



Modelación de la Distribución K en MATLAB para Aplicaciones de Radar

José Raúl Machado Fernández, Jesús de la Concepción Bacallao Vidal

RESUMEN / ABSTRACT

El clutter marino es una señal interferente que aparece en los sistemas de radar que operan en escenarios oceánicos o costeros. Los autores modelaron en MATLAB la distribución K, ampliamente reconocida como el mejor modelo para clutter marino ya que tiene en cuenta los dos mecanismos de fluctuación presentes en la señal de fondo. La implementación permite un acceso fácil a la manipulación de las funciones de distribución de probabilidad, funciones generadoras, momentos, algoritmos de bondad de ajuste y estimación de parámetros de la distribución. Expresiones matemáticas y notaciones de funciones informáticas fueron definidas para contribuir a la unificación de los estudios de clutter de radar relacionados al modelo K. La solución es un punto de partida para el desarrollo de nuevos esquemas de detección por parte del grupo de investigación de radares de la Facultad de Eléctrica de la CUJAE y contribuye a la conformación de la biblioteca MATE-CFAR 2.

Palabras claves: Distribución K de la Amplitud, Distribución K de la Potencia, Función de Densidad de Probabilidad, Clutter de Radar.

The sea clutter is an interfering signal that appears in radar systems operating in oceanic or offshore scenarios. The author simulated in MATLAB the K distribution, widely recognized as the best model for sea clutter because it takes into account the two fluctuation mechanisms coexisting in the background signal. The implementation allows an easy access to the handling of probability distribution functions, generating functions, moments, goodness of fit algorithms and parameter estimation techniques. Mathematical expressions and notations for computer functions were defined as a contribution to the unification of radar clutter studies related to the K distribution. The solution is a start point for the development of new detection schemes by the Radar Research Team from the CUJAE's Electrical Engineering Faculty and contributes to the realization of the MATE-CFAR 2 library.

Key words: K distribution of Amplitude, K distribution of Power, Probability Density Function, Radar Clutter.

Modeling the K Distribution in MATLAB for Radar Applications

1. -INTRODUCCIÓN

Los radares son dispositivos dedicados comúnmente a la detección y rastreo de objetivos próximos en ambientes terrestres, marítimos y aéreos [1]. En un entorno ideal, los radares obtienen solamente el eco de la señal emitida cuando ella encuentra un blanco. En cambio, en escenarios reales el carácter reflector de la superficie que rodea al blanco produce una señal adicional ruidosa conocida como clutter [2].

El clutter es una señal aleatoria cuyo aporte no puede ser deducido por mecanismos puramente determinísticos. Consecuentemente, su modelación cae en el campo de las probabilidades y estadística de los procesos aleatorios. Múltiples han sido las distribuciones de probabilidad sugeridas en diversas publicaciones para la modelación de clutter, entre ellas las distribuciones Weibull [3] y Log-Normal [4] están entre las más populares. Desarrollos basados en estas distribuciones han sido concebidos por el Grupo de Investigación de Radar de la Facultad de Eléctrica de la Universidad Tecnológica de La Habana José Antonio Echeverría (CUJAE) [5-7].

Actualmente, el grupo dirige sus esfuerzos al trabajo con muestras de clutter subordinadas a la distribución K, ampliamente reconocida como el modelo preferencial para clutter marino [2]. Esto se debe a que representa el comportamiento de los ecos de la superficie marina como el resultado de la unión de dos contribuciones. La primera de ellas, conocida como el *speckle* (capilaridad), caracteriza el aporte de variación rápida; mientras que la segunda, bajo el nombre de textura, representa el aporte de variación lenta con un tiempo de decorrelación del orden de los segundos.

Dado su carácter compuesto, la distribución K tiene una definición matemáticamente más complicada que sus contrapartes Weibull y Log-Normal. Mientras que las últimas se ajustan a datos de clutter marino por mecanismos meramente empíricos, la distribución K proporciona un modelo mucho más fiel al comportamiento de la superficie del mar. Precisamente, el oleaje se considera compuesto por dos tipos de olas: capilares y pesadas [2, 8]. Las primeras cambian muy rápidamente y las segundas tienen una mayor inercia.

La definición matemática de la distribución K precisa de un total de cuatro distribuciones probabilísticas. Cuando se utilizan detectores de ley cuadrada (dominio de la potencia), como en [9-11], se emplea la distribución Gamma para el componente de textura y la Exponencial para el *speckle*. En el caso de utilizar detectores lineales (dominio de la amplitud), como en [12, 13], la distribución root-Gamma es utilizada para la textura y la Rayleigh para la capilaridad. Nótese que las distribuciones root-Gamma y Rayleigh resultan de hallar la raíz cuadrada de variables distribuidas Gamma y Exponencial respectivamente.

La distribución K no tiene ninguna función propia asociada en el *Statistics Toolbox* de MATLAB, al contrario de lo que ocurre para alternativas más ampliamente difundidas como la Weibull o la Log-Normal. Por ello, su modelación en esta herramienta informática adquiere especial importancia.

Entonces, los autores del presente artículo se trazaron como objetivo la modelación de la distribución K en MATLAB. Específicamente, se concentraron en reproducir el PDF (Probability Density Function, Función de Densidad de Probabilidad), CDF (Cumulative Density Function, Función de Densidad Acumulada), Momentos Centrados y Algebraicos, Función Generadora de Variables Aleatorias y Método para la Obtención de los Parámetros K, tanto en el dominio de la amplitud como en el de la potencia.

El artículo se desarrolla como sigue. A continuación, en la sección “Materiales y Métodos”, se describen los fundamentos matemáticos de la modelación de la distribución K tanto en su versión para el dominio de la amplitud, como para el dominio de la potencia. Más adelante, también en “Materiales y Métodos”, se presentan las funciones informáticas que fueron implementadas en MATLAB para la conformación de una mini-biblioteca de modelación K. La sección consecutiva, denominada “Resultados”, ofrece prueba de la validez de cada una de las funciones informáticas, mostrando gráficos que revelan detalles de la implementación. Luego en “Discusión” se valora el aporte del presente artículo en el marco de los desarrollos actuales de radar. Por último, en “Conclusiones” se indican los logros fundamentales de la labor realizada y en “Recomendaciones” los autores describen algunas líneas futuras de investigación.

2. –MATERIALES Y MÉTODOS

La presente sección está dedicada a la caracterización de la distribución probabilística K tanto en su versión de la amplitud como en la de la potencia. Además, se presentan las funciones matemáticas que fueron implementadas, ofreciendo una breve descripción del objetivo perseguido por cada una. La información presentada está concebida para facilitar la reproducción del estudio por terceros.

2.1.- DISTRIBUCIÓN K EN EL DOMINIO DE LA AMPLITUD

El clutter de alta resolución fue representado por la distribución K por primera vez por K.D. Ward en [14]. En este modelo compuesto, el clutter marino consiste de dos componentes que caracterizan la amplitud de la envolvente de los retornos del clutter marino.

El primer componente representa la variación del nivel promedio, frecuentemente llamado proceso modulador o textura, que puede ser relacionado con el perfil de la superficie del mar. El nivel medio se asocia frecuentemente con las olas grandes del mar y la estructura del oleaje [15]. Este componente tiene un tiempo de correlación elevado (varía solo lentamente con el tiempo) y no está afectado por el empleo de esquemas con agilidad en frecuencia [15].

El segundo componente es llamado *speckle* y ocurre debido a la naturaleza múltiple de las olas capilares en cualquier celda de distancia. Se decorrelaciona en unos pocos milisegundos debido al movimiento relativo de estas olas o por el uso de agilidad en frecuencia [15, 16]. El período de decorrelación está entre 5 y 10 ms [16].

Por tanto, la envolvente compleja de la amplitud del clutter conforme al modelo de la distribución K se obtiene por la multiplicación de estos componentes según la siguiente ecuación [2]:

$$\tilde{X} = Y \times \tilde{S} \quad (1)$$

En la expresión anterior, \tilde{X} es la envolvente compleja del clutter marino, \tilde{S} la envolvente compleja del componente de capilaridad que se descompone en variables Gaussianas en fase y en cuadratura con media cero y varianza (σ^2):

$$\tilde{S} = S_1 + S_Q, \quad S_1 \sim N(0, \sigma^2) \text{ y } S_Q \sim N(0, \sigma^2) \quad (2)$$

La media local (Y) tiene una función de densidad $f_Y(y)$ modelada por la distribución Root-Gamma (también conocida como generalizada Chi) [17]:

$$f_Y(y) = \frac{2b^{2v}}{\Gamma(v)} y^{2v-1} \exp(-b^2 y^2) \quad (3)$$

Donde b y v son los parámetros de escala y forma respectivamente de la distribución Root-Gamma, y el término $\Gamma(\cdot)$ es la función Gamma.

La amplitud de la envolvente compleja de los retornos del clutter marino está dada por:

$$|\tilde{X}| = Y \times |\tilde{S}| \quad (4)$$

La amplitud de la componente de capilaridad \tilde{S} posee distribución de amplitud Rayleigh con parámetro α . La función densidad de probabilidad de $|\tilde{S}|$ está dada por [17]:

$$f_s(s) = \frac{s}{\alpha^2} \exp\left(-\frac{s^2}{2\alpha^2}\right) \quad (5)$$

La variable s puede ser escrita como:

$$s = \frac{x}{y} \quad (6)$$

Si las variables se arreglan de acuerdo a la expresión anterior, la Función de Densidad de Probabilidad Condicional $f_{X|Y}(x|y)$ será [17]:

$$f_{X|Y}(x|y) = \frac{x}{y^2 \alpha^2} \exp\left(-\frac{x^2}{2y^2 \alpha^2}\right) \quad (7)$$

Luego de varias transformaciones, detalladas en [17], se obtiene la PDF K:

$$f_x(x) = \frac{4c}{\Gamma(v)} (cx)^v K_{v-1}(2cx) \text{ donde } c = \sqrt{\frac{\pi}{4}} b \quad (8)$$

Donde b y c son los parámetros de escala y v es el parámetro de forma que depende de las condiciones del mar y de los parámetros del radar. Además, $K(\cdot)$ es la función modificada de Bessel de segunda especie y de orden $v - 1$.

Los momentos algebraicos de orden n de y son descritos según [17] como:

$$E\langle y^n \rangle = \frac{1}{b^n} \frac{\Gamma(v+n/2)}{\Gamma(v)} \quad (9)$$

Mientras que los momentos algebraicos de orden n de la distribución K están dados por [18]:

$$E\langle x^n \rangle = \frac{1}{c^n} \frac{\Gamma(v+\frac{n}{2})}{\Gamma(v)} \Gamma\left(\frac{n}{2} + 1\right) \quad (10)$$

La media del nivel del clutter $E\langle x \rangle$ y la media del cuadrado del clutter $E\langle x^2 \rangle$ calculadas a partir de (10) son:

$$E\langle x \rangle = \frac{\sqrt{\pi} \Gamma(v+\frac{1}{2})}{c \Gamma(v)} \quad E\langle x^2 \rangle = \frac{v}{c^2} \quad (11)$$

Las ecuaciones de transformación a los momentos centrales pueden tomarse de [19], y la forma de calcular los parámetros de asimetría (*skewness*) y curtosis (*kurtosis*) de [20].

La Función de Distribución Acumulativa K de la amplitud es:

$$F_X(x) = \int_{-\infty}^x \frac{4c}{\Gamma(v)} (ct)^v K_{v-1}(2ct) \partial t = 1 - \frac{2c^v}{\Gamma(v)} x^v K_v(2cx) \quad (12)$$

La relación de (11) permite fijar la media de las muestras generadas a 1 si se hace:

$$c = \sqrt{v} \quad (13)$$

Por último, en [21] se afirma que, luego de varias comparaciones, la mejor expresión para la estimación de los parámetros se basa en el segundo y cuarto momentos. Dicha expresión, luego de hacer un pequeño ajuste producto de diferencias encontradas en las notaciones aplicadas, queda como:

$$\tilde{v} = \frac{1}{\left(\frac{m_4}{2m_2^2} - 1\right)} \tilde{c} = \sqrt{\frac{v}{m_2}} \quad (14)$$

Donde m_4 es el momento algebraico de cuarto orden y m_2 es el momento algebraico de segundo orden.

Típicamente, el valor de v para clutter marino puede oscilar en el rango entre 0,1 e ∞ . Cuando tiende a infinito, la distribución K se reduce a la distribución Rayleigh, mientras que los valores pequeños de v ($v < 1$), representan un clutter de cola muy alargada en el que ocurren valores de magnitud elevada con frecuencia (este tipo de clutter se denomina en inglés como *spiky*)[15].

2.2.- DISTRIBUCIÓN K EN EL DOMINIO DE LA POTENCIA

En algunas aplicaciones de radar, el detector opera en la zona cuadrática, razón por la cual se requiere obtener la distribución estadística de la potencia de la envolvente compleja. Nuevamente, la potencia de la distribución K puede darse como la multiplicación de dos componentes, el cuadrado de la media local y el speckle, como se muestra a continuación [17]:

$$|\tilde{X}|^2 = Y^2 \times |\tilde{S}|^2 \quad W = Z \times R \quad (15)$$

Donde $W = |\tilde{X}|^2$, $Z = Y^2$ y $R = |\tilde{S}|^2$. La media local, Y , tiene una PDF $f_Y(y)$ dada de acuerdo a (3). Por su parte, la PDF de Z está distribuida gamma, según fue desarrollada en :

$$f_Z(z) = \frac{b^{2v}}{\Gamma(v)} z^{v-1} \exp(-b^2 z) \quad (16)$$

La amplitud del componente de capilaridad \tilde{S} posee una distribución Rayleigh y su versión al cuadrado (R) tiene una PDF exponencial:

$$f_R(r) = \frac{1}{2\sigma^2} \exp\left(-\frac{r}{2\sigma^2}\right) \quad (17)$$

De forma similar, r puede escribirse como:

$$r = \frac{w}{z} \quad (18)$$

Si las variables son cambiadas de acuerdo a (18), entonces la PDF condicional $f(x | y)$ queda como:

$$f(w|z) = \frac{1}{2\sigma^2 z} \exp\left(-\frac{w}{2\sigma^2 z}\right) \quad (19)$$

Luego de varias transformaciones abordadas en [17], la PDF de la distribución K de la potencia del clutter queda:

$$f_W(w) = \frac{2c^{v+1}}{\Gamma(v)} w^{\frac{v-1}{2}} K_{v-1}(2c\sqrt{w}) \quad \text{donde } c = \sqrt{\frac{\pi}{2}} \quad (20)$$

Donde b y c son los parámetros de escala y v es el de forma.

Los momentos algebraicos de orden n de la distribución K de la potencia son:

$$E\langle w^n \rangle = \frac{1}{c^{2n}} \frac{\Gamma(v+n)}{\Gamma(v)} \Gamma(n+1) \quad (21)$$

Entonces, el promedio de la potencia del clutter y la media del cuadrado de la potencia del clutter son:

$$E\langle w \rangle = \frac{v}{c^2} \quad E\langle w^2 \rangle = \frac{4}{c^4} (v+1) \quad (22)$$

Conjuntamente, para generar muestras de media igual a uno se puede hacer:

$$c = \sqrt{v} \quad (23)$$

La función de distribución acumulada K de la potencia se puede obtener luego de adaptar la notación utilizada en [11]:

$$F_W(w) = 1 - \frac{2}{\Gamma(v)} (c^2 w)^{\frac{v}{2}} K_v(2c\sqrt{w}) \quad (24)$$

Por último, la expresión para la estimación del parámetro de forma en el dominio de la potencia de la distribución K se puede tomar de [22]:

$$\frac{m_4}{m_2^2} = \frac{6(v+2)(v+3)}{v(v+1)} \quad (25)$$

2.3.- FUNCIONES IMPLEMENTADAS

La tabla 1 muestra las funciones que fueron implementadas en MATLAB para la modelación de la distribución K. Cada una de ellas es comentada un poco más adelante. Nótese que las funciones anteceditas por *k_* corresponden al modelo K de la amplitud y las precedidas por *k2_* corresponden al de la potencia. En lo adelante, cuando se haga referencia a una función sin especificar el atributo *k_* o *k2_* debe asumirse que se indican ambas alternativas (por ejemplo *_chi_squared* hace referencia tanto a *k_chi_squared* como a *k2_chi_squared*).

Las funciones *_pdf* y *_cdf* permiten graficar la PDF y la CDF K de amplitud y de la potencia para cualquier valor de los parámetros aplicando (8), (12), (20) y (24). La función *_gen* genera muestras usando (4) y (15) a partir de los mecanismos de generación de variables aleatorias Rayleigh, Exponencial, Gamma y Raíz-Gamma. La función *_gen_plot* utiliza *_gen* para generar muestras y luego grafica el resultado en una serie de tiempo. Algo parecido hace *_gen_hist* que organiza las muestras en un histograma, ofreciéndose así otra forma de visualización. Conjuntamente, *_gen_compare* genera muestras y grafica una comparación del histograma y la función PDF teórica. Esta última función es quizás la más ilustrativa de las tres en cuanto a la correspondencia de los datos al modelo.

El código colocado en *_residual* permite calcular los residuos del PDF obtenido a partir de muestras generadas con respecto al PDF ideal. También se grafica este residuo, en lo que constituye un medidor básico de la desviación por el uso de un conjunto finito de muestras. Una alternativa a esta medición es *_chi_squared* que realiza una prueba de ajuste muy conocida y que es detallada en [23].

La función *_ideal_moments* calcula los momentos teóricos centrales de los cuatro primeros órdenes mostrados en (10) y (21), a la vez que ofrece estimaciones de la asimetría y la curtosis. Por su parte, *_real_moments* estima los momentos centrales a partir de un conjunto finito de muestras. La comparación entre los resultados de estas funciones revela la distorsión en los momentos centrales por el uso de un conjunto finito de muestras. Similar a *_ideal_moments* y a *_real_moments* se contruyeron las variantes *_alg_ideal_moments* y *_alg_real_moments* que operan de manera igual forma pero para los momentos algebraicos en lugar de los centrales.

Tabla 1
Funciones de modelación K y K2 implementadas en MATLAB.

Funciones	Objetivo
<i>k_pdf k2_pdf</i>	Mostrar PDF
<i>k_cdf k2_cdf</i>	Mostrar CDF
<i>k_gen k2_gen</i>	Generar muestras
<i>k_gen_plot k2_gen_plot</i>	Generar y graficar muestras
<i>k_gen_hist k2_gen_hist</i>	Generar muestras y organizarlas en un histograma
<i>k_gen_compare k2_gen_compare</i>	Comparar el histograma con la curva PDF teórica
<i>k_residual k2_residual</i>	Calcular el residuo de muestras a partir de PDF teórica
<i>k_chi_squared k2_chi_squared</i>	Realizar la prueba Chi-Squared de bondad de ajuste
<i>k_ideal_moments k2_ideal_moments</i>	Calcular los momentos centrales ideales
<i>k_real_moments k2_real_moments</i>	Calcular los momentos centrales reales
<i>k_alg_ideal_moments k2_alg_ideal_moments</i>	Calcula momentos algebraicos ideales
<i>k_alg_real_moments k2_alg_real_moments</i>	Calcula momentos algebraicos reales
<i>k_estim_par k2_estim_par</i>	Estimar los parámetros a partir de grupos de muestras
<i>k_gen_sets k2_gen_sets</i>	Genera conjuntos de muestras

La función *_estim_par*, implementa las expresiones (14) y (25) para estimar el parámetro v de las distribuciones K y K2 respectivamente. Esta función permite apreciar la convergencia de la estimación con respecto al aumento del tamaño muestral.

Por último, la función *_gen_sets* permite generar varios conjuntos en una matriz bidimensional que almacena un conjunto en cada columna. La lógica implementada no utiliza ningún algoritmo nuevo; se limita a llamar en múltiples ocasiones a *_gen* para lograr una generación fácil de múltiples conjuntos. La conformación de este tipo de conjuntos es deseada cuando se entrenan clasificadores basados en redes neuronales como los desarrollados previamente por el Grupo de Radares de la CUJAE [7, 24].

3. –RESUTADOS

La presente sección muestra la aplicación de cada una de las funciones implementadas. Primeramente, se valida la implementación de la PDF por comparación con gráficas proporcionadas por otros autores. Luego, el resto de las funciones son validadas a través de su interacción con la PDF. La sección se divide en varios apartados que agrupan las funciones con propósitos comunes.

3.1.- FUNCIONES DENSIDAD DE PROBABILIDAD Y PROBABILIDAD ACUMULATIVA (PDF Y CDF)

Las funciones de PDF y CDF para la distribución K de la amplitud y de la potencia fueron implementadas a través de *_pdf* y *_cdf*. Cada una de las cuatro funciones creadas toma dos parámetros de entrada correspondientes a los parámetros de la distribución. Además, es necesario especificar los puntos del eje de las abscisas donde se desea realizar el gráfico.

La figura 1 muestra resultados obtenidos utilizando la función *k_pdf*. La gráfica (a) revela el efecto de la variación del parámetro de escala (c) y la gráfica (b) hace lo mismo para el parámetro de forma (v). La distribución K de la amplitud converge con la distribución Rayleigh cuando $v \rightarrow \infty$; mientras que valores pequeños de v exhiben colas muy alargadas.

El mismo efecto en la variación de los parámetros mostrado en la figura 1, se puede apreciar en la figura 2, generada con *k2_pdf*, para el caso de la distribución K de la potencia. Ahora la K se convierte en una exponencial para $v \rightarrow \infty$. Adviértase que en la gráfica b, el parámetro de forma fue alterado para apreciar su efecto sobre el trazo de la curva y el de escala se mantuvo en correspondencia con la variación de v para mantener la media igual a la unidad de acuerdo a (23).

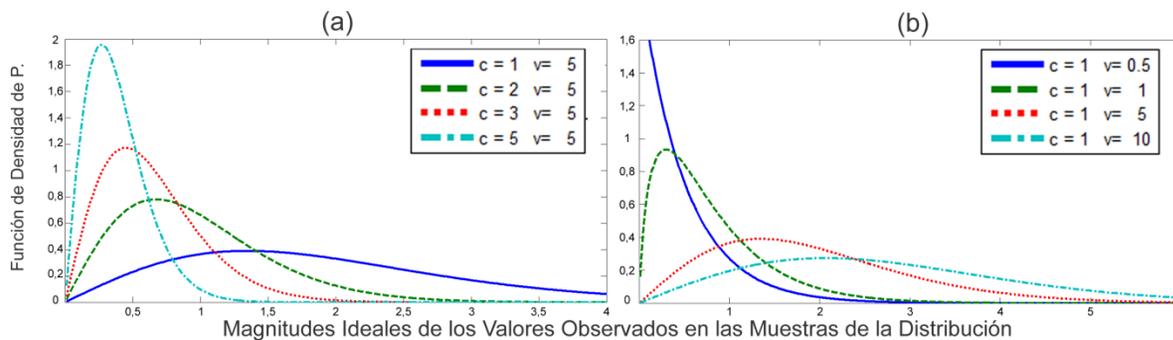


Figura 1

Efecto de la variación de los parámetros en la forma de la curva PDF K.

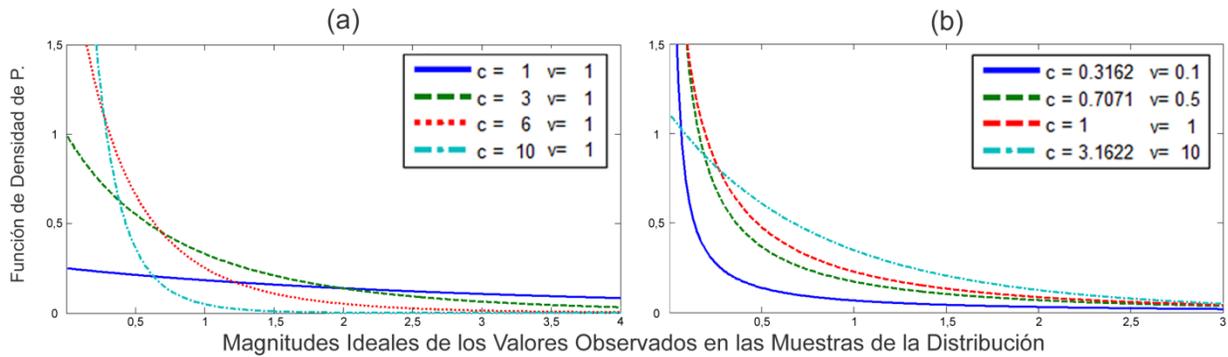


Figura 2

Efecto de la variación de los parámetros en la forma de la curva PDF K2.

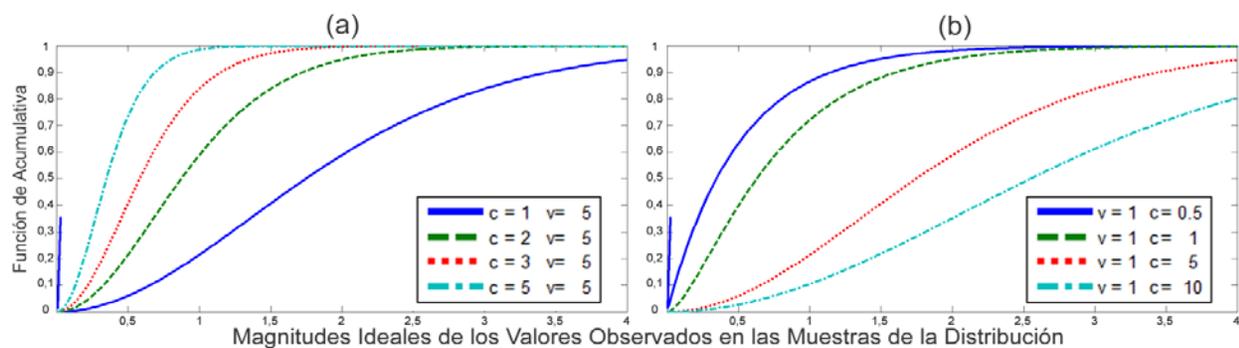


Figura 3

Efecto de la variación de los parámetros en la forma de la curva CDF K.

Las figuras 3 y 4 despliegan las salidas respectivas de las funciones k_cdf y $k2_cdf$. Obsérvese que la variación seleccionada para los parámetros coincide con aquella de las figuras 1 y 2. Es importante notar que las funciones acumulativas se acercan a uno cuando las de densidad están muy próximas a cero.

Los resultados ofrecidos en esta sección son básicos para la modelación cabal de la distribución K. La comparación de las curvas obtenidas con aquellas ofrecidas en [10, 17] permite comprobar que se implementó la distribución K tal y como se describe en la literatura. Los siguientes apartados intertuarán con las funciones de PDF para asegurarse que los resultados alcanzados quedan validados.

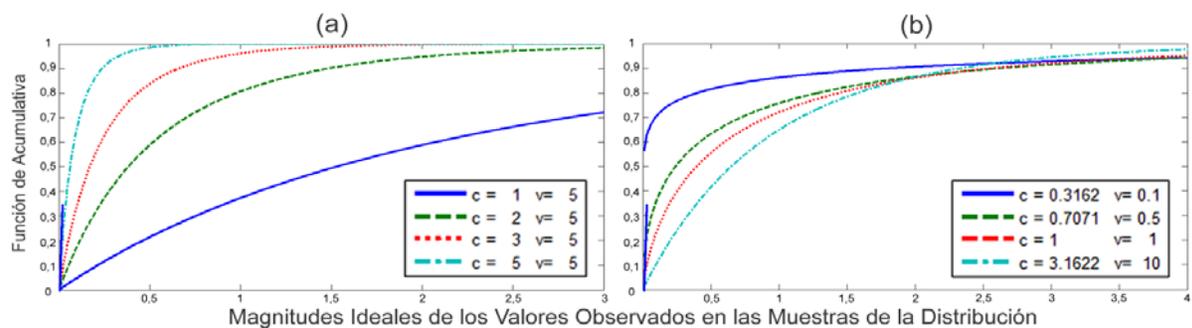


Figura 4

Efecto de la variación de los parámetros en la forma de la curva CDF K2.

3.2.- GENERACIÓN DE VARIABLES DISTRIBUIDAS K

Las variables distribuidas K se obtienen de la síntesis de una distribución Rayleigh y una Root-Gamma para el caso de la amplitud. Cuando se opera en el dominio de la intensidad se utilizan en su lugar una Exponencial y una Gamma. En total, se crearon cuatro pares de funciones asociadas a la generación de variables K. El par de funciones básicas está formado por k_gen y $k2_gen$ que permiten generar muestras K y K2 respectivamente, solicitando la configuración de los parámetros c y v y la cantidad de muestras a generar.

Los tres pares de funciones restantes llaman a k_gen y $k2_gen$ y grafican su contenido de diferentes maneras. El par $_gen_plot$ grafica las muestras en una secuencia, según se muestra en la figura 5. En la gráfica (a) se ilustran las muestras K y en la gráfica (b) las K2, en un esquema que será seguido en el resto de las figuras.

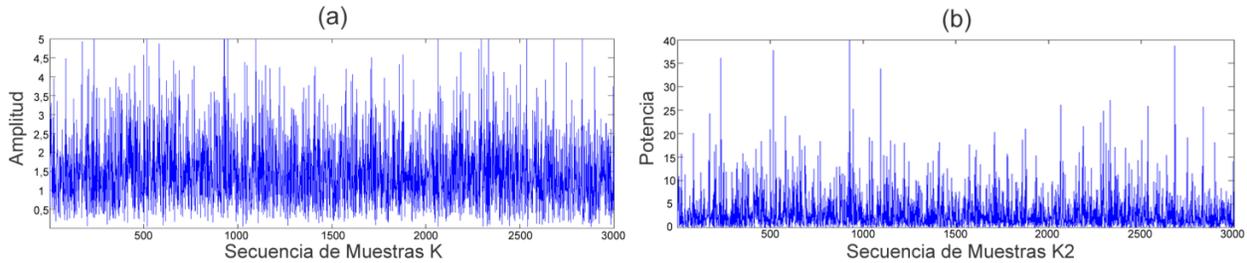


Figura 5

Secuencia de 3000 muestras de las distribuciones K y K2.

El par de funciones $_gen_hist$ arregla las muestras en histogramas. Así proporciona una manera mucho más cómoda de estudio del conjunto generado. En la figura 6 se grafican dos conjuntos de 3000 muestras K y K2 para los parámetros $c = 1$ y $v = 3$.

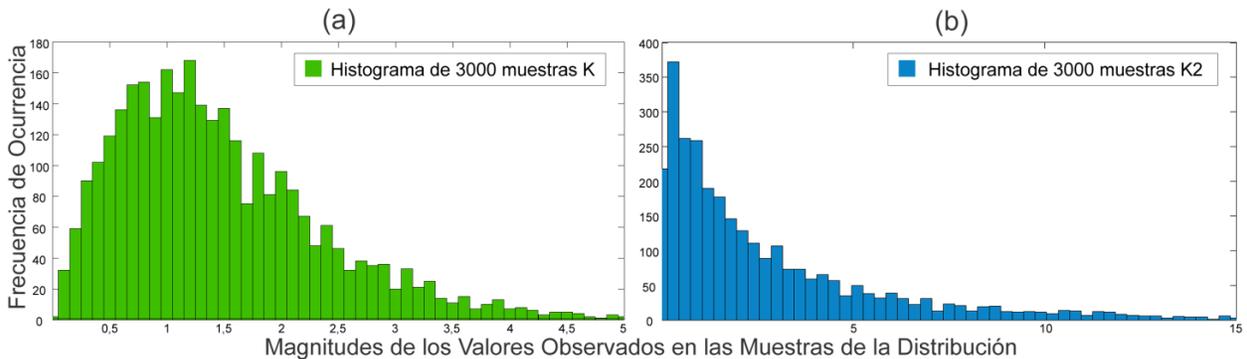


Figura 6

Histogramas generados a partir de dos conjuntos de 1000 muestras K y K2.

Adicionalmente, las funciones $k_gen_compare$ y $k2_gen_compare$ están concebidas para comparar los histogramas con las curvas teóricas de PDF. De esta forma se logra la validación del componente de generación de la variable K. Obsérvese en la figura 7 que dos conjuntos de 1000 muestras fluctúan alrededor del valor ideal del PDF de amplitud y potencia K. Este es el comportamiento esperado para un tamaño muestral reducido.

3.3.- AJUSTES DE LAS MUESTRAS AL MODELO

El presente apartado incluye dos pares de funciones que permiten comprobar la pertenencia de las muestras generadas al modelo. Los dos mecanismos clásicos implementados son: la medición de los residuos y la prueba de ajuste chi-cuadrado (*chi-squared*).

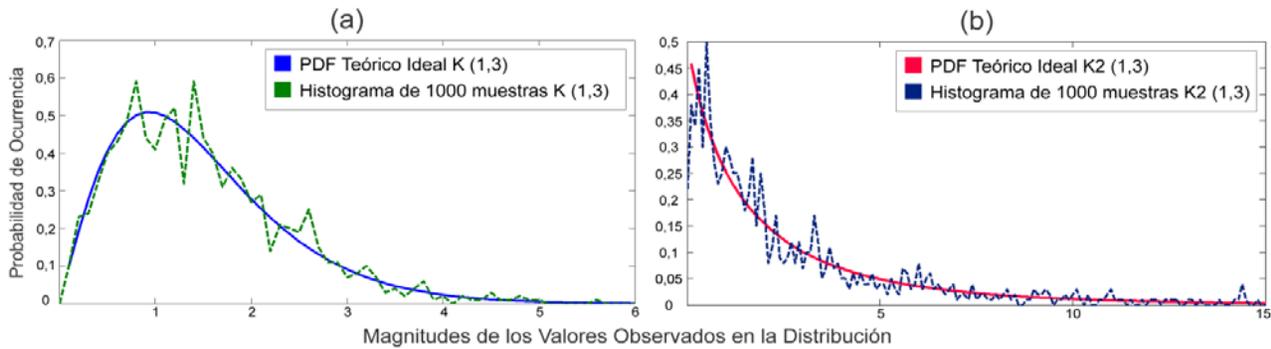


Figura 7

Comparación entre el PDF teórico y la densidad de probabilidad observada en histogramas K y K2.

Las funciones $k_residuals$ y $k2_residuals$ miden la diferencia entre el PDF teórico y el histograma generado a partir de muestras, recorriendo las curvas con la longitud del paso solicitada. La figura 8 presenta un gráfico de los residuos de un conjunto de 1000 muestras K y otro K2. Como puede observarse, los errores cometidos por la aproximación del histograma se distribuyen tanto en valores positivos como en negativos, en lo que constituye un resultado esperado. Obsérvese que la parte posterior de los gráficos presenta magnitudes inferiores de residuos. Esto está en correspondencia con la cercanía entre histograma y modelo observada en la figura 7 para la región de la cola.

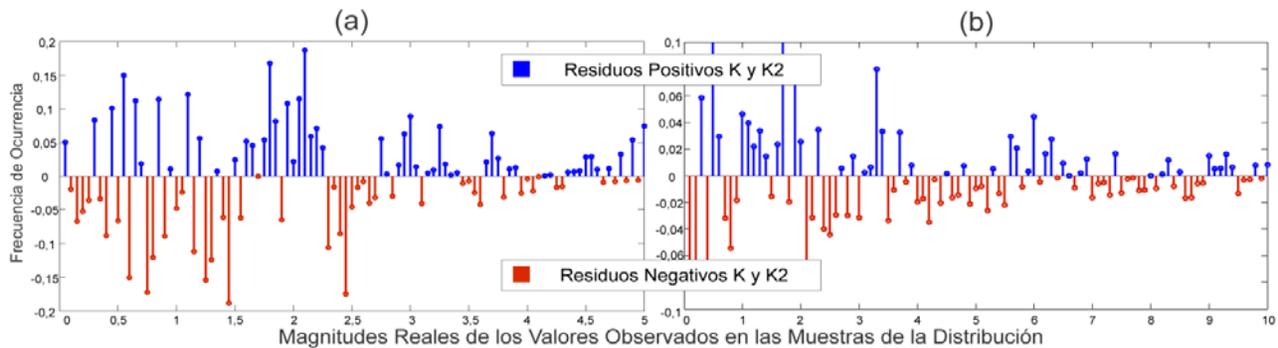


Figura 8

Medición de los residuos de histogramas con respecto a la PDF teórica K y K2.

En ocasiones es difícil deducir si un determinado valor medio o máximo en los residuos niega la pertenencia de un conjunto de muestras a una distribución. Precisamente, la prueba Chi-cuadrado brinda un algoritmo de bondad de ajuste que sintetiza la decisión sobre la pertenencia de las muestras a la distribución en un valor único. El valor, comúnmente denotado p , acepta la hipótesis de pertenencia siempre que esté por encima de 0,05.

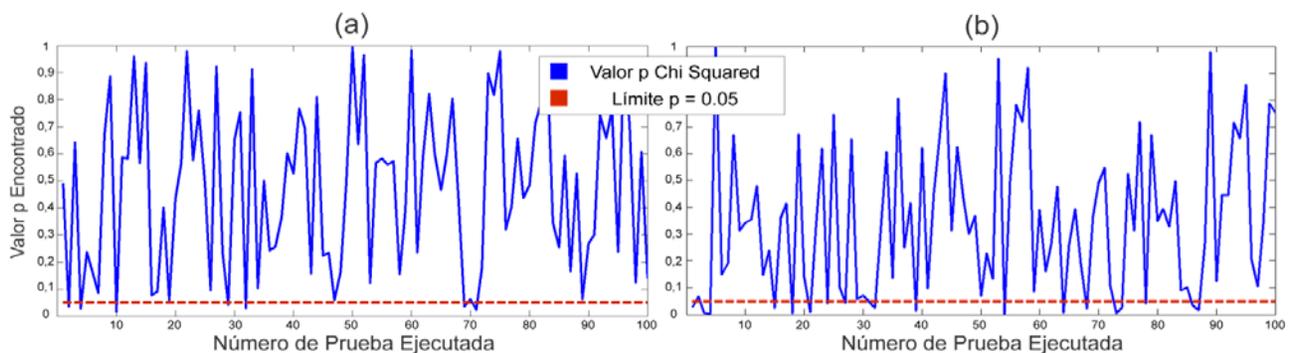


Figura 9

Valores p obtenidos luego de ejecutar 200 pruebas Chi-Cuadrado sobre un conjunto de 1000 muestras.

La figura 9 grafica 200 pruebas Chi-Cuadrado realizadas sobre 100 conjuntos K (gráfica a) y 100 conjuntos K2 (gráfica b). Obtenidos utilizando la función `_chi_squared`, los gráficos revelan una diferencia importante entre las distribución K y K2. La K2 parece ser más *spiky* porque rechaza la hipótesis de pertenencia más seguido para un tamaño muestral común.

3.4.- ESTIMACIÓN DE MOMENTOS Y PARÁMETROS

El último grupo de funciones está destinado a la estimación de los momentos algebraicos y centrales de la distribución K en el dominio de la amplitud y de la intensidad, así como a la estimación del parámetro de forma que tiene especial importancia en la detección según se ha comprobado en [6].

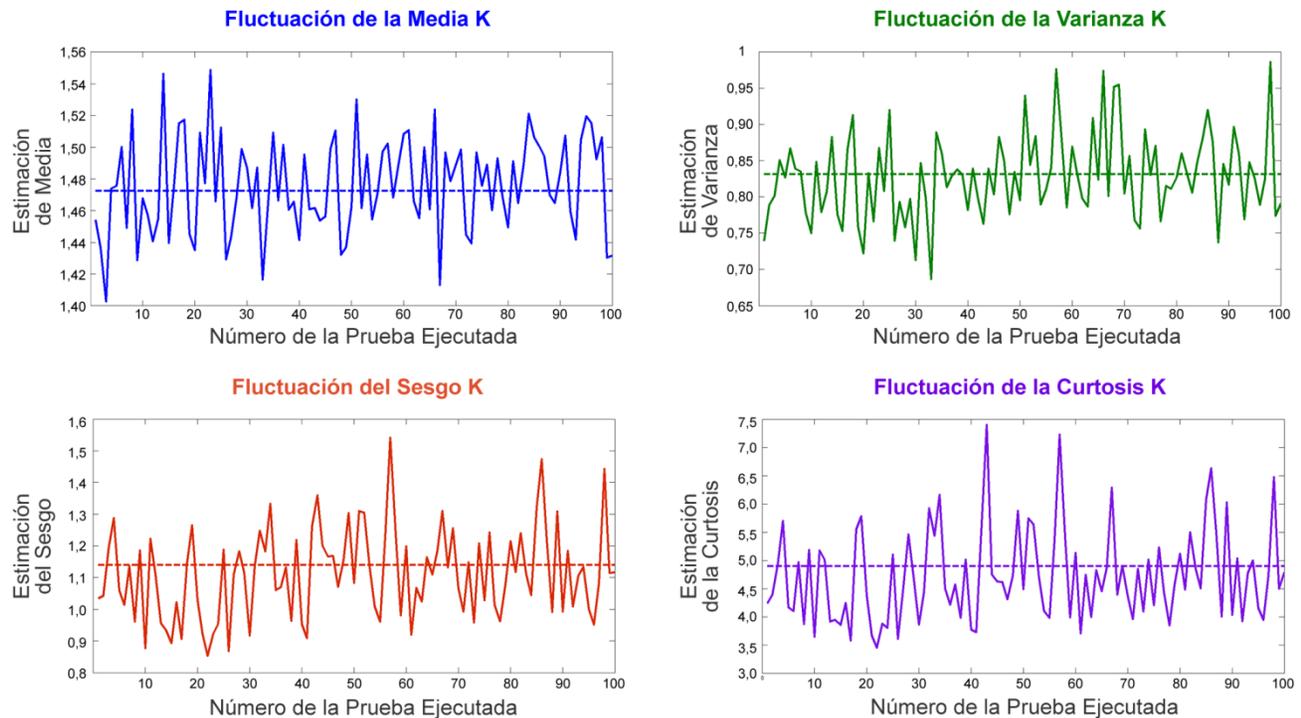


Figura 10

Fluctuación de variables estadísticas de la distribución K alrededor del valor teórico.

Un ejemplo del uso de las funciones `k_ideal_moments` y `k_real_moments` se muestra en la figura 10. En ella fueron estimados la media, varianza, asimetría y curtosis ideales y reales de 100 conjuntos de 1000 muestras distribuidas K. Como puede apreciarse, las funciones están correctamente implementadas pues los valores reales calculados de los conjuntos oscilan alrededor de los teóricos. Un comportamiento similar fue obtenido para los pares de funciones `(k2_ideal_moments, k2_real_moments)`, `(k_alg_ideal_moments, k_alg_real_moments)` y `(k2_alg_ideal_moments, k2_alg_real_moments)`.

El siguiente par de funciones, `k_estim_pary` y `k2_estim_par`, implementan la estimación del parámetro de forma de las distribuciones K y K2 utilizando los momentos segundo y cuarto. La estimación no es exacta y su precisión mejora con el aumento del tamaño muestral. La figura 11 grafica un ejemplo de la estimación para 100 conjuntos de 1000 muestras cada uno, generados con $v = 3$. Es fácil apreciar que el esquema de estimación funciona correctamente pues los valores obtenidos fluctúan alrededor de la magnitud de v con que fueron generadas las muestras.

Por último, las funciones `k_gen_sets` y `k2_gen_sets` permiten generar conjuntos de muestras especificando los parámetros deseados. Por ejemplo, para conformar la figura 11 fue necesario generar los conjuntos apropiados con `_gen_sets` antes de llamar a `_estim_par`.

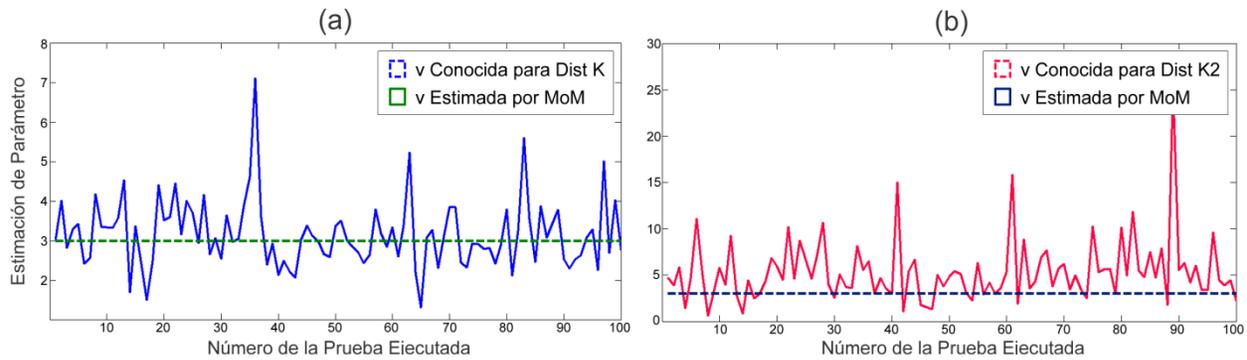


Figura 11

Estimación del parámetro de forma K y K2 usando el método de los momentos.

4. –DISCUSIÓN

El presente artículo contribuye al desarrollo del Grupo de Radares de la CUJAE pues define el conjunto de funciones necesarias para la modelación, en MATLAB, de la distribución estadística K, tanto en su versión de la amplitud como de la potencia. El desarrollo aquí logrado viene a ser una continuación de lo implementado en para las distribuciones Weibull, Log-Normal, Gamma, Exponencial y Rayleigh. La implementación de las funciones relacionadas con los momentos y con la estimación de parámetros facilita el desarrollo de estudios vinculados a los detectores DRACEC[25] (Detección de blancos de Radar por Análisis y Clasificación Estadística de la emisión Celular) y NATE-CFAR (Neural Adaptive Threshold Estimation- Constant False Alarm Rate, Estimación Neuronal Adaptativa del Umbral con Razón de Falsa Alarma Constante)[7, 26], dos de las líneas principales de investigación del grupo de radares del ISPJAE.

La implementación cabal de la distribución K lograda por los autores viene a suplir las carencias encontradas en investigaciones anteriores donde los resultados estuvieron limitados por el incompleto entendimiento y modelación de la distribución. El presente documento tiene además la ventaja de sintetizar una descripción teórica importante y unificada de las expresiones relacionadas con la distribución K. Los autores encontraron en la revisión de la literatura, que ninguna investigación previa había recogido y relacionado todas las expresiones referentes a la distribución K y que al mismo tiempo el modelo está entre los más usados en investigaciones actuales [13, 27]. El aporte actual cobra también una significación educativa porque ayuda a entender el mecanismo compuesto de generación de muestras de amplitud del clutter marino y puede ser incluido en análisis y experimentos conducidos en los cursos de postgrados de radares impartidos en el ISPJAE.

El aporte de [28] se une al actual para comenzar a formar una pequeña biblioteca de modelación de clutter de radar. El objetivo perseguido por los estudios es el de contribuir a la creación del producto informático MATE-CFAR 2 (MATlab Test Environment- Constant False Alarm Rate detectors 2, Ambiente de Pruebas en MATLAB para detectores de Razón de Falsa Alarma Constante que a su vez sería una progresión del MATE-CFAR versión 1 presentado en [5]. La nueva herramienta pretende incluir modelación de diversos tipos de clutter, blancos y detectores CFAR de radar, así como técnicas de estimación de parámetros.

5. –CONCLUSIONES

Se logró implementar la distribución K en MATLAB para la modelación de clutter de radar tanto en su versión de la amplitud como para el caso donde se opera con mediciones de potencia. La modelación adquiere un carácter novedoso pues la aplicación MATLAB no cuenta con ninguna función de modelación K, a diferencia de lo que sucede con otras distribuciones estadísticas. Las funciones implementadas incluyen la modelación de la PDF, la CDF, la función generadora de variables aleatorias, algoritmos de bondad de ajuste, momentos y estimación de parámetros, entre otros. La validez de la solución fue comprobada mediante la comparación con curvas de PDF y CDF dadas por otros autores, y mediante la interacción entre las diferentes funciones implementadas.

El aporte de este artículo pretende unirse a otros para la constitución de una biblioteca informática de simulación de radares que se denominará MATE-CFAR 2 y tiene un aporte educativo e investigativo. Conjuntamente, la implementación actual

viene a suplir defectos de investigaciones anteriores que no contaban con una definición y modelación cabal de la distribución K como la lograda. Consecuentemente, se viabiliza la realización de estudios relacionados con convergencia de los momentos K y técnicas de estimación de parámetros de distribuciones estadísticas.

Los autores de la presente recomiendan la modelación de otras distribuciones de radar como la Pareto e Inversa Gaussiana, y la simulación de procesadores CFAR. Sería provechoso también agregar al conjunto de funciones simuladas métodos alternativos de estimación de parámetros, un medidor de residuos con respecto al CDF y la prueba de bondad de ajuste K-S (Kolmogorov-Smirnov) en sus versiones clásica y mejorada. Igualmente, es necesario lograr modelar las propiedades de correlación del clutter marino en el ánimo de evaluar su influencia en los esquemas de detección. Por último, los autores recomiendan la realización de estudios bibliográficos sobre el rango de parámetros que pueden tomar las distribuciones abordadas cuando se ajustan a datos de clutter.

REFERENCIAS

1. Melvin WL, Scheer JA. Principles of Modern Radar, Vol III Radar Applications. Scitech Publishing; 2014)
2. Ward K, Tough R, Watts S: Sea Clutter Scattering, the K Distribution and Radar Performance. 2da ed. The Institution of Engineering and Technology; 2013.
3. Ping Q. Analysis of Ocean Clutter for Wide-Band Radar Based on Real Data. Proceedings of the 2011 International Conference on Innovative Computing and Cloud Computing, Wuhan, China. 2011; pp. 121-124.
4. Ishii S, Sayama S, Mizutani K. Effect of Changes in Sea-Surface State on Statistical Characteristics of Sea Clutter with X-band Radar. Wireless Engineering and Technology. 2011; 2(3): 175-183.
5. Machado Fernández JR, Bacallao Vidal JC. MATE-CFAR: Ambiente de Pruebas para Detectores CFAR en MATLAB. Telem@tica. 2014; 13(3): 86-98.
6. Machado Fernández JR. Estimation of the Relation between Weibull Distributed Sea clutter and the CA-CFAR Scale Factor. Journal of Tropical Engineering. 2015; 25(2): 19-28.
7. Machado Fernández JR, Bacallao Vidal JC, Chávez Ferry N. A Neural Network Approach to Weibull Distributed Sea Clutter Parameter's Estimation. Inteligencia Artificial. 2015; 18(56): 3-13.
8. Skolnik MI. Radar Handbook. 3ra ed. McGraw-Hill; 2008.
9. Abraham DA, Preston JR. Statistical Analysis of Monostatic and Bistatic Echoes from ShipWreck. IEEE Journal of Oceanic Engineering. 2010; 25(1): 1-8.
10. Oluwale Oyedokun T. Sea Clutter Simulation MsC Thesis. University of Cape Town; 2012. Disponible en: <https://open.uct.ac.za/handle/11427/14557>
11. Dong Y. Clutter Spatial Distribution and New Approaches of Parameter Estimation for Weibull and K-Distributions. Reporte Técnico DSTO-RR-0274; 2012. Disponible en: <http://dSPACE.dsto.defence.gov.au/dSPACE/handle/1947/3335>
12. Kumar Yadav A, Kant L. Moving Target Detection using VI-CFAR Algorithm on MATLAB Platform. International Journal of Advanced Research in Computer Science and Software Engineering. 2013; 3(12): 915-918.
13. Watts S, Rosenberg L. A Review of High Grazing Angle Sea Clutter. Proc IEEE 2013 International Conference on Radar; 2013.
14. Ward KD. Compound Representation of High Resolution Sea Clutter. Electronic Letters. 1981; 17(16): 561-563.
15. Ward KD. Sea-clutter: Scattering, the K distribution and Radar Performance. The Institute of Engineering Technology; 2006.
16. Watts S. A Practical Approach to the Prediction and Assessment of Radar Performance in Sea Clutter. Proc. IEEE International Radar Conference; 1995.
17. Cetin A. CFAR Detection in K-Distributed Sea Clutter. MsC Thesis. The Graduate School of Natural and Applied Sciences of Middle East Technical University; 2008. Disponible en: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.634.617&rep=rep1&type=pdf>
18. Nohara TJ, Haykin S. Canadian East Coast Radar Trials and the K-Distribution in Editor. IEE Proceedings on Radar and Signal Processing. 1991; 138(2): 80-88.

19. Walck C. On Moments and their Estimation. Reporte Técnico, (Internal Note SUF-PFY/91-01. Particle Physics Group, Department of Physics, University of Stockholm, Stockholm, Sweden; 1991.
20. O'Connor AN. Probability Distributions Used in Reliability Engineering. University of Maryland; 2011.
21. Yim JZ, Chou CR, Wong WK. A Study of the Statistics of Sea Clutter in the Northern Coast of Taiwan. Proceedings of the Seventeenth International Offshore and Polar Engineering Conference, Lisbon, Portugal; 2007.
22. Redding NJ. Estimating the Parameters of the K Distribution in the Intensity Domain. Reporte Técnico DSTO-TR-0839. Surveillance Systems Division , Electronics and Surveillance Research Laboratory, Department of Defence, Defence Science and Technology Organisation, Salisbury, South Australia; 1999. Disponible en: www.dtic.mil/get-tr-doc/pdf?AD=ADA368069
23. Marques de Sá JP. Applied Statistics using SPSS, STATISTICA, MATLAB and R. 2da ed. Springer; 2007.
24. Machado Fernández JR, García Delgado B, Machado Gil A. Reconocimientos de Parámetros de la Distribución K Asociados a Mediciones de Clutter Marino usando Redes Neuronales Artificiales. VIII Congreso Internacional de Telemática y Telecomunicaciones, Convención Científica de Ingeniería y Arquitectura '14, La Habana, Cuba; 2014.
25. Chávez Ferry N, Bacallao Vidal, JC. Análisis de la Calidad de la Detección de Blancos de Radar por el Método DRACEC. VIII Congreso Internacional de Telemática y Telecomunicaciones, Convención Científica de Ingeniería y Arquitectura '14, Palacio de las Convenciones, La Habana, Cuba, 2014.
26. Machado Fernández JR, Bacallao Vidal JC. Improved Shape Parameter Estimation in K Clutter with Neural Networks and Deep Learning. International Journal of Interactive Multimedia and Artificial Intelligence. 2016; 3(7): 96-103.
27. Rosenberg L. The effect of temporal correlation with K and KK distributed Sea-Clutter. Proc. IEEE Radar Conference; 2012.
28. Machado Fernández JR. Modelación de las Distribuciones Rayleigh y Exponencial en MATLAB para Aplicaciones de Radar. Telem@tica. 2016; 15(2): 1-15.

AUTORES

José Raúl Machado Fernández(josemf@electrica.cujae.edu.cu), Ing. en Telecomunicaciones y Electrónica, Profesor e Investigador, Aspirante a Doctor, Grupo de Investigación de Radares, Departamento de Telecomunicaciones y Telemática, Facultad de Telecomunicaciones y Electrónica, Universidad Tecnológica José Antonio Echeverría (CUJAE), La Habana, Cuba. Sus intereses de investigación incluyen la teledetección, el procesamiento digital de señales, el modelado del clutter marino y la aplicación de la inteligencia artificial a la resolución de diversos problemas de ingeniería.

Jesús de la Concepción Bacallao Vidal(bacallao@electrica.cujae.edu.cu), Doctor en Ciencias Técnicas, Profesor Titular e Investigador, 2do Jefe del Grupo de Investigación de Radares, Departamento de Telecomunicaciones y Telemática, Facultad de Telecomunicaciones y Electrónica, Universidad Tecnológica José Antonio Echeverría (CUJAE), La Habana, Cuba. Sus intereses de investigación incluyen la teledetección, la evaluación del desempeño de procesadores CFAR y la creación de esquemas de detección alternativos basados en el procesamiento de la señal en el dominio de los momentos.