A Neural Network Approach to Weibull Distributed Sea Clutter Parameter’s Estimation

José Raúl Machado-Fernández¹, Jesús de la Concepción Bacallao-Vidal² and Nelson Chávez-Ferry³

¹²³ Instituto Superior Politécnico José Antonio Echeverría (ISPJAE), La Habana, Cuba.
¹ m4ch4ndo@hispavista.com, ² bacallao@electrica.cuajae.edu.cu, ³ nelson@electrica.cuajae.edu.cu

Abstract The main problem faced by sea radars is the elimination of an undesirable signal that appears mixed with target information: sea clutter. One of the most popular probability distributions in clutter modelling is the Weibull distribution. Helpful in efficient detectors’ design, a system able to recognize the Weibull shape parameter knowing a priori that the mean of the distribution is equal to one is proposed. The solution achieves a more precise estimation than the traditional method of moments, and it’s appropriate for real time operating conditions as it is based on a neural network approximation.

Keywords: Sea Clutter, Artificial Neural Networks, Weibull Distribution, CFAR.

Resumen El principal problema que enfrentan los radares marinos es la eliminación del clutter que es una señal indeseable que aparece mezclada con la información del blanco. Una de las distribuciones probabilísticas más populares en la modelación del clutter es la distribución Weibull. Beneficioso en el diseño de detectores eficientes, es propuesto un sistema capaz de reconocer el parámetro de forma de la distribución Weibull conociendo de antemano que la media de la distribución es igual a uno. La solución logra una estimación más precisa que el tradicional método de los momentos, y es aplicable en condiciones de operación en tiempo real pues se basa el uso de redes neuronales.

Palabras Claves: Clutter Marino, Redes Neuronales Artificiales, Distribución Weibull, CFAR

1 Introduction

The task of primary radars is to detect objects in the observation area and to estimate their position [1]. Target detection would be an easy task if the objects that produce echoes were located in a non-reflecting background. In that case, the echo signal could simply be compared with a fixed threshold, and targets would be detected when the received signal exceeds the threshold [2].

However, in real life radar applications, targets always appear embedded in a background filled with clutter, which is a random signal. Frequently, the clutter signal is subject to time and position variations. Therefore, the application of adaptive processing techniques becomes necessary to calculate constantly changing detection thresholds that correspond with the clutter’s local situation [3]. The techniques are even more needed in widely variable backgrounds such as sea clutter, which is the signal obtained from the radar’s echo reflected at sea surface [4].

In order to obtain the required local information, schemes with sliding windows around the analyzed cell are commonly used [5, 6]. According the application, the number of cells to be used in the window may vary, being the larger amounts responsible for a better estimate of the clutter’s average and the smaller more effective in
eliminating critical situations such as: the presence of multiple nearby targets and the occurrence of abrupt changes in the background’s level. When such situations occur, the clutter is said to be heterogeneous. Otherwise, it is categorized as homogeneous [7].

When detectors are designed for situations where targets appear inside sea clutter, the well-known Neyman-Pearson theorem is applied. This means that the designer first seeks to guarantee a given probability of false alarm (Pf) and then tries to maximize the probability of detection (Pd). Thus, the most popular clutter level estimation mechanisms are known as CFAR (Constant False Alarm Rate) because they ensure that the detection will occur under the guarantee of a constant false alarm [3].

Conceived at first under the assumption of Gaussian distributed clutter, several types of CFAR algorithms can be found, all based on the sliding window mechanism. The most popular are the CA-CFAR (Cell Averaging CFAR), the GO-CFAR (Greatest-Of CFAR), the SO-CFAR (Smallest-Of CFAR) and OS-CFAR (Ordered Statistics CFAR). These detectors have been treated in the literature by several authors [5, 8, 9] and they are often used as a reference on recent researches [10-13]. In addition, each year new alternatives and contributions appear in multiple international journals. Some proposals try to introduce new processing methods [13, 14], while others focus on improving the existing ones [15, 16]. However, all CFAR implementations have in common that they allow the adjustment of the false alarm probability by means of the modification of a scale or adjustment factor (K), which has an inverse relationship with the probability of detection [5, 8]. Figure 1 shows the structure of the classical CA-CFAR detector, where the Adjustment Factor (K) is used to calculate the target detection threshold.

The preliminary statement that certified that the clutter was Gaussian distributed was quickly proven as false in several papers [17-19]. Specifically in the case of sea clutter, numerous studies have shown that the family of heavy-tailed distributions is the best suited for representing the measurements made on the sea surface. While many others have been proposed, the following distributions are generally the most accepted by the community: Rayleigh, Log-Normal, K, Weibull and Log-Weibull [19-22].

2 Motivations and Objectives

Recent studies have reinforced the theory that the Weibull distribution is one of the best sea clutter models [23, 24]. Jointly, it has been noted that the average height of the waves influences the choice of the distribution’s shape parameter. Likewise, the shape parameter (β) varies when using S-band radars instead of X-band. In the case of S-band radars, the β shape parameter remains around 4.5 while for X-band the mean is approximately 2.5. In addition, the height of the waves, and other weather factors not yet fully specified, make β vary from 1.75 to 6.25. The formula for the Weibull Probability Density Function (PDF) is given below [25].

![Figure 1: Block diagram of a CA-CFAR detector.](image-url)
\[ f(x|\alpha, \beta) = \frac{\beta x^{\beta-1}}{\alpha^\beta} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right] \]  

(1)

The previous statements brought doubts on whether the selection of a fixed scale factor truly allowed maintain a constant \( P_f \) during the entire operation period of a CFAR detector. So, taking as a priori information that the \( \beta \) parameter varies in a known range [23], the ISPJAE (Instituto Superior Politécnico José Antonio Echeverría) radar research group proved, by performing several experiments in MATLAB software, that a detector employing a fixed scale factor must operate inefficiently to ensure a constant \( P_f \) [26]. On the contrary, if the adjustment factor would change according to the variation of the Weibull \( \beta \) parameter the inefficiency will disappear.

Figure 2 illustrates the dependence of \( K \) on \( P_f \) and clutter states [26] by showing calculations performed at the limits of the range of possible \( \beta \) values [23, 24] using a CA-CFAR detector. As it is visible, the behavior is altered for different clutter states proving that \( K \) is not only defined by the selected \( P_f \) but also by \( \beta \). Indeed, differences between the plots for \( \beta = 1.75 \) and for \( \beta = 6.25 \) are remarkable, displaying changes superior to 2.5 units for \( K \)'s magnitude.

![Figure 2: Comparison of the behavior of the false alarm probability for different clutter conditions [26].](image)

Results illustrated at figure 2 emphasize the need for a system capable of identifying the Weibull \( \beta \) parameter through the analysis of radar readings. That is the goal that the author is pursuing in this paper.

Neural networks are the selected classifier for performing the task. The choice is logical if the necessity of achieving real time efficient operation is taken into account. Neural solutions are a very popular choice when it comes to digital processing and CFAR detectors' design [27-29], and they have been used particulary in investigations related Weibull radar clutter [30-32]. Moreover, previous efforts in sea clutter parameter identification have used moments of the distribution as distinctive features [33]. However, this paper follows a different approach which will prove to be successful: the employ of PDF plots as distinctive features. The effect of the variations of Weibull parameters on the shape of the distribution ideal histograms is shown in figure 3.
Once the intended design is complete, the estimator would be able to integrate into a CA-CFAR detection scheme. Then, a second block will be necessary in order to translate the Weibull $\beta$ into a viable correction of the K scale factor. Note that the solution may be adaptable to any system employing a K factor, which is very popular in CFAR detectors [34-36]. Moreover, a neural based system that computed the relation between K and Pf, has already been proposed [37], but unfortunately it was based in a theoretical assumption that did not took into account the clutter’s influence, which means that it can’t be use as the second block.

3 Design and Training the Neural Network

When designing a neural network, many parameters affect the outcome, most of them under non-fully predictable laws. According to this, for each new proposed configuration, results may improve or get worse. Therefore, the ideal design would try all possible configurations to check which is the best suited. Unfortunately, this approach consumes an enormous amount time, so common design choices are usually based on previous experience acquired during investigations of similar nature. In this case, references can be obtained from recently published articles which indicate that the preferred neural network will use a Multilayer Perceptron structure [38, 39], and will be trained under a back-propagation rule [37-39]. The rest of the network’s parameters are described in table 1, taken from successful investigations related to the sea modeling under K [29, 33] and Weibull [39] distributions.

Table 1: Selected variables for the network’s design.

<table>
<thead>
<tr>
<th>Design Variables</th>
<th>Choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network Type</td>
<td>Feed Forward Network (Multilayer Perceptron)</td>
</tr>
<tr>
<td>Training Function</td>
<td>BackPropagation (Levenberg-Marquardt)</td>
</tr>
<tr>
<td>Number of Layers</td>
<td>3 (Input Layer – Hidden Layer – Output Layer)</td>
</tr>
<tr>
<td>Transference or Activation Function</td>
<td>Hyperbolic Tangent Sigmoid (Hidden Layer), Lineal Transference Function (Output Layer)</td>
</tr>
<tr>
<td>Activation Order</td>
<td>Topological (Asynchronous Activation)</td>
</tr>
<tr>
<td>Error Measurement</td>
<td>mean square error</td>
</tr>
<tr>
<td>Training Set Division</td>
<td>70% Training - 15% Validation - 15% Test</td>
</tr>
<tr>
<td>Samples’ Presentation</td>
<td>Batch Training</td>
</tr>
</tbody>
</table>

For the full understanding of the meaning of the parameters shown in table 1, the reader is referred to specialized literature [40]. However, there is a variable that is worth of a comment in favor of clarity.
When neural networks are designed for pattern recognition applications, transference functions for both the hidden and output layers are selected as Hyperbolic Tangent Sigmoid to concentrate results in the range between 0 and 1. Following this tendency, some networks have already been designed by the author and applied to sea clutter [41]. However, the use of Hyperbolic Tangent didn’t bring good results in this investigation. Therefore, the Linear Transfer Function was selected for the output layer, improving the performance. As a consequence, there are no limits in the possible values at the network’s output and the nature of the solution changes from pattern recognition to value estimation.

3.1 Preparation of the Training Set

When training neural networks, the designer must select whether to use supervised or unsupervised training technique [42]. The supervised alternative is selected by the author because is the common choice for radar associated issues, where solutions usually search to relate a signal to a given class. This technique requires the use of two training sets: one for the inputs and another for the corresponding outputs.

In the current investigation, the employed training set consisted of 20 000 Weibull clutter input groups, where each group contained 50 values. Every group of 50 values resulted from the plot of Weibull clutter histograms composed by 3000 samples generated using the inversion method [43].

The 20 000 Weibull clutter input groups were created modifying the Weibull $\beta$ after every 20 groups. So, 1000 $\beta$ intervals were covered from 1.75 to 6.25. Note that even if there are 20 input groups for the same $\beta$, it doesn’t mean that they are all equal. They would be equal if an infinite amount of samples were used in the confection of the histograms, but in this case only 3000 samples were employed. Figure 4 shows several generated histograms.

![Four Weibull histograms generated with the same $\beta$.](image)

The $\alpha$ scale parameter of the Weibull distribution was maintained consistent to the variation of parameter $\beta$ so the average value of the clutter’s level will remain close to one. This is a viable process because it was found in preliminary surveys that $\alpha$ has no repercussion on the CA-CFAR K factor. The formula for the mean or average of the Weibull distribution is given below [25].

$$mean = a\Gamma \left(1 + \frac{1}{\beta}\right)$$  \hspace{1cm} (2)

3.2 Trial and Error Training

Even if the variable for error measurement is configured to use the mean square error, a different approach was used to select the best network at the end of the training. While a small error is acceptable because it only slightly modifies the response, a large error is definitely unacceptable in radar applications. Thus, neural networks making
a bit more errors are preferred as long as the error’s magnitude remains small, while others making very few but of great magnitude are undesirable. So, the quality of the network in this research will be defined by the confidence interval instead of the mean of the error’s magnitude.

Table 2: Estimated relation between network’s size and confidence interval.

<table>
<thead>
<tr>
<th>Amount of neurons</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence Interval</td>
<td>0.1480</td>
<td>0.1519</td>
<td>0.1522</td>
<td>0.1465</td>
<td>0.1510</td>
<td>0.1519</td>
<td>0.1519</td>
<td>0.1494</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amount of neurons</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence Interval</td>
<td>0.1491</td>
<td>0.1472</td>
<td>0.1453</td>
<td>0.1504</td>
<td>0.1394</td>
<td>0.1418</td>
<td>0.1454</td>
<td>0.1473</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amount of neurons</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>45</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence Interval</td>
<td>0.1401</td>
<td>0.1432</td>
<td>0.1441</td>
<td>0.1398</td>
<td>0.1436</td>
<td>0.1401</td>
</tr>
</tbody>
</table>

On another note, even though most of the design variables of the network are assumed beforehand, the size of the hidden layer is left free to be estimated by trial and error. Table 2 shows several hidden layer sizes that were considered during the search of the ideal value.

As it can be seen, good results were obtained for sizes of 18, 25, 40 and 50 but the smaller and more promising size was 17 neurons. Knowing this, 500 additional trainings were conducted with networks of 17 neurons in the hidden layer, yielding an improved value of 0.1378 for the confidence interval. The final structure of the network is offered in figure 5.

![Figure 5: Structure of the proposed network.](image)

4 Results and Discussion

The proposed network, with 17 neurons in the hidden layer, has 50 entries where 50 values will be placed. Each group of 50 inputs correspond to the plot of histograms put together with 3000 Weibull samples. As a response, the network will show at its output an estimate of the $\beta$ shape parameter with a maximum error of 0.1378. The achieved effect reduces the initial confidence interval of 4.50 (6.25-1.75) in more than 32 (4.50/0.1378) times, achieving a result whose variation corresponds only to a 3.1% of the initial interval. Figure 6 illustrates the reduction.

![Figure 6: Effectiveness of the proposed neural network solution.](image)
Figure 7: Response of the proposed network.

Figure 7 provides a comparison between the ideal response (red line) and the real response provided by the trained neural network. The reader may notice that the estimation is more accurate for the first input groups which were generated using smaller $\beta$'s. The situation is most evident in figure 8 where the committed error is plotted: equivalent to the subtraction of the ideal and the real behavior.

Figure 8: Errors made by the neural network.

It is evident in figure 8 that the network’s performance is much more favorable for the first half of the graphic, which includes 10 000 input groups in the $\beta$ range from 1.75 to 4. This behavior may be the result of two possible factors: specific characteristics of the Weibull distribution or influence of training choices. The precise determination of these factors is left for future researches. However, it is important to specify that if the behavior were a sequel of the training, a solution could be to replace the used batch mechanism.

Figure 9 shows a histogram of the error made by the network. In what constitutes a positive feature, the pattern is found to be close to Gaussian indicating that the mistakes committed are not only small but are also equally distributed between positives and negatives excess. The exact quantities for the limits of errors are 0.1358 for the negative mistakes and 0.1405 for the positive ones.
One important characteristic of the network’s performance that becomes evident after analyzing figure 9 is the low frequency of high magnitude errors. This leads to the idea of operating the network under its maximum peak of efficiency, reducing thus the confidence interval. Following this reasoning, figure 10 was calculated in order to offer a relation between the reductions of the confidence interval at networks’ effectiveness.

![Figure 9: Histogram of the network’s mistakes.](image)

Note that for the previously given confidence interval of 0.1378 the effectiveness was of 100% meaning that the estimator never made mistakes greater than 0.1378. In figure 10, it is visible that the interval can be reduced down to 0.09 maintaining the percentage on 99. In fact, an even greater reduction can be considered if it is perceived that the percentage is over 90 even when the interval is at 0.05 (a reduction of more than half of the initial interval).

![Figure 10: Relationship between confidence interval and the effectiveness percentage.](image)
4.1 Comparing the Neural Solution with the Method of Moments

A comparison between the well-known Method of Moments (MoM) and the proposed neural solution was performed by the authors. Regarding processing time, the MoM is slower since it requires a binary search to obtain the Weibull parameters. Expression (3) shows the equation used in [44] for estimating the $\beta$ parameter with the MoM. The mean ($E[Z]$) and the variance ($E[Z^2]$) are calculated by processing the samples.

$$\frac{E[Z^2]}{E[Z]} = \frac{\Gamma\left(\frac{2}{\beta}+1\right)}{\Gamma^2\left(\frac{1}{\beta}+1\right)}$$

Note that in (3) is not possible to obtain $\beta$ by isolating the variable because there is no way to take $\beta$ out of the gamma ($\Gamma$) function. To solve (3), a binary search must be executed by placing $\beta$ values to the right of the equal and comparing the result with $E[Z^2]/E^2[Z]$.

So, the MoM involves estimating the mean and the variance of the sample set and executing a search. The neural solution is faster because it only requires rearranging the samples into histograms. The rest of the process is really fast since the proposed neural network has only three layers where the neurons may be executed in parallel by placing the scheme into a proper device such as an FPGA (Field Programmable Gate Array) kit.

Besides the speed comparison, which does not indicates clearly the supremacy of the neural solution, a much more revealing evaluation was performed regarding the precision of the output. In the previous section, figure 9 provided a histogram of the error committed by the neural alternative, establishing the confidence interval at 0.1378 which signified and admissible variation of 3.1% of the initial interval. A new histogram is shown in figure 11, this time for the error committed when applying the MoM to 740 input groups. The result was obtained solving equation (3) by successive iterations until the difference between expressions at both sides of the equal was less than 0.01.

![Figure 11: Histogram of the MoM’s mistakes.](image)

It can be seen in figure 11 that the MoM has a longer interval of error than the proposed neural solution. The exact quantities for the limits of errors are 1.2750 for the negative mistakes and 1.3 for the positive ones, which results in a confidence interval of 1.2875 (28% of the initial interval). It’s also visible in figure 11 that most of the mistakes are inferior to |0.5|, but this bring no advantage when compared to the neural solution that also displays a Gaussian-like histogram of errors.

5 Conclusions and Future Research

A neural network able to recognize the $\beta$ shape parameter of the distribution was designed in MATLAB, under the assumption of Weibull distributed clutter. After conducting experiments on various hidden layer sizes, it was
concluded that an artificial neural network with 17 neurons was ideal for solving the problem. The proposed solution manages to work perfectly for a confidence interval of 0.1378, while it operates with an effectiveness of 99.9% for a range of 0.1150, with 99% for 0.09, and with 90% for a 0.05 interval. The constructed scheme exceeds by far the results achieved with the traditional method of moments. The obtained estimator has two main applications: (1) improvement of target discrimination by identifying more accurately clutter states, and (2) recognition of clutter anomalies related to natural disasters such as oil spills.

As a future research line, the author recommends the creation of a second system that transforms $\beta$ into the right choice of the CA-CFAR $K$ adjustment factor, in order to adapt the operation of the popular detector to the variation of clutter states. It is also recommended to develop similar systems based on neural networks under assumptions of other types of clutter, such as: Log-Normal, Log-Weibull and $K$ distributions.

6 References


