the McMaster University IPX radar, described in Section 4.1. We use the last 1024 realizations (the first 1024 were used to estimate the process coefficients). For a first test, to control the SCR, we consider that the realizations extend up to cell 45 to exclude the cells that contain the real target, and we add a synthetic target with the model proposed in Section 4.2.

Based on the results of the former section we consider a GARCH detector of orders (0, 1). From (18) the threshold is given by

\[
\lambda' = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^{\infty} e^{-\frac{1}{2}x^2} dx
\]

where \(\hat{\beta}_1\) and \(\hat{\beta}_2\) are the coefficient estimates presented in the previous subsection.

We ran the GARCH detector, the Gaussian detector and a third detector based in a CFAR algorithm for Weibull background [23]. For the latter detector the threshold is given by [23]

\[
\lambda_w' = b_\alpha^{1/\hat{\beta}}
\]

where \(b\) and \(\hat{\beta}\) are the maximum likelihood estimates of the parameters of the Weibull pdf and the coefficient \(\rho\) is a function of the desired probability of false alarm and of the number of samples used in the estimation procedure, see [23] for more details.

For the Gaussian (19) and Weibull (22) detectors, their parameters were estimated using the same data set used to estimate the GARCH process coefficients. The detection threshold is set for \(P_{fa} = 10^{-3}\) and the SCR is set to 5 dB. Fig. 6 shows detection maps for each of the tests. The black pixels correspond to a range cell where the detection statistic is larger than the threshold. The target location is indicated on the right side of the figures by the marker “•”.

We observe that the GARCH detector finds the true position of the target with only a few false detections despite the strong clutter echoes, which can be appreciated from Fig. 6(a). On the other hand, the Gaussian detector has lower target detection rate and higher false alarm rate, as it was anticipated. We mentioned that the higher false alarm rate is due to the clutter impulsivity not contemplated by the Gaussian model. These lead us to conclude that the GARCH model is a better model than the Gaussian for the clutter. The Weibull distribution models better the heavy tails than the Gaussian, but to account for the impulsivity associated to these heavy tails, the detection threshold takes a larger value than the Gaussian threshold. Thus the Weibull detector reduces the probability of false alarm at the expense of reducing the detection probability. Table 4 shows the empirical false alarm and detection rates for each detection algorithm. We have defined the empirical false alarm rate as the number of detected pixels, excluding the target pixels, over the total pixels of the detection map, again excluding the target pixels. Analogously, we have defined the empirical detection rate as the number of detected pixels in the range cell where the target is located over the total pixels of this cell.

We performed a second test also using the last 1024 realizations and all the range cells, including the real target. We ran the GARCH detector, the Gaussian detector and the Weibull detector, with the same parameters as those of the former test. Fig. 7 shows detection maps and Table 4 shows the empirical false alarm and detection rates for each of the tests. In this case we observe that GARCH and Gaussian detectors have a good target detection rate but the GARCH detector has a lower false alarm rate than the Gaussian detector. This is a consequence of the SCR being higher than SCR used in the previous test. One again, we observe that the Weibull detector presents a lower false alarm rate than the others detectors, at the expense of a detection rate reduction.

On the other hand, we can see that the target affects more than one cell and in the resulting detection map of the GARCH detector looks as if target masking occurs. This is due to the way in which we evaluate the threshold after the statistic exceeds it. However, it is not a problem because the target affects two or three range cells since the radar resolution is 30 m and the range sampling interval is 15 m. But it may be a drawback when two targets are near each other, in which case it may be convenient to use a different decision rule for selecting the threshold in the range cell after a presumed detection.

Note that the empirical false alarm rates deviate from the theoretical specification. Since the real clutter is not actually generated by any of the tested models this is expected. The Weibull distribution has a good fit on the tail of the distribution so it is the model with better agreement between the empirical and specified false alarm probabilities. However, as it also has the lowest detection rate.
probability it is not necessarily the best distribution for the detection problem. A fair comparison among the proposed model should not be made based on quality of fit of the probability distributions, but on a comparison of the overall performance on the solution of the detection problem. In order to achieve this, we compute the empirical ROC, i.e. the plot of $P_D$ versus $P_{FA}$ calculated from the data set [25].

The way in which the proposed GARCH detector works has two time scales; a slow adaptability given by the estimation of the model parameters which can be performed as a batch process, and the fast adaptability inherent to the model, in which the threshold for each detection is affected by the sample of the previous range cell. To make a fair comparison with other classical detectors we include two versions of the Gaussian and Weibull detectors: a slow adaptive and a fast adaptive version. The slow adaptive detectors estimate the Gaussian or Weibull parameters in the same way as the GARCH detector, i.e. using a large block of data of many radar returns, and use the resulting thresholds for detection while gathering the data for the next model update. On the other hand, the fast adaptive detectors estimate the Gaussian or Weibull distribution parameters from a single return with a set of secondary range cells around the cell under test, with a guard cell around it. These fast adaptive detector attempt to mimic the time varying threshold of the GARCH detector.

Fig. 8 shows the empirical ROC curves for actual clutter data and a synthetic target with different SCR. Fig. 9 shows the empirical ROC curve obtained for actual clutter data and target. The ROC curves of the Gaussian and Weibull slow adaptive detectors overlap. In fact, the same empirical curve would be obtained for a detector related to any time invariant distribution which involves a comparison with a threshold that is fixed (at least during a time interval). The theoretical ROC curves may differ, but the empirical curve can be obtain simply by varying the threshold, regardless of the theoretical false alarm probability associated to each threshold value. The Gaussian and Weibull fast adaptive thresholds change in each decision but their performance is degraded since the distribution parameters are estimated poorly by a small number of samples. For the Weibull detector two parameters must be estimated from the same small data set, so this effect is aggravated. From Figs. 8 and 9 we can again conclude that the GARCH model is a better model than the others presented in this work, if the detection performance is used for the comparison.

5. Discussion

Whitecaps of the sea waves often produce amplitude peaks. This sort of impulsive response is not well modeled by a Gaussian process. In this paper we proposed to apply a GARCH process for a better characterization of the sea statistics. Additionally, we developed a detection scheme based on this data model.

Since the proposed model considers the realizations in range, the number of samples required to obtain a reasonable estimation error is not enough. To solve this problem we presented an alternative estimation method for short data sequences, without losing performance. It is based on the use of several independent process
realizations. We showed that this procedure presents a behavior similar to that of the original estimation method.

We developed a detection algorithm based on a GARCH model of clutter. One of its advantages is that it is an adaptive detection scheme that depends on the previous clutter samples for each cell under test, thus it is able to pick up the fast time behavior of the noise. Also, note that if the clutter is Gaussian then the GARCH under test, thus it is able to pick up the fast time behavior of the clutter. One of its advantages is that it is an adaptive detection similar to that of the original estimation method.

We evaluated the detection and false alarm probabilities by means of Monte Carlo simulations. We showed that for low values of SCR our detector presents a better performance than the Gaussian one, i.e. a higher detection probability with a lower false alarm rate. In addition, our detector has the advantage of deciding pulsewise, meaning that it does not need to wait for the signal returns from multiple pulses to take a decision. However, we believe that if we extend these ideas using information of several pulses the performance will further improve, without need of standard stationarity assumptions. This may be accomplished for example using a 2D-GARCH model [26,27].

Regarding the results for actual clutter data, we showed that the GARCH detector performs better than Gaussian and Weibull detectors, except for high SCR and high false alarm probability, under this favorable conditions any detector achieves more or less the same detection probability. The success of the GARCH detector is explained by its time variability: the threshold for each comparison depends on the previous sample magnitude, but the parameters of the model do not change rapidly. This behavior cannot be obtained with classical detectors which involve a trade off between quality in the estimation of the distribution parameters and rate of adaptability.

Finally, for a real time application, the detection algorithm has a practical amount of computational load. Although we need the signal returns from multiple pulses for updating the coefficient estimates, we showed that the detector is robust against small errors in the parameters. Then, there is no need to estimate the parameters very often. Actually, the coefficients change when environments conditions change drastically. This flexibility allows us to wait the necessary signal returns to perform the estimation, while the detector is running. Therefore, comparing only the detection algorithms the GARCH detector has a slightly computational load than the Gaussian and Weibull slow adaptive detectors because their thresholds are constant. However, it has a lower computational load than the Gaussian and Weibull fast adaptive detectors because the computation of their adaptive thresholds requires an estimation of the distribution parameters for each range cell.

Acknowledgments

This work was supported by ANPCyT PICT 2011-11-0909, UNLP 11-I-166, and Cic-pBA.

References


Juan P. Pascual has received hisEng. (2006) degree in Electronics Engineering from the National University of La Plata (UNLP), Buenos Aires, Argentina. He is currently a Ph.D. candidate in the UNLP. His research interests are statistical and array signal processing with radar and communication applications.

Nicolas von Ellenrieder is a Professor at the Universidad Nacional de La Plata, where he received his Eng. (1998) and Ph.D. (2005) degrees. His postdoctoral experience includes research visits to the Washington University in St. Louis (2006), the Cuban Neuroscience Center (2006), and the Montreal Neurological Institute of McGill University (2010). His research interests include statistical and digital signal processing with applications in the fields of biomedicine and radar.

Martin Hurtado received the B.Eng. and M.Sc. degrees in electrical engineering from the National University of La Plata, Argentina, in 1996 and 2001, respectively. He received his Ph.D. degree in electrical engineering from Washington University in St. Louis in 2007. Currently, he is a Professor at the National University of La Plata. His research interests are in the area of statistical signal processing, detection and estimation theory, and their applications in sensor arrays, communications, and remote sens-
ing systems. In addition, he is interested in electromagnetic theory and antenna design.

Carlos H. Muravchik was born in Argentina in 1951. He graduated as an Electronics Engineer from the National University of La Plata (UNLP, 1973) and received the M.Sc. in Statistics (1983) and the M.Sc. (1980) and Ph.D. (1983) degrees in Electrical Engineering, from Stanford University, USA. He is a Professor at the Electrical Engineering Department (UNLP) and was Visiting Professor to Yale University, University of Illinois at Chicago and Washington University in St Louis. His research interests are statistical and array signal processing with biomedical, communications and control applications.