

Tesis Doctoral

Para acceder al Título de “Doctor en Ciencias Exactas e Ingeniería”

Modelo de Detección Automática para Ecos de Radar



Autora: María Graciela Molina

Director: Dr. Miguel A. Cabrera

**Facultad de Ciencias Exactas y Tecnología
Universidad Nacional de Tucumán
Tucumán, Argentina
2017**

Resumen.....	3
Motivaciones.....	6
Objetivos de la tesis	9
Organización de la tesis	10
CAPITULO I	12
Modelo computacional aplicado al procesamiento de señales de radares	12
1.1. Computación científica	12
1.2. Modelado computacional.....	16
1.3. Desarrollo de software científico	18
1.3.1. Algoritmos.....	20
1.3.2. Algoritmos Numéricos	25
1.4. Algoritmos para Procesamiento de Señales.....	25
1.4.1. Detección automática de señales de radar como una aplicación computacional.....	27
CAPITULO II	28
Radar.....	28
2.1. Introducción	28
2.2. Funciones básicas de un radar.....	30
2.3. Señales de radar	35
2.3.1. Ecuación de radar.....	39
2.3.2. Radar Cross Section	43
2.3.3. Modelado de la sección transversal	44
2.4. Radar de onda continua y radar pulsado	48
2.5. Técnica de compresión de pulso.....	53
2.5.1. Decodificación del Eco	55
2.5.2. Codificación.....	71
CAPITULO III.....	79
El medio de estudio.....	79
3.1. Introducción	79
3.2. La ionósfera	79
3.2.1. Estratos ionosféricos.....	80
3.2.2. Regiones ionosféricas regulares	82
3.2.3. Región ionosférica esporádica	84
3.2.3. Irregularidades ionosféricas.....	85
3.2.4. Variabilidad espacial y temporal de la ionósfera.....	85
3.4. Las comunicaciones y la ionósfera	88
3.5. Métodos de estudio y monitoreo de la ionósfera terrestre	90
3.5.4. Mediciones ionosféricas.....	90
CAPITULO IV.....	92
Detección de Ecos de Radar	92
4.1. Introducción	92
4.2. Detectores	93
4.3. Fundamentos de detección	95
4.4. Estimación del umbral	100
4.5. Umbral variable o adaptativo (CFAR).....	104
4.5.1. Pérdida CFAR.....	108

4.5.2. Algoritmo por promedio de celdas (CA-CFAR)	109
4.5.3. Algoritmo por reordenación de celdas (OS-CFAR)	114
4.5.4. Otros Algoritmos Basados en CA-CFAR	118
4.6. Mapa de Clutter	119
CAPITULO V	121
Modelo de Detección Automática de Ecos de Radar	121
5.1. Introducción	121
5.2. Sondaje vertical de la ionósfera mediante un radar ionosférico	121
5.3. Motivación para el estudio de la detección de estratos ionosféricos	126
5.4. Modelo de detección de estratos ionosféricos	132
5.5. Primera etapa del modelo: Codificación	134
5.6. Segunda etapa del modelo: Recepción de señales y muestreo	136
5.7. Tercera etapa del modelo: Correlación	139
4.8. Cuarta etapa del modelo: Mejoramiento de la relación señal/ruido	142
5.9. Última epata del modelo	162
CAPITULO VI	163
Etapa de Detección Automática de Ecos de Radar: Algoritmos	163
6.1. Introducción	163
6.2. Implementación de la Funcionalidad Automática de Detección	165
6.3. Evaluación de la capacidad de detección	167
6.4. Consideraciones respecto a los parámetros usados en los algoritmos	171
6.5. Implementación de los tres algoritmos propuestos	173
6.6. Análisis del desempeño de los algoritmos propuestos	186
6.6.1. Análisis de Complejidad Algorítmica	188
CAPITULO VII	191
Conclusiones Generales y Perspectivas	191

Resumen

Dentro de las numerosas aplicaciones del procesamiento de señales digitales, se encuentra el procesamiento de señales de radares. El término radar originalmente proviene del acrónimo de Radio Detection and Ranging. Su funcionamiento se basa en la medición del tiempo de viaje de la señal que se transmite, que luego de dispersada y/o reflejada por el blanco de interés, se recibe como un conjunto de ecos.

Este trabajo de tesis versa sobre el desarrollo de un modelo computacional para la detección de ecos de radar que permita la funcionalidad de la detección automática.

Este modelo se diseñó y desarrolló a partir del análisis a las condiciones específicas del blanco tales como las características estacionarias y/o dinámicas del mismo, la relación señal ruido (SNR) de las señales, los tiempos de respuesta del sistema propuesto, entre otras. Esto se debe a que, la capacidad de detección varía entre los diferentes tipos de radares por lo que no es posible contar con un modelo universal de detección. Donde cada radar está diseñado para detectar un tipo específico de blanco. En primera instancia se caracterizó el blanco que este equipo es capaz de detectar. La ionósfera es un blanco complejo por lo que fue necesario realizar un enfoque estadístico.

A lo largo de esta Tesis se describen las características de detección de un radar para aplicaciones geofísicas denominado sondador ionosférico. Este instrumento tiene como principal objetivo el estudio de los estratos ionizados de la alta atmósfera terrestre denominados “Ionosfera”, mediante la emisión de ondas de radio con barrido de frecuencias en banda de HF, que permiten obtener lo que se denomina “alturas virtuales”, de este medio.

El diseño fue orientado a los sistemas de radares modernos los que utilizan baja potencia de radio para iluminar el blanco usando esquemas de codificación como modulación digital de portadora de radio. Esta técnica se denomina “compresión de pulso”. Esta técnica requiere del uso de herramientas matemáticas más complejas para el procesamiento para finalmente lograr la detección del eco producido por el blanco y que permite lograr mejores resultados de sensibilidad del equipo.

Dado que el desarrollo del modelo requiere de la determinación de diferentes etapas de detección, se diseñaron cada una de éstas. Se determinaron las técnicas de procesamiento

de señales más adecuadas y se diseñaron e implementaron las herramientas matemáticas necesarias en cada caso.

En función del tipo de blanco a detectar, en nuestro caso los estratos ionosféricos, se seleccionó el esquema de codificación y la modulación de la portadora de radio frecuencia.

Se estudió y diseñó un algoritmo de correlación de la señal recepcionada con el código. En esta etapa a partir de la técnica matemática para realizar la correlación se realizó un análisis cómo deben tratarse las señales muestreadas para optimizar las capacidades en la etapa de correlación.

Durante la etapa de recepción de la señal, se estimó mediante un análisis empírico la cantidad de pulsos necesarios durante la integración coherente de la señal. Para ello se analizaron las características cuasi-estacionarias del blanco y la dinámica típica de la ionósfera durante periodos de tiempo cortos. Esto permitió, además, establecer una cota de tiempo de respuesta requerido por el sistema.

Finalmente se propone una solución algorítmica como modelo computacional para la detección automática basada en teoría de radares y que se ajusta a las funcionalidades requerida. Se utilizaron técnicas de umbral adaptativo que permitieron establecer el diseño algorítmico más adecuado.

El modelo computacional propuesto fue implementado y posteriormente validado utilizando señales adquiridas por un radar ionosférico digital modelo AIS-1 en el Istituto Nazionale di Geofisica e Vulcanologia de Roma, sección Roma 2. Se mediante experimentos de detección de estratos ionosféricos en diferentes condiciones. Los resultados de las simulaciones numéricas del modelo propuesto y utilizando diferentes configuraciones de señales relacionadas con diferentes estados de la ionósfera, mostraron un gran avance en las capacidades de detección del mencionado radar.

Tradicionalmente el ajuste en la configuración de detección de los radares se realizaba principalmente de manera empírica y requería de la expertitud del operador del mismo para mejorar la calidad de detección. Por este motivo el modelo computacional propuesto permite otorgar fundamentación teórica, mayor generalidad e implementar la funcionalidad de detección automática.

En síntesis, a partir del modelo propuesto fue posible obtener un enfoque formal y más adecuado que se basa tanto en teoría de detección como en la experimentación.

Los resultados de este trabajo son aplicables a un vasto campo del sensado remoto por ondas de radio o de sonido, tales como sistemas de vigilancia aérea, tránsito vehicular, seguridad marítima y aplicaciones en las ciencias médicas del diagnóstico por imágenes.

Motivaciones

Tradicionalmente el procesamiento de señales se circunscribía al análisis de señales analógicas utilizando principalmente complejos sistemas de hardware. Con el advenimiento de los sistemas de cálculo asistidos por computadores de mejores prestaciones, junto con el desarrollo de teorías de procesamiento digital de señales, el hardware evolucionó a sistemas que permiten un uso flexible de la electrónica mediante el desarrollo de software específico (Bianchi et al, 2003; Mahafza, 2008; entre otros). Así la posibilidad de almacenar señales digitales y procesarlas sin pérdida significativa de información, permite el desarrollo de algoritmos numéricos para ser implementados en diferentes sistemas (Oppenheim, 2005; Roberts, 2004). Actualmente, contar con un modelo computacional que potencie todos los aspectos recién mencionados requiere del conocimiento de herramientas y teorías de las Ciencias de la Computación. Este es el caso del desarrollo de la tecnología radar. Tanto el diseño de un radar como la mayoría del procesamiento de las señales involucradas se fundan en el desarrollo de un modelo computacional, y en la ejecución de simulaciones para finalmente tener un software que sea implementado en el equipo.

Esta tesis está basada en tres tipos de motivaciones científicas que corresponden a disciplinas complementarias.

1) Motivación respecto al aporte en la tecnología radar:

Tradicionalmente la habilidad para detectar de los radares en general y los radares ionosféricos en particular se sustentaba en la expertitud y pericia de los técnicos que los operaban. En su gran mayoría el ajuste de parámetros como el umbral de detección recaía en el conocimiento del operador respecto al blanco y al medio por el que se realizaba la transmisión. Esto significa que se basaba en datos empíricos, difíciles de estimar a priori, difícil de generalizar, de escalar a otros niveles, y hacía que el equipo sea altamente dependiente de las capacidades y experiencia del operador, así como también de la locación particular del equipo. Por lo tanto, se detectó la necesidad de desarrollar un sistema que pueda salvar la mayoría de estos problemas mediante una formulación científica, más general y que permita ser adaptada a diferentes circunstancias. Para ello se

propone en este trabajo de tesis obtener un modelo computacional para la detección automática de ecos de radar. Este aporte original tiene como objetivo mejorar el desarrollo tecnológico de equipos radares para aplicaciones geofísicas como es el caso del sondador ionosférico.

2) Motivación desde el punto de vista geofísico:

Desde el punto de vista geofísico “detectar” los diferentes estratos ionosférico y estimar a que altura se encuentran en la alta atmósfera terrestre, requiere de un estudio del medio físico por el que se propagan las ondas electromagnéticas emitidas por el radar. Dado que el problema de detectar un eco radica en decidir si un determinado valor representa un blanco o simplemente es interferencia, es una tarea muy compleja. Por ejemplo, un pájaro que vuela en un ambiente lluvioso puede ser el blanco para un determinado radar, pero por ejemplo si se trata de un radar meteorológico lo que se desea detectar es la lluvia. En el primer caso, el pájaro es el blanco y la lluvia es ruido. Mientras que en el segundo caso el blanco es la lluvia y el pájaro puede reflejar señales de interferencia. La detección no es trivial ya que no solo se debe tener un buen conocimiento del blanco, sino que además el mismo está inmerso en un ambiente que produce ruido, interferencia, u otras señales no deseadas. Si el umbral de detección está muy ajustado, las señales no deseadas no son detectadas pero se corre el riesgo de perder algunos blancos que sí se querían detectar. Por otro lado, si el umbral es bajo, existe una gran probabilidad de tener falsas alarmas (detectar alguna información erróneamente). Para cada blanco inmerso en un medio con ciertas características es necesario encontrar un compromiso en el valor del umbral. Si además estas características varían, entonces el umbral debe ser capaz de adaptarse.

En el caso de radar ionosférico, dado que: a) a partir del valor estimado de la altura de la ionosfera se puede calcular la densidad de electrones presentes; y b) el conocimiento de la morfología y dinámica de la ionósfera impacta directamente en el desarrollo de las comunicaciones que utilizan a la ionosfera para reflejar ondas de radio, o en las comunicaciones transionosféricas (como es el caso de los sistemas GNSS, otros). Entonces, es indispensable contar con un sistema que permita acotar la cantidad de falsas alarmas según con un criterio determinado.

En términos concretos para la aplicación geofísica, significa que los errores cometidos en la detección se propagan de manera que afectan considerablemente la calidad de los datos adquiridos por el radar. Se considera necesario optimizar la detección para aportar confiabilidad a los datos arrojados por el radar ionosféricos y que servirán como entrada a modelos para la estimación de la densidad de electrones.

3) Motivaciones relacionadas al desarrollo de software científico:

Las motivaciones para el desarrollo de este trabajo de tesis engloba, además, la necesidad de contar con un sistema de procesamiento de señales que se pueda implementar en un equipo real. Es decir se propone el diseño de un algoritmo que se pueda plasmar mediante el desarrollo de un software científico de calidad, que atienda los requerimientos respecto a los tiempos de respuesta esperados y que permita realizar el procesamiento digital de las señales del radar eficientemente.

En conclusión, la contribución central de esta tesis es la de llevar a cabo un diseño de un modelo computacional para la detección automática de ecos, que sea acorde a las características del blanco y que permita ser implementado en sistemas de sensado remoto por ondas de radio en un medio ionizado.

Objetivos de la tesis

El objetivo principal de este trabajo de tesis es proponer un modelo de detección automática para detección de ecos de señales de radar acorde a características específicas del blanco. En particular diseñar un modelo computacional para la detección automática de ecos de un radar ionosférico o sondador ionosférico.

Estudiar, analizar y caracterizar el blanco, la ionósfera, según su morfología y dinámica para establecer sus características estadísticas.

Determinar el esquema de codificación binaria más apropiado para la generación de pulsos codificados por parte del radar.

Determinar las estrategias de correlación para la decodificar las señales recepcionadas y establecer técnicas para el mejoramiento de la calidad de la relación señal/ruido (SNR).

Establecer criterios concretos para establecer un umbral adaptativo para la detección

Diseñar y evaluar un algoritmo específico de detección automática adecuado al modelo propuesto.

Realizar una evaluación del software numérico utilizado realizando un estudio de los diferentes productos disponibles en la actualidad.

Realizar un análisis de los resultados arrojados por las simulaciones numéricas para diferentes casos de estudio.

Realizar una evaluación del desempeño computacional.

Organización de la tesis

A lo largo de esta tesis se presentan fundamentos teóricos sobre modelos computacionales, sobre radares y sobre ionósfera que son los pilares para el diseño de un modelo de detección automática de ecos de radar para un radar de aplicaciones geofísicas.

En el Capítulo I se presenta el concepto de modelo computacional, y su relación con el modelo conceptual. También se presentan conceptos sobre algoritmos y análisis de eficiencia de los mismos. Finalmente se introduce la noción de la detección automática de ecos de radar como una aplicación computacional.

Los fundamentos de teoría para el modelado de un radar, sus funciones, ecuaciones que lo gobiernan, características estadísticas de blancos conocidos y el modelado de la sección transversal de un radar son tratados en el Capítulo II. Aquí se presenta una técnica muy importante denominada compresión de pulso, así como también otras técnicas importantes. Se muestran los principales tipos de codificación y sus propiedades matemáticas en vista a la detección del blanco.

A partir de estos fundamentos se inicia el estudio de las características del blanco para un radar ionosférico, la ionósfera. Esto se trata en el Capítulo III donde se realiza un estudio desde el punto de vista geofísico para entender la morfología y variabilidad de la ionósfera. Se subraya la importancia de la ionósfera para el desarrollo de las telecomunicaciones dada su capacidad de refractar ondas de radio. Se destaca la importancia de poder caracterizar el comportamiento típico de la ionósfera para posteriormente entenderla como un complejo blanco a detectar.

En el Capítulo IV se presenta la teoría de detección como un problema de optimización para encontrar un umbral adaptativo acorde. Se describen las estrategias tradicionales para la detección de ecos y se ahonda en los algoritmos basados en mantener constante la probabilidad de falsa alarma.

En el Capítulo V se presenta el modelo propuesto para la detección automática de ecos de radar para un sondador ionosférico. Se describe la técnica de sondaje vertical utilizada por este tipo de instrumento.

En el Capítulo VI se analizan las diferentes etapas en el proceso de detección.

Los algoritmos implementados para señales reales adquiridas de un radar son utilizados para validar el modelo propuesto. Allí se muestran los resultados obtenidos para diferentes casos de estudio basados en datos reales. Las señales utilizadas fueron obtenidas experimentalmente durante una pasantía en el Istituto Nazionale di Geofisica e Vulcanología de Roma. Durante esa estancia se diseñaron y llevaron a cabo los experimentos que se utilizaron en este trabajo de tesis y que permitieron publicar tres trabajos en revistas internacionales de alto impacto. Se analizó, además, el desempeño del modelo computacional propuesto.

En el Capítulo VII se presentan las conclusiones de la tesis y se muestran los trabajos originales publicados producto de la misma.

CAPITULO I

Modelo computacional aplicado al procesamiento de señales de radares

1.1. Computación científica

Tradicionalmente las predicciones o nuevos desarrollos en la ciencia o en la ingeniería seguían el paradigma clásico. Este paradigma se basa en la combinación de la teoría matemática y la experimentación para generar nuevos descubrimientos. Así, fue posible por ejemplo, la predicción del movimiento de los cuerpos celestes. En 1687 Sir Isaac Newton publica “*Philosophiæ naturalis principia matemática*” o “*principia*” (obra que marcó un punto de inflexión en la historia de la ciencia) donde vuelca sus descubrimientos en mecánica y cálculos matemáticos, observaciones y experimentos y donde enuncia leyes que llevan a describir el movimiento de los cuerpos celestes. Este enfoque clásico entre teoría y experimentación es el fundamento del desarrollo de la ciencia y la tecnología y dio como resultados una inmensa cantidad de descubrimientos y aplicaciones. Siguiendo esta metodología los hermanos Wright pudieron estimar y diseñar un control de vuelo que les permitió desarrollar en 1903 la primera aeronave auto impulsada y controlable (<https://wright.nasa.gov/overview.htm>, 2017).

Sin la teoría y la experimentación no ocurrirían nuevos descubrimientos. Sin embargo, a medida que los diseños fueron más complejos los cálculos numéricos comenzaron a ser más intensivos, la capacidad de cálculo humana se vio limitada. Fue claro, entonces, que se necesitaba encontrar una manera de realizar cálculos de manera intensiva y automática. Una muestra es el trabajo que supuso descifrar la máquina Enigma durante la Segunda Guerra Mundial por el creador del primer computador eléctrico (Colossus) y padre de los modernos algoritmos, Alan Turing. Entre 1942 y 1951, John von Neumann dirigió un equipo en la Universidad de Princeton que diseñó y construyó la primera computadora digital y que en cuya arquitectura se basan las computadoras modernas.

Con el crecimiento en las capacidades de cálculo de las computadoras modernas, el paradigma clásico en el desarrollo científico agregó un componente más. Actualmente, el

paradigma interrelaciona tres pilares importantes: *la teoría, la experimentación y la computación.*

Este paradigma de la actual investigación y desarrollo permite obtener ventajas de las nuevas capacidades de cómputos que se ven reflejadas en predicciones y desarrollos a gran escala y que no podrían ser resueltas de otra manera. La posibilidad de simular computacionalmente el comportamiento de un sistema, su evolución en el tiempo y tomar decisiones a partir de estas predicciones tiene un gran impacto en la vida humana moderna. Los pronósticos meteorológicos y climáticos, por ejemplo, requieren de simulaciones computacionales que requieren capacidades de cómputo intensivas para poder probar y verificar los diferentes modelos. A partir de los resultados que estas simulaciones arrojan es posible activar un sistema de alertas, como por ejemplo un alerta de Tsunami.

El sistema de predicciones meteorológicas es uno de los más complejos y que, para tener un resultado acorde, debe ser capaz de realimentar cada uno de los pilares del paradigma. Entonces el modelado de la realidad, los cálculos numéricos ejecutados en las simulaciones computacionales y la retroalimentación para validar y verificar el sistema son claves para obtener resultados positivos.

Si bien, agregar la dimensión de computación en principio significa mejorar la capacidad de cálculo, también significa agregar una dimensión en complejidad. No solo es importante tener más capacidad física de cómputo (computadoras con más memoria, mayor velocidad, etc.) sino que requiere de nuevas estrategias y diseños algorítmicos eficientes que sean capaces de sacar ventajas reales principalmente en cuanto a velocidad de procesamiento. Esto significa que es necesario el desarrollo de nuevas técnicas de programación, de manejo de datos y de computación en general.

Un ejemplo interesante de cómo fue migrando la comunidad científica en torno a la computación es la experiencia del “Conseil Européen pour la Recherche Nucléaire” CERN. Este es el mayor laboratorio de investigación en física de partículas del mundo y es mundialmente conocido por desarrollar un experimento denominado “The Large Hadron Collider” (LHC) que actualmente es el más grande acelerador de partículas en el mundo. Este experimento se caracteriza por generar una inmensa cantidad de datos que deben ser almacenados, analizados, procesados y enviados a diferentes centros de

investigación en el mundo. Por ejemplo en un momento determinado uno de los instrumentos generaba alrededor de 2.38 gigabits por segundo, estos datos debían ser analizados y ser transportados entre Sunnyvale y Ginebra (vía Chicago) mediante alguna red de comunicación. Se hicieron los cálculos para a la infraestructura de la red, las distancias se debía diseñar un sistema que permita transmitir 23.888,06 terabit-metros/segundo (Williams, 2004). Este es un ejemplo que muestra los desafíos reales planteados por este sistema. Esto significó, en términos de evolución del cómputo intensivo, que sea necesario encontrar nuevas herramientas en las Ciencias de la Computación para poder asistir a las necesidades de este centro. A partir de estos requerimientos tan exigentes se desarrollaron técnicas para gestionar grandes volúmenes de datos (Big Data) y que cambió los tradicionales paradigmas de computación, de administración de bases de datos, etc.

Este ambicioso proyecto del CERN no solo tuvo un impacto en el desarrollo de la computación en el ámbito científico, sino también sobre la sociedad. Tim Berners-Lee, científico británico que trabajaba en el CERN, inventó Internet o la “World Wide Web” (WWW) en 1989. Esta red originalmente fue concebida y desarrollada para satisfacer las demandas de compartir de manera automática información entre los diferentes científicos y universidades alrededor del mundo.

La evolución en la electrónica de las computadoras modernas además soporta este esquema en el desarrollo científico donde juega un rol preponderante la computación.

El modelo de von Neumann se basaba en la idea de que un programa podía ser representado en la memoria de la computadora, en forma digital junto con los datos, y que la aritmética decimal podía ser sustituida por una aritmética binaria. Las computadoras actuales poseen una arquitectura basada en este modelo pero con nuevas extensiones.

Desde 1986 al 2002, el rendimiento de cada microprocesador aumentaba en promedio, 50% por año. Esta diferencia en el aumento de rendimiento estaba asociada a un cambio en el diseño de los procesadores. Esta evolución se inició en el año 1971 cuando la empresa Intel logró fabricar su primer microprocesador como un circuito integrado para una calculadora, iniciándose así una revolución en la tecnología de la computadoras: el Intel 4004. Gran parte del enorme aumento en el rendimiento de un solo procesador ha

sido impulsado por la cada vez mayor densidad de transistores en los circuitos integrados. A medida que el tamaño de los transistores disminuye, la velocidad total del circuito integrado se incrementa. Sin embargo, así como la velocidad de conmutación de los transistores aumenta, también lo hace el consumo de energía. La mayor parte de esta energía se disipa en forma de calor, y cuando un circuito integrado se calienta demasiado se vuelve poco confiable. Este límite para seguir aumentando la velocidad de los circuitos integrados de manera confiable impactó en el rendimiento de un solo procesador que comenzó a decrecer alrededor a un 20% por año a partir del 2002.

Para el 2005, la mayoría de los principales fabricantes de microprocesadores habían empezado a integrar varios procesadores sobre un solo circuito integrado en vez de seguir desarrollando procesadores más rápidos monolíticos. Esta decisión permitió que el desempeño aumentara rápidamente en dirección al paralelismo (Pacheco, 2011).

Los algoritmos evolucionaron simultáneamente permitiendo generar programas que se ejecuten en paralelo en diferentes unidades de procesamiento. Para esto, se determinaron protocolos y librerías que permiten el pasaje de mensajes entre procesos y organizan la comunicación y sincronización entre estos.

Además, dado que en muchas aplicaciones numéricas el requisito de cálculo intensivo se focaliza en cálculos algebraicos, como operaciones de álgebra lineal, las arquitecturas fueron adaptándose a este escenario. Surgen así los aceleradores de cómputo.

Junto al desarrollo de las capacidades de cómputo se mejoraron las capacidades de conectividad, que sumado a características especiales de hardware llevaron al advenimiento de las supercomputadoras.

Desde 1986 Hans Meuer, profesor de la Universidad de Mannheim, publica un listado de los principales fabricantes de supercomputadoras, sus características y los institutos donde se encuentran instaladas. El principal objetivo de la lista TOP500 es proporcionar una lista de los 500 sistemas informáticos más potentes de propósito general, que son de uso común para aplicaciones de alta gama. Se realiza una actualización dos veces al año desde junio de 1993, que requiere la ayuda de expertos en informática de alto rendimiento, de científicos computacionales, fabricantes, y la comunidad de Internet en general. La lista está disponible gratuitamente en <http://www.top500.org> (2017).

Desde los inicios de la investigación científica las necesidades en el volumen de cálculos se han expandido de tal manera que pensar en un modelado para un sistema físico requiere pensar, además, en un modelado computacional.

1.2. Modelado computacional

Se puede decir entonces que es necesario contar con un medio para poder explicitar el rol de la computación en el modelado de la realidad. Es así que nace el concepto de modelado computacional que establece los mecanismos para diseñar una solución computacional. Para ello se diseñan algoritmos que posteriormente serán implementados o programados en algún lenguaje de computación.

La figura 1.1 muestra un esquema que se basa en las definiciones dadas por la “Society for Computer Simulation” (SCS) en 1979 y que identifica dos tipos de modelos: el modelo conceptual y el modelo computacional.

El modelo conceptual está compuesto por datos y ecuaciones matemáticas que describen de una manera simplificada un sistema o proceso físico de interés. Este modelo se produce mediante el análisis y la observación de los sistemas físicos, hasta lograr una representación que incluya las variables de interés, la interacción entre éstas y las condiciones de borde.

El modelo computacional hace referencia a los algoritmos y programas que implementan lo propuesto en el modelo conceptual. El modelo computacional debe ser capaz de encontrar un algoritmo apropiado que modele eficientemente las necesidades presentadas por el modelo conceptual. Este modelo debe ser un diseño eficiente, fácil de verificar y validar y que sea escalable. Se agrega así una capa más de abstracción. Estas tres capas: realidad, modelo conceptual y modelo computacional deben poder retroalimentarse continuamente (Oberkampf et. al., 2010).

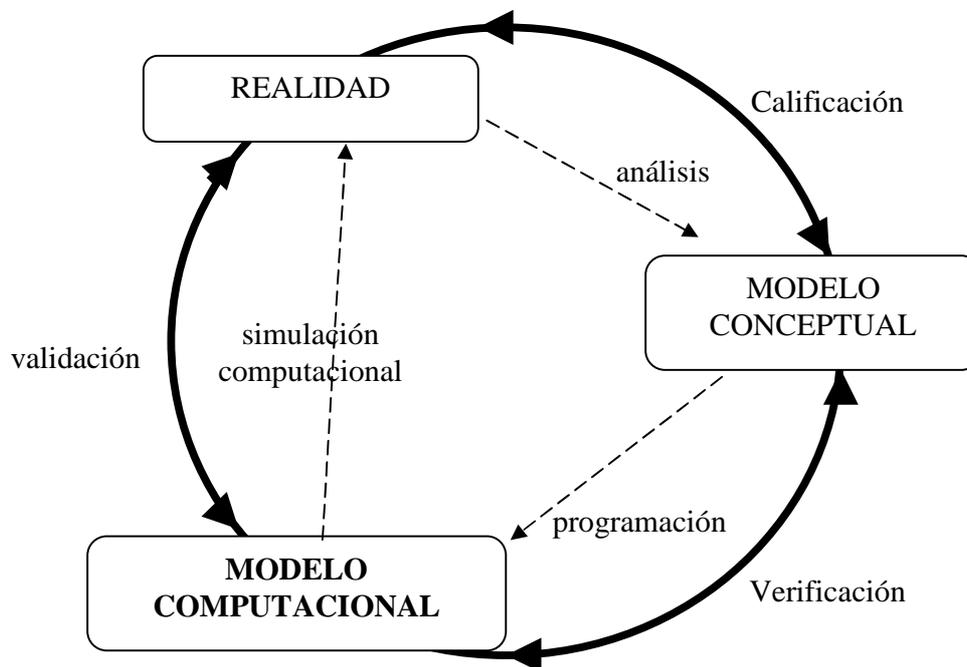


Figura 1.1 Esquema de los modelos conceptual y computacional, sus interrelaciones y su conexión con el problema real.

La “traducción” de un sistema de la realidad a un modelo conceptual necesita de un proceso constante de calificación. Se puede definir la calificación de un sistema como el proceso de determinación de que tan adecuado es un modelo de manera que provea un nivel aceptable de concordancia para el dominio de una aplicación dada. Esto significa que el proceso de “calificar” al sistema real debe tener en cuenta cómo este se describe mejor en términos de los efectos que provoca en su ambiente, el tipo de proceso físico que se trate, cuáles son las variables más relevantes. Esto se logra mediante el “análisis” del sistema y sus interacciones (Oberkampf et. al., 2010).

Una vez determinado un modelo conceptual, es necesario poder modelarlo computacionalmente. Para ello es necesario diseñar algoritmos que se ajusten al modelo conceptual, que tengan en cuenta las capacidades y limitaciones del hardware donde serán implementados, y que permita predecir el comportamiento en tiempo y uso de recursos cuando el problema original escale. A partir de este diseño, inicia la etapa de

programar el modelo en algún lenguaje de programación adecuado. La verificación constante de que el modelo computacional se acerque al modelo conceptual es necesaria. Además el proceso de verificación también funciona en el sentido inverso. Es muy común que durante la determinación del modelo computacional surjan limitaciones o incluso nuevas relaciones que no estaban planteadas inicialmente y que sirven para ajustar el modelo conceptual.

A partir del modelo computacional obtenido, se ejecutan las simulaciones que serán comparadas con la realidad. Durante un proceso continuo de validación el modelo computacional es puesto a prueba y de la misma manera la información sobre la realidad permite conseguir un modelo computacional más adecuado.

1.3. Desarrollo de software científico

La manera de obtener un modelo computacional para un determinado modelo conceptual es mediante el desarrollo de software científico.

La manera convencional en el proceso de desarrollo de software está pensada fundamentalmente para implementar programas de propósitos generales y van de la mano de las técnicas y estrategias de la ingeniería de software.

Según la (IEEE93a) la ingeniería de software es: La aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo, operación y mantenimiento de software; es decir, la aplicación de la ingeniería al software (Pressman, 2010).

El proceso de desarrollo de software se inicia con la definición de los requisitos del sistema, que es provisto por el cliente (o persona interesada). Estos requisitos comprenden características, propiedades, condiciones de borde, etc.

La estrategia típica es la de diseñar un enfoque global del sistema (diseño de algoritmos, decisión sobre el paradigma a usar, etc.) y tomar una decisión respecto a cómo se realizará la implementación (programación). A continuación se procede a traducir los requerimientos en sub tareas individuales y mediante el uso de metodologías para la gestión de proyectos se puede establecer una planificación para la implementación, validación y entrega. Luego de completar estas etapas, el software está en condiciones de

ser entregado y el proyecto finaliza. En términos generales este proceso garantiza la calidad del software que se entrega.

En principio se podría pensar que el desarrollo de software científico puede llevarse a cabo siguiendo las mismas etapas que en el enfoque convencional. Si bien las características fundamentales para un software de calidad son las mismas, el software científico tiene algunas particularidades. La primera gran diferencia es que generalmente los requerimientos no están bien definidos. Esto ocurre porque en la investigación científica no es posible identificar, en la mayoría de los casos, cuales serían las salidas esperadas. Típicamente los requerimientos van variando a medida que se profundiza la investigación. La validación del sistema no siempre es factible o fácil de realizar. Esto se debe a que muchas veces los cálculos alcanzan las limitaciones naturales de la representación de número de punto flotante en las computadoras. Otra razón está relacionada a la naturaleza caótica de las soluciones. También ocurre que una implementación determinada solo sea necesaria una única vez, ya que por ejemplo las condiciones del sistema cambian. Un gran problema y que comúnmente es subestimado es que, muy pocos científicos son programadores, esto hace que el desarrollo se torne dificultoso y que resulte en un sistema monolítico, difícil de depurar, de entender y de escalar. Esto va de la mano con que es común que la implementación sea realizada por estudiantes sin mucha experiencia en el ámbito científico y mucho menos experiencia en programación. Suponiendo que se cuente con un equipo multidisciplinario que cuente con programadores experimentados, aún queda el problema de que es muy complejo evaluar la correctitud de los resultados. Es muy probable que muchas de las simulaciones que se realizan, se hagan debido a que no es factible realizar un experimento determinado. Por ejemplo sólo se pueden realizar medidas indirectas, el tiempo que llevaría el experimento lo hace prohibitivo, es difícil construir un equipo solo para probar una teoría, etc.

Queda claro que escribir código científico para un determinado modelo computacional es una tarea compleja. Sin embargo, es una tarea ineludible en la mayoría de las ciencias.

En este punto es importante recordar algunos puntos sobresalientes del enfoque del desarrollo de software tradicional y que sirven para aportar calidad al software científico.

Es necesario:

- Escribir código reusable y modular

- Aprovechar las capacidades del hardware disponible (implica tenerlo en cuenta a la hora de programar).
- Aprovechar el software científico ya probado como el caso de librerías altamente optimizadas.
- Analizar los lenguajes de programación (usar lenguajes compilados, interpretados o una combinación de estos).
- Escribir software que se encuentre constantemente en prueba y re validación.
- Utilizar y escribir código bien documentado, aprovechar y desarrollar software colaborativo.

En síntesis el desarrollo de software científico para implementar un modelo computacional debe estar organizado en tareas que requiere de especialistas entrenados, dedicación en tiempo y recursos para producir resultados confiables.

1.3.1. Algoritmos

Una vez que el modelo conceptual fue determinado por un dado problema, es necesario encontrar una formulación para el modelo computacional acorde. El objetivo es encontrar una solución en forma de algoritmo. Un algoritmo es una secuencia de instrucciones, cada una de las cuales tiene un significado claro, y que puede ser realizado en una cantidad finita de esfuerzo para una cantidad finita de tiempo (Aho, 2001).

Estas instrucciones deben tener un propósito y donde, incluso cuando una tarea se deba repetir, la cantidad de tiempo que las instrucciones insumen debe ser finito.

Comúnmente los algoritmos se presentan usando pseudo-lenguaje que no es más que una construcción sintáctica similar a un lenguaje de programación para describir las sentencias del algoritmo.

La primera versión de un algoritmo generalmente es una secuencia de pasos simples que necesitan posteriormente ser refinados en pasos más específicos e instrucciones más definidas. Este ejercicio se repite hasta conseguir un algoritmo en pseudo-lenguaje lo suficientemente detallado como para que pueda ser traducido a cualquier lenguaje de programación. El diseño algorítmico es el paso previo a la traducción (implementación) mediante algún lenguaje de programación. Este diseño es clave para determinar la lógica

del programa. Si el algoritmo es el adecuado y se puede demostrar que es eficiente, entonces la etapa de programación solo agregará detalles específicos para permitir que sea ejecutado en una determinada computadora. Se puede decir que en el diseño del algoritmo se encuentra representado el modelo computacional.

No solo se diseña la lógica que tendrá el futuro programa, sino que es necesario definir las estructuras de datos que contengan la información. Estas estructuras de datos no son simples contenedores sino que además establecen el comportamiento o las operaciones que pueden tener los datos que en ellas se almacenen. Así, determinar que un dato es un vector de enteros indica que se trata de una secuencia de números enteros a los que se puede acceder mediante un índice, establece la relación entre sus elementos (son secuenciales en este caso), y determina que tipo de operaciones pueden realizarse (por ejemplo del álgebra lineal). Existen tipos de datos más complejos como aquellos asociativos que se pueden pensar como un par (key, value), donde el valor (value) es accedido mediante una clave (key). Se pueden incluso definir nuevos tipos de datos abstractos que se basan en tipos de datos simples e identifican complejas estructuras (árboles, grafos, pilas, listas, etc).

Definir la lógica para un determinado algoritmo requiere del conocimiento de las estructuras de programación básicas y además de diferentes técnicas de diseño. Existen un gran número de estas técnicas tales como, estrategias divide & conquer, programación dinámica, técnicas greedy, backtracking, búsqueda local, etc.

Diseñar algoritmos no es fácil, requiere de entrenamiento y estudio de estas técnicas, y de conocimiento de estructuras de datos abstractas. Una vez escrito el algoritmo, e implementado en algún lenguaje, aún hay que depurar el programa (determinar que errores se presentan) y verificar que el programa se ajuste a los requisitos (en este caso al modelo conceptual).

Además se deben cumplir dos objetivos:

- a. Es necesario que el algoritmo sea fácil de entender, de codificar y de depurar.
- b. Es necesario que el algoritmo haga un uso eficiente de los recursos de computación, especialmente que éste se ejecute de la manera más rápida posible

A primera vista estos dos objetivos no están relacionados. Sin embargo el costo de cumplir el objetivo a) excede por mucho el costo de ejecutar un programa. Si el programa

será ejecutado pocas veces entonces es conveniente poner mucho más énfasis en lograr cumplimentar el objetivo a). Pero, si el programa debe ejecutarse múltiples veces, entonces es conveniente invertir en optimizar el tiempo de ejecución.

En computación científica, típicamente una simulación puede ejecutarse pocas veces pero la cantidad de cálculos que realiza hace que el tiempo de ejecución escale mucho con el aumento de los valores de entrada. Esto significa que es importante sobremedida el tiempo de ejecución (en algunos casos hay simulaciones que duran días). Esto hace que además los algoritmos se vuelvan más dependientes de la arquitectura. Para acelerar las aplicaciones, el primer paso y más importante es re pensar el algoritmo y si aún así la velocidad no mejora recién invertir en nueva tecnología. Lamentablemente no suele ser el caso más usual lo que lleva en muchas ocasiones a realizar gastos innecesarios.

Muchas veces una solución algorítmica utilizando las técnicas mencionadas anteriormente resulta en algoritmos más complejos pero con eficiencia considerablemente mayor al momento de implementarlos. Un ejemplo interesante es el algoritmo que resuelve la transformada de Fourier. Es conocido en el ámbito científico que la manera de transformar el dominio del tiempo al dominio de la frecuencia (en tiempo discreto) es utilizando la transformada discreta de Fourier. Este problema se resuelve realizando una sumatoria de productos por cada elemento de la serie temporal discreta. Algorítmicamente no es un trabajo complejo y sólo involucra realizar sumas y multiplicaciones. Este primer enfoque tuvo muy buenos resultados pero a medida que el tamaño de las muestras aumentaba, la cantidad de cálculos hacía prohibitiva su implementación. No fue sino hasta que J. W. Cooley y J. W. Tukey (1965), re diseñaran el algoritmo que realmente pudo ser aplicado eficientemente. La nueva versión del algoritmo utiliza la técnica de Divide & Conquer aprovechando algunas propiedades de los coeficientes que utiliza la serie. Este nuevo algoritmo se denomina transformada rápida de Fourier o *Fast Fourier Transform* (FFT) y actualmente es muy popular.

Para dimensionar las ventajas de este nuevo algoritmos podemos hacer el siguiente ejercicio: Supongamos que cada operación que se realiza demora 1 ns, y si tengo $N=10^9$ muestras, el algoritmo tradicional de transformada de Fourier realiza 10^{18} operaciones mientras que el algoritmo FFT realiza 30×10^9 operaciones algebraicas. En términos de tiempo esto significa que para ese tamaño de muestras o entrada el primer algoritmo

demora alrededor de 31.2 años mientras que el algoritmo FFT demora 30 seg. Esta diferencia abismal hace que continuamente los investigadores en el área de las Ciencias de la Computación investiguen nuevas técnicas y re planteen otros algoritmos para encontrar una solución más eficiente y poder aprovechar correctamente los recursos computacionales disponibles.

Dado que no existe una única solución algorítmica para un problema, surge una pregunta natural. ¿Cómo sabemos que un determinado algoritmo es el adecuado? Más importante aún, ¿Podemos estimar cómo será la eficiencia en tiempo de ejecución y en uso de recursos, aún cuando el tamaño del programa escale? Si podemos responder a estas preguntas podemos elegir el mejor algoritmo que resuelva el problema planteado.

Podemos decir que el tiempo de ejecución de un programa depende de factores como

1. La entrada al problema (tipo, tamaño)
2. La calidad del código generado por el compilador que se utilizó
3. La naturaleza y velocidad de las instrucciones de la computadora que se utiliza para ejecutar el programa
4. De la complejidad de tiempo del algoritmo que se programó (Aho., 2001).

A priori es imposible determinar factores como 2) y 3) dado que dependen de la arquitectura, lenguaje, compilador, etc particulares que se utilice. Se puede hacer un estudio de la entrada, cómo ésta crece con el problema. A partir de esto se puede estimar cómo es el comportamiento del algoritmo en la medida que varía el tamaño de la entrada (por ejemplo cantidad de muestras que se utilizan). Este estudio sobre la eficiencia algorítmica es independiente del computador donde se programe y está asociado a la lógica con la que éste se encuentra estructurado. Esto permite hacer un estudio de eficiencia a priori y además permite comparar diferentes soluciones algorítmicas.

Para hablar de la razón de crecimiento del tiempo respecto al tamaño de la entrada al algoritmo, se define la notación O-grande. Esta notación analiza el peor caso en tiempo de ejecución de un dado algoritmo.

Se dice que el tiempo de ejecución de un programa $T(n)$ pertenece a $O(f(n))$ si existen las constantes positivas c y n_0 , tal que para todo $n \geq n_0$ el tiempo de ejecución se encuentre

acotado como $T(n) \leq c f(n)$ (Aho, 2001). Se dice que el tiempo de ejecución crece a razón de $f(n)$, donde n es el tamaño de la entrada.

Supongamos un problema que tenga un $T(n) \in O(1)$, esto indica que para cualquier valor de la entrada la cantidad de operaciones que se realiza siempre es constante. Si $T(n) \in O(n^2)$, esto indica que en el peor de los casos el tiempo de ejecución crece cuadráticamente con el tamaño de la entrada. La Figura 1.2 muestra comparativamente diferentes funciones O-grande, se puede observar que para valores pequeños de la entrada (n) la mayoría tiene un tiempo de ejecución (cantidad de operaciones) relativamente aceptables. Es notable, también, que los problemas en cuanto al tiempo de ejecución se presentan cuando el tamaño de la entrada escala.

Este es el problema de la transformada discreta de Fourier tradicional que tiene un tiempo de ejecución $T(n) \in O(n^2)$ mientras que el tiempo de ejecución de la FFT es de $T(n) \in O(n \log_2 n)$. Para n muestras, con n muy grande y potencia de 2, el algoritmo FFT es muy rápido y por ello representó una gran revolución en el área del procesamiento digital de señales (Tomasi, 2003).

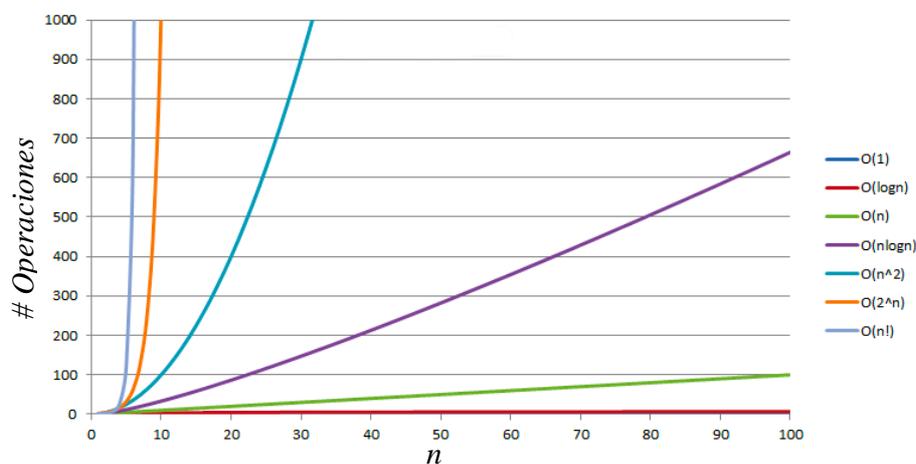


Figura 1.2. Comparación entre cantidad de operaciones y tamaño de la entrada según la función del tiempo de ejecución.

En síntesis, cuando se diseña un algoritmo, se desea que éste tenga un tiempo de ejecución que sea aceptable a pesar que la cantidad de datos de entrada crezca. Este es un objetivo de diseño que debe estar presente al momento de desarrollar el modelo computacional

1.3.2. Algoritmos Numéricos

El análisis algorítmico al momento de determinar el modelo computacional cobra mayor relevancia si se trata de algoritmos numéricos. Resolver complejas aproximaciones numéricas para ecuaciones de evolución, descomposición de matrices u otras funciones del algebra lineal, o métodos estadísticos, etc requieren de una gran cantidad de operaciones. En muchas ocasiones además existe el requisito de que el programa debe resolver estos cálculos en tiempo real y no basta con que las acciones del sistema sean correctas. La respuesta del sistema debe determinarse en un tiempo muy acotado siendo así crucial un buen diseño algorítmico.

Estos sistemas usualmente deben correr simulaciones donde las condiciones de borde no siempre están claras, que son difíciles de depurar y que insumen una inmensa cantidad de recursos de cómputos.

1.4. Algoritmos para Procesamiento de Señales

Una de las características subyacentes en el diseño de sistemas que deben realizar procesamiento de señales es el nivel significativo de pruebas y experimentación que normalmente se requiere antes de llegar a una solución aceptable. Esta característica implica que es necesario tener la habilidad para formular enfoques que rápidamente constituyan prototipos candidatos para la solución general. El objetivo principal es el de reducir tanto los costos como el tiempo para la implementación del sistema final.

Con el advenimiento de las nuevas tecnologías para cómputo, los tradicionales sistemas para procesar señales analógicas comenzaron a ser reemplazados por nuevos sistemas de procesamiento digital.

El procesamiento digital de señales se aplica a un amplio número de campos: comunicaciones, procesamiento de imágenes médicas, tecnologías radar, reproducción de música de muy alta fidelidad, prospección petrolera, entre muchas otra aplicaciones. Se basa en una gran cantidad de herramientas matemáticas y algoritmos específicos (Smith, 1999).

En el caso particular de las telecomunicaciones el procesamiento digital de señales es particularmente relevante, participando en casi todas las etapas del desarrollo de estos sistemas.

El objetivo fundamental de un sistema electrónico de comunicaciones, es transferir información de un lugar a otro. Por consiguiente, se puede decir que las comunicaciones electrónicas son la transmisión, recepción y procesamiento de información entre dos o más lugares, mediante circuitos electrónicos. La fuente original de información puede estar en forma analógica (continua), como por ejemplo la voz humana o la música, o en forma digital (discreta), como por ejemplo los números codificados binariamente o los códigos alfanuméricos. Sin embargo, todas las formas de información se deben convertir a energía electromagnética antes de ser propagadas a través de un sistema electrónico de comunicaciones (Tomasi, 1996).

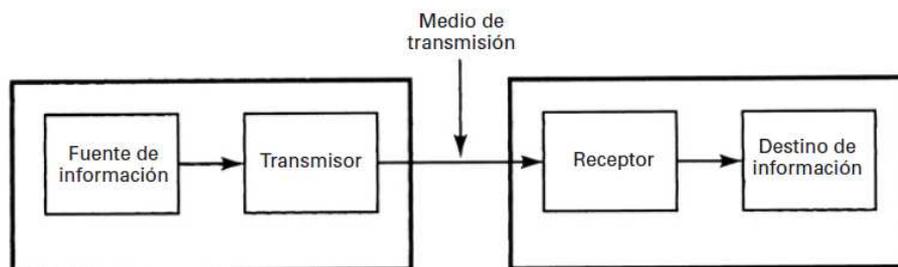


Figura 1.3. Diagrama simplificado de bloques de un sistema de comunicaciones electrónicas

En la Figura 1.3 se muestra un diagrama simplificado de bloques de un sistema de comunicaciones electrónicas. Consta básicamente de un transmisor y un receptor que se comunican vía algún medio de transmisión. Diferentes funcionalidades y características de los sistemas de comunicación como la generación de una señal, filtrado, transmisión, recepción, codificación y decodificación (en ciertos casos), transformación entre dominios de una señal (FFT), etc; requieren del conocimiento de herramientas matemáticas y técnicas de procesamiento digital.

1.4.1. Detección automática de señales de radar como una aplicación computacional

En particular, un radar es un sistema de telecomunicación electrónica que permite enviar señales que serán reflejadas en algún objeto de interés denominado blanco. La función principal de este equipo es obtener información adecuada, en forma de ecos que deberán ser procesados para extraer información del blanco. Esta función se conoce como detección y comúnmente es complementada con información del tamaño, posición y velocidad del blanco.

Las tendencias actuales en el diseño de radares es la de simplificar el hardware del sistema poniendo el foco del procesamiento de las señales en sistemas de software. Esto se debe principalmente al bajo costo de las computadoras de propósitos generales que tienen capacidad de procesamiento suficiente para realizar cómputo intensivo y en algunos casos dar respuesta cuasi en tiempo real.

Un sistema con detección automática de blancos debe recurrir a técnicas y algoritmos del procesamiento digital de las señales involucradas para alcanzar su objetivo.

Esto significa que es necesario encontrar un modelo computacional específico que permita implementar la funcionalidad automática en un determinado radar.

CAPITULO II

Radar

2.1. Introducción

La palabra “radar” proviene del acrónimo RADAR (Radio Detection and Ranging) que puede traducirse como “Detección y medición de distancia por radio”.

Skolnik (2008) y puede definirse como un sensor electromagnético para la detección y localización de objetos reflectantes”.

Mientras que según Barton y Leonov (1998), “un radar es un dispositivo que transmite señales electromagnéticas y recibe ecos provenientes de objetos de interés (objetivo) dentro de su volumen de cobertura. La presencia del objetivo es revelada por la detección de su eco. La información adicional sobre el objetivo que provee el radar puede incluir distancia (rango), tiempo de retardo entre la transmisión y la recepción de la señal, dirección mediante patrones de antenas directivas, velocidad radial o angular del blanco mediante técnicas Doppler, y descripción o clasificación del blanco mediante el análisis del eco.

Según Melvin y Scheer (2013), los modernos sistemas de radar son sistemas muy complejos que aprovechan los últimos avances de la tecnología recurriendo a sofisticados algoritmos y técnicas de procesamiento. Estos radares se diseñan para detectar, localizar, caracterizar, y en algunos casos, rastrear y seguir blancos de interés. Las aplicaciones de radares son muy diversas. Por ejemplo, los radares son usados en aeronaves, misiles, satélites, navíos, vehículos terrestres, para aplicaciones geofísicas, meteorología, entre otras aplicaciones. Y sus funciones pueden incluir detectar blancos móviles o no, realizar el seguimiento por ejemplo de aeronaves, misiles, personas, u otros objetos. Con tan amplia variedad de plataformas de radares y blancos, el proceso de clasificación un radar específico muchas veces se transforma en una tarea desalentadora

La historia del radar tiene su comienzo desde el desarrollo de la moderna teoría electromagnética. En 1886, Hertz demostró la reflexión de ondas de radio y en 1900 Tesla describió el concepto de detección electromagnética y medición de velocidad durante una entrevista. En 1903 y 1904, el ingeniero alemán Hülsmeier experimentó con la detección de una nave mediante reflexión de ondas de radio, una idea impulsada por Marconi en 1922. En el mismo año, Taylor y Young del Laboratorio de Investigación Naval de Estados Unidos (NRL), detectó una aeronave (presuntamente accidentalmente) por radar, logrando una investigación más profunda en el tema lo que llevo a una patente en 1934 del denominado radar de onda continua (CW por sus siglas en Inglés).

El desarrollo y divulgación de la tecnología radar se vio acelerada alrededor de fines de los años 30s con grandes desarrollos independientes por parte de Estados Unidos, Gran Bretaña, Francia, Alemania, Rusia, Italia y Japón. Este crecimiento se vio acentuado con el advenimiento de la II Guerra Mundial teniendo principalmente aplicaciones bélicas (Barton y Leonov, 1998; Richards, 2005; Skolnik, 1980; entre otros).

Aún hoy el desarrollo de nuevos equipos y técnicas de radares tienen mayormente aplicaciones militares que incluyen tareas de vigilancia, navegación y guía de proyectiles por tierra, mar y en vehículos aéreos.

Sin embargo, los sistemas de radares son utilizados cada vez más para una gran variedad de aplicaciones. Entre las más difundidas se encuentran: el control del límite de velocidad de vehículos, de adquisición de información meteorológica para predicción del clima, control de tráfico aéreo y de naves, e incluso existen diversos equipos radares de aplicaciones geofísicas (sensado de la atmósfera terrestre, para mediciones de profundidad de glaciares, aplicaciones para la prospección geográfica y espacial, etc).

Existen diversas formas de clasificar los sistemas radar. En contraposición a los radares de onda continua, existen los radares pulsados. Estos últimos emiten pulsos electromagnéticos de una determinada duración. Actualmente la tecnología de radares pulsados tiene preponderancia sobre los radares de onda continua. Los radares de onda continua se utilizan generalmente para sistemas de altimetría de naves aéreas y prospección territorial desde satélites, los pulsados en control y vigilancia. Este trabajo de tesis se centrará en el estudio de radares pulsados.

Otra categorización que se suele hacer es respecto a las antenas que se utilizan en transmisión y recepción. Un radar puede poseer un dispositivo o antena transmisora y otra antena receptora, que se encuentran separadas a una distancia significativa comparable con el blanco previsto, estos equipos se denominan radares biestáticos. Si por el contrario a transmisión y recepción se realiza en una misma antena se dice que el radar es monoestático.

En este capítulo se estudiarán las funciones y características fundamentales de los equipos de radar. Se analizarán algunos tipos de radares, las señales que transmiten y reciben, los diferentes modelos de blancos y cómo éstos deben ser estudiados al momento de realizar la detección.

2.2. Funciones básicas de un radar

En la mayoría de los casos las funciones requeridas por un sistema radar pueden clasificarse en “detección”, “rastreo” (Tracking) y “visualización” (Imaging) (Richards, 2005).

La detección es la principal funcionalidad y refiere a la capacidad del equipo de discriminar si la señal de radio reflejada (el eco) por un determinado objeto corresponde a un blanco. Cada eco recepcionado, en un primer orden, puede corresponder a uno de los siguientes casos:

- a. El eco corresponde a una reflexión en el objeto de interés o blanco, ó
- b. Corresponde a energía dispersada por otro objeto que circunda al blanco, normalmente llamado clutter; ó
- c. Se recibe una señal de interferencia (Jamming) proveniente de otros sistemas que generan ondas de radio; ó
- d. La energía recibida es originada por ruido ambiente o producido por el hombre.

Lo más probable en un ambiente real es obtener un eco producto de la combinación de algunos de estos casos o incluso de todos ellos.

Realizar la tarea de detección significa discriminar a cuál de las situaciones mencionadas (a,b,c, ó d) corresponde una determinada señal recepcionada en el radar.

Una vez detectado el objetivo o blanco, es preciso obtener información asociada a este. Para ello, el radar debe realizar operaciones adicionales para realizar mediciones de diferentes parámetros del blanco como por ejemplo distancia o rango, velocidad del blanco, sección transversal del objetivo, entre otros.

Adicionalmente puede ser necesario contar con la funcionalidad denominada “reconocimiento e identificación” que consiste en poder clasificar los diferentes tipos de blancos que están siendo detectados e incluso discriminar aquellos que sean de mayor interés (Barton y Leonov, 1998). Finalmente, muchos de los sistemas de radares, deben ser capaces de rastrear el objetivo, analizando su posicionamiento en diferentes estados de tiempo. Un ejemplo es el de los radares para control de tráfico aéreo

En cualquier caso, el principal parámetro a estimar es el rango que significa establecer a qué distancia se encuentra un objeto del sistema.

Supongamos que la energía recepcionada en la antena del radar corresponde solamente a un blanco. La presencia del mismo es percibida mediante la detección del eco que se recibe como respuesta a la reflexión de la señal emitida sobre dicho blanco. Bajo la hipótesis de que la onda de radio en el vacío se desplaza a la velocidad de la luz, la distancia a la que se encuentra el blanco respecto al equipo se denomina rango (R) y se calcula midiendo el tiempo que transcurre desde la transmisión de la señal hasta su recepción. El rango del radar se puede calcular como:

$$R = \frac{ct}{2} \tag{2.1}$$

Donde R es el rango, t es el intervalo de tiempo desde que se transmite la señal hasta que se la recibe y c es la velocidad de la luz en el vacío. El cálculo se divide en dos debido a que el tiempo de viaje de la señal incluye el tiempo desde el radar hacia el objeto y el tiempo de retorno al radar.

Dado que la velocidad de propagación de una onda electromagnética en la atmósfera es próxima a la luz en el vacío, por lo tanto, el valor $c = 3 \times 10^8 \text{ m/s}$ es una buena aproximación.

Tanto el transmisor como el generador de ondas del equipo juegan un rol preponderante en la determinación de la sensibilidad y resolución de rango del radar, Los sistemas de radar operan a frecuencias que varían entre los 2 MHz y los 220 GHz (Skolnik, 2001). Los radares lasers operan a frecuencias del orden de 10¹² a 10¹⁵ Hz, que corresponde a longitudes de onda de 0.3 a 30 mm (Jelalian, 1992). Sin embargo, la mayoría de los radares operan en regiones de frecuencias de microondas (~200 MHz) hasta cerca de 95 GHz y que corresponde a longitudes de onda de entre 0.67 a 3.16 mm. La Tabla 2.1 resume la nomenclatura usada para las diferentes bandas utilizadas por sistemas de radar (IEEE, 1976).

Banda	Frecuencias	Longitud de onda
HF	3 - 30 MHz	100 – 10 m
VHF	30 - 300 MHz	10 – 1 m
UHF	300 MHz - 1 GHz	1 – 30 cm
L	1 - 2 GHz	30 – 15 cm
S	2 - 4 GHz	15 – 7.5 cm
C	4 - 8 GHz	7.5 – 3.75 cm
X	8 – 12 GHz	3.75 – 2.5 cm
K _u	12 – 18 GHz	2.5 – 1.67 cm
K	18 – 27 GHz	1.67 – 1.11 cm
K _a	27 – 40 GHz	1.11 cm – 7.5 mm
mm	40 – 300 GHz	7.5 – 1 mm

Tabla 2.1. Nomenclatura para bandas de frecuencia nominal para radares

Radares de bajas frecuencias son los más utilizados para aplicaciones de vigilancia debido a la baja atenuación atmosférica para capacidades de altas potencias. En cuanto a las frecuencias más altas, éstas se prefieren para diseñar radares con resolución más fina, mayor atenuación y menos capacidad de potencia. Las condiciones climáticas también tienen efectos en la propagación de las ondas de los radares. Por ejemplo, para frecuencias en banda X y menores las ondas son afectadas significativamente por lluvias severas (Richards, 2005).

En líneas generales un radar consta de un transmisor de ondas, de un receptor y de un sistema de sincronización.

En la Figura 2.1 se muestra un esquema de cómo opera un radar. La sincronización entre el transmisor y el receptor consiste en mantener una referencia de fase estable en todo el proceso, desde el inicio de la transmisión hasta la recepción del eco.

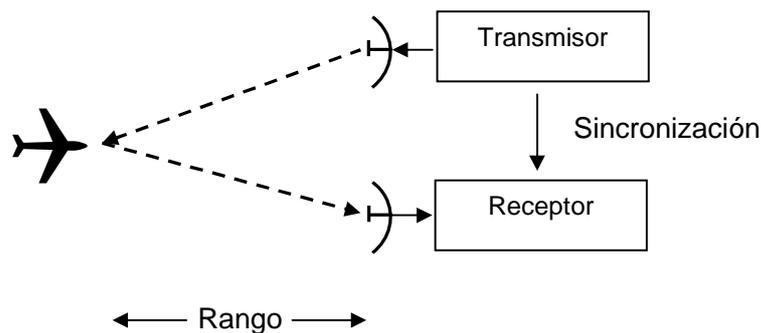


Figura 2.1. Esquema del funcionamiento básico de un radar

La detección y rastreo de blancos por parte de un equipo radar son las funciones más conocidas. Sin embargo, actualmente los radares están siendo utilizados también para generar imágenes bidimensionales de algún área. En general estas imágenes son utilizadas para vigilancia, análisis de topología geográfica, análisis de terreno, etc. La importancia de estas imágenes radica en la capacidad de los radares de “ver” a través de nubes, precipitaciones y otros eventos meteorológicos cuando se utilizan frecuencias de microondas. (Richards, 2005).

Un ejemplo interesante es el equipo instalado en el satélite SAC-D/ Aquarius que fue puesto en órbita el 10 de junio de 2011 como un proyecto de colaboración entre CONAE (Comisión Nacional de Actividades Espaciales) de Argentina, y NASA. Un ejemplo de los primeros resultados obtenidos en esta misión se muestra en la Figura 2.2 (Le Vine, D. et al, 2012) que muestra el primer mapa global de la salinidad en superficie del mar. Este mapa es el resultado de numerosos instrumentos instalados en el satélite entre los que se cuenta una *sonda de radio ocultación* para la atmósfera (instrumento ROSA) y que

aprovecha la tecnología radar de onda continua para obtener datos relacionados con perfiles atmosféricos de temperatura, humedad y presión.

Todas estas funciones dependen fundamentalmente de la calidad del proceso de detección, lo cual repercute en la calidad del sistema radar. Analizar el desempeño de la detección se transforma en un trabajo prioritario y conlleva a la utilización de diversas técnicas dependiendo en gran medida del tipo de blanco que sea objeto de estudio. Esto significa que es preciso concentrarse en los aspectos del procesamiento de señales para intensificar la relación entre las señales provenientes de ecos y las provenientes de interferencias o ruido. En término del procesamiento de señales esto significa incrementar la relación señal a ruido o Signal to Noise Ratio o SNR (por sus siglas en inglés). Para esto, el conocimiento de las características particulares del blanco, son un objeto de estudio indispensable. Mediante herramientas matemáticas tales como la correlación, la integración coherente de pulsos en algunos casos, o la codificación de la señal portadora, es posible mejorar la detección.

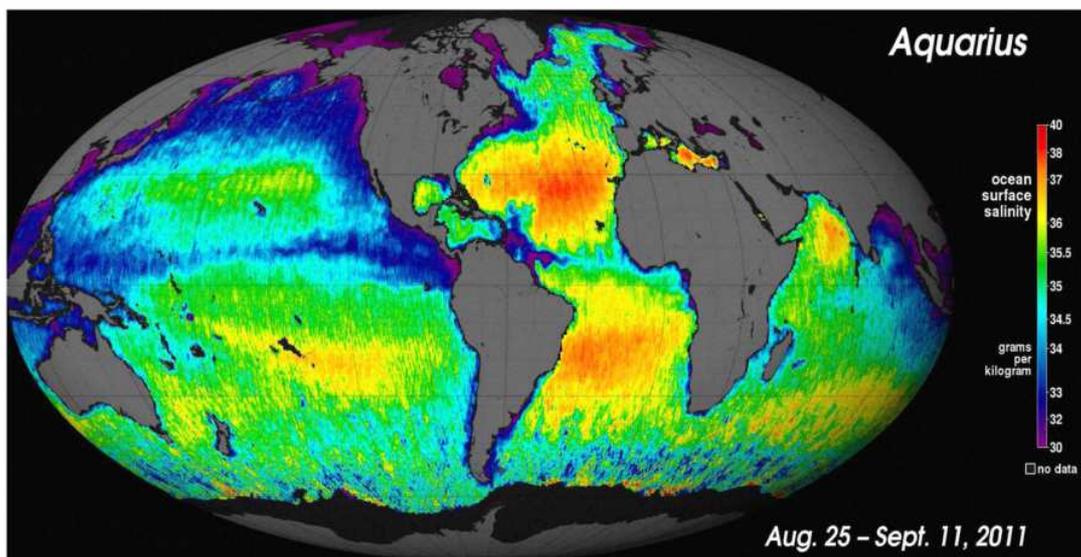


Figura 2.2. Mapa global de la salinidad en la superficie marina generado por datos obtenidos del satélite SAC-D/Aquarius. (Le Vine, D. et al, 2012)

2.3. Señales de radar

Un blanco u objetivo es un objeto de interés para el radar, en líneas generales se puede tratar de cualquier objeto que dispersa energía y parte de esta retorna como eco. Parte de la energía contenida en el eco proviene del área que circunda al objetivo, a esta se la denomina señal del clutter. Así el eco recepcionado se encuentra conformado por la señal de interés a la cual se encuentran adicionadas señales de ruido. Denominadas en líneas generales así por no ser de interés para la detección del blanco. Aunque en términos más exactos puede estar conformada por ruido, clutter o Jamming.

Establecer criterios para el diseño de un detector en un radar significa establecer cuál es el tipo de blanco que se espera detectar.

Dependiendo del uso que se le dará al radar un mismo eco puede representar un blanco o ruido. Por ejemplo, un eco puede corresponder a un objeto móvil como es el caso de una aeronave o a nubes, lluvia u otros eventos meteorológicos. De esta manera para algún tipo radar la aeronave constituye un blanco mientras las precipitaciones o nubes constituyen clutter. En otros casos, por ejemplo en un radar meteorológico, lo que se desea detectar son los eventos como la lluvia, nieve, o vientos siendo en estos casos una montaña el clutter de base.

Los blancos pueden clasificarse como simples o complejos. A su vez estos últimos como:

- a. Discretos, la señal en el receptor es resultante de la reflexión en un único punto como es el caso de los blancos tradicionales como aeronaves, navíos, etc;
- b. Distribuidos, ocurre cuando el eco es producto de la dispersión distribuida en una superficie como el caso de una porción de clutter de tierra o incluso distribuida en un volumen como es el caso de la lluvia. (Barton y Leonov, 1998).

Uno de los mayores desafíos es la de poder caracterizar o modelar los blancos u objetivos. En particular, se debe estudiar y en modelar las señales recepcionadas resultante de la dispersión de la señal transmitida en el objetivo.

Supongamos que $x(t)$ es la señal transmitida, tal que:

$$x(t) = a(t) \sin(2\pi f_i t + \theta(t)) \quad (2.2)$$

Donde $a(t)$ representa la amplitud de la señal, f_i es la frecuencia de la portadora y $\theta(t)$ puede ser constante o representar una modulación en fase. Usualmente se asume que $a(t)$ es, idealmente, la envolvente de un pulso cuadrado de amplitud A y de duración τ .

La potencia instantánea de esta señal es $P_s = A^2/2$.

La señal en el receptor será una combinación de ecos de $x(t)$ producto de reflexiones en el blanco, de clutter, ruido y posiblemente jamming. Tanto los blancos como el clutter producen señales atenuadas y retardadas del pulso transmitido. Además de tener adicionado ruido o tener una señal modulada de interferencia.

Por lo tanto una señal recepcionada de un único pulso proveniente de un objeto dispersivo a un rango $R = ct_0/2$ puede escribirse como:

$$y(t) = b(t - t_0) e^{j[2\pi f_i(t-t_0) + \theta(t)]} + n(t) \quad (2.3)$$

Donde

$n(t)$ = ruido

$b(t)$ = amplitud del eco

$\theta(t)$ = modulación de fase del eco

Los parámetros mas importantes de esta ecuación son el tiempo de retardo (t_0), la amplitud del eco $b(t)$ y la componente relativa al ruido $n(t)$. (Richards, 2005).

Es importante destacar la importancia del tipo de radar que se está diseñando y del tipo de blanco que se trate para establecer la frecuencia “ f ” de la portadora utilizada. Usualmente van desde los kHz hasta los GHz. Así la longitud de onda de exploración se corresponderá con la naturaleza del objetivo a detectar.

Las funciones de modulación de amplitud y fase también determinan la resolución de rango ΔR de una medición. La resolución de rango del radar está relacionada con la capacidad de identificar dos respuestas consecutivas de un mismo blanco.

Si dos dispersiones de igual energía producen dos señales identificables separadas como salida del sistema se consideran “resueltas” en contraposición a encontrarse combinadas en una única salida.

El concepto de resolución también se aplica a otras dimensiones. Por ejemplo, en el caso de radares que miden rango y velocidad. Ambas dimensiones son independientes, e incluso pueden estar resueltas en una de ellas (por ejemplo el rango) y no estarlo en la otra (velocidad). Este es el caso por ejemplo de radares Doppler.

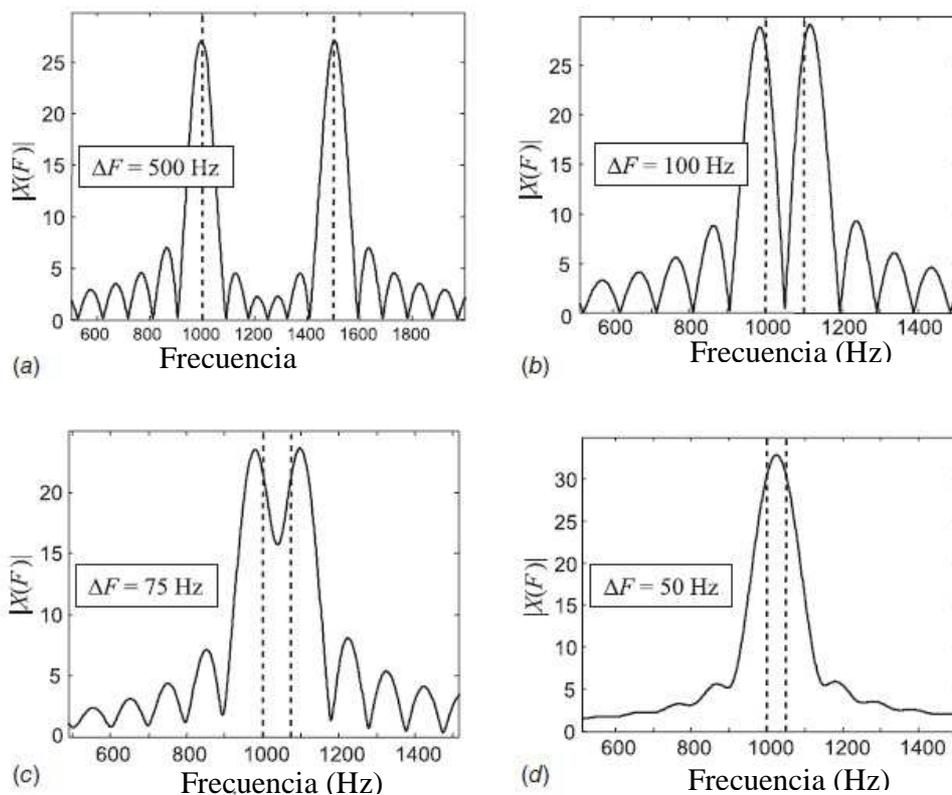


Figura 2.3. Resolución de dos señales sinusoidales en frecuencia a) Señales bien resueltas en rango con un espaciado de 500 Hz. b) También se encuentran resueltas con un espaciado de 75 Hz. d) Señales marginalmente resueltas con un espaciado de 75 Hz. d) Señales no resueltas (espaciado de 50 Hz). (Richards, 2005)

La Figura 2.3 ilustra el concepto de resolución en este caso en frecuencia. Se muestran dos señales cosenoidales. Esta señal podría representar el espectro Doppler de dos blancos que se mueven con la misma amplitud de eco pero a diferentes velocidades radiales. Las líneas verticales punteadas marcan las frecuencias de ambos cosenos. Se pueden observar dos picos bien distintivos, bien separados en frecuencia.

La resolución de un radar, a su vez, determina el tamaño de la celda de resolución. Una celda de resolución en rango, velocidad, o ángulo es el intervalo en estas dimensiones que contribuye al eco recibido por el radar en cualquier instante de tiempo. Es decir, para que dos ecos puedan ser resueltos por el radar, éstos deben estar separados por un espacio de

$ct/2$ de manera que dichos ecos no se solapen. Este valor de $ct/2$ se denomina resolución en rango (ΔR).

De manera análoga se puede definir celda de resolución en 2D y 3D considerando simultáneamente la resolución en, por ejemplo, rango, ángulo azimuth y ángulo de elevación.

El procesamiento de señales de radares utiliza muchas de las técnicas y conceptos usados en otras áreas del procesamiento de señales, tales como filtrado lineal, procesamiento en tiempo discreto, en el dominio de la frecuencia, diferentes técnicas con transformadas de Fourier, etc. Al mismo tiempo tiene características únicas que lo diferencia de otras áreas, como el caso de las teorías estadísticas de detección, técnicas de modulación, integración coherente en algunos tipos de radares, etc.

2.3.1. Ecuación de radar

La denominada ecuación de rango de radar o simplemente ecuación de radar (Eaves y Reedy, 1987; Skolnik, 2001) es un modelo determinístico que permite relacionar la energía de un eco recepcionado con la energía transmitida en términos de un número de parámetros para el diseño del sistema. Esta relación no solo sirve con el propósito de estimar el rango de un radar en función a las características del blanco sino que es muy útil como guía en el diseño y análisis de sistemas radares. Este simple modelo se basa en la ecuación de transmisión de radio diseñada por Harald Friis en 1954 para establecer la relación entre potencias y distancias en un enlace de radio.

Este es un modelo simplificado donde la primera suposición es que se posee un radiador isotrópico que transmite una onda de potencia P_t [W] en un medio sin pérdida. Un radiador isotrópico es una antena ideal teórica que radia igual energía exactamente en todas direcciones. Esta antena estaría en el centro de una esfera en la cual todos los puntos de la superficie recibirían la misma energía. Debido a que la transmisión es isotrópica en un medio sin pérdida, la densidad de la potencia a un rango R es el total de la potencia P_t dividido en el área de la superficie de una esfera de radio R , y por lo tanto:

$$\text{densidad de potencia isotropica transmitida} = \frac{P_t}{4\pi R^2} \quad [W / m^2] \quad (2.4)$$

Si se considera que la antena tiene una ganancia de potencia G , la energía puede concentrarse en un área menor y así aumentar la densidad de potencia en el área de interés. Por lo tanto, en la dirección de intensidad máxima de radiación, la densidad de potencia en el rango R se convierte en:

$$\text{pico de densidad de potencia transmitida} = Q_t = \frac{P_t G}{4\pi R^2} \quad [W / m^2] \quad (2.5)$$

Esta es la densidad de potencia incidente sobre el objetivo si este está alineado con los ejes de la antena.

Cuando una onda electromagnética de densidad de potencia como la descrita en la ecuación anterior (Eq 2.4) incide sobre un único objeto dispersor discreto, a un rango R la energía incidente será dispersada en varias direcciones, parcialmente absorbida por el blanco y otra porción será reflejada nuevamente hacia el radar. Podemos imaginar que el blanco pudiese recolectar toda la energía incidente sobre un colector de área σ en metros cuadrados y la re irradia isotrópicamente. La potencia de esta retornada o re-dispersada es entonces:

$$\text{potencia redispersada} = P_b = \frac{P_t G \sigma}{4\pi R^2} \quad [W] \quad (2.6)$$

(Richards, 2005)

Donde σ se denominada *Radar Cross Section (RCS)* ó *Sección transversal del radar* para un objetivo, esta medida en unidades de área y es una medida de la energía re irradiada por el blanco hacia el radar (Skolnik, 2008).

Una característica importante de la RCS es que σ no es igual al área física de la sección transversal del blanco.

En la tabla 2.2 se muestran algunos valores típicos de RCS para frecuencias de microondas (Skolnik, 2001; Richards, 2005). Como la RCS fue definida bajo la

suposición de que la energía es refractada o dispersada isotropicamente, la densidad de potencia de esta dispersión de retorno un rango R se puede encontrar dividiendo la potencia de la ecuación 2.6 por el área de superficie de la esfera de radio R como en el caso en la ecuación 2.4, obteniéndose así la densidad de potencia de la dispersión de retorno hacia el receptor del radar como:

$$\text{densidad de potencia redispersada} = Q_b = \frac{P_i G \sigma}{(4\pi)^2 R^4} \quad [W / m^2] \quad (2.7)$$

Blanco	RCS, m ²	RCS, dBsm
Misil no tripulado con alas convencional	0.5	-3
Aeronave pequeña de un solo motor	1	0
Pequeña aeronave de batalla 4 pasajeros	2	3
Gran aeronave de batalla	6	8
Bombardero mediano o jet	20	13
Bombardero grande o jet	40	16
Jumbo jet	100	20
Bote pequeño	0.02	-17
Pequeño yate civil	2	3
Yate	10	10
Gran nave en ángulo rasante cero	10000+	40+
Pickup	200	23
Automóvil	100	20
Bicicleta	2	3
Hombre	1	0
Pajaro	0.01	-20
Insecto	0.000001	-50

Tabla 2.2. Valores típicos de RCS o sección transversal de diferentes tipos de blanco

Si el área de apertura eficaz de la antena del radar es A_e (metros cuadrados), el total de potencia recibida en la antena será:

$$\text{potencia recepcionada} = P_r = \frac{P_i G A_e \sigma}{(4\pi)^2 R^4} \quad [W] \quad (2.8)$$

Se puede probar que la apertura eficaz de una antena esta relacionad con su ganancia y de acuerdo a la longitud de onda (λ) operativa como $A_e = \lambda^2 G / 4\pi$ (Kraus, 1950; Richards, 2005). Entonces:

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \quad [W] \quad (2.9)$$

Esta ecuación describe la potencia ideal que se obtendría en el receptor en un radar ideal en el espacio libre y sin usar ninguna técnica para mejorar la sensibilidad del equipo. Si se tienen en cuenta varios factores de “perdida” y de “ganancia”, estos pueden ser agregados al modelo. Por ejemplo puede agregarse un factor de perdida asociado a la atenuación atmosférica ($L_a(R)$) que es función del rango. Otro factor de perdida es el asociado al sistema (L_s). Agregando estos factores, la ecuación de radar queda expresada finalmente como:

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4 L_s L_a(R)} \quad [W] \quad (2.10)$$

Esta es la forma más simple de ecuación de rango de radar. La misma, relaciona la potencia del eco y parámetros del blanco tales como la potencia transmitida, la frecuencia de operación, la ganancia de la antena, la RCS y el rango.

Debido a que la potencia de la señal de radar es proporcional al cuadrado de la amplitud del campo eléctrico, la ecuación del rango sirve para modelar la amplitud de las componentes de la señal correspondiente al eco y al clutter. Una observación importante es que todas las variables en la ecuación (2.10) están expresadas en unidades lineales, no en decibeles, aun cuando algunos parámetros puedan estarlo. Otra observación es que P_r es una potencia instantánea, y no promediada.

Una consecuencia importante de esta ecuación es que, para un blanco puntual o discreto, la potencia decrece con una potencia de cuatro del rango desde el radar al objetivo. En consecuencia, la capacidad de detectar un blanco para una determinada RCS decrece

rápida con el rango. El rango puede ser incrementado con el aumento de la potencia de transmisión, pero debido a la dependencia de R^4 , la potencia debería incrementarse en un factor de 16 (12dB) solo para alcanzar a duplicar el rango efectivo. Una alternativa es la de incrementar la ganancia de la antena en un factor de 4 (6 dB). En contraposición, por ejemplo, en aplicaciones militares, muchos diseñadores de aeronaves de vigilancia u otros vehículos reducen su RCS por ejemplo en un factor de 16 con el objetivo de reducir a la mitad el rango al cual puedan ser detectados por un determinado sistema de radar.

La ecuación básica del radar, como se expresó anteriormente, tiene gran utilidad para el diseño de sistemas de radar y además permiten realizar la calibración del equipo. Diferentes técnicas de procesamiento de señales ayudan a incrementar el valor de rango y en consecuencia mejor performance en la detección.

2.3.2. Radar Cross Section

Como se menciona anteriormente se puede describir a la RCS σ de un determinado blanco como el área imaginaria del mismo que un determinado sistema de radar puede detectar. Si bien en la tabla 2.2 se muestran algunos valores de RCS típicos en el caso de algunos objetivos o blancos estudiados, en la realidad estos modelos no se adaptan a blancos más complejos. En general, la RCS es una función compleja del ángulo de aspecto (posición con respecto a un eje determinado), frecuencia y polarización, incluso en objetos dispersivos relativamente simples.

El comportamiento complejo observado en blancos moderadamente complejos conlleva a que la Radar Cross Section sea descripta estadísticamente (Levanon, 1988; Nathanson, 1991; Skolnik, 2001). Esto significa que la RCS σ de los blancos dentro de una celda de resolución se considera como variables aleatorias con una función de densidad de probabilidad (fdp o pdf de su nombre en Inglés). La ecuación de rango del radar (ecuación 2.10) es utilizada para estimar el valor de RCS promedio mientras que alguna de una variedad de pdfs es usada para describir el comportamiento estadístico de la sección transversal.

Una celda de resolución (o elemento de resolución) es la región del espacio que contribuye al eco recepcionado por un radar en un instante de tiempo (Figura 2.4). Si la

celda contiene un eco se denomina “celda blanco” y si contiene clutter es llamada “celda clutter”. (Barton y Leonov, 1998).

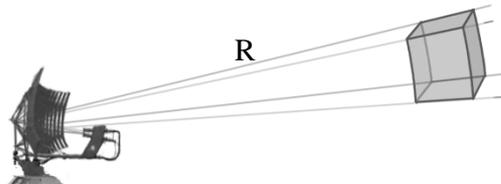


Figura 2.4. El volumen sombreado esquematiza una celda de resolución para el rango R

La tabla 2.3 sintetiza diferentes tipos de modelos estadísticos para RCS mas comunes. Una de las diferencias fundamentales entre los diferentes modelos de RCS de la tabla 2.3 es respecto a si la función de densidad de probabilidad tiene uno o dos parámetros libres. Los modelos, no fluctuante, Rayleigh/exponencial y todos los chi-cuadrados (una vez determinado el orden) son distribuciones de un único parámetro. Ese único parámetro es el promedio de la RCS, $\bar{\sigma}$. Los modelos Rice, Weibull y long-normal son distribuciones de dos parámetros. En el caso de los modelos de distribución de un solo parámetro, la estimación del promedio implica además la estimación de la varianza. Mientras que en el caso de dos parámetros, el promedio y la varianza deben estimarse de manera separada. Esta distinción es de suma importancia en el diseño de algoritmos de detección.(Richard, 2005; Skolnik, 2001; Barton y Leonov, 1998).

2.3.3. Modelado de la sección transversal

Gran parte de los resultados de la teoría de detección de ecos de radares se basa en utilizar alguno de los denominado modelos Swerling para la RCS de blancos fluctuantes (Swerling, 1960; Mayer, 1973; Nathanson, 1991; Skolnik, 2001). La principal característica de estos modelos es que intentan resolver el problema de la detección

mediante un conjunto de M muestras correspondientes a un eco dentro de una celda de resolución, en vez de utilizar una sola muestra.

Nombre del modelo	Pdf para RCS σ	Observaciones
Blanco no fluctuante, Marcum, Swerling 0 o Swerling 5	$P_{\sigma}(\sigma) = \delta_D(\sigma - \bar{\sigma})$ $\text{var}(\sigma) = 0$	Potencia de eco constante, reflector completamente estacionario sin movimiento del radar y sin movimiento del blanco.
Rayleigh/exponencial, chi-cuadrado de grado 2	$P_{\sigma}(\sigma) = \frac{1}{\bar{\sigma}} \exp\left(-\frac{\sigma}{\bar{\sigma}}\right)$ $\text{var}(\sigma) = \bar{\sigma}^2$	Varios blancos, distribuidos aleatoriamente, sin blanco dominante. Usado en casos 1 y 2 de modelos Swerling.
Chi-cuadrado de grado 4	$P_{\sigma}(\sigma) = \frac{4\sigma}{\bar{\sigma}^2} \exp\left(-\frac{2\sigma}{\bar{\sigma}}\right)$ $\text{var}(\sigma) = \bar{\sigma}^2/2$	Aproximación al caso donde se poseen pequeños blancos con un blanco dominante, con RCS dominante. Utilizado en modelos Swerling 3 y 4
Chi-cuadrado de grado 2m, Weinstock	$P_{\sigma}(\sigma) = \frac{m}{r(m)\bar{\sigma}} \left(\frac{m\sigma}{\bar{\sigma}}\right)^{m-1} \exp\left(-\frac{m\sigma}{\bar{\sigma}}\right)$ $\text{var}(\sigma) = \bar{\sigma}^2/m$	Generalización de los dos casos anteriores.
Rice o Rician, chi - cuadrado no central de grado 2	$P_{\sigma}(\sigma) = \frac{1}{\bar{\sigma}} (1 - \alpha^2) \exp\left[-\alpha^2 - \frac{\sigma}{\bar{\sigma}}(1 + \alpha^2)\right]$ $xI_0[2\alpha\sqrt{(1 + \alpha^2)(\sigma/\bar{\sigma})}]$ $\text{var}(\sigma) = \frac{(1 + 2\alpha^2)}{(1 + \alpha^2)^2} \bar{\sigma}^2$	Solución exacta para un blanco dominante mas otros mas pequeños.
Weibull	$P_{\sigma}(\sigma) = CB\sigma^{C-1} \exp(-B\sigma^C)$ $\bar{\sigma} = r\left(1 + \frac{1}{C}\right)B^{-\frac{1}{C}}$ $\text{var}(\sigma) = B^{-\frac{2}{C}} \left(r\left(1 + \frac{2}{C}\right) - r^2\left(1 + \frac{1}{C}\right)\right)$	Ajuste empirico para muchos objetivos medidos y distribuciones de clutter.
Log-normal	$P_{\sigma}(\sigma) = \frac{1}{\sqrt{2\pi s\sigma}} \exp\left[-\ln^2(\sigma/\sigma_m)/2s^2\right]$ $\bar{\sigma} = \sigma_m \exp(s^2/2)$ $\text{var}(\sigma) = \sigma_m^2 \exp(s^2)[\exp(s^2) - 1]$	Otro ajuste empírico para muchos objetivos medidos y distribuciones de clutter.

Tabla 2.3. Modelos para valores de RCS.

En general los radares son diseñados para transmitir ráfagas de pulsos. La frecuencia de repetición de pulsos (PRF) se puede definir como la cantidad de pulsos por unidad de tiempo, generalmente expresado en segundos. La elección del valor para la PRF es muy importante para el diseño con el objetivo de evitar la ambigüedad. Para el caso de altos valores de RPF, puede ocurrir que en la misma ventana de tiempo se este transmitiendo y recibiendo pulsos. Los ecos en ambos casos se solapan causando ambigüedad (Barton y Leonov, 1998).

En el caso de estos equipo que envían ráfagas de pulsos, se puede calcular el denominado intervalo de procesamiento coherente determinado por M / PRF . Donde M corresponde a un bloque de mediciones de la sección transversal del blanco.

Cada modelo Swerling es la combinación de una función de densidad de probabilidad y el intervalo de correlación para una RCS σ .

Swerling considero dos casos extremos para las propiedades de correlación de un bloque M de mediciones de σ . El primero asume que todos los ecos están perfectamente correlacionados, y por lo tanto, todos los M contribuyen a la detección en un mismo valor para el primer barrido de la antena. En el siguiente barrido, también se podrán colectas M mediciones idénticas. Sin embargo, los valores correlacionados entre los ecos son independientes entre los diferentes barridos. Este caso se denomina correlación scan-to-scan. El segundo caso asume que cada pulso individual en cada barrido posee un valor independiente de σ . E esta última situación se habla de correlación pulse-to-pulse.

Las dos funciones de densidad de probabilidad consideradas en estos modelos son las funciones de Rayleigh/exponencial y la Chi-cuadrada de grado cuatro. La función de Rayleigh/exponencial describe el comportamiento de un blanco complejo que consiste en numerosos puntos de dispersión, ninguno de ellos particularmente dominante. Mientras que la función Chi-cuadrada de grado cuatro modela blancos que poseen puntos de dispersión de similar energía más un punto de dispersión dominante.

Los modelos Swerling se conforman, entonces, por la combinación de las dos opciones de pdf para σ y las dos opciones de correlación mencionadas anteriormente. Se obtienen, así, cuatro modelos denominados Swerling1, Swerling2, Swerling3 y Swerling4. En diferente bibliografía los blancos no fluctuantes se los identifica como modelos Swerling0 o Swerling5. (Richards, 2005)

La Figura 2.5 muestra un esquema de los modelos Swerling.

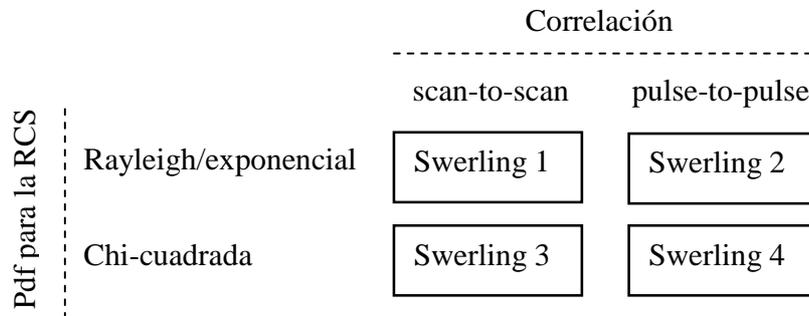


Figura 2.5. Esquema de los modelos Swerling de acuerdo a la pdf de la RCS y las características de correlación

2.4. Radar de onda continua y radar pulsado

Desde los inicios de las investigaciones sobre la tecnología radar, se usaron repetición de pulsos junto a técnicas de detección por envolvente para detectar débiles ecos inmersos en ruido. Esta técnica se basa en pulsos de portadora de radio sin modular. Esto es posible sólo emitiendo ondas de radio de gran potencia (Rishbeth y Garriot, 1969; Jursa, 1985). Estos equipos tienen como desventaja el alto costo que implica elaborar un transmisor de alta potencia. Estas limitaciones respecto a las potencias que se deben emitir tienen gran impacto especialmente en frecuencias altas. Un caso típico son los radares para las aeronaves de vigilancia que llegan a tener un pico de potencia de 1 MWatt, para un pulso de 1 μ s, una resolución de rango de 150 mts, siendo la energía de un pulso de 1 Joule. Para obtener una resolución de 15 cm y restringir la energía por pulso a 1 Joule, esto implica que un pulso de 1 μ s requiere un pico de potencia de 1 GWatt (O'Donnell, 2010). La resolución de rango establece la capacidad del radar para detectar distinguir dos objetivos que están cercanos en una determinada distancia (rango).

Supongamos que un radar tenga una resolución de rango ΔR determinada, si dos blancos se encuentran a una distancia menor que la resolución del radar, el mismo no será capaz de detectar ambos blancos sino que detectará un único objetivo. Esto significa que es un

objetivo de diseño tener una buena resolución y esto se logra analizando las características del blanco o blancos que se desean detectar.

Esto se ilustra en la figura 2.6 que esquematiza la salida de un receptor para un único pulso transmitido de Δt tiempo de duración, y que es reflejado por dos blancos separados por una determinada distancia. Si la distancia de separación es lo suficientemente grande, los dos ecos se pueden observar en el receptor (Figura 2.6 b). Se dice en tal caso que los dos blancos fueron “resueltos” en rango. En el caso de la la Figura 2.6 c, los dos blancos se encuentran muy próximos y por lo tanto los ecos se solapan. Dependiendo de el espaciado exacto de los dos blancos, sus ecos pueden combinarse constructivamente (aumentar la energía recepcionada), destructivamente o en algún término intermedio. El resultado es un sistema muy sensible a cambios en el espaciado de los blancos, y por lo tanto las no se puede resolver de manera confiable (Richards et. al, 2010). El caso intermedio se muestra en la Figura 2.6 d, que ocurre cuando:

$$\Delta R = \frac{c \Delta t}{2} \quad (2.11)$$

En conclusión, ΔR es la resolución de rango y dos blancos separados por una distancia mayor a ΔR .

A partir de esta ecuación se puede analizar cómo debe diseñarse un sistema con el objetivo de mejorar la resolución de rango.

Si la duración del pulso (Δt) aumenta, entonces ΔR también aumenta. Si se quiere mejorar la resolución, o sea disminuir ΔR entonces hay que disminuir Δt .

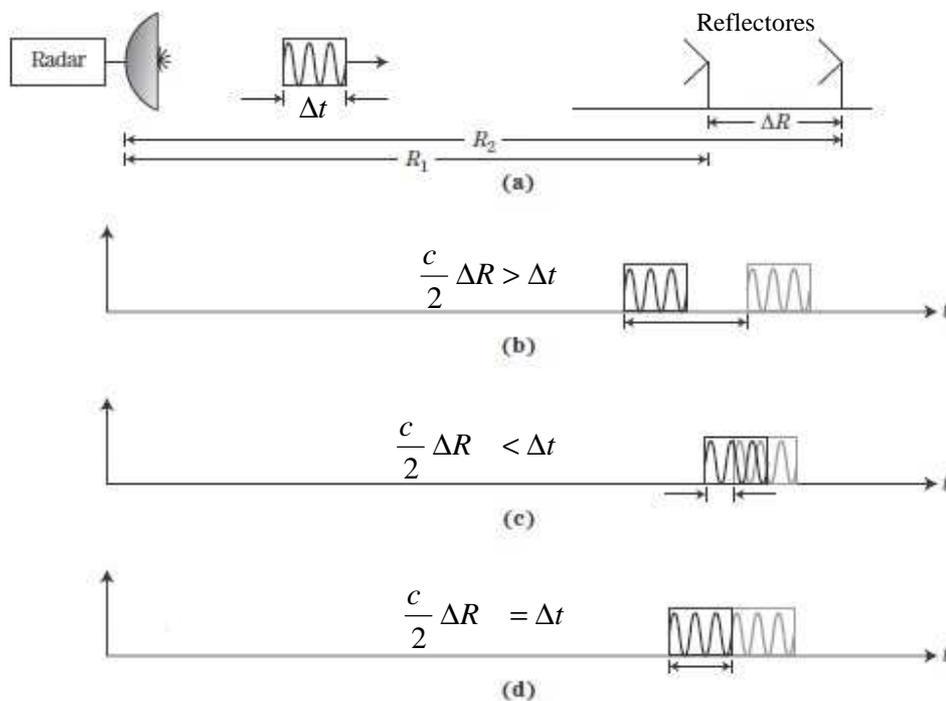


Figura 2.6. Esquema de: Un pulso transmitido de Δt tiempo de duración (a). Dos blancos resueltos en rango (b). Dos blancos muy cercanos donde la detección no es confiable (c). Situación intermedia que permite calcular la resolución de rango ΔR

Por lo tanto se puede decir que, la resolución de rango es proporcional al ancho del pulso y es inversamente proporcional al ancho de banda. Esto explica por que para tener mejor resolución, es necesario contar un mayor ancho de banda lo que, a su vez, significa tener pulsos más cortos. La Figura 2.7 muestra la relación entre el ancho de un pulso (T) y el ancho de banda de su espectro ($B = 1/T$).

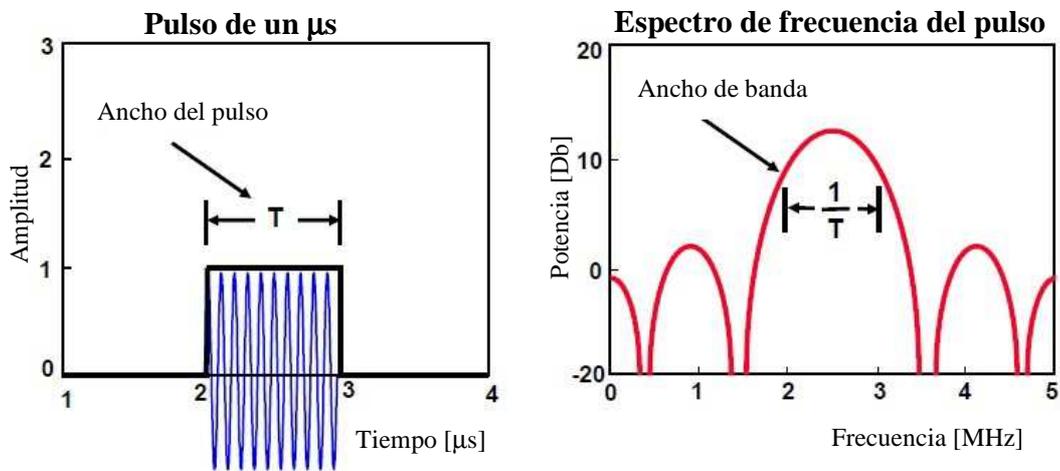


Figura 2.7. Relación entre el ancho de un pulso y el ancho de banda de su espectro (O'Donnell, 2010).

Contar con una alta resolución en rango es una de los aspectos más importantes de los radares. Esto permite obtener una mejor caracterización de los blancos que ilumina el equipo y permite que las mediciones sean más precisas. Por lo tanto, las restricciones respecto a la potencia que pueda emitir el instrumento son determinantes en el momento del diseño del mismo en cuanto al alcance y la resolución.

Dado que la naturaleza reacciona a la energía, donde la energía puede escribirse como:

$$E = P t \quad (2.12)$$

Donde E es la energía, P es la potencia y t es el tiempo.

Supongamos un pulso de corta duración cuya energía transmitida es:

$$E_1 = P_1 t_1 \quad (2.13)$$

Esto indica que cuanto más corto es el pulso, mayor es la potencia requerida para transmitir la misma energía (Figura 2.8).

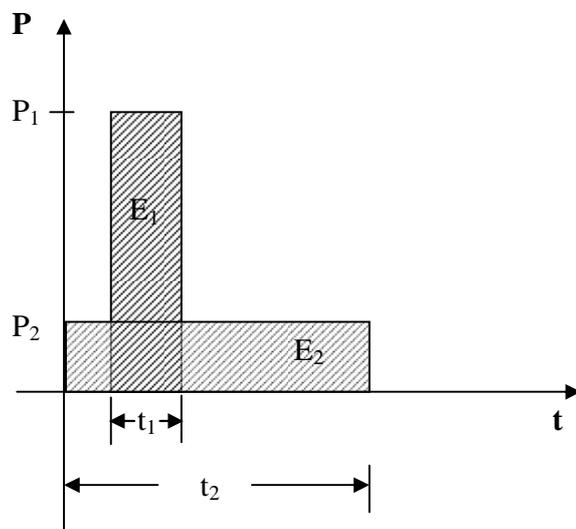


Figura 2.8. Esquema de la relación entre un pulso corto y uno largo respecto a la potencia emitida

Si queremos mantener la misma resolución de rango pero emitir menos potencia, lo que significa que deseamos mantener la misma energía transmitida, entonces:

$$P_1 t_1 = P_2 t_2 \quad (2.14)$$

Tal que P_2 corresponde a la potencia del segundo pulso, con $P_2 < P_1$, y t_2 es su duración.

Para que esta igualdad se mantenga tiene que ocurrir el segundo pulso de mayor se encuentre modulado.

Actualmente esta opción se encuentra implementada en los denominados radares pulsados modulados, los que mediante la transmisión de una ráfaga de pulsos codificados logra obtener excelente performance en detección emitiendo baja potencia. Este tipo de sistemas se basa en la denominada técnica de compresión de pulso, y utiliza una serie de herramientas matemáticas y de procesamiento de señales para discriminar el eco a la entrada del receptor. Así esta técnica permite emitir menos potencia pero complica la etapa de recepción.

2.5. Técnica de compresión de pulso

Los radares modernos en su mayoría utilizan técnicas de esparcimiento de frecuencia o “spread-spectrum” para mejorar la resolución de rango transmitiendo bajos niveles de potencia (Barry, 1971; Poole, 1985; Reinisch, 2000; Bianchi and Altadill, 2005; entre otros). Esto significa que se puede lograr igual ancho de banda que en la técnica de detección por envolvente típica de los radares pulsados sin modular. Esta técnica se denomina compresión de pulso y es posible obtener con ella alta resolución de rango transmitiendo señales moduladas de gran longitud. Así, se generan ondas de larga duración, se transmiten estas como portadoras codificadas y se procesa la señal recibida (recepcionada) mediante algoritmos numéricos para poder realizar la detección del eco, como es el caso de la correlación de señales (Rastogi, 1990; Mohamed, 1991; entre otros).

Una medida de la calidad de la señal afectada por el ruido es el cociente entre la potencia de la señal y la potencia del ruido, conocido comúnmente como relación señal ruido o por sus siglas en inglés SNR. Es posible mostrar que para un determinado ancho de banda, a mayor densidad espectral de energía en las señales transmitidas, mayor es la relación señal ruido (SNR) de la señal en el receptor. Cuando la potencia del ruido aumenta, hay una disminución gradual de la inteligibilidad de la señal y a cierto nivel de ruido la señal se vuelve completamente ininteligible (Roberts, 2004).

En las siguientes gráficas se comparan, a modo cualitativo, las dos técnicas mencionadas. La Figura 2.9 muestra un pulso corto de gran potencia y su respectivo espectro de energía recepionado. Esta señal tiene una longitud de 30 μs y un ancho de banda aproximado de 60 Khz. La Figura 2.10 muestra un tren de pulsos codificados de 480 μs para el mismo ancho de banda (Cabrera y Molina, 2007). La respuesta espectral en este último caso es muy similar al de la Figura 2.7. De este modo se puede ver que la compresión de pulsos logra obtener un espectro esparcido con propiedades similares al caso del pulso corto.

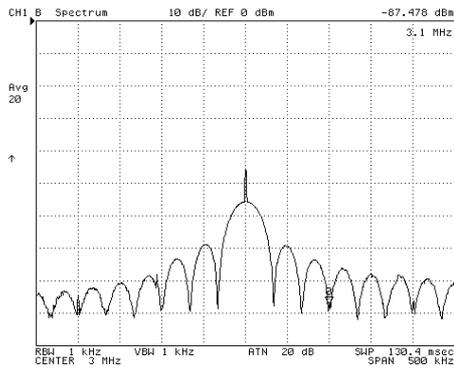


Figura 2.9

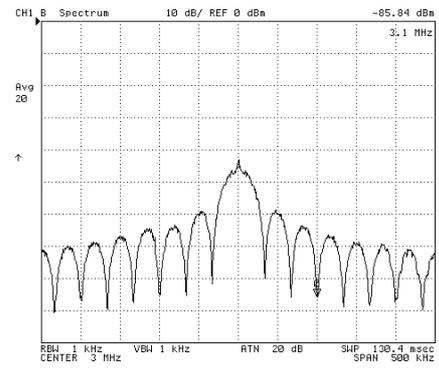


Figura 2.10

Imágenes de prueba de laboratorio obtenidas de un analizador espectral.
(Laboratorio de Ionósfera. 2007)

Este resultado se debe a que los pulsos fueron codificados de acuerdo a alguna secuencia determinada. Para ello, se codificó una portadora con un código binario. La Figura 2.11 muestra un pulso largo modulado. Cada vez que el valor del código cambia, la portadora cambia de fase, este tipo de modulación se llama modulación bifásica o BPSK.

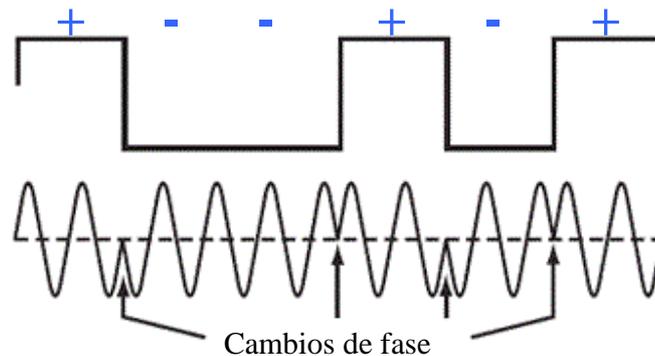


Figura 2.11. Codificación de un pulso largo usando modulación bifásica

La portadora no modulada es una senoide. La modulación se realiza en este caso desplazando la fase (se desfasa en π) de la portadora según el código binario (Tomasi,

1996). Esta es una modulación bifásica y es un caso especial. La Figura 2.12 muestra un esquema de bloques del generador de pulsos codificados del radar. El código utilizado durante la modulación es almacenado para ser usado en la etapa de decodificación de la señal receptionada.

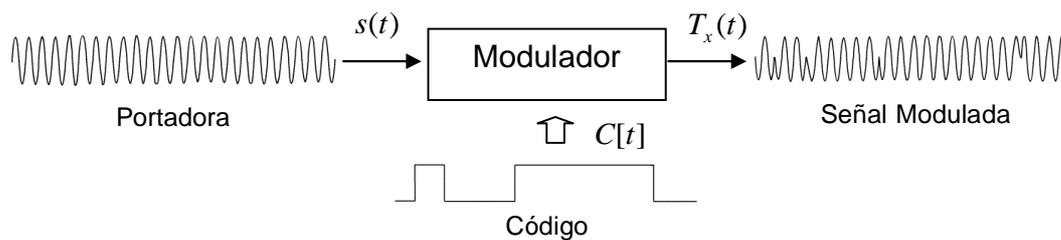


Figura 2.12. Esquema de bloque del generador de pulsos codificados

Las claras ventajas que posee la compresión de pulsos hacen que sea la técnica más ampliamente difundida para aplicaciones generales de radares.

Los dos elementos fundamentales a tener presente al momento de implementar la compresión de pulso son la codificación de la portadora y, durante el proceso de decodificación, la correlación con el código utilizado. Ambos temas se desarrollan en las subsecciones siguientes.

2.5.1. Decodificación del Eco

Una vez que los pulsos son transmitidos, el radar inicia su periodo de escucha. Es decir, el equipo receptiona señales provenientes del medio para iniciar el proceso de detección del eco reflejado por el blanco, en caso de que algún blanco se encuentre presente.

En cualquier caso, la señal receptionada está compuesta por una combinación de posibles ecos, de ruido del medio ambiente y/o introducido por el mismo equipo, tal vez contenga señales de interferencia casuales e incluso señales emitidas deliberadamente para degradar la detección (Jamming).

Como primera instancia en la detección, es necesario utilizar alguna herramienta matemática para procesar la señal receptionada. El objetivo es determinar dentro de esta señal, el instante de tiempo en el cual se puede encontrar el pulso codificado que fue emitido

Una señal codificada inmersa en una señal de interferencia o ruido puede ser “decodificada” correlacionándola con el código utilizado. De manera intuitiva, si es posible encontrar el instante de tiempo donde la correlación entre el código y la señal dispersada por un blanco es máxima, entonces será posible determinar (a) la presencia del blanco a pesar del ruido, (b) la distancia a la cual se encuentra el blanco. Esta capacidad se ve disminuida frente a ambientes con alto niveles de ruido, algunas señales de interferencia y/o con ecos de clutter, y a la posibilidad que el blanco se encuentre, además, en movimiento.

Desde el punto de vista del radar, una característica importante es la “firma” de un blanco es la denominada correlación en tiempo, frecuencia y ángulo. El “intervalo de correlación” es el cambio en el tiempo, frecuencia o ángulo que provoca que la amplitud del eco no correlacione más en un grado específico. Si se trata, por ejemplo, de un blanco estático como puede ocurrir con un edificio que es iluminado por un radar con una serie de pulsos idénticos y no hay movimiento entre el eco y el blanco, se espera que en cada instante se obtenga la misma energía correspondiente a cada pulso (ignorando el ruido en el receptor). En este caso, mientras se mantenga esta situación se mantendrá la correlación. Sin embargo, si se permite el movimiento entre el radar y el blanco, la distancia relativa de camino entre el radar y las diferentes dispersiones en el objetivo pueden causar la fluctuación en la amplitud del eco. Dependiendo de la velocidad con la que se mueve el blanco este intervalo de correlación será mayor o menor.

Supongamos que se emite un pulso Tx modulado con algún código, al cabo de un tiempo τ correspondiente al tiempo de viaje de la señal desde y hacia el radar, dicho pulso es receptionado. Intuitivamente, para determinar el retardo temporal τ que sufrió la señal es necesario comparar la señal receptionada con el código. En el instante τ la concordancia será máxima (Figura 2.13). Para poder realizar esta comparación entre señales es necesario recurrir a una herramienta matemática denominada correlación.

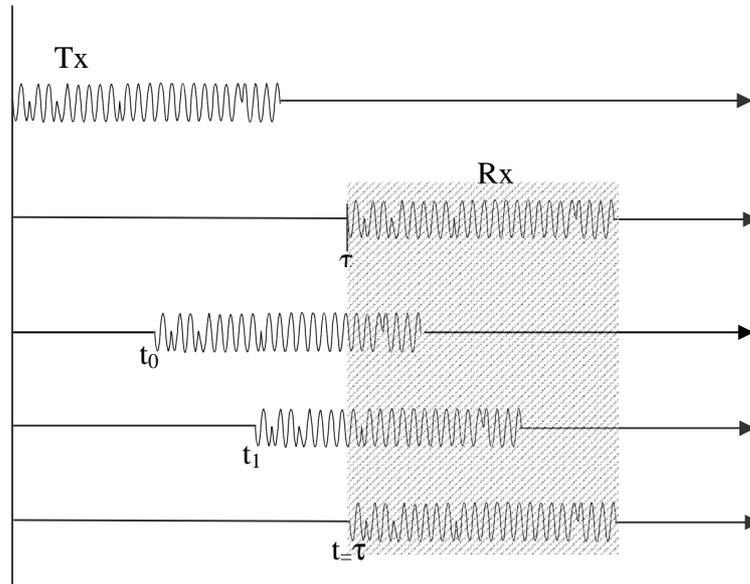


Figura 2.13. Esquematiza cómo la señal codificada y transmitida (Tx) es comparada con la señal recepcionada (Rx). Desplazando Tx para diferentes intervalos de tiempo se puede determinar el instante de tiempo (τ) para el cual la concordancia entre ambas señales es máxima.

2.5.1.1. Correlación

Desde el punto de vista del procesamiento de las señales de radar, la correlación es una herramienta matemática que permite comparar y determinar como se relacionan dos señales. Dichas relaciones indican si los fenómenos físicos que las causan se relacionan o si una señal es una versión modificada de la otra. Este método se aplican a todo tipo de señales: en tiempo continuo y en tiempo discreto, determinísticas o aleatorias.

Intuitivamente, dos señales cuanto más similares sean, más correlacionadas se encuentran. Si dos señales corresponden a la energía dispersada por un mismo blanco en una posición determinada, los ecos se encuentran altamente correlacionados.

Dadas dos señales es posible graficar una en función de la otra y el gráfico resultante se denomina correlograma. Si las señales se encuentran altamente correlacionadas, el correlograma tiende a formar una línea recta. La Figura 2.14 muestra dos señales en x_1 y x_2 tal que $x_1 = -x_2$, y su correspondiente correlograma (Figura 2.15). Las señales

x_1 y x_2 están altamente correlacionadas de manera negativa porque su correlograma forma una recta de pendiente negativa.

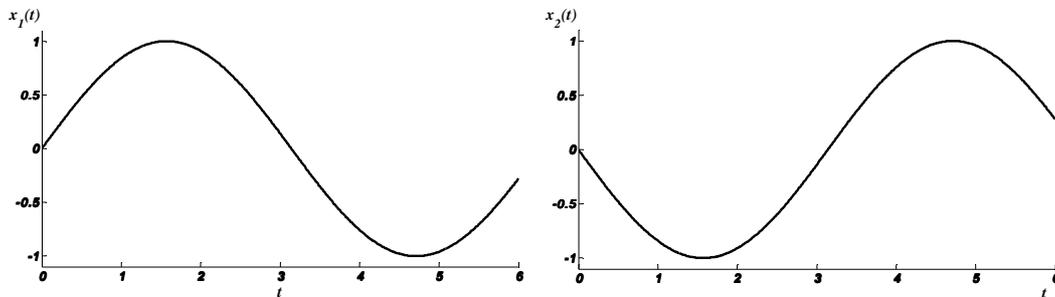


Figura 2.14. Dos señales senoidales.

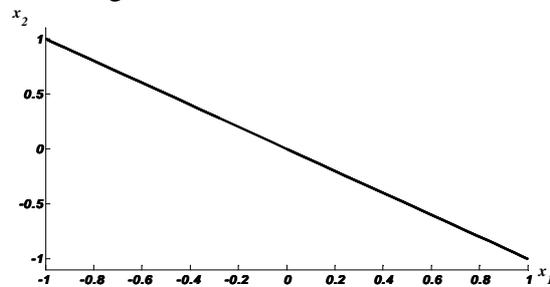


Figura 2.15. Correlograma

Si al realizar el correlograma de un par de señales de tiempo se obtiene una recta con pendiente positiva, se dice que las señales se encuentran altamente correlacionadas positivamente. Si por el contrario la pendiente es negativa, se encuentran altamente correlacionadas de forma negativa. En el caso de tratarse de señales de tiempo discreto, el correlograma será un gráfico con valores discretos distribuidos de manera que si éstos tienden a agruparse en una línea recta se dice que las señales están altamente correlacionadas.

Dadas dos señales discretas es posible realizar un análisis estadístico de cómo se distribuye la nube de puntos que representa a una señal en función de la otra. Si es posible ajustar una recta que aproxime a los puntos entonces se podrá decir que las dos señales están correlacionadas linealmente, e incluso se podrá obtener un coeficiente de correlación que indique qué tanto se encuentran correlacionadas.

Para los valores de la señal, puede realizar el siguiente análisis estadístico, calculando el coeficiente de correlación lineal según la fórmula,

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (2.15)$$

Donde σ_{xy} es la covarianza, y donde σ_x y σ_y son las desviaciones estándar. Si r es próximo a 1 o -1 la correlación es muy fuerte, si el valor del coeficiente de correlación es cero entonces los datos no están correlacionados linealmente. Para los datos del ejemplo el coeficiente de correlación es 0.94, lo que indica una fuerte correlación negativa que confirma la apreciación visual que se obtiene al observar el correlograma (Figura 2.15).

En el caso que el correlograma no presente una distribución de puntos similar a una recta o que el coeficiente de correlación sea próximo a cero, no quiere decir que las señales no se encuentren correlacionadas. Puede ocurrir que la relación entre estas no sea una proporcionalidad simple.

Un caso interesante es el de aquellas señales de tiempo continuo x_1 y x_2 tales que

$$x_2(t) = kx_1(t - \tau) \quad (2.16)$$

donde k es una constante y τ representa el desplazamiento de la señal x_1 en el tiempo. De modo que, si se desplaza la segunda señal un poco después en el tiempo (si $k = 1$ y $\tau < 0$) se obtendría un correlograma de línea recta con pendiente positiva que indicaría una fuerte correlación positiva.

Un caso de particular interés se presenta cuando se tienen dos senoides con un cambio de fase de 90° . Esta situación se grafica en la Figura 2.16, donde se observa $x_1(t) = \sin(t)$ y $x_2(t) = \sin(t - \pi/2)$ y el correlograma correspondiente a dichas funciones. El correlograma muestra que las funciones no están correlacionadas, aunque la disposición de los puntos alrededor de cero sugiere que existe algún tipo de relación, aunque esta no sea lineal. En este caso,

$$x_1(t) = x_2(t + \pi / 2) \text{ donde } k = 1 \text{ y } \tau = \frac{\pi}{2},$$

si se desplaza la segunda función τ unidades, se obtendrían dos señales que se correlacionan positivamente.

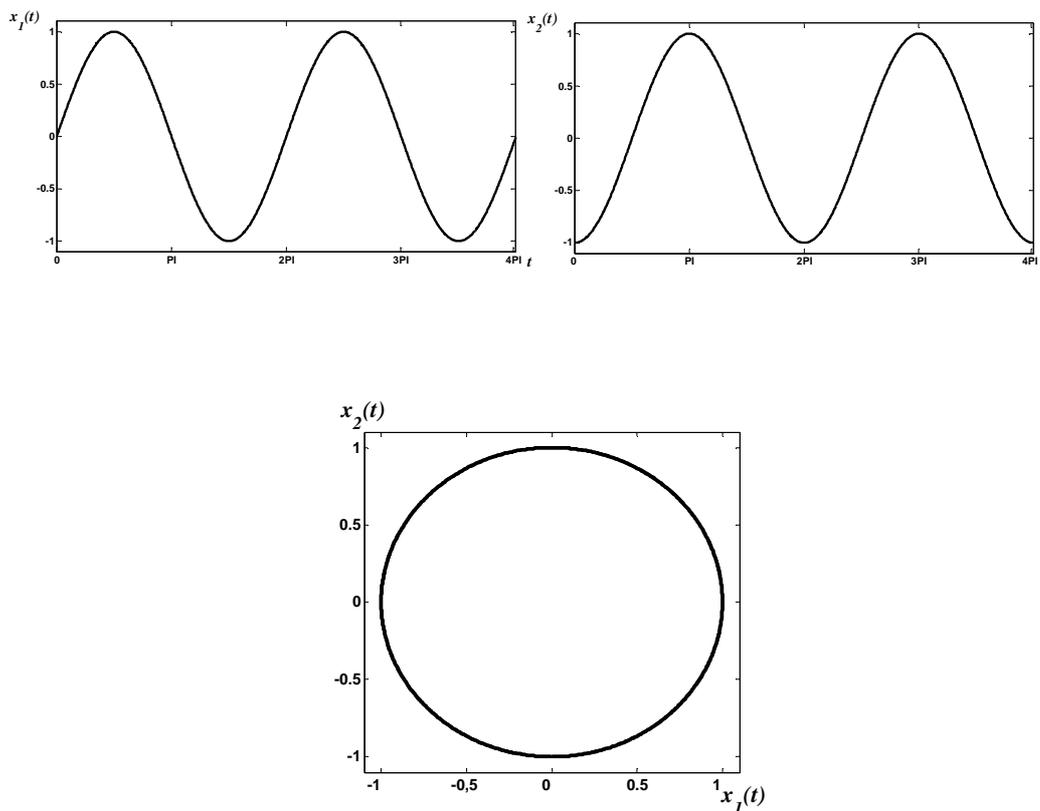


Figura 2.16. Correlograma correspondiente a las señales x_1 y x_2 .

Si bien el correlograma es una herramienta de visualización muy útil, es necesario poder expresar la relación entre dos señales de una manera matemática precisa.

La idea de correlación se basa en el análisis de si dos señales tienden a moverse juntas. Si las dos señales se mueven en la misma dirección al mismo tiempo, están correlacionadas durante ese tiempo, y su correlación es positiva. Si, por el contrario, las dos señales se

mueven en dirección opuesta durante un largo periodo de tiempo, se dice que están correlacionadas negativamente. En el caso de las dos senoides, una 90° fuera de fase respecto de la otra, que no están correlacionadas, estas señales se mueven la mitad del tiempo en la misma dirección y la otra mitad en direcciones opuestas.

Para incorporar estos conceptos acerca de cómo se mueven las señales una en relación a la otra, se debe observar el valor promedio del producto de las funciones.

Dadas dos señales (analógicas o digitales) x_1 y x_2 , hay dos casos para analizar: 1) El valor promedio de cada una de las señales es cero, 2) El valor promedio de cada una de las señales es distinto a cero.

Caso 1: El valor promedio de cada una de las señales es cero.

Si las dos señales se correlacionan positivamente (se mueven en la misma dirección durante un período largo de tiempo) entonces el promedio de sus productos es mayor que cero. Esto ocurre cuando ambas señales son positivas o ambas son negativas (Figura 2.17).

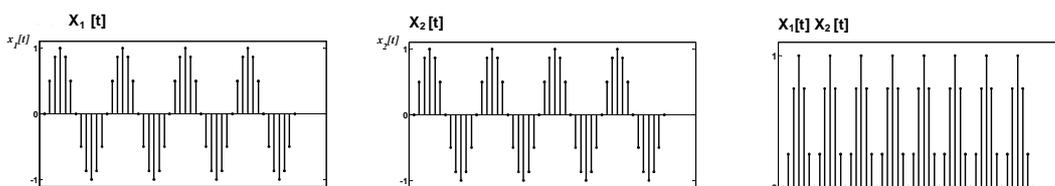
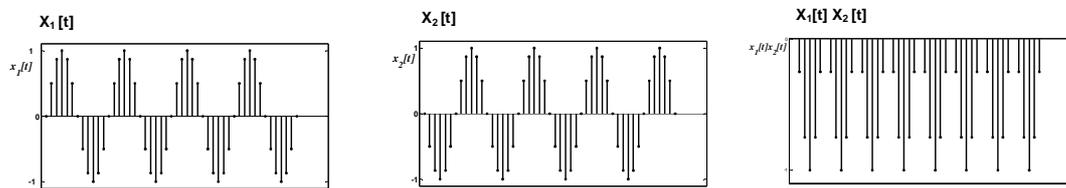


Figura 2.17. Señales digitales correlacionadas positivamente

Si las dos señales, en cambio, se correlacionan negativamente entonces el promedio de sus productos tenderá a ser negativo. Esto ocurre cuando las señales se mueven en sentidos opuestos (Figura 2.18).



2.18 Señales digitales correlacionadas negativamente

0

Por último, si las señales no están correlacionadas ocurrirá que el promedio del producto de x_1 y x_2 será cero, o muy próximo a cero. (Figura 2.19)

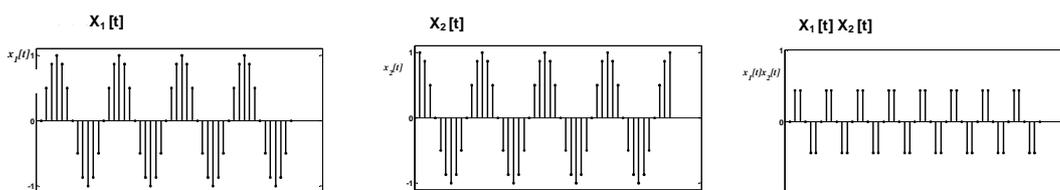


Figura 2.19. Señales digitales no correlacionadas

Caso 2: El valor promedio de cada una de las señales es distinto a cero.

En este caso solo se agregará un sesgo al producto, pero la variación alrededor del sesgo seguirá indicando si sus variaciones se mueven en direcciones iguales u opuestas. Las relación entre el promedio del producto y el producto de los promedios se cumplen como en el caso 1. Es decir:

Si el promedio del producto de x_1 y x_2 es mayor que el producto de los valores promedios de las señales individuales, las dos señales están correlacionadas positivamente; caso contrario, las dos señales están correlacionadas negativamente.

En el caso que el promedio del producto de las dos señales sea igual al producto de los promedios implica que las señales no se encuentran correlacionadas. Aunque no sean

exactamente iguales, pero sean muy cercanos (esto ocurre en un periodo corto de tiempo). Cuando se incrementa el tiempo, estos dos valores se acercan al mismo límite. En las Figuras 2.20.a, 2.20.b, 2.20.c se muestran las correlaciones de señales con valor promedio distinto a cero.

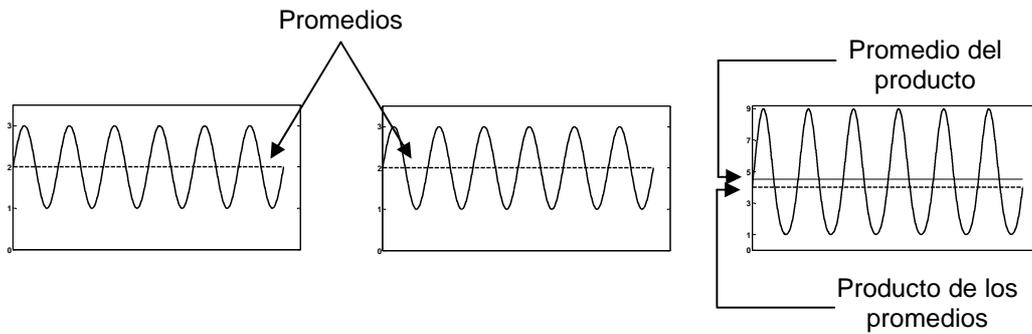


Figura 2.20.a. Señales en tiempo continuo correlacionadas positivamente

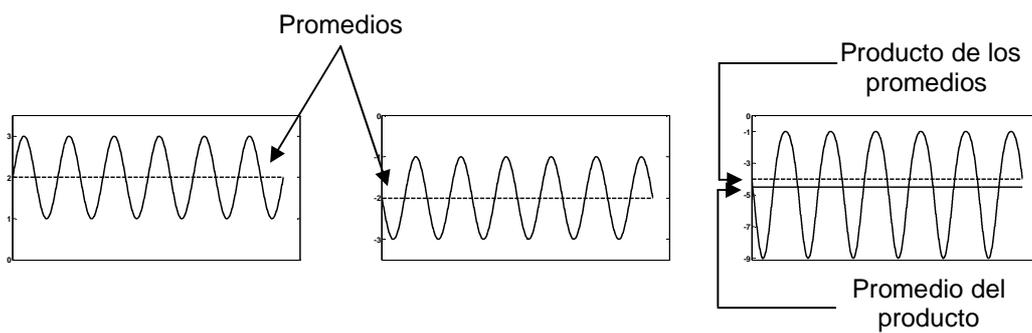


Figura 2.20.b. Señales en tiempo continuo correlacionadas negativamente

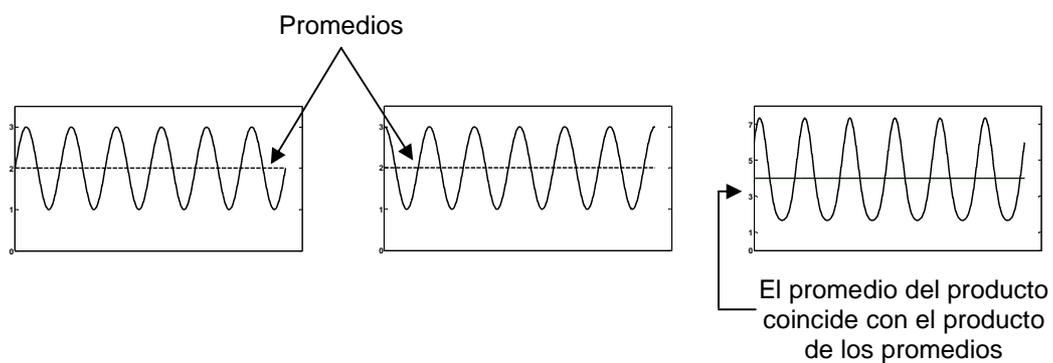


Figura 2.20.c. Señales en tiempo continuo no correlacionadas

La correlación entre dos señales da por resultado un valor numérico que representa cuantitativamente que tan correlacionadas se encuentran dos señales. La función de correlación es una expresión matemática de que tanto están correlacionadas dos señales en función de la cantidad de desplazamiento. La idea es obtener un valor de correlación entre las señales para cada valor de desplazamiento que sufran.

La función de correlación es mucho más general que sólo la correlación debido a que es una función que depende del grado en que se desplaza la segunda función. Algunas funciones no se correlacionan para un desplazamiento determinado pero lo hacen en gran medida para otro. Un ejemplo es la función de correlación de las funciones seno y coseno. Para el desplazamiento cero las funciones no están correlacionadas pero para un desplazamiento de 90° de una de las funciones, están altamente correlacionadas (ya sea positiva o negativamente).

La definición matemática de la correlación depende del tipo de señal que se analiza. Hay dos definiciones aceptadas de manera común, una para señales de energía y una para señales de potencia. En ambos casos se denota $R_{xy}(\tau)$ a la función resultante de correlacionar las funciones x e y para distintos valores de desplazamiento τ . El resultado es una función donde la abscisa del punto máximo representa el desplazamiento donde la correlación es máxima. Si las señales son digitales, la notación para la función de correlación discreta es $R_{xy}[m]$ donde m representa el desplazamiento.

Se pueden definir la correlación de señales de energía (a) o de potencia (b)

a) Correlación de señales de energía

Para dos señales de energía $x(t)$ e $y(t)$ en tiempo continuo se define la correlación mediante la ecuación,

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y^*(t + \tau)dt \quad (2.17)$$

o

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t-\tau)y^*(t)dt \quad (2.18)$$

Dependiendo de cual sea la función que se desplace.

Si ambas señales son reales la expresión es,

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y(t+\tau)dt = \int_{-\infty}^{\infty} x(t-\tau)y(t)dt \quad (2.19)$$

Si las señales son discretas la ecuación es,

$$R_{xy}[m] = \sum_{n=-\infty}^{\infty} x[n]y^*[n+m] = \sum_{n=-\infty}^{\infty} x[n-m]y^*[n] \quad (2.20)$$

Y en el caso que ambas señales discretas sean reales,

$$R_{xy}[m] = \sum_{n=-\infty}^{\infty} x[n]y[n+m] = \sum_{n=-\infty}^{\infty} x[n-m]y[n] \quad (2.21)$$

b) Correlación de señales de potencia

La función de correlación entre dos señales de potencia $x(t)$ e $y(t)$ en tiempo continuo, se define como

$$R_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T x(t)y^*(t+\tau)dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T x(t-\tau)y^*(t)dt \quad (2.22)$$

Si $x(t)$ e $y(t)$ son reales,

$$R_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T x(t)y(t + \tau)dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T x(t - \tau)y(t)dt \quad (2.23)$$

La correlación de una función consigo misma se denomina autocorrelación y es un caso particular de la función de correlación.

Dada una función de energía, su autocorrelación es

$$R_{xx}(t) = \int_{-\infty}^{\infty} x(t)x(t + \tau)dt \quad (2.24)$$

si se trata de una señal de tiempo continuo, y

$$R_{xx}[m] = \sum_{n=-\infty}^{\infty} x[n]x[n + m] \quad (2.25)$$

Si se trata de señales discretas. Para el caso que el desplazamiento sea cero, las ecuaciones anteriores se reescriben como,

$$R_{xx}(0) = \int_{-\infty}^{\infty} x^2(t)dt \quad (2.26)$$

y

$$R_{xx}[0] = \sum_{n=-\infty}^{\infty} x^2[n] \quad (2.27)$$

respectivamente.

Si la señal es de potencia, entonces las ecuaciones de las funciones de autocorrelación en tiempo continuo y discreto son análogas a las funciones de correlación de señales de potencia, con la salvedad de que lo que se busca es el promedio del producto de dos funciones idénticas, una de las cuales se encuentra desplazada. Si el desplazamiento es cero la autocorrelación de señales de potencia se calcula como,

$$R_{xx}(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_T x^2(t)dt \quad (2.28)$$

y

$$R_{xx}[0] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=N} x^2[n] \quad (2.29)$$

para señales en tiempo continuo (Ecuación 2.28) y en tiempo discreto (Ecuación 2.29).

Como el valor máximo de autocorrelación de cualquier señal se alcanza cuando ambas señales coinciden, se puede decir que este valor se alcanza con el desplazamiento cero. Esto implica que,

$$R_{xx}(0) \geq R_{xx}(\tau) \quad , \forall \tau$$

y

$$R_{xx}[0] \geq R_{xx}[m] \quad , \forall m$$

Además se puede demostrar que para cualquier función $x(t)$ o $x[t]$ (sea par o no) la autocorrelación de ésta será una función par, o lo que es lo mismo,

$$R_{xx}(-\tau) = R_{xx}(\tau)$$

o

$$R_{xx}[-m] = R_{xx}[m]$$

Otra propiedad de la función de autocorrelación es que un desplazamiento en el tiempo de una señal no cambia su autocorrelación. Por ejemplo, sean las señales de energía $x(t)$ y sea $y(t) = x(t + \tau_0)$, sus funciones de autorrelación serían,

$$R_{xx}(t) = \int_{-\infty}^{\infty} x(t)x(t + \tau)dt \quad (2.30)$$

y

$$R_{yy}(\tau) = \int_{-\infty}^{\infty} y(t)y(t+\tau)dt = \int_{-\infty}^{\infty} x(t+\tau_0)x(t+\tau_0+\tau)dt = R_{xx}(\tau) \quad (2.31)$$

Lo que demuestra que las funciones de autocorrelación de $x(t)$ y $y(t)$ son independientes del valor de τ_0 .

En el siguiente ejemplo se analiza la función de autocorrelación de la señal $x(t) = \text{sen}(t)$. La Figura 2.21 muestra la gráfica de la función de autocorrelación para $x(t)$ y una porción de ésta ampliada.

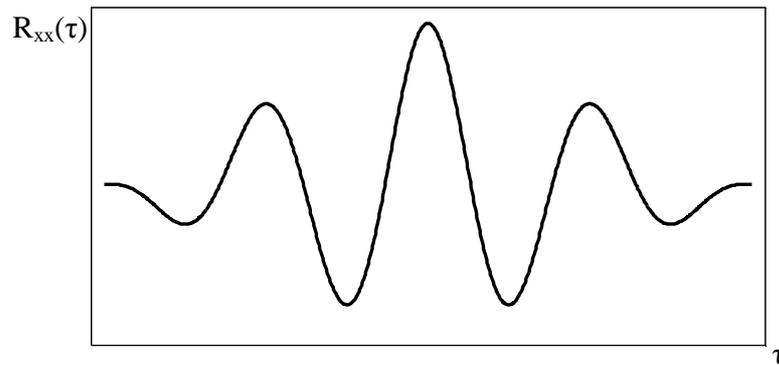


Figura 2.21. Función de autocorrelación de $\text{sen}(t)$

La gráfica es simétrica respecto al origen y presenta no solo un pico definido para el valor máximo sino que es notable la presencia de otros picos menos pronunciados. Estos valores máximos representan los instantes t donde hay mayor correlación entre las señales.

Se puede observar claramente que para un desplazamiento cero la gráfica alcanza el máximo valor de autocorrelación ($R_{xx}(0)$ es máximo), esto ocurre porque el área bajo el producto es máxima. Cuando se desplaza una versión de la señal a la mitad del periodo fundamental de la senoide no desplazada, los picos positivos de la función desplazada se

alinean con el pico negativo de la senoide subyacente. El producto para ese desplazamiento es negativo y se obtiene una gran área negativa. Cuando se produce un nuevo desplazamiento de la mitad del periodo fundamental, los picos positivos se alinean de nuevo, pero ahora los picos de los extremos opuestos de las dos versiones se multiplican por cero. Por consiguiente, aunque el área bajo el producto alcanza un pico positivo, este es más pequeño. Conforme crece el desplazamiento de los picos, tanto positivos como negativos, el área tiende a cero debido a la reducción del traslape entre las porciones distintas de cero de las señales. En la Figura 2.22 se ilustra esta situación.

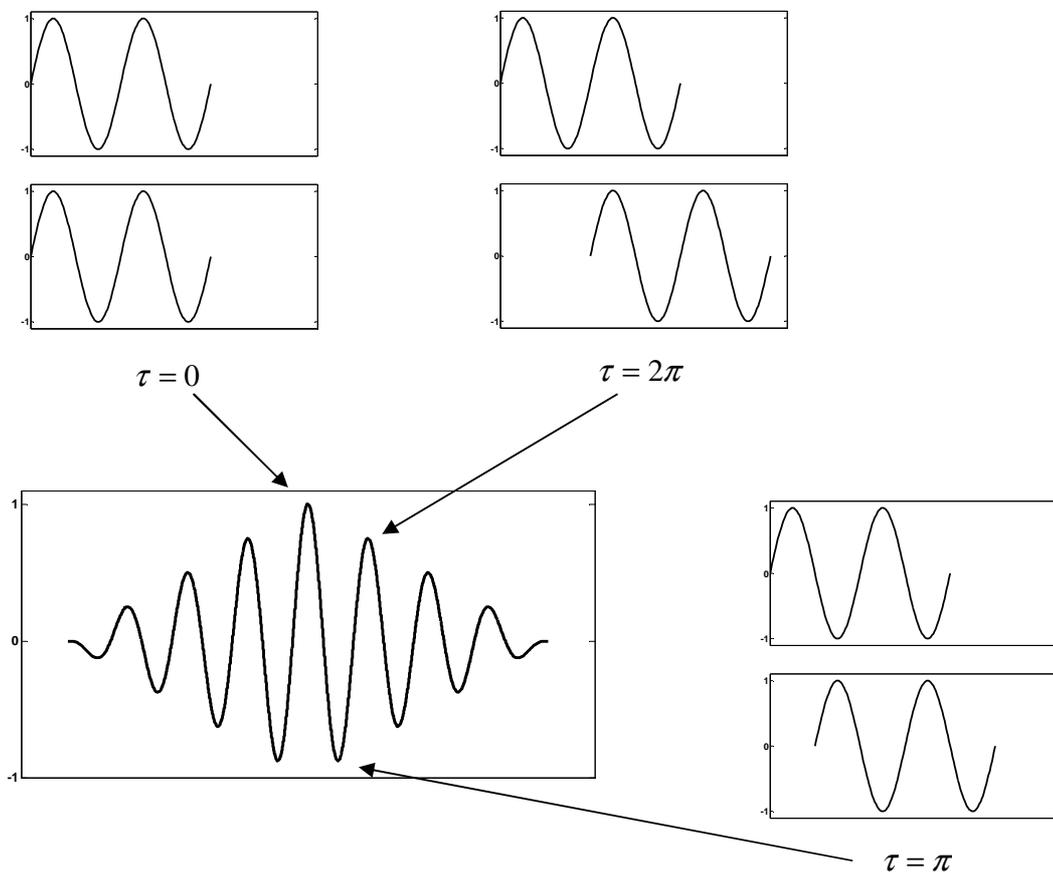


Figura 2.22. Valores de la función de autocorrelación para $\sin(t)$ para diferentes desplazamientos (τ)

Este método es la base para el diseño de un receptor que pretende reconocer una señal. Cuando se intenta detectar una señal, emitida por la antena de un radar, se recibirá en realidad una versión retardada en el tiempo de la misma. Mediante la detección del pico de la correlación de estas dos señales se podrá estimar el valor del desplazamiento donde esto ocurre, que será cuando ambas señales coincidan completamente. Esta etapa de análisis de la señal se realiza usando software apropiado que permite ejecutar numéricamente la operación de correlación. (Roberts, 2004)

Un término común para la función de correlación entre dos señales diferentes es correlación cruzada para distinguirse de la autocorrelación.

En la Figura 2.23 se muestra la correlación cruzada entre las funciones $x_1(t) = \sin(t)$ y $x_2(t) = \cos(t)$. La función de correlación muestra claramente un pico más pronunciado que el resto. Esto implica que las dos señales están correlacionadas para el desplazamiento donde se encuentra este pico.

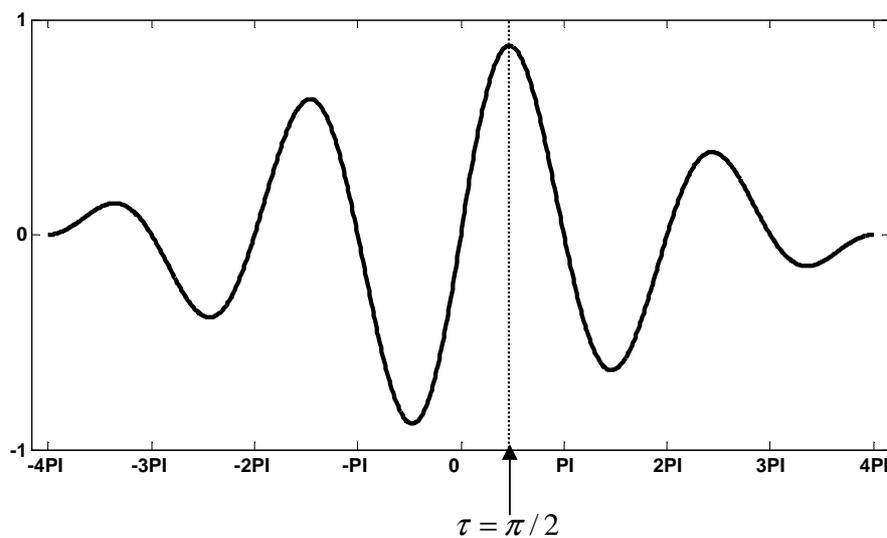


Figura 2.23. Función de correlación cruzada entre $\sin(t)$ y $\cos(t)$

Para el caso del ejemplo, la correlación cruzada es máxima para el desplazamiento $\tau = \pi/2$, lo que significa que si a una de las funciones se la desplaza en $\pi/2$ ambas funciones coincidirán. De hecho,

$$x_2(t) = \cos(t) = \text{seno}\left(t + \frac{\pi}{2}\right) = x_1\left(t + \frac{\pi}{2}\right)$$

Establecer la correlación entre dos señales de radar permite entonces discriminar el retardo temporal en el cual la correlación es máxima, y en consecuencia constituye una etapa en la detección del blanco.

2.5.2. Codificación

Durante el proceso de detección, el uso de técnicas de codificación de la señal transmitida favorecen a las señales que fueron codificadas, y puede llegar a suprimir aquellas señales que no lo fueron. El grado de supresión de las señales no codificadas depende de las características espectrales de las señales de interferencia y también del código usado (Ghebrebrhan, 2004).

Existen dos tipos de códigos, bifásicos y polifásicos. Los primeros, también llamados binarios, se basan en la modulación de una portadora usando dos fases discretas (0° y 180°). Los códigos polifásicos usan pequeños incrementos de fase para codificar la portadora. Los códigos bifásicos son los más usados en sistemas de radares.

Un código binario para implementar aplicaciones de radares es idéntico a una cadena de bits de datos excepto que está diseñado con un patrón específico. Estos códigos son tales que poseen un patrón con una característica deseable para su función de autocorrelación.

A los bits que componen un código binario se los llamará chips (y no bits) porque en el contexto de las comunicaciones por Spread-Spectrum es necesario emitir múltiples chips para transmitir un bit de información.

Entre los códigos que se propusieron y probaron para diferentes aplicaciones de radares se encuentran, código Barker (Barker, 1953), Secuencias Huffman (Huffman, 1962), Códigos de Convolución (Coll, 1961), M-codes (Sarwate y Pursley, 1980), Secuencias complementarias de Golay o código complementario (Golay, 1961), entre otros.

En muchos casos un criterio para seleccionar un tipo de esquema de codificación, es usar códigos que posean una función de autocorrelación con un lóbulo (máximos relativos presentes en la función de correlación) principal bien definido y con lóbulos laterales

mínimos. Sultzer y Woodman (1984) mostraron que, la energía presente en los lóbulos laterales de las correlaciones deben ser 20% o menos que la del lóbulo principal para alcanzar una buena calidad en la detección.

Sin embargo, la elección un esquema de codificación para un radar particular es una decisión de diseño asociada a las funciones que se espera que dicho radar posea, al blanco que se desea detectar, a la capacidad de resolución de rango que se desee, a la capacidad de emisión de altas potencias, etc.

En este trabajo se trataran los códigos binarios. Dos de los esquemas de codificación más utilizados son el esquema de codificación Barker y de Código Complementario.

2.5.2.1. Codificación Barker

El código Barker es un código binario con la propiedad de que el pico (lóbulo principal) de su función de autocorrelación es igual a la longitud del código. Además, los lóbulos laterales tienen una estructura tal que, en teoría, poseen la mínima energía posible distribuída uniformemente entre estos lóbulos. Por este motivo, muchas veces, el código Barker es denominado como código perfecto.

Si la longitud del código es N , y el pico de su función de autocorrelación es este valor, los lóbulos laterales poseen un valor unitario.

La Figura 2.24 muestra la gráfica de la función de correlación para el código Barker donde se puede observar la diferencia entre los valores de los lóbulos laterales y el lóbulo principal.

Las únicas secuencias de código Barker conocidas en la actualidad pueden ser de longitudes de 2, 3, 4, 5, 7, 11 y 13 bits. La Figura 25 muestra las secuencias de código Barker conocidas.

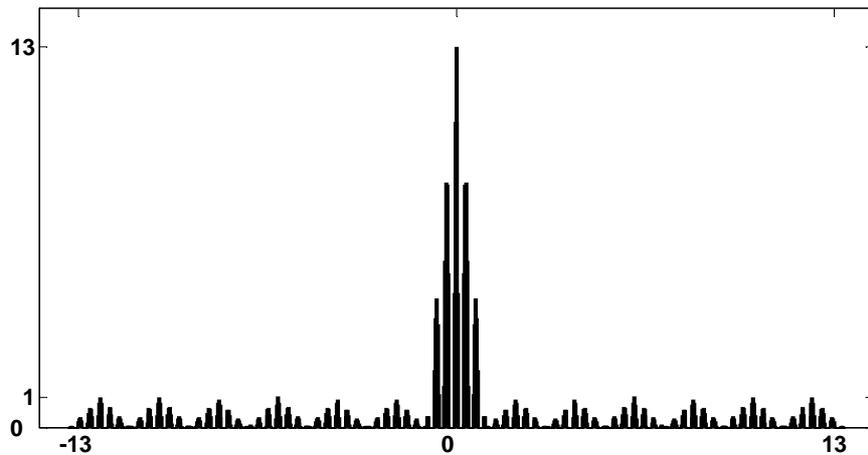


Figura 24. Función de autocorrelación de un código Barker de longitud de 13 bits

Long. código	Secuencias Barker
2	+ - , + +
3	+ + -
4	+ + - + , + + + -
5	+ + + - +
7	+ + + - - + -
11	+ + + - - - + - - + -
13	+ + + + + - - + + - + - +

Figura 25. Secuencias de código Barker

Como se mencionó anteriormente, el objetivo de utilizar la técnica de compresión de pulso es alcanzar un gran rango de resolución usando un pulso de larga duración. En general se utiliza el código Barker de mayor longitud que es de 13 bits. Sin embargo, las ondas moduladas con éste no siempre lograrían aprovechar completamente las ventajas de la compresión de pulso (Nathanson y Reilly, 1991). Para ello, muchas veces se recurre a otros esquemas de codificación más largos.

2.5.2.2. Código Complementario

El código complementario introducido por Golay en 1961, permite trabajar con secuencias largas de chips y mantener la ventaja de reducir los lóbulos laterales de su función de autocorrelación. Sin embargo, esta técnica de codificación solo se puede aplicar a blancos de gran tamaño, cuya velocidad de movimiento no sea alta en el período de tiempo que demora la emisión y recepción de los pulsos.

Si bien la técnica de codificación mediante código complementario se utiliza ampliamente en sistemas de radares, actualmente se usa, también, en otros sistemas de comunicaciones e incluso está especificado como un esquema de modulación en el estándar IEEE 802.11, para la tasa de transmisión de datos de 5.5 y 11 Mbps en la banda de 2.4 GHz.

Un código complementario binario es un subconjunto de una clase más general de códigos, conocida como códigos complementarios polifásicos.

Definición 1: (Ghebrebrhan, 2004) Un par de secuencias de código complementario consiste en dos secuencias de igual longitud finita C_0 y C_1 que se pueden representar como, $C_0 = \{c_0, c_1, \dots, c_N\}$ y $C_1 = \{c'_0, c'_1, \dots, c'_N\}$ tal que $c_i, c'_i \in \{1, -1\}$, $i = 1, 2, \dots, N$ donde N es el número de elementos de cada secuencia y determina la longitud del código. Estos pares de códigos complementarios tienen la importante propiedad que la suma de las funciones de autocorrelación de cada secuencia del par es cero para todos los desplazamientos, excepto para el desplazamiento cero, donde valen 2N (Figura 2.26) (Golay 1961).

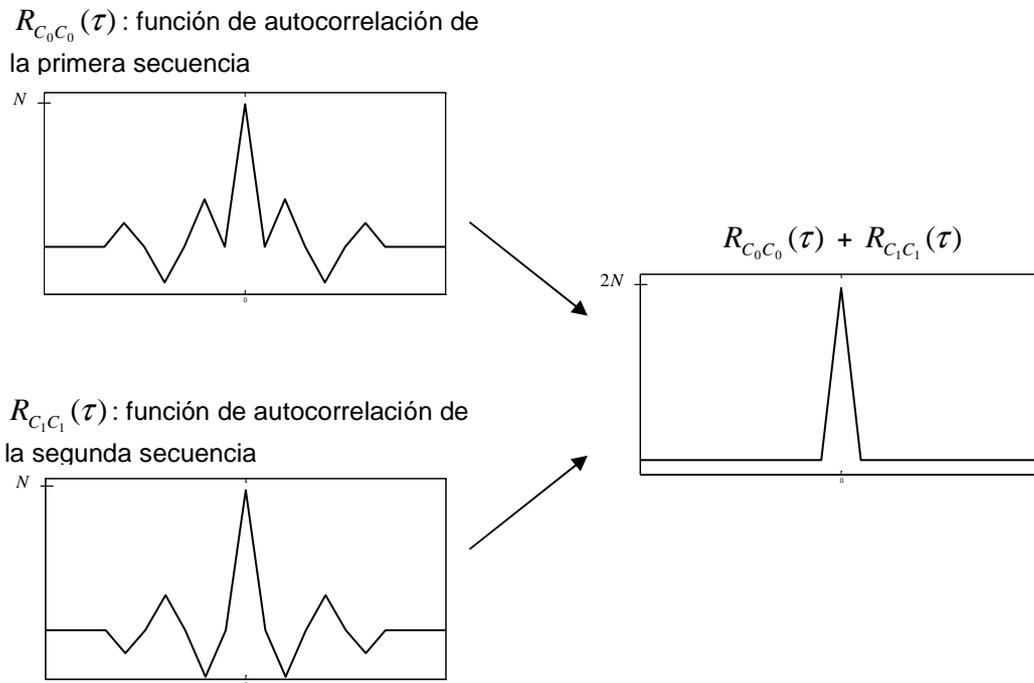


Figura 2.26. Función de autocorrelación de dos secuencias de código complementario

Definición 2: El Código Complementario, también llamado secuencia o serie complementaria binaria, esta compuesto por un par de secuencias binarias de longitud finita, que tienen la propiedad de que el número de elementos iguales para cualquier separación de bits en una de las secuencias es igual al número de elementos desiguales con la misma separación en la otra secuencia (Sivaswamy, 1978).

La simetría que implica la definición anterior no es intuitiva pero es fácil de ver a través de un ejemplo (Golay, 1061):

Dadas las secuencias $C_0 = \{-1, -1, -1, 1, 1, 1, -1, 1\}$ y $C_1 = \{-1, -1, -1, 1, -1, -1, 1, -1\}$

Secuencia C_0 -1 -1 -1 1 1 1 -1 1

$\underbrace{\quad}_{=}$ $\underbrace{\quad}_{=}$ $\underbrace{\quad}_{\neq}$ $\underbrace{\quad}_{=}$ $\underbrace{\quad}_{=}$ $\underbrace{\quad}_{\neq}$ $\underbrace{\quad}_{\neq}$

Secuencia C_1 -1 -1 -1 1 -1 -1 1 1

$\underbrace{\quad}_{=}$ $\underbrace{\quad}_{=}$ $\underbrace{\quad}_{\neq}$ $\underbrace{\quad}_{\neq}$ $\underbrace{\quad}_{=}$ $\underbrace{\quad}_{\neq}$ $\underbrace{\quad}_{\neq}$

Para una separación de un solo elemento, la primera secuencia cuenta con 4 pares de elementos iguales y 3 desiguales, mientras que la segunda secuencia cuenta con 3 pares de elementos iguales y 4 desiguales. Del mismo modo se puede mostrar que esta simetría se cumple para bits apareados para distintos valores de separación (2,3, etc).

Si bien la segunda definición nos da una visión de la estructura interna de las secuencias de código complementario, la primera definición muestra claramente las ventajas de este tipo de codificación.

Una observación importante respecto a ambas definiciones es que no especifican la longitud que debe tener el código. Se puede demostrar que aquellas secuencias de códigos complementarios, cuya longitud se pueda expresar como potencias de 2, (N=4, 8, 16, 32,...) elegida apropiadamente se denominan secuencias óptimas de código complementario. (Ghebrebrhan y Spano, 1996).

Golay introdujo varias maneras de construir secuencias de código complementario recursivamente e incluso una forma de construcción explícita para diferentes longitudes de códigos.

En principio se pueden generar de esta manera secuencias de código complementario para cualquier valor de n. Esto lleva a pensar que se pueden generar secuencias de código tan largas como se desee, aunque no siempre tendrán aplicación práctica.

Sin embargo, a los fines útiles de la codificación de señales de radar, basta con seleccionar una secuencia de código de una longitud adecuada que servirá como referencia al momento de correlacionar la señal recepcionada en el equipo.

Dicho esto, es claro porqué no es necesario contar con un generador de código complementario para obtener un par de secuencias nuevas por cada par de pulsos que se emita. Por esta razón, basta con tener un único par de secuencias de código complementario bien formadas.

Las secuencias C_0 y C_1 , son usadas para codificar dos pulsos consecutivos. La transmisión de estos pulsos se realiza de manera alternada, enviando primero un chip de la primera secuencia y a continuación un chip de la segunda secuencia. Este proceso se repite hasta que se envían los dos pulsos completos. La diferencia de tiempo entre los chips de las 2 secuencias se denomina distancia interpulso, mientras que la diferencia entre dos chips consecutivos de una misma secuencia se llama frecuencia de repetición de pulso.

En la Figura 2.27 se esquematiza el proceso de transmisión de ráfagas de código complementario.

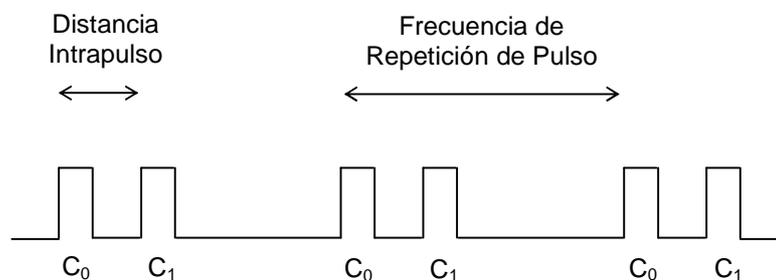


Figura 2.27. Esquema de la transmisión de un par de secuencias complementarias de longitud 3

Si la longitud del código es demasiado larga es posible que en el momento que retorne la primera ráfaga aún se estén enviando otros, lo que lleva a pérdida de información. Si la longitud, por el contrario, es chica se puede perder la ventaja que provee la técnica de compresión de pulsos. Entonces, es necesario, encontrar un equilibrio que aproveche todas las ventajas.

Como la longitud de un código puede influir en la capacidad de detección de cualquier radar pulsado, es de suma importancia el estudio de los diferentes esquemas de codificación.

CAPITULO III

El medio de estudio

3.1. Introducción

Cuando se aborda el diseño de cualquier tipo de instrumento, en nuestro caso un proceso algorítmico para detectar señales de un radar ionosférico en banda de HF, es necesario conocer el medio sobre el que actuará a fin de determinar cuales son los parámetros relevantes que describen el medio. Esto permitirá conocer los órdenes de magnitud de las respuestas esperadas del medio ante estímulos de las señales que se utilizaran para su estudio. En este caso particular el diseño se aplicará como parte de un procesamiento de señales para obtener respuestas de una parte de la alta atmósfera terrestre denominada Ionosfera. Las telecomunicaciones, en general, utilizan las propiedades refractivas de la ionósfera, en la que existen cargas libres como resultado de la absorción de la energía solar incidente por parte de las moléculas o átomos que afectan a la propagación de las ondas de radio (Zolesi y Cander, 2014, entre otros).

3.2. La ionósfera

Según la definición de la IEEE Std 211-1997, la ionósfera es una parte de la alta atmósfera terrestre donde iones y electrones están presentes en cantidades suficiente como para afectar la propagación de ondas electromagnéticas.

La ionósfera es la región de espacio localizada aproximadamente entre los 50 y 2000 km por encima de la superficie de la tierra y es la porción más alta de la atmósfera. Absorbe grandes cantidades de la energía radiante del sol. En este proceso se generan iones y electrones libres, además de absorber la radiación letal para la vida en la tierra. A pesar de que los mismos procesos físicos gobiernan en toda la ionósfera terrestre, existen diferencias significativas entre baja y altas latitudes debido principalmente a la orientación del campo magnético. Estas diferencias cobran mayor significado al

momento de desarrollar o poner a punto determinados equipos para realizar mediciones de parámetros ionosféricos.

3.2.1. Estratos ionosféricos

La ionósfera no es uniforme, existiendo gradientes de temperatura y densidad. Se puede clasificar en regiones que varían en altura y en densidad de ionización con la hora del día, fluctúan con un patrón cuasi cíclico todo el año y de acuerdo con la actividad solar, y efectos de circulación global (Rishbeth and Garriot, 1969). No existe una separación total entre regiones de la ionósfera, sino un cambio en la densidad de cargas, lo que marca la diferencia entre ellas. Así se puede clasificar las distintas regiones como sigue: la región D está comprendida entre los 60 y los 90 km de altura, el orden de magnitud de la densidad electrónica es de 10^2 a 10^4 cm^{-3} , la región E, entre 105 y 160 km con densidad electrónica del orden de 10^5 cm^{-3} y la región F sobre los 180 km, con densidad del orden de 10^5 a 10^6 cm^{-3} . Esta última región tiene un pico de densidad máxima a una altura del orden de los 300 km. Sin embargo la densidad de electrones posee una fuerte variabilidad debido a diferentes factores lo que normalmente altera estos valores típicos. La Figura 3.1 muestra un esquema de la estratificación de la ionósfera. El gráfico representa el perfil de densidad electrónico de la ionósfera como la altura en función de la densidad electrónica. Se pueden observar dos perfiles típicos, correspondientes al día y a la noche. Se puede apreciar claramente una variación en la estructuración alcanzando valores máximos locales de densidad electrónica diferentes de acuerdo a la hora del día. Se pueden observar, entonces, diferentes regiones rotuladas.

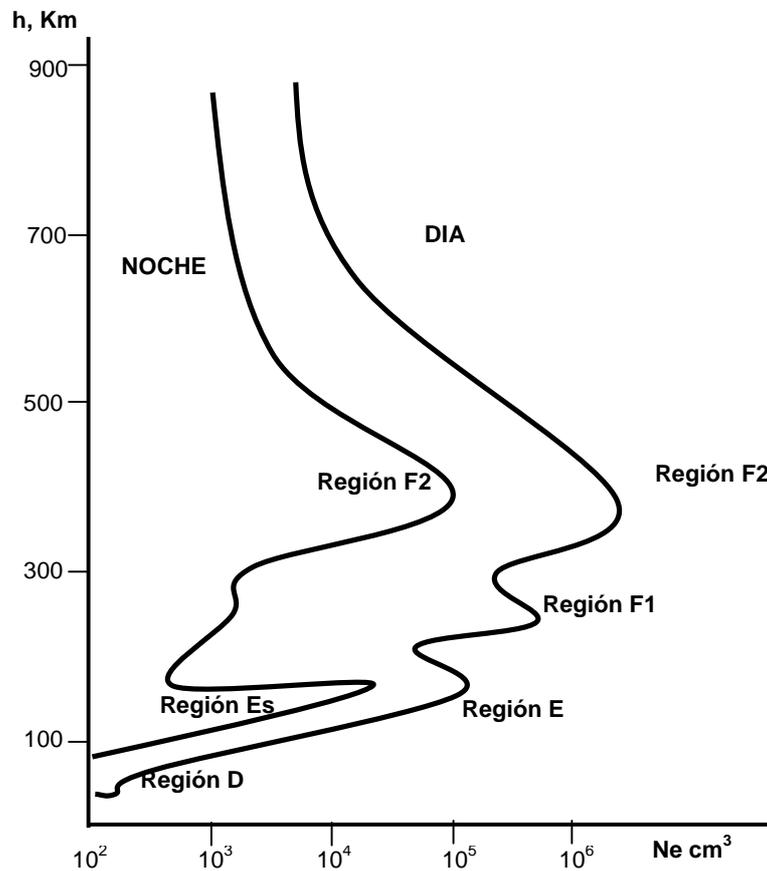


Figura 3.1. Morfología de la ionósfera diurna y nocturna. Se muestra la altura en función a la densidad de electrones.

La estructura y dinámica de la ionósfera terrestre esta sujeta tanto a grandes cambios espaciales y a variaciones temporales frecuentes, los cuales pueden ser tanto periódicos tanto como irregulares. Los cambios que ocurren en la ionósfera son diferentes a diferentes altitudes debido a la variación relativa a ionización, pérdida y fenómenos de transporte. En consecuencia, la ionósfera terrestre se puede dividir en regiones regulares y regiones esporádicas de acuerdo a las propiedades termales y químicas de los componentes gaseosos neutrales e ionizados.

3.2.2. Regiones ionosféricas regulares

Las investigaciones modernas experimentales y teóricas dividen a la ionósfera en tres regiones: D, E y F. Bajo algunas condiciones de interacción solar-terrestre, estas regiones se subdividen en: D, E, F1 y F2 (Figura 3.2).

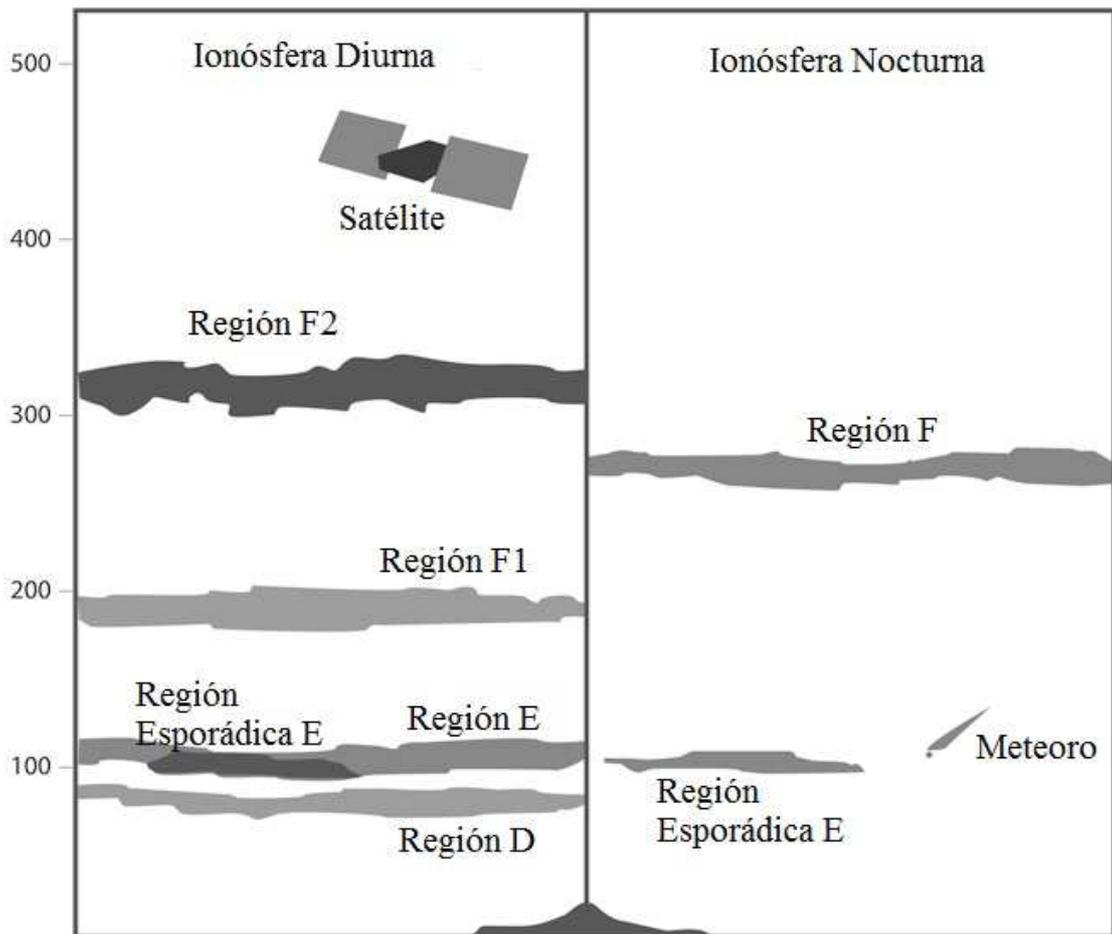


Figura 3.2. Gráfico comparativo de la estructura de la ionósfera diurna y de la ionósfera nocturna (Zolesi y Cander, 2014).

En principio, la baja ionósfera (hasta los 100 km) es la zona donde los procesos fotoquímicos tienen mayor influencia en su formación y en el balance de ionización. Las fronteras de la ionósfera media (100-170 km) marcan el límite de procesos de ionización-recombinación junto a procesos termales y dinámicos. La ionósfera superior es la denominada región F, caracterizada por la transferencia de partículas cargadas en el plasma por difusión ambipolar, vientos termosféricos, e interacción ionósfera-magnetosfera. Las alturas reales de las capas ionosféricas varían con el tiempo del ángulo zenith solar, momento del día, estaciones, ciclos solares y actividad solar.

La región F es el mayor segmento de la ionósfera terrestre y la más importante desde el punto de vista de las comunicaciones por radio y de los sistemas de navegación. Se encuentra entre los 140 y los 600 km de altura que ocasionalmente se extienden por arriba de los límites de la ionósfera. Como resultado a los complejos mecanismos físicos involucrados en su formación, la radiación solar puede hacer que la región F se divida en dos estratos denominados capas F1 y F2. La capa F2 esta siempre presente durante la noche, de modo que ambas capas se fusionan unas horas después del atardecer, reorganizando la región F.

La Figura 3.3 muestra los valores de densidad de electrones típicos para diferentes alturas, así como también los elementos constituyentes dominantes en cada estrato.

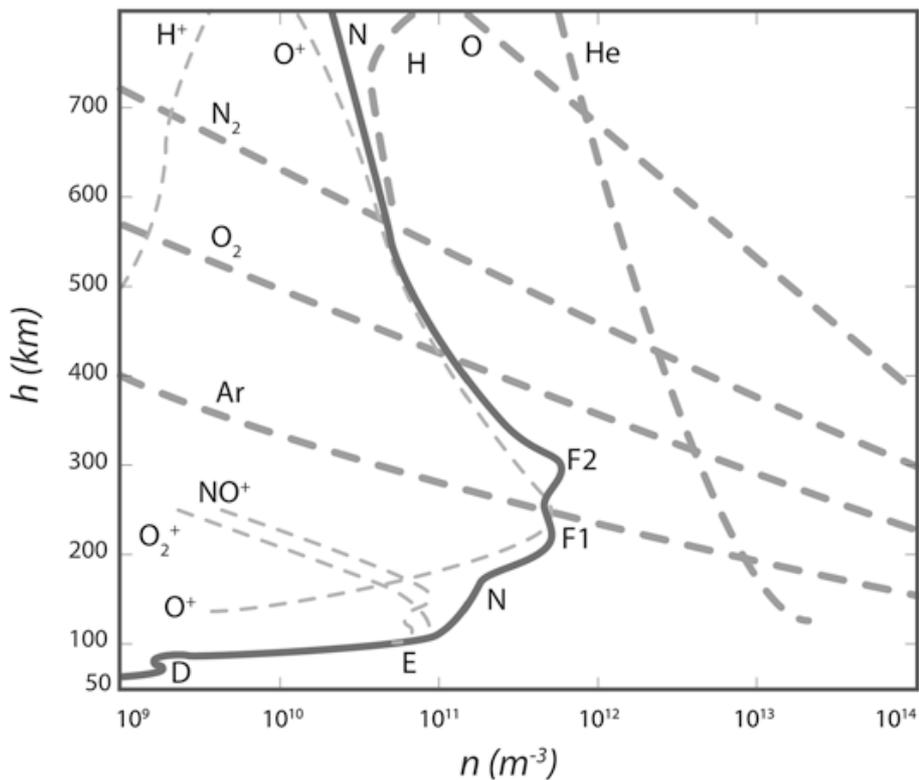


Figura 3.3. Valores típicos de densidad de electrones en cada región de la ionósfera. Se muestran además las especies dominantes en cada capa. N corresponde a electrones libres. (Dominici, 1998)

3.2.3. Región ionosférica esporádica

La capa esporádica E o Es se genera aproximadamente entre los 90 y 140 km y usualmente es considerada independiente a la región normal E de la ionósfera. La razón principal es que la capa Es tiene patrones muy diversos tanto espacialmente (puede estar muy esparcida en un área o por el contrario puede encontrarse confinada a una pequeña área), como patrones diurnos (puede aparecer tanto de noche como de día), y patrones

estacionales en alta, media o baja latitud generada por diversos procesos físicos. Una capa Es en baja o media latitud ocurre generalmente de día y es mas frecuente durante el verano. En altas latitudes aparece mayormente durante la noche y esta asociada generalmente a la ocurrencia de auroras. Una de las características más importantes de la capa Es es que posee una densidad electrónica similar a la capa F. Sin embargo, su ocurrencia aleatoria en un lugar determinado, hace difícil la predicción de su ocurrencia.

3.3.3. Irregularidades ionosféricas

En general existen irregularidades ionosféricas que incluyen condiciones que no pueden ser descritas de manera precisa por los modelos ionosféricos estándar y por eventos que no siguen patrones convencionales. Existen dos fenómenos muy estudiados, entre otras irregularidades, conocidos como “perturbaciones ionosféricas viajeras” (Travelling Ionospheric Disturbances - TIDs) y “esparcimiento de la capa F” (SpreadF). Estos fenómenos afectan la región F de la ionósfera.

Las TIDs juegan un rol predominante en la dinámica de la termosfera y la ionósfera, y son irregularidades de la región F expresadas como oscilaciones similares a ondas del contorno de una región de densidad de electrones constante a medida que transcurre el tiempo. Un ejemplo son las ondas de gravedad en la alta atmósfera forzadas desde el sistema meteorológico de la estratósfera y la tropósfera (Hines, 1960; Zolesi y Cander, 2014).

En algunas ocasiones la región F se encuentra difusa y diversas irregularidades dispersan las ondas de radio incidentes. Esta irregularidad se conoce como SpreadF.

3.2.4. Variabilidad espacial y temporal de la ionósfera

La estructura de la ionósfera presenta una gran variabilidad diaria debido a la variación del ángulo cenital solar. Sin embargo no es la única variabilidad presente en la estructura de la ionósfera, sino que la misma es altamente variable tanto en espacio como en tiempo. Esta morfología afecta la propagación mediante la ionósfera y está caracterizada por

variaciones geográficas y geomagnéticas por un lado, y variaciones diarias, estacionales y del ciclo solar por otro. Las variaciones diurnas, estacionales, con el ciclo solar y en altura, pueden ser deducidas del monitoreo de la ionósfera para una determinada posición geográfica.

El comportamiento de la ionósfera desde el punto de vista de regiones geográficas se puede dividir en:

Alta latitud: donde la latitud geomagnética es de $90^\circ > \lambda > 60^\circ$ a cada lado del ecuador magnético.

Latitud media: localizado en latitudes geomagnéticas de $60^\circ > \lambda > 20^\circ$ aproximadamente a ambos lados del ecuador magnético

Baja latitud: donde la latitud geomagnética es $0^\circ < \lambda < 20^\circ$, también a ambos lados del ecuador magnético.

Hay que tener en cuenta que normalmente el ecuador magnético y el ecuador geográfico no coinciden.

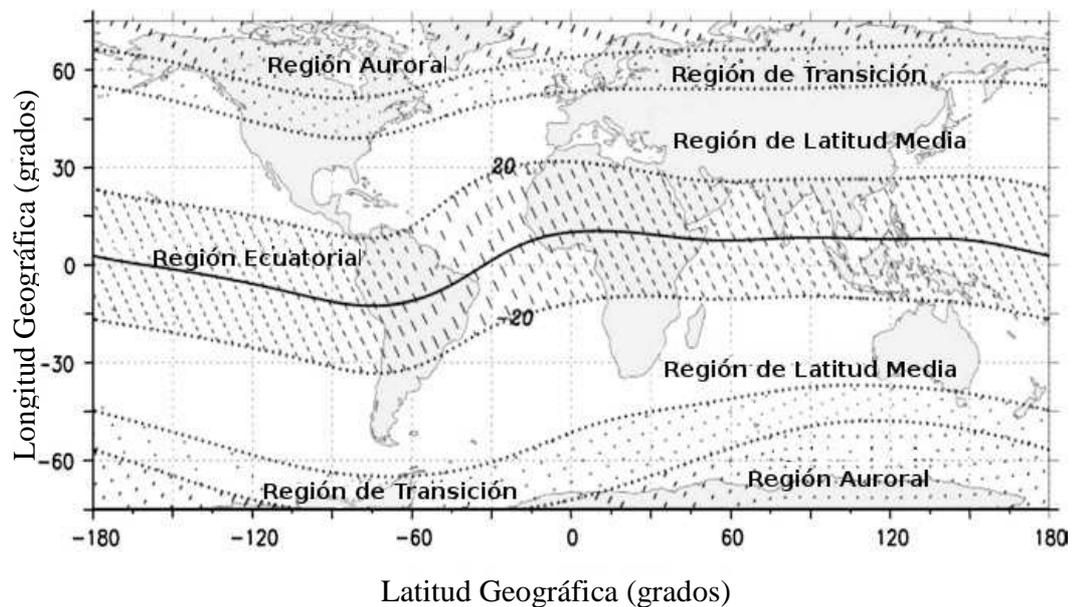


Figura 3.4. Principales regiones geográficas de la ionósfera (Zolesi y Cander, 2014).

Algunas de estas regiones se subdividen, a su vez, en subregiones. Tal es el caso de los denominados casquetes auroral y polar en el caso de altas latitudes, y para bajas latitudes la subdivisión es en región ecuatorial y región de anomalía ecuatorial.

La figura 3.4 muestra la extensión geográfica aproximada de las principales regiones bajo condiciones ionosféricas tranquilas.

Durante condiciones geomagnéticas perturbadas la región auroral se extiende en dirección al ecuador, reduciendo así el tamaño de la región de latitud media. Usualmente la actividad auroral es un buen indicador del incremento de la conexión entre la magnetósfera terrestre, ionósfera y atmósfera.

En general la ionósfera de latitudes medias es menos compleja que otras regiones bajo condiciones no perturbadas y se caracteriza por gradientes de variación espacial relativamente pequeños.

En altas latitudes prevalecen irregularidades con una marcada ionización y gradientes asociados.

Esta clasificación se refiere a condiciones de ionósfera no perturbada. Además suelen existir irregularidades transitorias originadas en el sistema sol-tierra que están asociadas a eventos de meteorología del espacio y que perturban a la ionósfera en sus diferentes estratos.

Los efectos de fenómenos de origen solar y su actividad geomagnética, magnetosférica e ionosférica asociada en plasmas cercanos a la tierra, varían ampliamente en importancia, escalas de tiempos, predictibilidad y consecuencias.

Estos efectos pueden ser:

1. Efectos directos, causados por rápidos cambios en la radiación solar UV y de rayos X que iluminan la ionósfera y atmósfera terrestre durante fulguraciones solares.
2. Efectos indirectos, causados por complejas interacciones entre el viento solar y el sistema acoplado magnetósfera-ionósfera-atmósfera.

El análisis exhaustivo de las perturbaciones ionosféricas regulares tanto como las originadas por eventos de meteorología del espacio se encuentra fuera del alcance de este trabajo de tesis.

3.4. Las comunicaciones y la ionósfera

En los sistemas de comunicación de radio, las ondas se pueden propagar de varias formas, dependiendo del tipo de sistema y del ambiente. Las ondas electromagnéticas sólo se propagan en línea recta en un medio isotrópico no dispersivo, así esta trayectoria puede ser alterada dependiendo del medio. Estas desviaciones en la trayectoria de la señales de radio influyen directamente en las radiocomunicaciones.

Las ondas electromagnéticas que se propagan a través de la ionósfera sufren un proceso de refracción. Cuando una onda de radio incide en la ionósfera, el campo eléctrico de la onda ejerce una fuerza en los electrones libres, haciéndolos que oscilen en modo forzado. Los electrones oscilantes reemiten esa onda y hacen que la trayectoria de las ondas electromagnéticas se curven alejándose de las regiones de alta densidad de electrones, hacia regiones de baja densidad de electrones. Entre más alta sea la densidad de iones, mayor será la refracción que sufre la onda de radio (Tomasi, 1996).

Cuando una señal es emitida hacia la ionósfera, en cada instante la señal sufre tres situaciones, parte de la energía de la señal es reflejada y devuelta a la tierra y parte es absorbida en el medio. Cuando la frecuencia de la onda incidente es igual a la frecuencia del plasma ionosférico, la reflexión es máxima.

En términos prácticos, y de acuerdo a la frecuencia de la onda incidente, puede ocurrir que:

La señal de radio es completamente absorbida por la ionósfera.

La señal es reflejada parcialmente hacia la tierra en algún estrato ionosférico

La señal de radio atraviesa la ionósfera hacia el espacio exterior.

La Figura 3.5 ilustra estos tres casos.

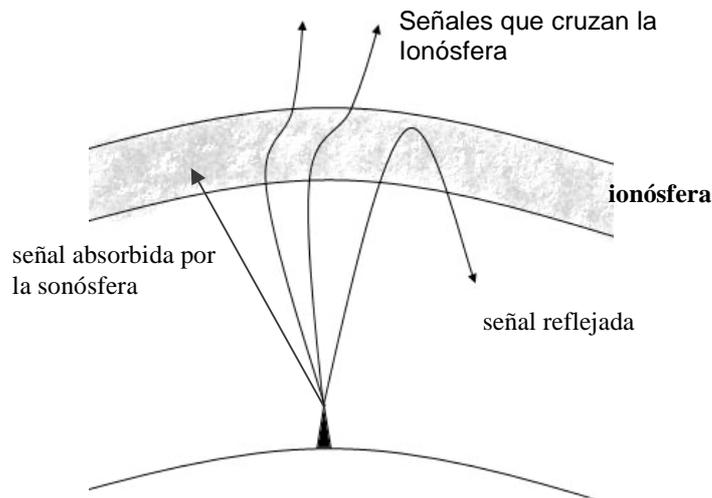


Figura 3.5. Casos de incidencia de ondas de radio en la Ionósfera

Dependiendo de cuán ionizada esté la ionósfera, las comunicaciones se ven afectadas en mayor o menor medida. Muchos sistemas de comunicación utilizan la ionósfera como medio para transmitir ondas de radio a grandes distancias. Algunas frecuencias de radio son absorbidas (desvanecimiento) y otras reflejadas, lo que hace que las señales fluctúen con rapidez y que sigan rutas de propagación inesperadas, esto se denomina multicamino. Las comunicaciones aero-terrestres, barco-puerto, las estaciones de radio de onda corta y algunas frecuencias de radio aficionados se ven afectados frecuentemente. Un ejemplo concreto son estos efectos sobre la aviación comercial. En casos de eventos extremos se producen grandes apagones en las comunicaciones de radio que afectan por ejemplo a aeronaves en tránsito que se ven impedidas de transmitir o recibir comunicaciones en HF en tierra por periodos incluso de horas. Además se producen efectos de fluctuaciones en las fases de las señales de comunicación en la banda L, como la de los sistemas de posicionamiento global, GNSS (Global Navigation Satellite System) que pueden llevar a una determinación equivocada de la posición de la aeronave.

3.5. Métodos de estudio y monitoreo de la ionósfera terrestre

El impacto que tiene el estado de la ionosfera terrestre en las comunicaciones y en consecuencia en el desarrollo de la vida humana actual, ha motivado a desarrollar teorías, métodos y sistemas para su estudio.

Existen dos metodologías complementarias para llevar a cabo estos estudios: (a) Realizar el modelado tanto de la estructura como la dinámica de la ionosfera. O (b) realizar mediciones sistemáticas para monitorear el estado de la alta atmósfera utilizando diversos instrumentos que permiten tanto mediciones directas o indirectas.

Tradicionalmente el objetivo era el estudio de corto plazo, con el objetivo de encontrar tendencias y características a corto plazo, y análisis de tendencias de largo plazo para investigar variaciones en grandes períodos de tiempo.

Actualmente la meteorología del espacio requiere de metodologías de monitoreo que permitan predecir comportamientos de los parámetros ionosféricos y sus consecuencias en las actividades tecnológicas. Estas predicciones pueden ser a corto plazo, denominada “nowcasting”, o predicciones de más largo plazo, llamadas “forecasting”. En consecuencia, tanto los modelos como el instrumental de monitoreo se encuentran en evolución constante. Los primeros para obtener una descripción de las condiciones ionosféricas futuras en base a la obtención de datos en tiempo cuasi-real de diversas fuentes instrumentales. Esto es posible gracias a que de las nuevas tecnologías de computación permiten resolver complejas simulaciones numéricas de forma cada vez más veloz. Junto a la posibilidad de tener datos disponibles y confiables en la web, numerosos centros de monitoreo de meteorología del espacio cuentan con posibilidad de generar alertas de diferentes fenómenos.

3.5.4. Mediciones ionosféricas

Las técnicas mas importantes usadas sistemáticamente para el monitoreo de los estratos ionizados de la alta atmosfera terrestre se basan en los efectos que esta produce sobre la propagación de ondas de radio. Tanto ionosondas y otros sistemas de radares de alta frecuencia, ya sea que actúen vertical u oblicuamente, son los sistemas mas ampliamente

usados para mediciones de parámetros ionosféricos. A estos sistemas tradicionales actualmente se los complementa con equipos GNSS (Global Navigation Satellite System) utilizando receptores en tierra. Además existen sistemas de radar de dispersión coherentes HF (SuperDARN) y VHF (STARE) así como también sistemas de radar de dispersión incoherentes. Desde el comienzo de la exploración espacial con cohetes y satélites, numerosas misiones han contribuido al estudio de la ionósfera con mediciones in-situ de las propiedades del plasma bajo diferentes condiciones solares-terrestres (Zolesi and Cander, 2014; entre otros).

En particular uno de los métodos de medición tradicional se basa en radares HF que permiten mediante sondaje vertical u oblicuo observar la morfología de la ionósfera. Estos instrumentos se encuentran instalados ampliamente en el globo y producen las mediciones con mayor disponibilidad, organizados siguiendo reglas estándares (Zolesi y Cander, 2014).

Este tipo de medición se basa en el principio de que una onda electromagnética que se propaga verticalmente en el plasma ionosférico alcanza la máxima altura de reflexión en el nivel en que la frecuencia de la señal iguala a la frecuencia natural del plasma ionosférico (Rishbeth y Garriot, 1969).

En este trabajo de tesis se propone un modelo numérico para estimar automáticamente la altura a la cual se produce esta reflexión para el caso de sondaje vertical. Este modelo permitirá obtener mediciones más confiables y robustas, en especial teniendo en cuenta la alta variabilidad ionosférica. Este modelo tendrá un especial impacto en las mediciones que normalmente utilizan los diferentes modelos. Este aporte, a su vez, permitirá estudiar y pronosticar con mayor exactitud las condiciones ionosféricas para aplicaciones operacionales en meteorología del espacio.

CAPITULO IV

Detección de Ecos de Radar

4.1. Introducción

Como se mostró en capítulos precedentes, la función primaria de un radar es la de detectar la presencia de un blanco mediante la discriminación del eco que este produce dentro de una señal atenuada, afectada por ruido o interferencia y retardada en el tiempo.

En términos de la tecnología radar, detección se refiere a la capacidad de decisión del sistema para determinar si una determinada medición corresponde a un eco de un blanco o solo es producto de ruido o interferencia.

Diferentes funciones colaboran con el proceso de detección y pueden aplicarse en diferentes etapas del procesamiento de las señales.

El proceso de detección puede ser aplicado en señales recibidas sin procesamiento previo tanto como en datos ya procesados tales como espectros Doppler o incluso a imágenes de radares de apertura sintética.

En definitiva la función de detección no se realiza como una única etapa sino que refiere a diversas técnicas en diferentes instantes del procesamiento de las señales recepcionadas. Estas técnicas varían ampliamente según el tipo de radar, las características del blanco y de acuerdo al tipo de aplicaciones del equipo.

Obtener un modelo de detección significa determinar qué técnicas y en que instantes se realizarán para un sistema de radar específico. Sin embargo, comúnmente en la bibliografía la detección se refiere a la etapa final cuya salida es una respuesta binaria, siendo positiva si se detecta la presencia de un blanco o negativa en caso contrario.

Como se vio en el capítulo III, tanto las señales provenientes de ecos como de interferencia son mejor descritas mediante modelos estadísticos. En consecuencia, la detección se convierte en un problema de carácter estadístico denominado “prueba de hipótesis estadística”.

La capacidad de *detección automática* del sistema viene dada por la implementación de algoritmos específicos acordes a diferentes criterios. Estos algoritmos reciben como entrada las señales recepcionadas en el radar, estiman un valor de umbral (contra el cual se comparará la amplitud de cada muestra) para finalmente dar una respuesta binaria (hay un blanco presente o no) como salida.

En este capítulo se presenta teoría sobre detección y las estrategias que llevan al concepto de “umbral”. Se describirán, además, los principales algoritmos tradicionales diseñados para estimar estos valores de umbral de acuerdo a las características de las señales involucradas.

4.2. Detectores

Tradicionalmente se llamaba detector al componente electrónico que realizaba las tareas de detección en un sistema radar. En la actualidad debido a la gran capacidad de procesamiento de las computadoras modernas la mayor parte de las funciones del detector se realiza mediante software implementado en computadores de propósitos generales. Sin embargo, en muchos casos aun se denomina detector a aquel componente ya sea de hardware o software que implementa la función de detección (Barton and Leonov, 1990; Skolnik, 2001; entre otros).

Varios autores han investigado diversos detectores y comparado su desempeño en detección para el caso óptimo. Estudiar el caso óptimo significa, establecer como hipótesis que se cuenta con un ambiente conocido y homogéneo. En escenarios más realistas, estos detectores no siempre logran un rendimiento aceptable y entonces es necesario recurrir a nuevas estrategias.

En general se utilizan tres enfoques:

- a. Umbral adaptativo (o *Adaptive Thresholding*)
- b. Detectores no paramétricos
- c. Mapa de Clutter

Se supone que cada señal está compuesta por un determinado número de muestras, cada una de las cuales puede ser una amplitud de eco dispersado por un blanco, interferencia, ruido, clutter o la combinación de cualquiera de estos casos.

Tanto los detectores de umbral adaptativo como no paramétricos se basan en la hipótesis de que la homogeneidad existe en pequeñas regiones alrededor de la muestra que se encuentre evaluando. En el caso (a), el método asume que la densidad del ruido es conocida excepto por algunos parámetros como por ejemplo el promedio o la varianza. De este modo, el objetivo principal recae en obtener una estimación aproximada de estos parámetros para determinar el valor del umbral, que en consecuencia irá variando de acuerdo a la muestra que se utilice. En el caso de detectores no paramétricos (b), se supone que todas las muestras son independientes entre sí. La estrategia a seguir es la de reordenar las muestras para estimar el umbral. Las muestras son analizadas en una ventana de tamaño finito que irá variando a lo largo de la señal y estimará en cada paso un valor de umbral diferente. En ambos casos (a y b), cada muestra (amplitud) que exceda el valor del umbral se interpreta como la amplitud correspondiente a un blanco (Skolnik, 2001).

Se denomina clutter a las señales de fondo que no representan un blanco. Por ejemplo en un radar meteorológico, para la medición de precipitaciones, una montaña representaría clutter. Es decir, para una onda emitida por el radar la montaña dará una respuesta en energía que no es de interés al equipo, que no varía con el tiempo y por lo tanto puede filtrarse. Es claro que determinar qué representa o no clutter depende fuertemente del tipo de radar y de lo que se pretenda detectar con éste. En algunos casos lo que es clutter para algún radar es blanco para otro. Los mapas de clutter se generan almacenando los niveles de amplitud del ambiente (de la montaña en el ejemplo anterior) y a partir de estos valores es posible discriminar si es un blanco o no (Skolnik, 2001).

El objetivo general es poder construir un detector óptimo que no produzca detecciones no válidas (falsas alarmas) y que siempre sea capaz de detectar él o los blancos. Para obtener un detector óptimo es necesario estudiar la teoría de detección estadística a partir de la evaluación de una prueba de hipótesis binaria. Si se utilizan cualquiera de los tres enfoques es necesario estudiar previamente la teoría de detección subyacente a fin de

evaluar que tan próximo a un detector óptimo se encuentran estas estrategias para un determinado radar (Richards et al, 2010).

4.3. Fundamentos de detección

El primer punto a tener en cuenta es que una señal receptionada por el radar está compuesta tanto de dispersiones provenientes de un blanco, de ruido, o de interferencia. La primera hipótesis es que ambos casos provienen de procesos aleatorios, y por lo tanto la detección se puede tratar como un proceso estadístico (Curry, 2005).

Los modernos sistemas de radar muestrean las señales analógicas obteniendo señales discretas de tiempo que forman conjuntos de datos (o arreglos de datos), los que posteriormente son procesados. Matemáticamente, asumimos que señal discreta está compuesta por M puntos o muestras $\{y_0 = y(0), y_1 = y(1), \dots, y_{m-1} = y(M - 1)\}$.

Esto significa que el problema de decisión se basa en el análisis de una serie de tiempo discreto. El problema de decisión sobre una señal evolucionó en un problema de decisión basada en datos finitos, lo que significa que puede estar sujeto a la evaluación de una hipótesis estadística (Kay, 2013).

Para cada medición proveniente de un sistema radar, se desea realizar la prueba de si ésta corresponde o no a un blanco. Esto significa tomar la decisión entre dos posibilidades y en consecuencia se transforma en un problema de prueba de *hipótesis binaria* (Kay, 1998).

Por lo tanto, se puede asumir que una de dos hipótesis siguientes es verdadera:

1. La medición es el resultado de ruido o interferencia únicamente.
2. La medición es una combinación resultante de ruido o interferencia y además de ecos provenientes de algún blanco

La primera hipótesis se denomina hipótesis nula y se denota con H_0 y la segunda se denomina H_1 . Si la prueba indica que se verifica la hipótesis nula, entonces el sistema declara que no existe ningún blanco en la medición. Mientras que si H_1 es verdadera, el

sistema declara la presencia de un blanco en el rango, ángulo o coordenada Doppler de la medición (dependiendo del tipo de radar) (Richards, 2005).

Decidir entre alguna de estas hipótesis puede llevar a uno de los siguientes casos de respuestas por parte del detector (Figura 4.1):

	H_0	H_1
No hay blanco	Sin detección	Falsa alarma
Hay blanco	Fallo	Detección positiva

Figura 4.1. Cuadro con posibles respuestas para cada una de las hipótesis

Si no existe un blanco presente en la señal, la respuesta óptima del detector debe ser la elección de la hipótesis nula. De lo contrario se produce un error en la detección y en consecuencia el sistema detecta un blanco que no es tal (falsa alarma).

Si por el contrario hay un blanco presente, la respuesta óptima es una detección positiva (Elegir la hipótesis H_1) y en caso de que no sea así se falló en la detección.

Debido a que las señales son mejor descritas estadísticamente, la decisión entre las dos hipótesis resulta en un trabajo estadístico de teoría de decisión (Kay, 1998). El análisis comienza con la descripción estadística de la función de densidad de probabilidad (pdf de su *nombre en inglés*) que describe la medición que será probada bajo estas dos hipótesis. Supongamos que la muestra sujeta a prueba se denota con y , es necesario obtener las siguientes funciones de densidad de probabilidad:

1. La función de densidad de probabilidad de y cuando el blanco no esta presente
2. La función de densidad de probabilidad de y cuando el blanco si se encuentra presente

Asumiendo que estas dos funciones de densidad de probabilidad pueden ser modeladas con éxito, se pueden definir las siguientes probabilidades:

P_d = Probabilidad de detección: Es la probabilidad de que un blanco sea declarado (esto significa que la hipótesis H_1 es verdadera) cuando el blanco esta realmente presente en la muestra.

P_{fa} = Probabilidad de falsa alarma: Es la probabilidad de que un blanco sea declarado por el sistema (lo que equivale a que la hipótesis H_1 es verdadera) cuando en realidad el blanco no esta presente.

P_m = Probabilidad de fallo: Es la probabilidad que el sistema no se detecte un blanco que si esta presente (H_0 es verdadera).

Una observación importante es que $P_m = 1 - P_d$, y por lo tanto obtener las probabilidades P_d y P_{fa} son suficientes para obtener toda la información de probabilidades necesarias.

El siguiente paso es seleccionar las reglas que se utilizaran para tomar la mejor decisión para la elección entre las dos hipótesis. En el ámbito de los radares, la manera mas común de llevar a cabo este trabajo es mediante el denominado “criterio de Neyman-Pearson”. Bajo este criterio de optimización, el proceso de decisión es ideado de tal manera que permita maximizar la probabilidad de detección P_d bajo la restricción de que la probabilidad de falsa alarma P_{fa} no exceda un valor constante predefinido (Richards, 2005; Skolnik, 2001, entre otros). La combinación de P_d y P_{fa} factibles esta afectada por la calidad del sistema de radar y el diseño del procesador de señales. Se puede probar que para un determinado diseño de radar, al incrementar la P_d también se incrementa la P_{fa} (Richards, 2005). De manera intuitiva, aumentar la P_d significa que el sistema aumenta el numero de ecos detectado pero no garantiza que algunos de ello no sean falsas alarmas aumentando así la P_{fa} .

En general el diseñador del sistema establece cual es el valor aceptable de P_{fa} para un determinado equipo. Algunos valores estándares de P_{fa} en sistemas de radares varían

entre 10^{-4} y 10^{-8} (Richards, 2005; Curry, 2005). A pesar de ser valores muy pequeños, esto no garantiza la completa ausencia de falsas alarmas. Una vez más, estos valores deben ser ajustados al tipo de blanco que se desea detectar y cuan preciso se requiere que sea el sistema.

En resumen, el criterio de Neyman-Pearson está motivado por el objetivo de obtener la mejor performance de detección posible (no solo se quiere detectar más, sino que las detecciones correspondan a blancos que efectivamente están presentes); mientras que se garantice que la P_{fa} no exceda un valor tolerable (α). Así, la regla de decisión de Neyman-Pearson es:

Maximizar P_d sujeto a que $P_{fa} \leq \alpha$

Se puede probar que es posible obtener un valor de umbral λ acorde al valor deseado de α (Richards, 2005). De esta manera pueden ocurrir dos situaciones para una determinada muestra:

La amplitud de la muestra supera el valor del umbral, y por lo tanto se declara la presencia de un blanco en dicha muestra (Se escoge la hipótesis H_1). Idealmente, la detección cumplirá además con que la P_{fa} este acotada por el valor de α .

La amplitud de la muestra no supera el valor del umbral y entonces *no* se declara la presencia de un blanco. Esto significa que la muestra es producto del ruido o de interferencia y corresponde a la elección de H_0 .

La elección de un valor de umbral puede afectar considerablemente el desempeño en detección de un equipo. Supongamos que se obtienen las probabilidades de detección y de falsa alarma y ambas poseen una distribución gaussianas. La Figura 4.2 muestra diferentes condiciones de detección de acuerdo al umbral elegido.

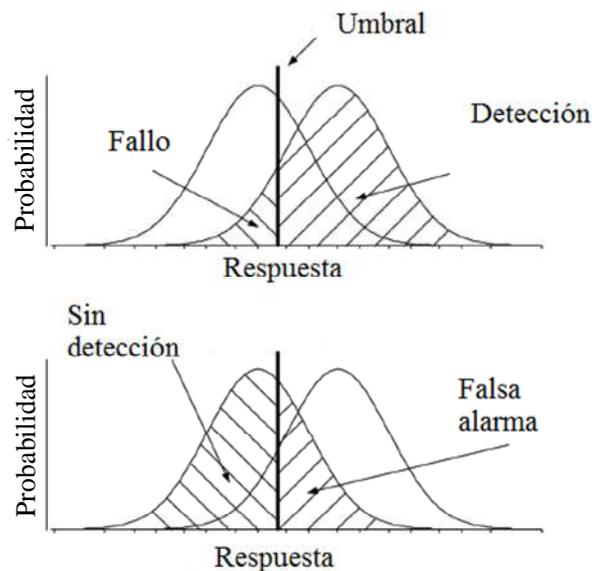


Figura 4.2. Posibles respuestas de un detector

La posición del valor de umbral puede generar diferentes situaciones (Figura 4.3), suponiendo las mismas distribuciones de probabilidad. Esta situación varía para diferentes funciones de distribución de las probabilidades. Una conclusión importante es que se puede observar que al incrementar la detección, también aumentan las falsas alarmas. Esto significa que el costo de bajar la cantidad de falsas alarmas impacta en la gran cantidad de detecciones fallidas bajando la calidad del detector.

En consecuencia, se debe determinar aun un valor de umbral apropiado. Como se explicó en el capítulo anterior, la señal recepcionada en el equipo es una combinación de señales reflejadas en el blanco, con ruido o interferencia, mas posiblemente la presencia de clutter. Entonces, es necesario encontrar un valor adecuado de umbral tal que al evaluar cada muestra contra dicho umbral, se pueda discriminar si se trata o no de un eco (si lo excede o no). Intuitivamente, se puede observar que el valor del umbral esta íntimamente

relacionado con el nivel de ruido o interferencia de la señal. La teoría basada en el criterio de Neyman-Pearsons lleva al concepto de umbral (thresholding en Inglés) y a las técnicas para la estimación del mismo.

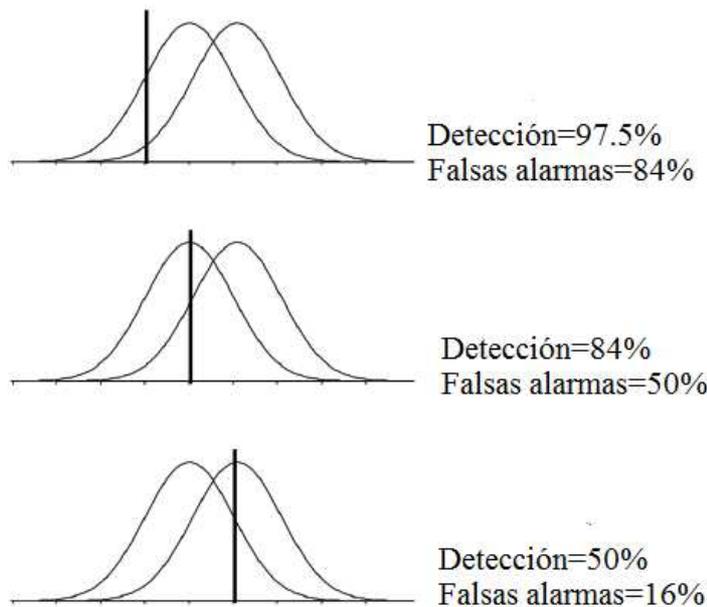


Figura 4.3. Diferentes resultados en performance para diferentes valores de umbral

4.4 Estimación del umbral

En principio, es necesario poder modelar las diferentes funciones de densidad de probabilidades (para el caso de falsa alarma y detección). Sin embargo, de acuerdo al criterio de Neyman-Pearsons se puede establecer un umbral para un valor de P_{fa} determinado en tiempo de diseño de manera que se maximice la P_d . Se puede probar que la probabilidad de detección alcanzable por un equipo puede ser determinada por la SNR de la señal (Richards, 2005). Por lo tanto, obtener un valor apropiado para el umbral depende fuertemente del conocimiento que se pueda tener del nivel de interferencia presente en la señal recibida.

Estimar el valor del umbral (T) es realmente complejo teniendo en cuenta que en la señal recibida es una señal atenuada, inmersa en ruido y retardada en el tiempo.

Supongamos que contamos con una señal $y[t]$ en el receptor del equipo compuesta por M muestras. Se puede pensar en $y[t]$ como un vector de M valores de amplitud de la forma $y[t] = [y_0, y_1, \dots, y_{M-1}]$.

El enfoque más sencillo y que se corresponde al concepto de detector óptimo es el de obtener un valor de umbral constante para la señal.

En este caso, se utiliza el total de muestras para estimar un único valor de umbral T , dependiente del nivel de ruido o interferencia de la señal, contra el cual se compararan cada una de las muestras. El valor de T será:

$$T = \alpha \hat{y} \quad (4.1)$$

(Richards, 2005)

Donde α es un valor de escala y \hat{y} es el nivel de interferencia o ruido obtenido estadísticamente a partir de las muestras. Utilizando esta ecuación se puede estimar un único valor de umbral T de tal manera que:

$$Deteccion = \begin{cases} 1 & \text{si } y_i > T \\ 0 & \text{si } y_i \leq T \end{cases} \quad (4.2)$$

Sin embargo, al utilizar un único valor de umbral, hay algunas situaciones que se deben contemplar:

- a. De acuerdo al valor de escala y el valor de interferencia, puede ocurrir que el valor del umbral sea relativamente bajo. En consecuencia, la probabilidad de detección aumentara y un gran numero de blancos serán declarados. La figura 4.4 muestra esta situación. Supongamos que en esta señal hay un par de blancos claramente diferenciados (los dos picos de mayor amplitud). Idealmente las dos únicas detecciones deberían corresponder a estos picos. Sin embargo, al ser el valor de T muy bajo, la evaluación de la ecuación 4.2 arroja un gran número de

detecciones que evidentemente no corresponderían con ecos reales (falsas alarmas).

- b. Una manera de solucionar este inconveniente sería “ajustar” el valor de T , de modo que este se encuentre en un valor mas alto y evitar que la cantidad de falsas alarmas sea grande.

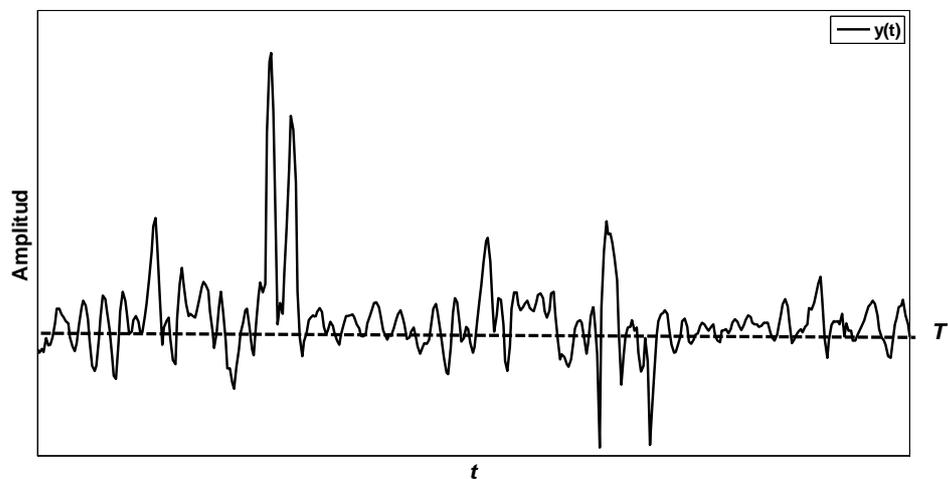


Figura 4.4. Alto número de detecciones debido a la determinación de un umbral muy bajo.

En el caso que el valor del umbral sea alto, la probabilidad de falsa alarma disminuye pero también lo hace la probabilidad de detección. La figura 4.5 muestra esta situación. Aquí se puede observar que disminuyen las falsas alarmas pero el umbral es demasiado alto evitando que se detecte un eco que correspondería a un blanco.

Este tipo de situación donde no se detecta un eco real (aumenta la probabilidad de fallo P_m) se denomina enmascaramiento de blanco (target masking). Es decir, existen casos donde se pueden encontrar ecos enmascarados o “escondidos” debido a un valor incorrecto para el umbral. Esto sucede cuando para obtener el valor del umbral se usan estadísticas que se encuentran sesgadas por los valores lejanos de amplitud de los

ecos. Por ejemplo, si se utiliza para estimar el nivel de ruido de la señal el promedio de las amplitudes, al ser un numero finito de muestras (y probablemente no muy grande), los valores de mayor amplitud (posiblemente correspondiente a ecos) afectaran el resultado logrando que el valor del umbral aumente.

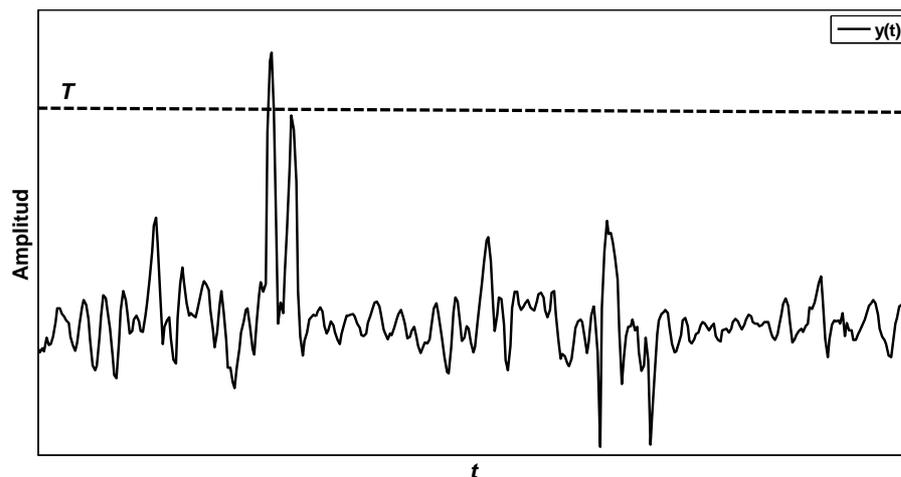


Figura 4.5. El umbral tan alto disminuye la probabilidad de falsa alarma pero tambien de detección.

De a) y b) se puede inferir que el valor del umbral es altamente dependiente del valor de interferencia. De acuerdo a la ecuación 3.1, fijados T , el nivel de interferencia es estimado mediante el promedio de las muestras y entonces el valor de escala (α) depende del valor de P_{fa} . Más aun, al variar el valor de la probabilidad de falsa alarma, varía la probabilidad de detección, lo cual infringe el criterio de Neyman-Pearson (Skolnik, 2001; Richard, 2005). Se podría suponer que si la P_{fa} deseada es del orden de 10^{-6} y que al usar un umbral fijo se alcanza en algunos casos una P_{fa} de 10^{-4} , esto no tendrá un impacto significativo en la performance del detector. Pero si consideramos, por ejemplo, que tenemos un radar que tiene una frecuencia de repetición de pulso de 10 kHz. Esto significa que el sistema debe realizar 10.000 decisiones de detección por segundo. Con una $P_{fa} = 10^{-6}$, una falsa alarma ocurrirá

en promedio una vez cada 100 segundos, mientras que si la P_{fa} aumenta a 10^{-4} en promedio se tendrá una falsa alarma por segundo. En términos prácticos esta alta sensibilidad significa que pequeños cambios en el valor del umbral tendrá un gran impacto en el desempeño del radar.

Es necesario, contar con alguna estrategia que permita encontrar algún valor de umbral pero manteniendo el valor de P_{fa} acotado y que logre un desempeño de detección consistente. Esto se logra mediante los denominados algoritmos para calculo de umbral variable o adaptativo o algoritmos CFAR (Constant False Alarm Rate).

4.5. Umbral variable o adaptativo (CFAR)

Con el fin de lograr un comportamiento y desempeño robustos en la detección, es preciso determinar el valor más ajustado a la realidad para la interferencia a partir del análisis online de los datos. El objetivo, es que el valor del umbral se ajuste para garantizar que la P_{fa} se mantenga un valor deseado. Un procesador para detección que mantiene constante el valor de P_{fa} se denomina procesador Constant False Alarm Rate (CFAR). Tanto los detectores de umbral variable en general como los no paramétricos se basan en este tipo de procesador. La diferencia entre ambos radica en la manera que se estiman los niveles de ruido o interferencia.

Dada la señal $y[t] = [y_0, y_1, \dots, y_{M-1}]$ compuesta por M muestras, cada muestra y_i , $i = 0, 2, \dots, M - 1$ será testeada utilizando un valor apropiado de umbral T_i particular para cada muestra. La muestra y_i se denomina celda de prueba. El valor de nivel de ruido o interferencia en el cual está inmersa la celda de prueba se calcula utilizando las celdas vecinas. Estas celdas forman una ventana móvil de un determinado número de muestras para cada de las celdas de interés.

El enfoque CFAR se basa en dos grandes hipótesis:

1. Las celdas continuas o adyacentes a la celda de prueba contienen interferencia con igual estadísticas que en ésta (esto se denomina interferencia homogénea). De este modo es suficientemente representativo como valores de interferencia que compiten con un blanco potencial.
2. Las celdas adyacentes a la celda de prueba solo contienen interferencia o ruido, no contienen blancos.

Bajo estos supuestos, el nivel de interferencia para encontrar un valor adecuado de T_i se utilizan los valores de amplitudes de las celdas adyacentes a y_i . (Richards, 2005)

La Figura 4.6 muestra un esquema de una ventana (unidimensional) de $N+1$ celdas utilizadas para evaluar T_i , para una celda de prueba y_i . Los valores de amplitud de las muestras dentro de la ventana son utilizados para estimar estadísticamente que valor de ruido circunda a la celda de interés. Ésta será evaluada contra el umbral determinado para ella. Para la próxima celda, la ventana se desplazará también obteniendo un nuevo valor de umbral.

Existen numerosos algoritmos basados en el concepto CFAR. La principal diferencia entre estos algoritmos radica en la manera que estos calculan el valor de la interferencia en las celdas adyacentes a la celda de interés (celda de prueba). Cada uno de estos algoritmos tiene sus ventajas y desventajas de acuerdo al tipo de señal que se trate, y al blanco y sus características. Determinar el algoritmo que mejor se ajuste a un tipo de aplicación de radar, forma parte del modelado del mismo. Estos algoritmos implementan la funcionalidad automática del radar mediante el análisis online de las señales y la estimación de un umbral variable para cada celda de prueba (Richards, 2005).

A continuación se describen dos de los algoritmos mas utilizados, y a partir de los cuales se deducen numerosas variantes que están fuera del alcance de este trabajo de tesis.

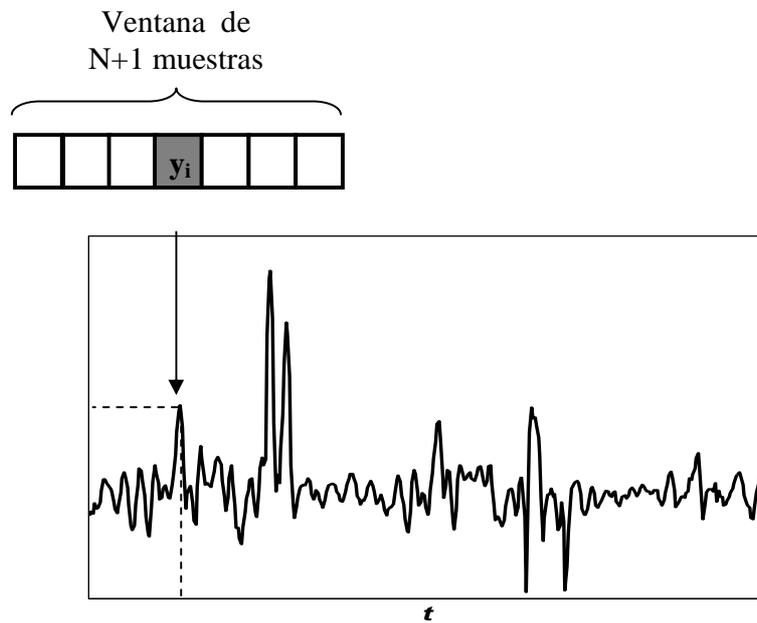


Figura 4.6. Esquema de funcionamiento del algoritmo CFAR.

El siguiente algoritmo en pseudo código muestra el procedimiento a seguir para implementar un algoritmo CFAR:

POCEDIMIENTO CFAR (y:arreglo; var M: numero total de muestras; N: numero de muestras en la ventana)

- (1) **begin**
- (2) **PARA** cada muestra y_i de y **HACER** (M veces)
- (3) Obtener T_i usando N celdas
- (4) **SI** $y(i) > T_i$ **ENTONCES**
- (5) **ESCRIBIR**(Deteccion positiva, i)
- (6) **FIN**
- (7) **FIN**
- (8) **FIN.**

El algoritmo inicia estimando para cada muestra de la señal un valor de umbral (paso 3 en el algoritmo). Si la amplitud correspondiente a la muestra en la celda de prueba supera el valor de su umbral, entonces se obtiene una detección (pasos 4-6 en el algoritmo). Al finalizar el procedimiento se obtiene un listado de detecciones realizadas (que incluso puede ser nula).

Se puede realizar un análisis del costo computacional de este procedimiento. Dado que la ventana móvil tiene un tamaño siempre inferior al total de muestras (se mostrará además que es realmente muy inferior) el costo computacional de obtener el umbral (paso 3) puede considerarse casi constante. Esto significa que no crece con el tamaño de muestras que posee la señal, una vez fijado el tamaño de la ventana.

Para cada asignación de variables, o evaluación de operaciones (matemáticas o lógicas) siempre se realizará la misma cantidad para cada muestra. O sea, el costo de las mismas es constante.

Está claro que el mayor costo radica en la repetición de la estimación de un umbral para cada muestra. A lo sumo se calculará un umbral por cada muestra, o sea a lo sumo se realizará M veces. Por lo tanto, si la cantidad de muestras aumenta, el costo computacional del algoritmo crece linealmente con este valor. Este es un costo computacional muy bajo, siendo el algoritmo un algoritmo eficiente (Aho, 1983).

Este algoritmo CFAR está diseñado para un procesador en una dimensión.

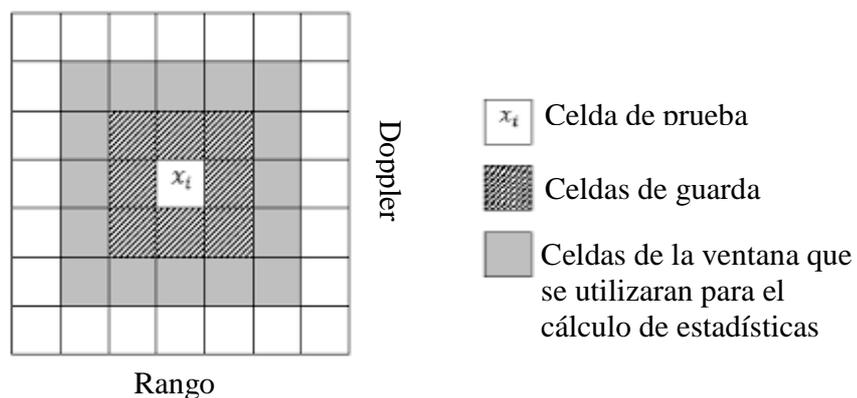


Figura 4.7. Concepto CFAR generalizado

Si se requiere de un procesador Doppler (en rango), entonces la evaluación del umbral de debe realizar en una ventana de dos dimensiones. En tal caso se puede realizar una generalización para el caso bidimensional (Figura 4.7)

4.5.1. Pérdida CFAR

Todos los algoritmos basados en el concepto CFAR parten del concepto de fijar un valor de P_{fa} deseable a partir del cual estimar el umbral variable. Sin embargo esto no garantiza que la P_d sea la adecuada, sino que es necesario poder establecer un compromiso entre estas dos probabilidades de manera que se pueda alcanzar una buena performance de detección. Lo que se espera tener es un valor alto de P_d y uno bajo para la P_{fa} (Crurry, 2005).

Para alcanzar este objetivo es requisito tener una buena relación señal ruido y por lo tanto es imprescindible estimar el nivel de interferencia de la señal (Richards, 2005).

Si el valor real del ruido o interferencia fuese conocido, entonces se podría estimar un umbral ideal. Esta situación no se da en escenarios reales, sino que las estadísticas sobre la interferencia se calculan sobre un número limitado de muestras o celdas. En general el valor del umbral en el caso CFAR suele ser mayor al valor del umbral ideal. Esto se debe a que es necesario un umbral mayor comparado con el valor de interferencia para compensar la obtención imperfecta del nivel de interferencia y garantizar que se mantenga el valor deseado de P_{fa} (Richards, 2005). Debido a que el multiplicador (α) se incrementa en CFAR, la probabilidad de detección promedio para un blanco para un determinado valor de SNR, disminuirá respecto al caso de que la interferencia fuera conocida. Por lo tanto para alcanzar una P_d para una dada P_{fa} es necesario que se incremente el valor de SNR. Este incremento en el SNR necesario para alcanzar los objetivos de las técnicas CFAR (baja P_{fa} y alta P_d) es llamado CFAR-loss o pérdida CFAR. La pérdida en desempeño del algoritmo CFAR depende fuertemente del tamaño de la ventana que se utilice para el cálculo de las estadísticas del ruido o interferencia. Por lo tanto es una función del número de muestras usadas. Si este número tiende a

infinito, la estimación del nivel de ruido converge al valor verdadero (Richards, 2005). Sin embargo, que el tamaño de la ventana tienda a ser infinito significa en términos prácticos tener una ventana que tienda a ser la cantidad total de muestras de la señal. Pero si se utilizan todas las muestras disponibles, se obtiene un único umbral y no un umbral por muestra. No tener un umbral adaptivo, sino fijo tiene graves consecuencias en la capacidad de detección del equipo. Supongamos que tenemos una señal con más de un eco (lejanos en el tiempo), al momento de calcular el único umbral las estadísticas para el ruido pueden estar sesgadas por la amplitud de un eco más lejano. Esto produce que el umbral sea mayor y que enmascare otros blancos. Es necesario realizar un balance respecto al número de celdas que van a constituir la ventana móvil.

Para ello, si se utilizan pocas celdas en la ventana CFAR, una estrategia para disminuir la pérdida de performance debido a CFAR-loss es la de mejorar sustancialmente el valor de SNR. Esto significa que una vez determinado el valor de P_{fa} , el SNR del sistema determinará la P_d (Richards, 2005).

Por esto, además de utilizar técnicas CFAR, muchos sistemas necesitan implementar técnicas adicionales para mejorar la SNR. Si por ejemplo, hay presentes señales emitidas deliberadamente para interferir con el radar (jamming), se suele utilizar un cancelador de lóbulos laterales. En otros casos usar un mapa de clutter puede ser útil para filtrar señales provenientes de clutter y con ello mejorar la SNR. Muchas veces la alternativa es utilizar integración coherente de múltiples pulsos. Estas técnicas para mejorar la SNR dependerán del sistema de radar particular. Diversas opciones son tratadas por Nathanson (1991).

4.5.2. Algoritmo por promedio de celdas (CA-CFAR)

El algoritmo por promedio de celdas denominado Cell Averaging–CFAR (CA-CFAR) es un método para estimar el nivel de interferencia necesario para calcular el umbral variable mediante el promedio de las amplitudes en la ventana. Es decir, en cada ventana de muestras usadas para la evaluación del ruido o interferencia, se calcula el valor promedio de las muestras adyacentes como medida de interferencia. (Richards, 2005). La

Figura 4.8 muestra un esquema de la ventana utilizada en un algoritmo CA-CFAR. En el esquema se muestra una ventana de tamaño $N+3$. La celda central es la celda de prueba y_i . Las celdas inmediatamente adyacentes a ésta se denominan celdas de guarda y no son tenidas en cuenta para el cálculo del promedio.

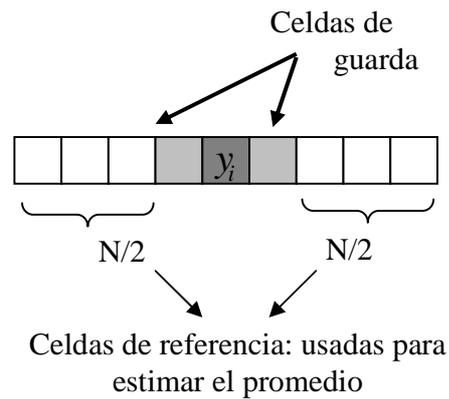


Figura 4.8. Esquema de una ventana para el algoritmo CFAR

La Ecuación 4.1 se puede reescribir en términos de posición de la celda de prueba como sigue:

$$T_i = \alpha \hat{y}_i \quad (3.3)$$

De tal manera que por cada celda y_i se tendrá un valor adaptativo T_i calculado mediante las estadísticas (\hat{y}_i) en la ventana. Es importante notar que el valor de escala α no depende de las muestras. De hecho, el valor de escala no es empírico y se computa a partir del valor predefinido de P_{fa} y del número de celdas en la ventana. (Richards, 2005). Esto garantiza que se mantiene la característica principal de los algoritmos CFAR.

El valor de escala se calcula como:

$$\alpha = N \left(P_{fa}^{\frac{-1}{N}} - 1 \right) \quad (4.4)$$

De este modo se puede calcular \hat{y}_i como:

$$\hat{y}_i = \frac{1}{N} \sum_{j=1}^N y_j \quad (4.5)$$

El procedimiento de estimación del umbral se repite por cada muestra de la señal, en cada caso calculando un valor apropiado de T_i .

La figura 4.9 muestra un ejemplo típico resultante de implementar un algoritmo CA-CFAR en una señal.

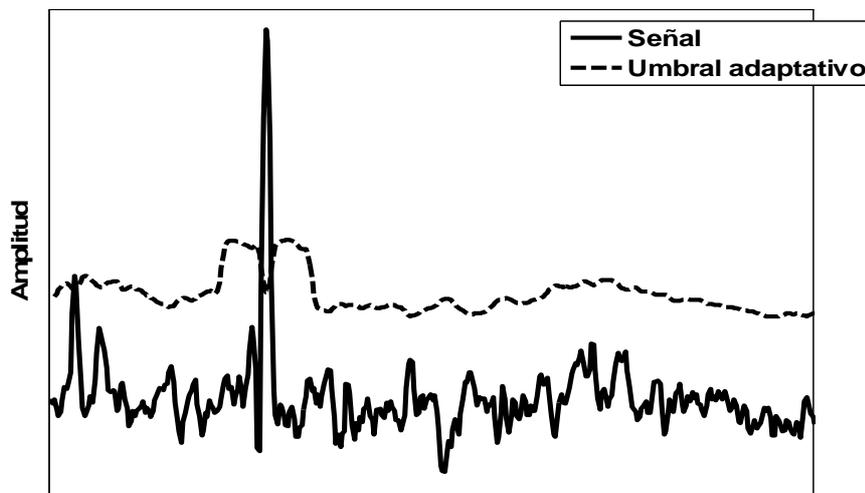


Figura 4.9. Grafica que muestra el resultado de aplicar el algoritmo CA-CFAR

Una característica del comportamiento del algoritmo CA-CFAR es que si la celda que se está evaluando no corresponde a un eco, el valor del umbral es alto. Mientras que si en la celda de evaluación o prueba hay presente un eco, el valor de T_i es menor permitiendo la detección. En la figura 3.8 se puede observar este comportamiento (líneas punteadas).

Se puede observar como en el caso de celdas cercanas al eco, el umbral aumenta significativamente. Esto se debe a que al evaluarse una celda cercana a un eco o blanco, la amplitud de este será computada en el promedio (haciendo que este valor se incremente). Esto se debe a que no se cumple la suposición de que todas las celdas de referencia poseen iguales estadísticas (se viola el supuesto de homogeneidad de la interferencia).

Un análisis importante es el correspondiente al uso o no de celdas de guarda para el cómputo del promedio. El motivo de la utilización de celdas de guarda radica en que la energía en las celdas inmediatamente adyacentes a y_i pueden contener no solo energía correspondiente a ruido o interferencia, sino que pueden contener además energía parcial de eco (bajo el supuesto de que el eco se encuentra en y_i).

Al utilizar como estadística, del nivel de ruido circundante en la celda de evaluación el promedio de las amplitudes en la ventana, y si existe en éstas energía parcial de algún eco, el promedio estará sesgado elevando el umbral. En consecuencia es posible que los ecos se encuentren enmascarados (self masking). Si la resolución en rango del sistema es tal que algunos blancos puedan esparcir su energía en más de una celda, entonces es conveniente utilizar más de una celda de guarda a cada lado de la celda de evaluación y_i .

La figura 4.10 muestra el desempeño de un algoritmo CA-CFAR si no se utilizan celdas de guarda y que genera que el eco se auto enmascare (self-masking). Además se compara con el caso de utilizar 3 celdas de guarda a cada lado de la celda de prueba.

Hay que prestar especial atención a seleccionar la cantidad adecuada de celdas totales en

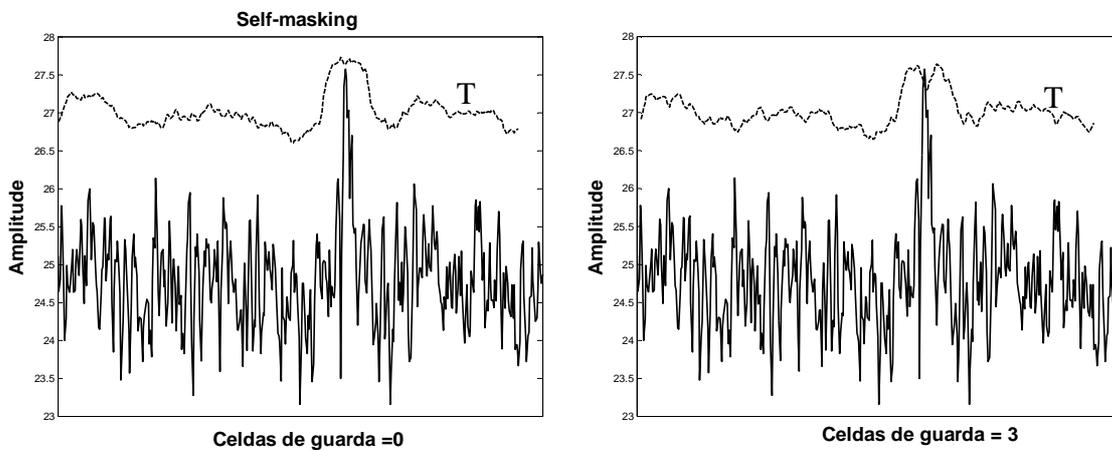


Figura 4.10. A la izquierda se muestra una señal (línea continua) y el umbral obtenido con el algoritmo CA-CFAR (línea punteada) usando una ventana de 20 muestras y sin celdas de guarda. Se puede observar que ninguna muestra supera el umbral y por lo tanto no hay detecciones (self-masking). A la derecha se muestra la misma señal y el umbral obtenido con el mismo algoritmo pero con un número de 3 celdas de guarda. En el segundo caso, es posible la detección.

la ventana. Mientras el número de celdas de referencia aumente (N), la estimación del promedio de la interferencia convergerá al valor verdadero de la misma. En contraposición, a medida que aumente el número de celdas de referencia N , puede ocurrir que en las celdas de referencia se encuentren más de un blanco. Esto tiene como consecuencia la posibilidad de tener ecos o blancos enmascarados (Figura 4.11).

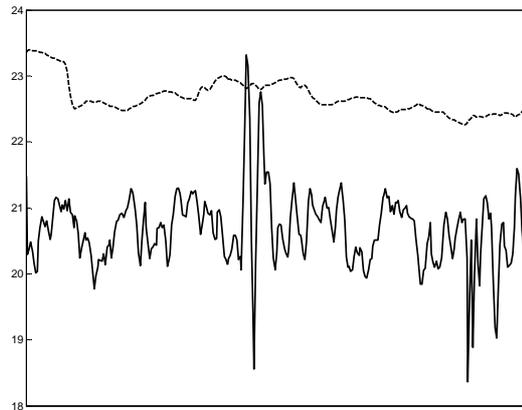


Figura 4.11. En esta gráfica se muestra una señal (línea continua) con dos aparentes blancos cercanos y el umbral obtenido con el algoritmo CA-CFAR (línea punteada). El algoritmo CA-CFAR (con una ventana de 50 muestras) falla en detectar el segundo eco (blanco enmascarado). Este es un ejemplo de una mala selección del tamaño de la ventana

El caso extremo es el de tomar todas las muestras disponibles y que significa llegar al caso de tener un umbral fijo y que fue analizado con anterioridad en este mismo capítulo. Con el fin de combatir la degradación de la detección por blancos enmascarados, existen numerosas variantes del algoritmo CA-CFAR que se ajustan a diferentes tipos de blancos.

4.5.3. Algoritmo por reordenación de celdas (OS-CFAR)

Este algoritmo denominado OS-CFAR (Order Statistics-CFAR o OS-CFAR) se basa en la reordenación de las muestras presentes en la ventana y su objetivo es combatir algunos casos de enmascaramiento de blanco. En este caso no se necesita de celdas de guarda.

Se basa en la hipótesis de que las muestras son estadísticamente independientes y están idénticamente distribuidas. Las muestras son reordenadas de manera ascendente tal que $y(1) < y(2) < \dots < y(N)$. Los índices denotan el ranking de las muestras, así la muestra en $y(1)$ tiene el menor valor de amplitud y $y(N)$ el máximo valor de amplitud (Rohling, 1983).

Una vez que las amplitudes fueron reordenadas, el objetivo principal es seleccionar un valor $y(k)$ de la secuencia ordenada que represente el nivel de ruido o interferencia para esa ventana. El valor de $y(k)$ es normalizado usando la desviación estándar. Si se selecciona el k -ésimo valor de la secuencia cercano a la posición $N/2$, entonces bajo los supuestos mencionados anteriormente se puede considerar como una especie de media de las amplitudes (Rohling, 1983). Si ocurre que más de un eco está presente en la ventana, al reordenarse la secuencia, las amplitudes correspondientes a estos ecos se encontrarán próximas al índice N . Esto garantiza que las amplitudes de los ecos no sean tomadas para la estimación de la interferencia. La figura 4.12 muestra un esquema que explica el funcionamiento del algoritmo OS-CFAR.

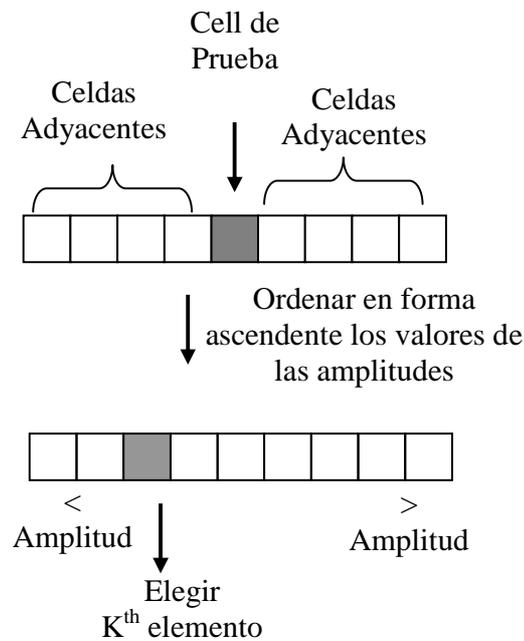


Figura 3.12. Comportamiento del algoritmo OS-CFAR

Como toda variante de algoritmo CFAR la diferencia radica en la manera que se estima el valor del ruido para luego obtener el umbral.

La ecuación 3.3 sigue siendo válida. Una consideración importante es que si bien es un algoritmo CFAR, es decir mantiene constante la P_{fa} , la obtención del valor de escala α se

debe calcular de otra manera. Rohling (1983) obtuvo una expresión para el cálculo de α :

$$Pfa = k \binom{N}{k} \frac{(k-1)!(\alpha + N - k)!}{(\alpha + N)!} \quad (4.6)$$

Dado el valor deseado de P_{fa} , y resolviendo esta ecuación iterativamente para valores de N y k se puede estimar el valor de escala. En el mismo trabajo se presenta una tabla con valores de α teniendo en cuenta valores comunes de P_{fa} , N y k para radares (ver tabla 4.1).

k	N=8	N=16	N=24	N=32
1	7999992.0	15999984.0	23999976.0	31999968.0
2	7475.8	15476.4	23471.2	31464.5
3	688.2	1482.8	2275.5	3067.9
4	196.0	442.7	688.1	933.3
5	86.4	206.7	326.0	444.9
6	46.7	120.4	192.8	265.0
7	27.8	79.4	129.5	179.2
8	16.8	56.6	94.1	131.3
9		42.4	72.1	101.4
10		32.9	57.3	81.4
11		26.1	46.8	67.2
12		20.9	39.1	56.7
13		16.9	33.1	48.5
14		13.7	28.3	42.2
15		10.9	24.5	37.0
16		8.3	21.3	32.7
17			18.6	29.2
18			16.3	26.1
19			14.3	23.5

20			12.5	21.2
21			10.8	19.2
22			9.3	17.4
23			7.9	15.8
24			6.3	14.4
25				13.1
26				11.9
27				10.7
28				9.6
29				8.6
30				7.6
31				6.6
32				5.4

Tabla 4.1. Factor de escala (α) para el algoritmo OS-CFAR para $P_{fa} = 10^{-6}$

El nivel de ruido (\hat{y}_i) se calcula con el k-ésimo valor $y(k)$, y el valor del umbral se puede estimar con la ecuación:

$$T_i = \alpha y(k) \quad (4.7)$$

Una observación importante es que en este caso no es necesario contar con celdas de guarda ya que si existen celdas que contengan interferencia y además energía parcial proveniente de algún eco, estas celdas serán reordenadas hacia el final y no serán computadas en la estimación del nivel de ruido.

Como en todos los algoritmos basados en CFAR, este proceso se repite por cada muestra en la señal.

Respecto a la manera que se comporta este algoritmo:

- Si todos los valores de amplitud que se encuentran en la ventana no varían demasiado (no hay un eco en la ventana, por ejemplo), la desviación estándar es cero, entonces el k-ésimo elemento no tendrá mucha variación en amplitud y por lo tanto el umbral será tan alto que no habrá detecciones.

- Si por el contrario, dentro de la ventana se encuentran ecos, las amplitudes van a variar significativamente y el k -ésimo elemento tendrá una amplitud baja comparada a los valores en el extremo N . Esto significa que el umbral será bajo y entonces ocurrirá la detección.

La elección del valor k -ésimo también tiene un gran impacto en la capacidad de detección del algoritmo. Si ocurre que existen muchos ecos en la ventana (o muestras que contengan parcialmente energía proveniente de ecos), puede ocurrir que éstos se vean enmascarado. Concretamente este riesgo se corre si la cantidad de ecos es mayor a $N-k$ y está relacionado con la capacidad en resolución de rango del radar.

En principio, el valor de $y(k)$ para un dado índice k del vector y ordenado representará una medida cercana a la media estadística de las muestras (si se eligen valores cercanos al medio del arreglo). Rholing (1983) mostró que una elección de $k > N/2$ (valor mayor a la media de la secuencia ordenada) y que $N-k$ no menor al doble de la longitud del blanco es necesario para un buen desempeño del algoritmo.

4.5.4. Otros Algoritmos Basados en CA-CFAR

En muchos casos además de ecos, ruido, interferencia o jamming, las señales contienen clutter. Es decir, se encuentra superpuesta una señal correspondiente a algún objeto estacionario que produce gran reflectividad como por ejemplo el mar, una montaña, etc. El clutter en general es altamente heterogéneo (Richards, 2005).

La figura 4.23 muestra un esquema de una señal con presencia de clutter. A la derecha de la línea roja se muestra que la señal se encuentra afectada por clutter donde se puede observar como la potencia se incrementa. $P1$, $P2$ y $P3$ denotan tres posibles ecos. Mientras que $P1$ y $P3$ serán detectados fácilmente debido a que la localía del cálculo de las estadísticas, $P2$ se encuentra enmascarado por encontrarse al borde del clutter que produce un incremento repentino de las amplitudes a la derecha de $P2$. En este tipo de casos CA-CFAR y OS-CFAR pueden fallar y muchas veces se utilizan algoritmos alternativos que surgen como una variante de CFAR

Estas extensiones del concepto CFAR en muchos casos son motivados por heurísticas y suelen ser difíciles de analizar exhaustivamente (Richards, 2005).

Dos ejemplos de estas variantes de algoritmos son SOCA-CFAR (smallest of cell averaging-CFAR) y GOCA-CFAR (greatest of cell averaging- CFAR).

En ambos casos se dividen las celdas de la ventana en celdas anteriores a la celda de prueba y celdas posteriores a ella. En ambos casos se calculan dos promedios separados, uno utilizando las celdas anteriores y el otro utilizando las celdas posteriores.

En el primer caso se elige para el cálculo del umbral el menor de los dos promedios. Al utilizar la mitad de celdas para el cálculo del promedio, el factor de escala α de la ecuación 3.3 debe ser recalculado utilizando otra expresión diferente al caso de CFAR.

En el caso de GOCA-CFAR se selecciona el mayor de los promedios y también requiere de una expresión diferente para el cálculo α (Richards, 2005).

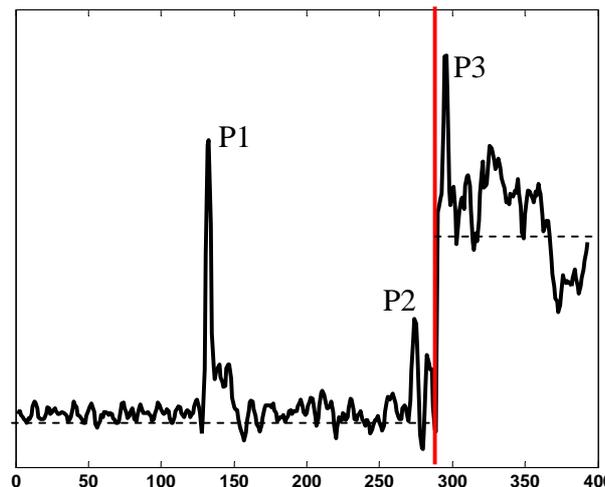


Figura 4.13. Muestra una señal con un eco enmascarado por el borde de clutter.

Existen otros algoritmos basados en CA-CFAR todos muy vulnerables ante modificaciones del clutter básicamente porque tienen una formulación empírica.

4.6. Mapa de Clutter

Un mapa de clutter es una tabla almacenada que obtiene magnitudes correspondientes a clutter con el objetivo de posteriormente remover fluctuaciones temporales que no correspondan a blancos (Barton y Leonov, 1998). Un ejemplo de clutter para un radar

primario de un aeropuerto estaría constituido por los diferentes accidentes geográficos alrededor del radar. Se puede generar un mapa de clutter que permita establecer mejorar la capacidad de detección sustrayendo a las señales recepcionadas las medidas de clutter. Típicamente un mapa de clutter forma parte de la configuración de un detector de blancos móviles (Skolnik, 1990). Para ello se debe realizar una descripción matemática de la reflectividad y otros parámetros del clutter (Barton y Leonov, 1998). La principal característica es que un mapa de clutter está adecuado al tipo de radar, pero fundamentalmente al emplazamiento que éste tendrá.

CAPITULO V

Modelo de Detección Automática de Ecos de Radar

5.1. Introducción

Desde los inicios del siglo pasado las mediciones mediante radio sondajes han jugado un rol significativo en el estudio de las capas altas de la atmósfera terrestre como medio para las comunicaciones a larga distancia. En la actualidad tiene un gran rol en aplicaciones de observación de meteorología del espacio y su relación con la actividad humana en diversos campos, tales como sistemas de posicionamiento global mediante satélites, comunicaciones via satélite, observaciones geodésicas, etc. Para tales propósitos se utilizan radares de alta frecuencia, denominados sondadores ionosféricos. Mediante el análisis de los llamados ionogramas se pueden extraer importantes características del medio bajo estudio.

Un sondador ionosférico es un radar y cómo tal su principal objetivo es la detección de un blanco. En particular el objeto de interés en este caso es la altura virtual de la ionósfera. Siendo éste un tipo de blanco particular y que no se encuentra caracterizado en los sistemas de radares típicos, es necesario encontrar un modelo adecuado de detección. En este capítulo se realizará una descripción de un sondador ionosférico, sus principales funciones y magnitudes que pueden ser estimadas. Se mostrarán las principales motivaciones para el estudio de métodos de detección y se realizará una propuesta de modelo computacional de detección automática para un sondador ionosférico.

5.2. Sondaje vertical de la ionósfera mediante un radar ionosférico

Un sondador ionosférico trabaja en banda de alta frecuencia (HF), típicamente entre 1 a 30MHz, por lo general tiene dos antenas tipo Delta montadas en un mástil. Mediante este equipo se pueden obtener diferentes magnitudes ionosféricas tales como frecuencias

críticas de los diferentes estratos ionosféricos, estructura y altura de las mismas, entre otras. Utiliza una técnica denominada sondaje vertical.

Esta técnica se basa en las características refractivas del medio respecto de las ondas de radio que se propagan en el plasma ionosférico. Mediante un barrido de frecuencia entre (1-30 MHz), emite pulsos hacia la ionosfera donde, debido a las características del plasma ionosférico (ver capítulo III), son reflejados de manera similar a un espejo para ser recepcionadas posteriormente en el mismo equipo. Al medir el tiempo transcurrido entre la emisión y recepción de un pulso, se puede estimar la altura en la cual se produjo la reflexión y consecuentemente la altura virtual de la ionosfera (Davies, 1969, Reinisch and Galkin 2011, Baskaradas et al, 2014, entre otros).

Es importante notar que como se explicó en el capítulo III la reflexión en la ionósfera no se produce a una única altura. La onda incidente que va penetrando en el plasma produciendo en cada instante de tiempo es parcialmente absorbida, parcialmente reflejada (posiblemente en caminos de reflexión diferentes) y otra parte continúa su camino de propagación. La máxima potencia reflejada (reflexión resonante) se produce cuando el índice de refracción de la onda alcanza al índice de refracción del plasma. Por lo tanto, la altura que se detecta en el equipo se denomina altura virtual. La altura real es menor a la altura virtual dado que se calcula de acuerdo a la trayectoria proyectada que podría haber tomado la onda reflejada y ser regresada a la tierra hacia el mismo lugar. La altura máxima que habría alcanzado esta onda reflejada hipotética es la altura virtual (Zolesi y Cander, 2014).

La Figura 5.1 ilustra esta situación. Se muestra como la onda de radio es emitida (TX) hacia la ionósfera, para luego de ser reflejada en los estratos inosféricos es recepcionada en el equipo (RX). Se esquematiza además la relación entre la altura virtual (h_v) y la altura real (h_r).

Si se grafica la altura virtual obtenida para cada frecuencia emitida, se puede obtener el denominado ionograma. Esta representación gráfica de la altura virtual de la ionosfera en función de la frecuencia es de suma importancia para el estudio de la alta atmósfera terrestre.

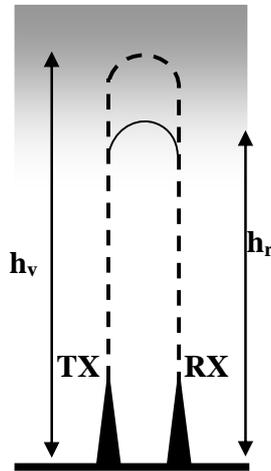


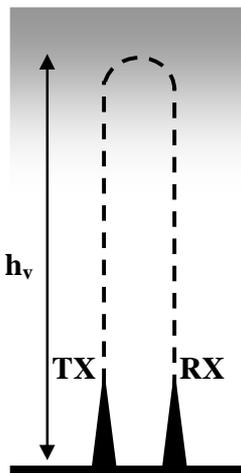
Figura 5.1. Esquema que muestra la relación entre la altura virtual (h_v) y la altura real (h_r)

La figura 5.2 muestra un esquema simplificado del proceso realizado por un sondador ionosférico desde el sondaje hasta la obtención del ionograma.

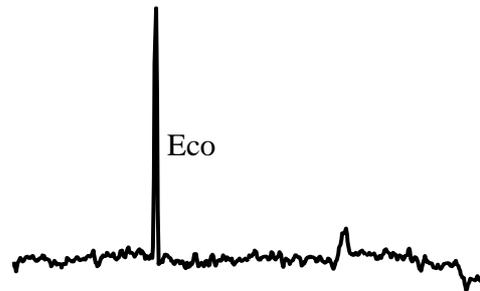
En primera instancia se transmite una serie de pulsos codificados hacia la ionósfera, se mide el retardo temporal que sufre esta señal para cada frecuencia entre 1 y 20 MHz. Utilizando algún modelo de detección se establece la altura de la máxima reflexión en el plasma ionosférico y se lo grafica en el ionograma.

Cada eco graficado en el ionograma es producto de una serie de etapas de detección en el radar que inicia con la codificación de la señal transmitida y que continúa a luego de la recepción de la reflexión en algún estrato ionosférico.

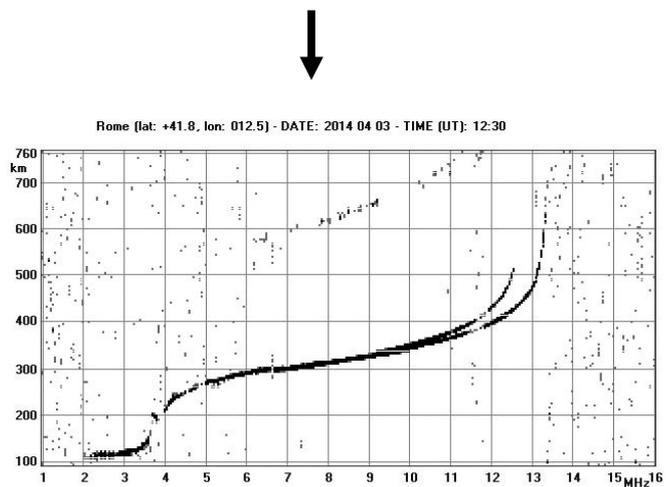
Eco obtenido a partir de un sondaje para una única frecuencia



Sondaje vertical: Se envía un pulso (TX) hacia la ionósfera donde el mismo es reflejado a una determinada altura (h_v) para posteriormente ser recepcionado por el equipo (RX).



Cada eco detectado para cada frecuencia, se grafica como un punto en el ionograma



Ionograma correspondiente al día 03 de abril de 2014 a las 12:30 UT en la estación ionosférica INGV-Roma. Se muestran las alturas virtuales detectadas para frecuencias entre 0 y 20 MHz

Figura 5.2. Esquema simplificado del proceso de obtención de un ionograma a partir de sondajes verticales por un equipo sondador ionosférico

Tradicionalmente los radares ionosféricos eran equipos analógicos que daban como resultado un ionograma en papel o en film. Se basaban principalmente en detección por envolvente emitiendo pulsos cortos con una resolución en altura de 15 km. Lo más destacable es que estos equipos necesitaban emitir un pico de potencia del orden de los KW. Actualmente los sondadores son equipos digitales que tienen como salidas archivos (binarios, e imágenes) que permite post procesamiento. La gran diferencia con los antiguos sistemas es se que utiliza la técnica de compresión de pulsos (ver capítulo II) utilizando pulsos más largos. La gran ventaja es que alcanza una resolución en altura de 5 km y necesita un pico de potencia del orden de cientos de Watts para transmitir (Zolesi y Cander, 2014).

Si bien existen diferentes diseños para radares ionosféricos, es posible establecer una serie de componentes generales que éstos poseen.

En la Figura 5.3 se muestra un esquema simplificado de los componentes de un radar ionosférico.

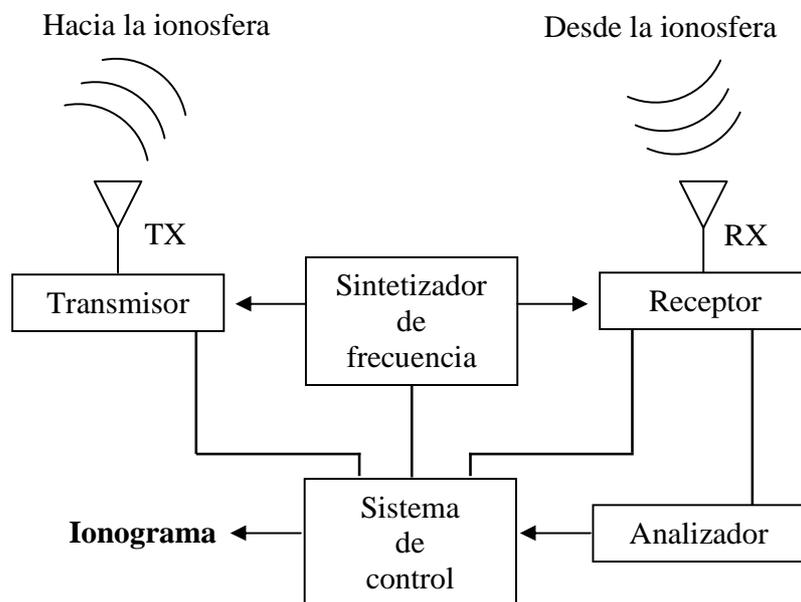


Figura 5.3. Esquema simplificado de un radar ionosférico

Mediante el sintetizador de frecuencia se genera una señal codificada que será transmitida hacia la ionósfera. Una versión local del código es utilizada por el receptor para ser comparada con la señal receptionada desde la ionósfera. Mediante una serie de pasos de detección se discrimina el retardo de tiempo sufrido por la señal y finalmente al cabo de un barrido completo de frecuencias se genera un ionograma. Este proceso es coordinado por el sistema de control del equipo.

Las decisiones de diseño de cada uno de estos componentes tendrá un impacto directo en las capacidades del equipo para la detección.

Obtener un modelo de detección automático para un radar implica encontrar la secuencia de etapas de detección que mejor ajusten a un blanco determinado.

La detección automática requiere de un gran conocimiento y caracterización del blanco y por tal motivo en muchos casos depende fuertemente de datos o ajustes empíricos. Esta situación hace que la obtención de un único modelo de detección sea un trabajo complejo.

5.3. Motivación para el estudio de la detección de estratos ionosféricos

La principal motivación para realizar la investigación desarrollada por este trabajo de tesis se centra en obtener un modelo de detección específico de los diferentes estratos ionosféricos utilizando un sondador. Para comprender mejor la problemática se describirán los componentes típicos en un ionograma y cómo las capacidades de detección del sondador repercuten en la calidad de las mediciones.

En un ionograma típico se pueden discriminar las alturas para las diferentes capas ionosféricas. La Figura 5.4 muestra un ionograma diurno idealizado (Zolesi y Cander, 2014) con sus trazas ordinarias y extraordinarias y sus características específicas.

Esquemáticamente se encuentra dividido en las regiones convencionales D,E, Es y F.

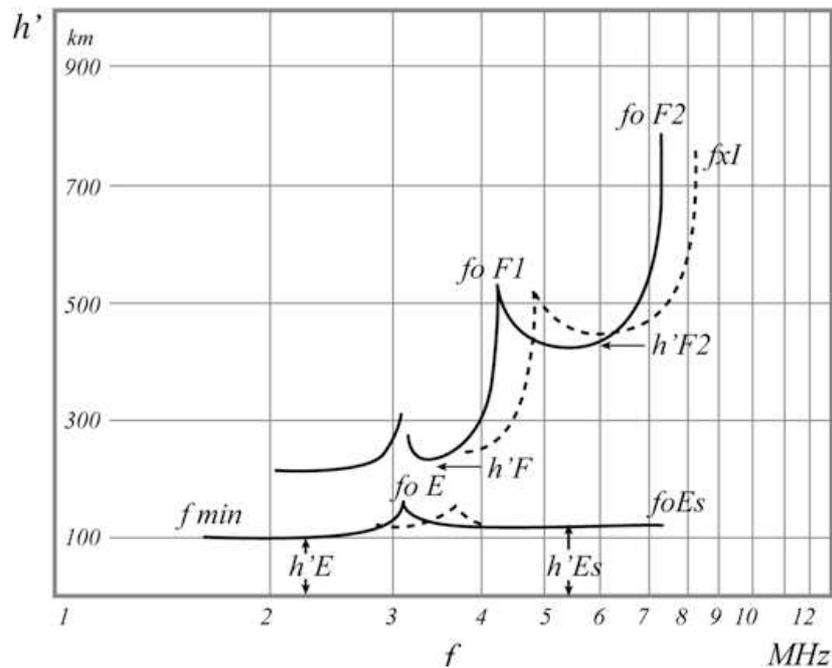


Figura 5.4. Esquema de un ionograma diurno ideal (Zolesi y Cander, 2014)

En la Región F se pueden observar los siguientes parámetros:

f_oF2 : Frecuencia crítica correspondiente a la traza ordinaria del estrato más alto de la región F, denominada capa F2 cuando se encuentra presente la capa F1.

f_xI : Máxima frecuencia detectada para la región F.

f_oF1 : Frecuencia crítica de la traza o rayo ordinario de la capa F1 cuando ésta se encuentra presente.

$h'F2$: Altura virtual mínima del rayo ordinario de la capa F2

$h'F$: Altura virtual mínima del rayo ordinario de la región F.

En la Región E a su vez se puede discriminar:

f_oE : Frecuencia crítica del rayo ordinario de la región E.

$h'E$: Altura virtual mínima del rayo ordinario de la región E.

Además en algunos casos también ocurre la aparición de la denominada capa esporádica E (E_s), y en estos casos se pueden determinar:

f_oE_s : Máxima frecuencia del rayo ordinario para la continua capa esporádica E.

$h'E_s$: Mínima altura virtual de la capa E_s (rayo ordinario).

f_bE_s : Frecuencia de cubrimiento o “blanketing” de la capa E_s .

f_{min} : Mínima frecuencia detectada en el ionograma.

A partir de estos parámetros y utilizando diferentes mecanismos de interpretación de ionogramas, ya sean manuales o automáticos (escalado de ionogramas), finalmente se pueden obtener magnitudes físicas de la ionósfera y construir un perfil de densidad electrónica (Zolesi y Cander, 2014).

Como se mencionó, este ionograma es un ionograma ideal. Los ionogramas reales varían mucho de acuerdo a condiciones estacionarias, diurnas y nocturnas, a condiciones solares, a características latitudinales, y a diversos eventos que ocurren en la alta atmósfera tanto de características globales como locales. La detección de los ecos provenientes de los diferentes estratos ionosféricos cobra una relevancia preponderante en el proceso de obtención de ionogramas confiables.

Ejemplos de ionogramas obtenidos por el sondador AIS-INGV del Observatorio Tucumán de Baja Latitud se muestran en la Figura 5.5.a (ionograma nocturno) y en la Figura 5.5.b (ionograma diurno).

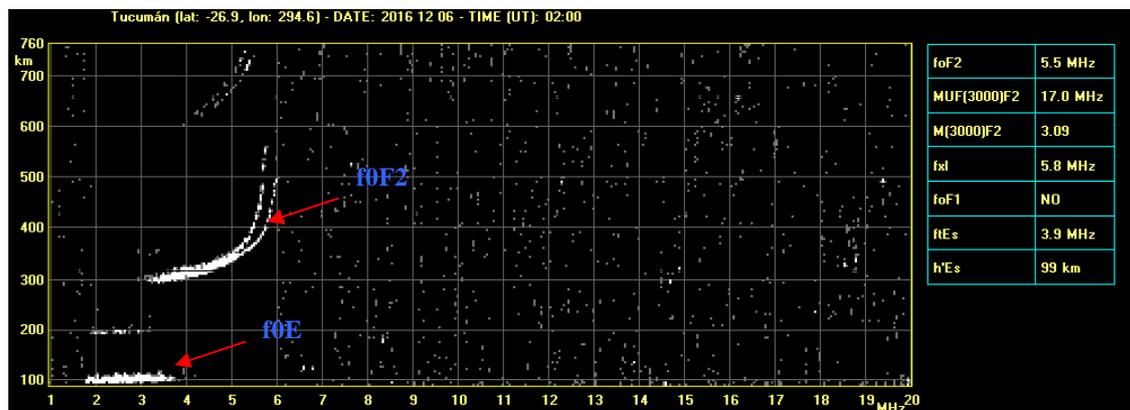


Figura 5.5.a. Ionograma correspondiente al día 06/12/2016 obtenido a las 02 hs UT (24 hs LT). El mismo se pueden observar los estratos ionosféricos típicos para la ionósfera nocturna en baja latitud. Nótese que los valores autoescalados para los estratos (tabla a la derecha), no coincide necesariamente con la exploración visual del gráfico. Estos valores son el resultado de aplicar un algoritmo de interpretación en una etapa posterior a la detección automática de las capas. Sobre el gráfico se encuentra indicada la interpretación manual del ionograma.

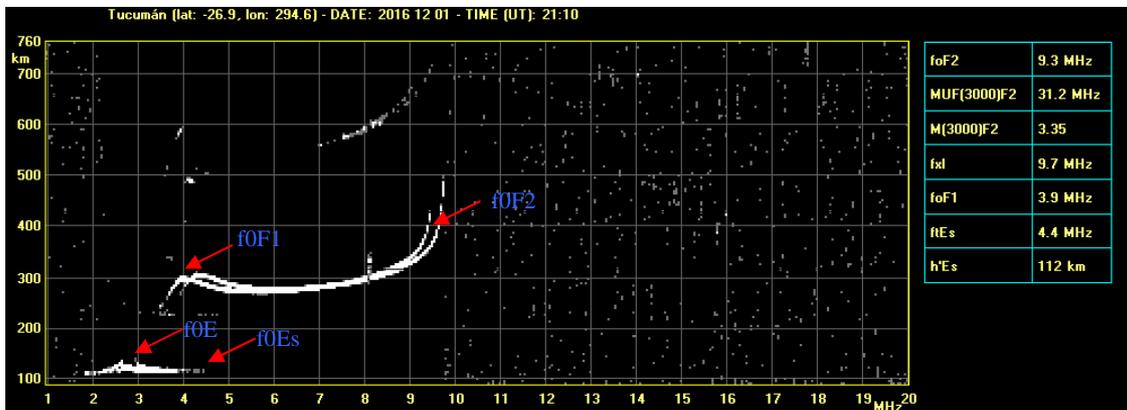


Figura 5.5.b. Ionograma correspondiente al día 01/12/2016 obtenido a las 21:10 hs UT (19:10 hs LT). El mismo se pueden observar los estratos ionosféricos típicos para la ionósfera diurna en baja latitud. En este caso valores escalados automáticamente (tabla a la derecha), coinciden con los valores escalados manualmente (señalados en el ionograma).

Si ocurre algún evento local que modifica la morfología o dinámica en la ionósfera, como por ejemplo perturbaciones viajeras, pasaje del terminador solar, etc, ó si ocurren eventos relacionados con perturbaciones provenientes del sistema Sol-Tierra; entonces los ionogramas serán muy variables e incluso en casos extremos imposibles de interpretar.

La interpretación de ionogramas se encuentra en una etapa posterior a la etapa de detección que se encarga de generar la representación gráfica. Cada punto blanco en el gráfico es el resultado de una detección positiva.

Como se mencionó en el capítulo II la función principal que debe llevar a cabo un radar es la función de detección del blanco. En el caso de un radar ionosférico el blanco es la ionósfera. Más precisamente, el sondador debe ser capaz de detectar los diferentes estratos ionosféricos.

En términos de detección, esto significa que cada muestra tomada de la señal recepcionada debe evaluarse de tal manera que se pueda decidir si corresponde o no a un eco reflejado en la ionósfera. Si esta decisión se realiza apropiadamente, cada vez que el sistema realice una detección, ésta corresponderá efectivamente a un eco y entonces el ionograma resultante será más confiable.

La operación de detección debe, por lo tanto, evitar falsas alarmas mientras que incrementa su capacidad de detección. La figura 5.6 muestra dos ionogramas sin post-procesamiento (antes de ser escalados) con dos configuraciones diferentes para sus parámetros de detección. Este experimento se llevó a cabo de la siguiente manera:

1. Se utilizaron dos valores de umbrales diferentes en cada uno. Un umbral bajo, para aumentar las detecciones y luego un umbral alto para evitar falsas alarmas.
2. El segundo ionograma fue tomado un minuto después del primero. El experimento se realizó en condiciones ionosféricas tranquilas, lo que significa que las éstas no varían significativamente durante un determinado lapso de tiempo (1 minuto en este caso) permitiendo comparar ambos casos (Rishbeth and Garriot, 1969)

Al analizar ambos ionogramas se puede observar que en el primero de ellos diferentes estratos ionosféricos son fácilmente reconocibles aunque además muestra un gran número de detección espurias (muchos de los puntos corresponden a ruido). Esta estrategia de detección introduce demasiada información adicional que en su mayoría no corresponden a ecos reales. Obviamente esta no es una situación deseable y en consecuencia esta estrategia no es la óptima.

Si por el contrario observamos el segundo ionograma, veremos que el ruido ha bajado y que los puntos graficados corresponden a ecos. A primera vista, la configuración para la detección en el segundo ionograma es satisfactoria. Sin embargo, no solo se eliminaron detecciones falsas sino que además se eliminó información útil. Un ejemplo de esto es que la segunda reflexión proveniente de la capa E no es visible, así como la traza de la capa F no está completamente determinada. Por lo tanto, se pierde información útil de la estructura de la ionósfera.

Es necesario obtener un compromiso entre detectar mucha información (elevando el ruido en el ionograma) y tener un ionograma mas claro (que presupone la pérdida de información valiosa). Con este experimento se muestra la necesidad de realizar un estudio exhaustivo para el diseño del sondador y determinar un modelo de detección acorde.

Si además, en una etapa posterior, se debe realizar el escalado (automática o manualmente), entonces el algoritmo de detección debe ser lo suficientemente robusto y sensible para responder adecuadamente a la mayoría de los casos.

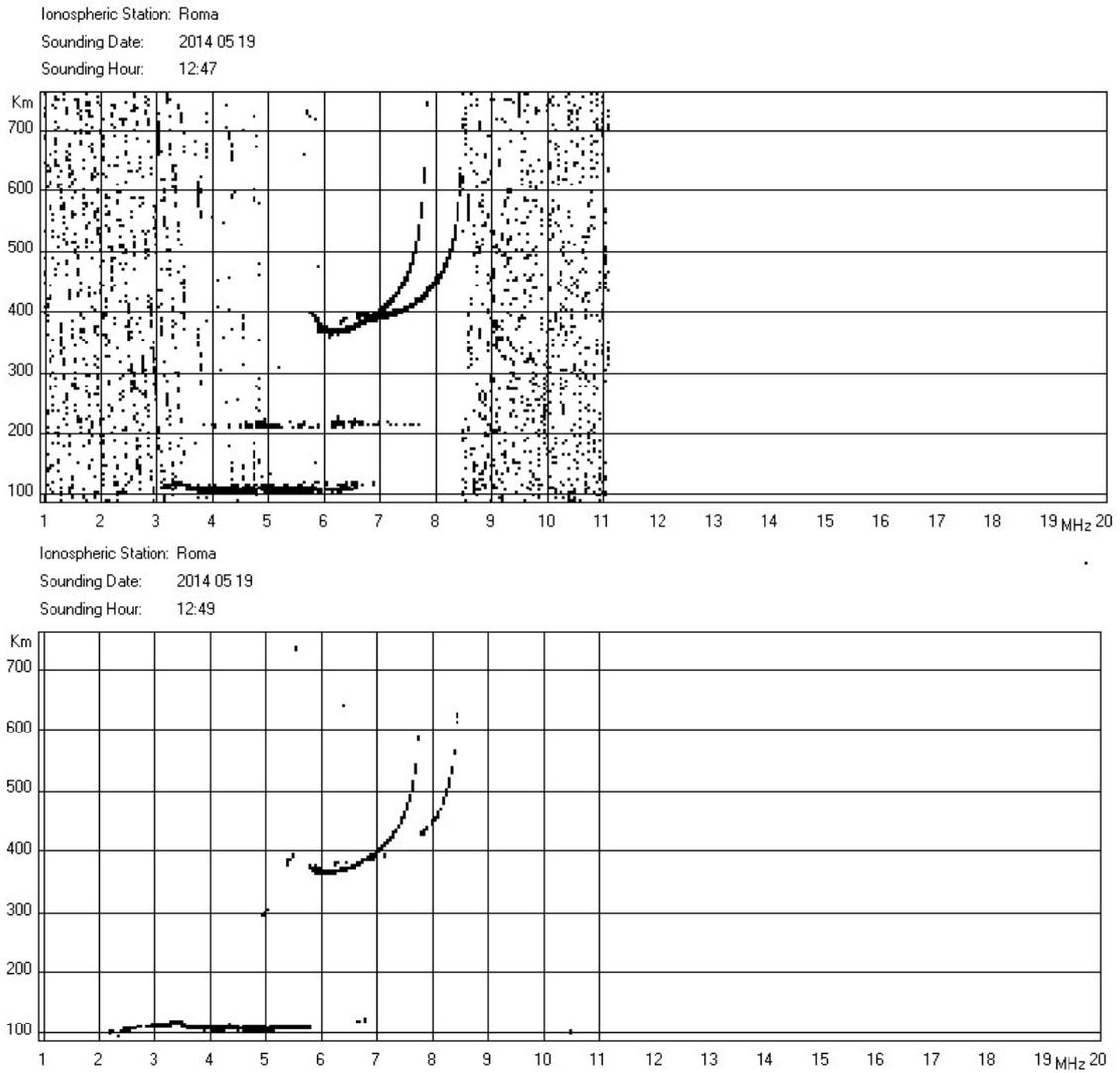


Fig. 5.6. Dos ionogramas con diferentes configuraciones para la detección.

Como se mencionó anteriormente, detectar significa reconocer un blanco dentro de una señal dada y distinguirlo del ruido interferencia, jamming o clutter en el que se encuentre inmerso.

Desde el punto de vista geofísico, la ionósfera puede ser considerada cuasi estacionaria por un período corto de tiempo (por ejemplo durante un sondaje) durante el cual no presenta variaciones significativas (Rishbeth and Garriot, 1969). En consecuencia durante ese período puede ser tratada como un blanco no fluctuante de acuerdo a la teoría de radares. Generalmente esta hipótesis se cumple, a menos que se presente una gran perturbación repentina. Este caso se analiza en secciones sucesivas cuando se trata el tema de integración coherente de pulsos.

A su vez, desde el punto de vista del radar, las capas ionosféricas son blancos complejos cuya RCD es descripta más precisamente estadísticamente utilizando para ellos su función de distribución de probabilidad (Curry, 2005; Richards, 2005). Por lo tanto el proceso de decisión (presencia o no del blanco) lleva a determinar una prueba de hipótesis estadística (ver Capítulo III).

Al finalizar el proceso de detección se obtendrá un valor de umbral (T) a partir del cual toda muestra (x_i) que lo supere será declarada como blanco. Así,

$$Detección = \begin{cases} \text{capa ionosferica} , & x_i > T \\ \text{no capa ionosferica} , & x_i \leq T \end{cases}$$

A partir de estas hipótesis sobre el blanco se realizará una propuesta de modelo de detección automática para un radar ionosférico.

En particular se realizará el estudio sobre la capacidad de detección del sondador ionosférico AIS-INGV y para ello se utilizarán datos adquiridos por el mismo. Las características de este equipo se encuentran descriptas por Zuccheretti et. al, 2003.

5.4. Modelo de detección de estratos ionosféricos

En este trabajo de tesis se propone un modelo de detección automática en base a una sucesión de etapas específicas para la estimación de la altura virtual de la ionósfera. Las

diferentes etapas que constituyen el modelo computacional de detección automática que realiza un sondador mediante sondaje vertical pueden enumerarse como sigue:

1. Generación de una señal portadora $s(t)$
2. Generación de un código de referencia, también denominado código local, $C[t]$ para codificar la portadora y para usar como patrón al momento de realizar la correlación de la señal transmitida $T_x(t)$ con la señal recibida $R_x(t)$.
3. Emitir múltiples pulsos hacia la ionósfera y posteriormente realizar la recepción.
4. Luego de la conversión a frecuencia intermedia en la antena receptora, muestrear la señal en cuadratura, $I[t]$ y $Q[t]$.
5. Correlacionar las señales muestreadas con el código local.
6. Mejorar la relación señal/ruido mediante diferentes técnicas.
7. Aplicar algún algoritmo apropiado para la detección automática.
8. Generar el ionograma.

La figura 5.7 muestra un esquema del proceso de detección para un sondador ionosférico en general. La señal portadora $s(t)$ es codificada con un código binario $C[t]$ para obtener una secuencia de pulsos $T_x(t)$ que es transmitido hacia la ionósfera. La etapa de recepción inicia con el muestro de la señal $R_x(t)$ en dos señales discretas $I[t]$ y $Q[t]$ que serán correlacionadas con el código local $C[t]$. Este proceso se repite para múltiples pulsos hasta obtener una señal Integral $[t]$ a la que se le aplicará el algoritmo de detección automática. Para cada detección se estimará la altura virtual correspondiente al retardo temporal sufrido por la señal emitida hacia la ionósfera.

Los 8 puntos anteriores pueden resumirse en las siguientes etapas que en conjunto constituye el modelo de detección para un determinado radar ionosférico:

1. Primera etapa: Determinación de un esquema de codificación
2. Segunda etapa: Recepción y Muestreo
3. Tercera etapa: Correlación
4. Cuarta etapa: Mejoramiento de la Relación Señal/Ruido.
5. Quinta etapa: Algoritmos de detección Automática

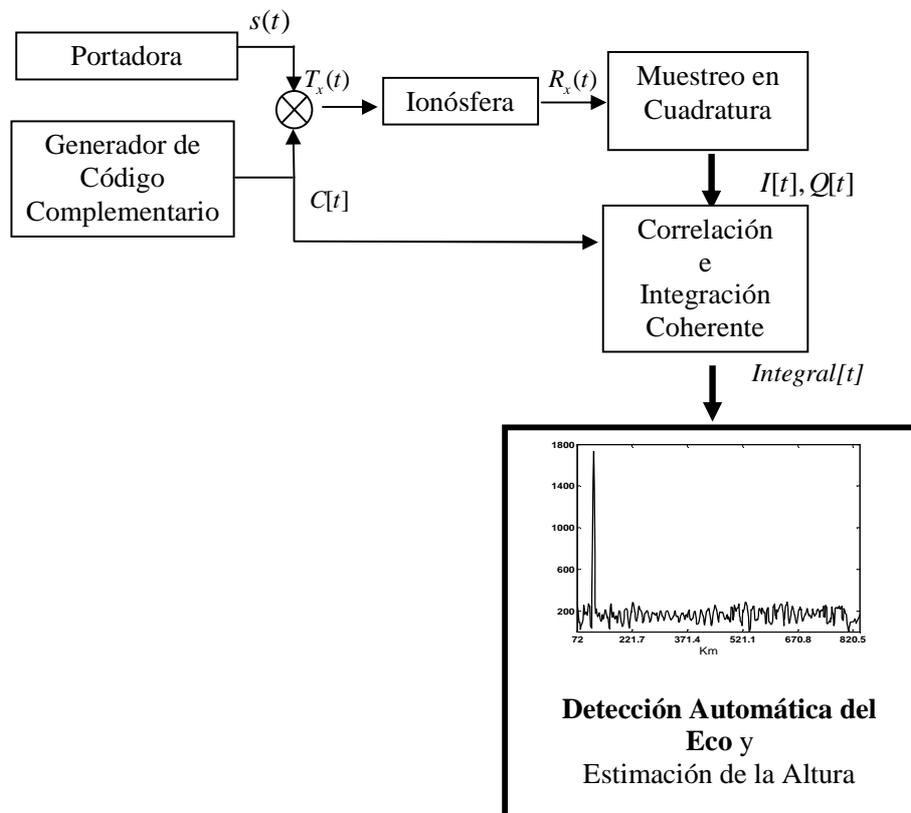


Figura 5.7. Esquema del proceso de detección para el AIS-INGV

5.5. Primera etapa del modelo: Codificación

Se mostró en el Capítulo II las ventajas de la técnica de esparcimiento de frecuencia sobre la tradicional técnica de detección por envolvente. Mediante el esparcimiento de frecuencia se obtiene una gran resolución de rango emitiendo baja potencia (Barry 1971, Poole 1985, Bianchi and Altadill 2005, entre otros). Para ello es preciso generar una señal portadora que deberá ser codificada con algún esquema de codificación adecuado. Esta técnica se denomina compresión de pulso. Cabe mencionar que la resolución de rango está relacionada a la mínima duración de un elemento del código, usualmente denominado subpulso (Patro et. al., 1990). En el caso del AIS-INGV se utilizan subpulsos de 30 ns de longitud lo que supone una resolución de rango de aproximadamente 4.5 km (Baskaradas et al. 2002, Zuccheretti et al. 2003) .

La primera etapa en la determinación del modelo entonces radica en la elección de código más apropiado. La manera más sencilla de codificación se realiza mediante la modulación en fase de la portadora, realizando cambios de fase entre dos valores posibles (0° y 180°) de acuerdo a una secuencia de dígitos binarios (Bianchi et. al., 2003). En el Capítulo II se mostraron algunos de los esquemas de codificación binarios más utilizados en estos tipos de radares.

El criterio de selección se basa en utilizar un esquema de codificación que permita utilizar una secuencia de gran longitud pero con la propiedad de que su función de autocorrelación tenga lóbulos laterales significativamente inferiores respecto al lóbulo principal. El código Barker, denominado código perfecto, posee las características adecuada respecto a su función de autocorrelación, sin embargo la mayor longitud de código Barker conocido es de 13 bits.

Como la longitud de un código puede influir en la capacidad de detección de cualquier radar pulsado, esta es una consideración importante que se debe tener en cuenta. Como se estudió, el código complementario introducido por Golay en 1961 (o secuencias de códigos complementarios) es un esquema de codificación binario que permite trabajar con secuencias largas de bits y mantener la ventaja de reducir los lóbulos laterales de su función de autocorrelación. Como esta técnica de codificación solo se puede aplicar a blancos de gran tamaño, cuya velocidad de movimiento no sea alta en el período de tiempo que demora la emisión y recepción de los pulsos es el esquema de codificación adecuado para el modelo propuesto.

Una copia del código generado se almacena para luego ser correlacionado posteriormente con la señal recepcionada.

En el caso del AIS-INGV no es necesario contar exactamente con un generador de secuencias de código complementario porque no es necesario modificar la información en cada pulso. El código es utilizado únicamente como referencia al momento de la recepción. Por este motivo se utiliza una secuencia fija de código complementario.

5.6. Segunda etapa del modelo: Recepción de señales y muestreo

Una vez codificada la portadora se emiten los pulsos hacia la ionósfera alternando los bits correspondientes a cada secuencia de código complementario.

La señal reflejada en la ionosfera es recepcionada y luego de ser filtrada, amplificada, y convertida a frecuencia intermedia, es finalmente muestreada en cuadratura. El resultado son dos señales discretas $I[t]$ y $Q[t]$ desfasadas una 90° respecto a la otra. Como a partir de ahora todas las señales involucradas serán señales discretas la notación para una señal discreta $S[t]$ será reemplazada por motivos de claridad simplemente por S .

El objetivo de realizar este muestreo es el de garantizar que se pueda detectar la fase de la señal original. En el peor caso de muestreo ocurre cuando para la señal I ó Q las muestras son tomadas en los instantes de tiempo correspondientes a amplitud cero. Pero si esto ocurre la muestra desfasada corresponderá al valor de la amplitud máxima (amplitud original de la señal), lo que implica que aún así no hay pérdida de información. En todos los otros casos, las muestras permitirán conseguir la fase de la señal original.

Además, se puede representar la señal como vectores bidimensionales en el plano I-Q. Estos vectores sirven para definir las amplitudes de las componentes en fase y cuadratura de la señal. De este modo para recuperar la amplitud (A) de la señal en banda base se puede realizar el siguiente cálculo (Ariokiasami, 2003):

$$A = \sqrt{I^2 + Q^2} \quad (5.1)$$

El código complementario posee dos secuencias de 16 bits, donde los bits de cada secuencia (cada bit representa un pulso) se emiten alternadamente. Es decir, cada pulso de la segunda secuencia es emitido a continuación del bit correspondiente de la primera secuencia.

Por lo tanto, se obtienen cuatro señales discretas resultantes del muestreo. Esto es, se obtienen, dos señales I (una para cada secuencia) que denominaremos I_1 e I_2 , y dos señales Q , a saber Q_1 y Q_2 .

En la siguiente etapa, los pulsos muestreados de I_1 y Q_1 deberán ser comparados con la primera secuencia de código complementario. El objetivo será poder determinar el retardo en tiempo que sufrió la señal. Del mismo modo, las señales I_2 y Q_2 deberán ser comparadas con la segunda secuencia de código complementario.

Es preciso remarcar que en el proceso de digitalizar una señal analógica mediante muestreo se puede incurrir en errores.

Los errores típicos relacionados con el muestreo de señales son aquellos que están asociados a la frecuencia de muestreo elegida. En este caso, como se conoce la frecuencia de la señal portadora es posible determinar una tasa de muestreo tal que permita la correcta reconstrucción de la señal receptionada en la antena, esto es, se muestrea a una tasa superior a la tasa de Nyquist.

Existe un segundo tipo de error que tiene directa relación con la técnica de compresión de pulso. Supongamos que se transmite una señal de modo que el instante en el cual se inicia la transmisión se denomina t_0 . Supongamos además que después de τ instantes de tiempo se obtiene el primer eco en la antena receptora. De manera ideal, el eco presente en la señal receptionada será una versión desplazada en el tiempo de la señal transmitida. La Figura 5.8 ilustra esta situación.

Si, además pensamos que la señal receptionada es una señal digital resultante de la discretización donde cada valor de la muestra fue tomada cada Δt instantes de tiempo; entonces podemos suponer que para la mayoría de los casos τ no es un múltiplo de Δt . Esto significa que no es probable poder recuperar exactamente el instante de tiempo que coincide con el desplazamiento real que sufrió la señal original.

Es importante saber cual es la cota de error que se puede tener en este caso. Dicho de otra manera, es necesario conocer cual es máximo error que se puede cometer en términos de distancia de la altura virtual que es el producto final, y si estos errores son o no aceptables.

Si las muestras se toman cada $10 \mu s$, o lo que es lo mismo $\Delta t = 10 \mu s$ y la altura virtual de la ionósfera se puede estimar con la ecuación:

$$h_v = \frac{c\tau}{2} \quad \text{con } c = 3 \times 10^8 \text{ m/s} \quad (4.2)$$

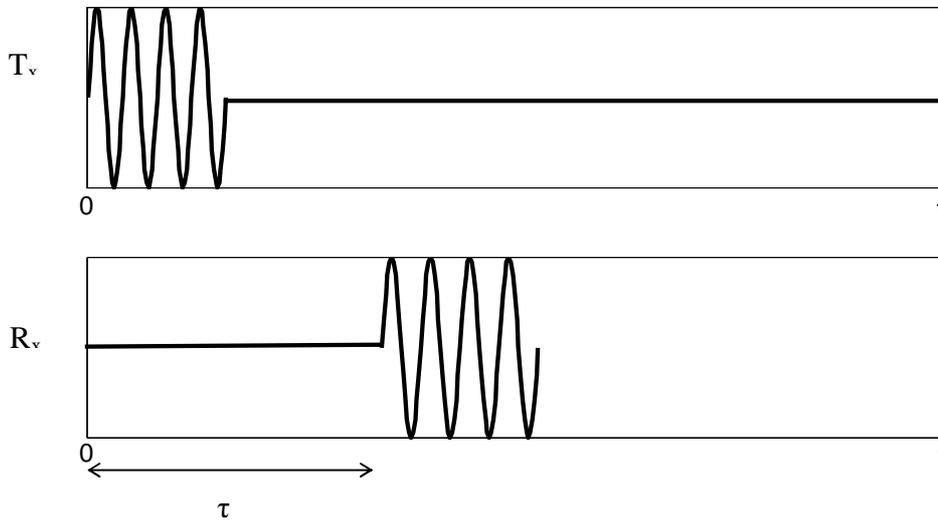


Figura 5.8. La señal transmitida Tx es receptionada al cabo de τ instantes de tiempo

Se puede estimar Δh_v como

$$\Delta h_v = \frac{c\Delta t}{2} \quad (4.3)$$

$$\Delta h_v = \frac{3 \times 10^5 \text{ Km/s} \times 10 \times 10^{-6} \text{ s}}{2} = 1,5 \text{ Km}$$

Por lo tanto se puede decir que la altura virtual de la ionósfera es $h_v \pm 1,5 \text{ Km}$.

En este caso la portadora se codifica con un código de $480 \mu\text{s}$ que tiene una longitud de 16 bits ($\Delta t = 30 \mu\text{s}$ de duración de cada chip de código). Esto significa que cada 30μ cambia la fase de la modulación si es que el chip del código cambia (de 0 a 1, o viceversa). Por lo tanto,

$$\Delta R = \frac{3 \times 10^8 \frac{\text{m}}{\text{s}} \times 30 \times 10^{-6} \text{ s}}{2} = \frac{3 \times 10^2 \times 30 \text{ m}}{2} = \frac{90 \times 10^3 \text{ m}}{2} = 4.5 \text{ Km}$$

Este valor representa la resolución de rango del radar. Por lo tanto, el radar no puede distinguir entre dos blancos diferentes si éstos se encuentran separados por menor distancia.

Una observación importante es que a pesar de trabajar con dos secuencias de códigos complementarias de $480\mu s$ de duración cada una, la distancia intrapulso y la frecuencia de repetición de pulsos permite tratar a cada secuencia de manera independiente. De este modo, ambas secuencias solo serán necesarias al momento de correlacionar la señal completa.

5.7. Tercera etapa del modelo: Correlación

Después de realizado el muestreo es necesario recurrir a diversas técnicas de procesamiento de señales con el objetivo de extraer la información presente en una señal recepcionada, que en este caso significa obtener el tiempo de retardo del eco (Bianchi et al., 2003, entre otros).

De manera ideal al “comparar” las muestras con el código local se puede determinar el desplazamiento en tiempo que sufrió la señal. A pesar de que en teoría es sencillo comparar estas dos señales, en la práctica resulta que la señal recepcionada se encuentra muy atenuada e inmersa en ruido proveniente de diversas fuentes (en muchos casos afecta por interferencia, jamming o incluso clutter).

La técnica utilizada para realizar esta comparación es la correlación de señales. En el Capítulo II se realizó un estudio pormenorizado de este método.

Se consideraron dos metodologías para realizar la correlación en el dominio del tiempo.

Module-Based Correlation:

El primero combina los canales I y Q para obtener un valor de amplitud compleja para posteriormente correlacionar con el código local (Skolnik, 1980). En este caso con el objetivo de mantener la información del código es necesario recuperar correctamente su signo antes de realizar a correlación. Para ello el signo es recuperado de uno de los canales, por ejemplo del canal I. A este método lo denominamos *Module-Based Correlation* (MBC).

Por lo tanto este método se puede escribir como:

1. Extraer el signo
2. Obtener las amplitudes correspondientes a cada secuencia de código complementario.

$$A_1 = \sqrt{I_1^2 + Q_1^2} \quad (5.4)$$

$$A_2 = \sqrt{I_2^2 + Q_2^2} \quad (5.5)$$

3. Calcular la correlación entre cada una de las amplitudes con su correspondiente secuencia de código complementario.

$$Corr_1 = A_1 * Cod_1 = KA_1 \quad (5.6)$$

$$Corr_2 = A_2 * Cod_2 = KA_2 \quad (5.7)$$

Donde K es la ganancia de correlación (Cabrera et.al., 2010).

4. Estimar la amplitud total con el fin de obtener los beneficios del código complementario de la siguiente manera:

$$MBC_amp = Corr_1 + Corr_2 = K(\sqrt{I_1^2 + Q_1^2} + \sqrt{I_2^2 + Q_2^2}) \quad (5.8)$$

Quadrature Component Correlation

El segundo método se denomina *Quadrature Component Correlation* (QCC) y se basa en realizar la correlación de cada canal, de manera separada, con la secuencia de código complementario que corresponda. La amplitud se obtiene al final del proceso (Skolnik, 1990). La principal diferencia con el método anterior es que no es necesario realizar ninguna corrección para obtener el signo puesto que la amplitud se calcula luego de correlacionar.

El método QCC se puede resumir en los siguientes pasos:

1. Cada uno de los canales I y Q se correlacionan con la secuencia de código local que le corresponda

$$CorrI_1 = I_1 * Cod_1 = K I_1 \quad (5.9)$$

$$CorrQ_1 = Q_1 * Cod_1 = K Q_1 \quad (5.10)$$

$$CorrI_2 = I_2 * Cod_2 = K I_2 \quad (5.11)$$

$$CorrQ_2 = Q_2 * Cod_2 = K Q_2 \quad (5.12)$$

2. Una vez realizadas las correlaciones, se realizan las sumas

$$SumI = CorrI_1 + CorrI_2 = K (I_1 + I_2) \quad (5.13)$$

$$SumQ = CorrQ_1 + CorrQ_2 = K (Q_1 + Q_2) \quad (5.14)$$

En este paso se obtienen las ventajas del código complementario. SumI y SumQ se obtienen mediante suma algebraica debido a que las muestras I1 e I2 tienen la misma fase, lo cual también se cumple para las muestras Q1 y Q2. Como la correlación es una operación lineal, es posible realizar estas sumas.

3. En el siguiente paso se combinan las correlaciones por canales SumI y SumQ con el fin de obtener la amplitud total

$$QCC_amp = (\sqrt{SumI^2 + SumQ^2}) \quad (5.15)$$

Esta operación es posible debido a que la correlación no produce un desfase de la información de fase original en los canales I y Q, manteniendo las propiedades del muestreo en cuadratura. Más explícitamente se puede reescribir la ecuación anterior como:

$$\begin{aligned}
QCC_amp &= (\sqrt{K^2 (I_1 + I_2)^2 + K^2 (Q_1 + Q_2)^2}) = \\
&= K\sqrt{I_1^2 + I_2^2 + 2I_1 I_2 + Q_1^2 + Q_2^2 + 2Q_1 Q_2}
\end{aligned} \tag{5.16}$$

Como se puede observar los métodos MBC y QCC no arrojan los mismos resultados y por lo tanto es necesario analizar el desempeño de cada uno de ellos.

En un caso ideal donde no exista presente ruido, un blanco puede encontrarse fijo entre dos pulsos consecutivos. Si consideramos este caso particular, podemos decir que podemos tener $I_1=I_2$ y $Q_1=Q_2$. Entonces:

$$MBC_amp = K(\sqrt{I_1^2 + Q_1^2} + \sqrt{I_2^2 + Q_2^2}) = K(2\sqrt{I_1^2 + Q_1^2}) = 2KA$$

y

$$\begin{aligned}
QCC_amp &= K\sqrt{I_1^2 + I_2^2 + 2I_1 I_2 + Q_1^2 + Q_2^2 + 2Q_1 Q_2} \\
&= K\sqrt{I_1^2 + I_1^2 + 2I_1 I_1 + Q_1^2 + Q_1^2 + 2Q_1 Q_1} \\
&= K2\sqrt{I_1^2 + Q_1^2} = 2KA
\end{aligned}$$

Así en el caso ideal ambos métodos obtienen el mismo valor de amplitud.

En general $I_1 \neq I_2$ y $Q_1 \neq Q_2$ y las señales se encuentran afectadas por diversas fuentes de ruido.

Se puede mostrar (Cabrera et. al., 2010) que en presencia de ruido y en escenarios más realistas, el métodos MBC puede introducir ruido adicional al momento de la recuperación del signo. En estos casos el método QCC muestra un comportamiento más robusto y por esto será el método que se utilizará en el modelado.

4.8. Cuarta etapa del modelo: Mejoramiento de la relación señal/ruido

Una vez realizadas las correlaciones y el cálculo de la amplitud de la señal en muchas ocasiones es necesario realizar tareas para mejorar las posibilidades de detección. Es

decir, el objetivo es potenciar aquellos posibles ecos dentro de la señal mientras se intenta atenuar el ruido o interferencia.

Si ocurre que el blanco no fluctúa durante un determinado lapso de tiempo, cada pulso emitido y recepcionado en dicho lapso será reflejado a igual distancia. Esto significa que todos estos pulsos recepcionados pueden ser sumados en fase.

Esta situación genera que los ecos se sumaran aportando mayor energía a la señal final, mientras que las reflexiones provenientes de ruidos tenderán a ser canceladas. Las características estocásticas del ruido o de la interferencia permiten que ocurra esta situación.

Esta técnica para el mejoramiento de la relación señal/ruido o “Signal to Noise Ratio” (SNR) se denomina integración coherente. En un esquema ideal de integración coherente, si se suman n pulsos, se logra una ganancia igual a n (Barton y Leonov, 1998).

Como se analizó con anterioridad, se puede suponer que la ionosfera permanece cuasi estacionaria durante un período corto de tiempo, en general en el orden del minuto (Rishbeth and Garriot, 1969). Si utilizamos esta hipótesis podemos utilizar la técnica de integración coherente de múltiples pulsos para mejorar la SNR para el caso particular del sondador ionosférico.

Con el objetivo de poder evaluar el aporte de la integración coherente durante el proceso de detección es necesario poder estimar un valor para el SNR.

Idealmente al finalizar la etapa de correlación, se obtiene una señal que en el caso de poseer un eco tendrá un pico sobresaliente correspondiente al retardo de tiempo donde ocurrió la reflexión de mayor energía. Éste corresponde a la presencia del blanco. La Figura 5.9 muestra esta situación.

Supongamos entonces que cada pico o valor máximo representa la presencia del blanco. Por lo tanto, se puede evaluar la relación que existe entre la amplitud de este valor máximo versus el ruido en la señal.

Para este cálculo se utilizará la ecuación que corresponde al valor de la relación señal ruido medida en dB,

$$SNR = 10 \log \left(\frac{S_0}{N_0} \right) \quad (5.17)$$

donde S_0 representa la potencia de la señal original y N_0 es la potencia de ruido.

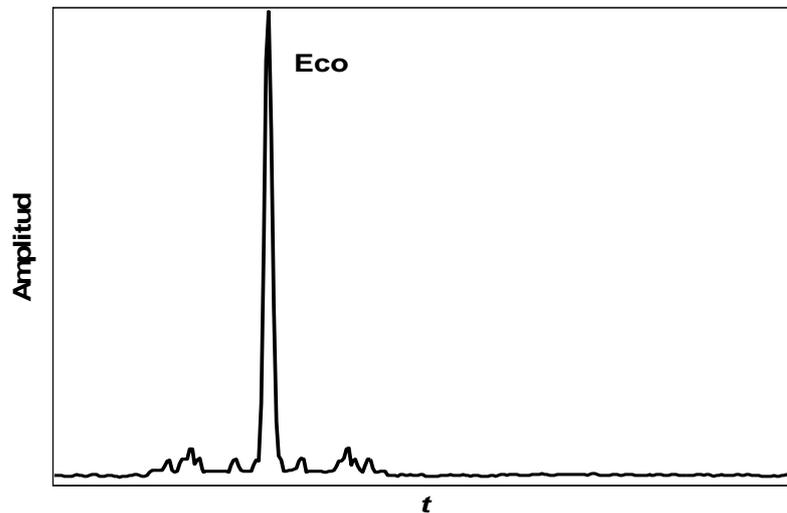


Figura 5.9. Grafica de una señal ideal con un eco presente obtenida al finalizar la correlación de un pulso.

La aplicación de esta ecuación resulta complicada por el hecho de que en la señal recepcionada no es fácil discriminar que es señal y que es ruido. De hecho, la intención es poder determinar si el valor de un pico determinado (en este caso el valor o pico máximo) representa la presencia del blanco o no. Puede ocurrir que no sea posible distinguir claramente cada pico máximo y que incluso aparezcan más de uno.

Por lo tanto, para realizar esta evaluación es preciso discriminar el pico del ruido. Si se quiere determinar si un valor particular representa la presencia de información, entonces podemos suponer que representa la amplitud de la señal. Y la ecuación anterior puede ser re escrita como:

$$SNR = 10 \log \left(\frac{(N-1)A^2}{\sigma^2} \right) \quad (5.18)$$

Donde,

$(N-1)A^2$ representa la potencia de la señal, y σ^2 representa la potencia del ruido.

Con N igual a la longitud del código, A la amplitud de la señal y σ es la desviación estándar de las muestras correspondientes al ruido. Debido a que la muestra que representa la señal útil no forma parte del ruido, este valor no es tenido en cuenta para realizar el cálculo de la desviación estándar. (Tomasi, 1996)

A partir de la ecuación 4.18 se evaluará la SNR para diferentes casos de estudio y se analizarán las mejoras que puede introducir la integración coherente de múltiples pulsos. Se realizó el estudio de diversos casos reales adquiridos mediante el sondador ionosférico AIS-INGV ubicado en la estación ionosférica de Roma para una frecuencia fija. Cada uno de estos casos responde a diferentes casos generales:

- a. Sondaje a lazo cerrado. Con el objetivo de poder comparar los diferentes resultados se generaron datos a lazo cerrado simulando el comportamiento de un eco pero sin tener presente ruido RF. La adquisición de este ejemplo a lazo cerrado se obtiene una señal a partir de un circuito interno en el sondador que genera un retardo programable para poder, bajo condiciones controladas, calibrar el radar.
- b. Datos reales adquiridos mediante el AIS-INGV. En particular se tratarán los siguientes casos:
 - i. Señal que contiene un eco de gran energía respecto al ruido.
 - ii. Señal que contiene un eco inmerso en altos niveles de ruido
 - iii. Señal que contiene un eco inmerso en una señal de interferencia.
 - iv. Señal que contiene más de un eco.
 - v. Señal que contiene únicamente ruido.

Los datos de cada uno de estos casos de estudio fueron siguiendo las etapas de detección anteriores (codificación, muestreo, correlación) y luego se procedió a integrar múltiples pulsos.

Las simulaciones que se ejecutaron para el caso de datos sintéticos arrojaron como resultado que la relación señal ruido está en el orden de 38 dB, que será utilizada como

patrón para evaluar el resto de los casos de estudio ya que se trata de un caso ideal sin ruido.

a. Sondaje a lazo cerrado

Se generó una portadora codificada con una secuencia de código complementario. Se realizó el sondaje en circuito cerrado en el equipo. Se muestreo en cuadratura, y se realizó la correlación siguiendo el método QCC. La figura 5.10 muestra el muestreo de la señal correspondiente a ambas secuencias de código. Para la codificación se usaron las siguientes secuencias de código complementario: Código1=[1101111010001011] y Código1=[1101111001110100].

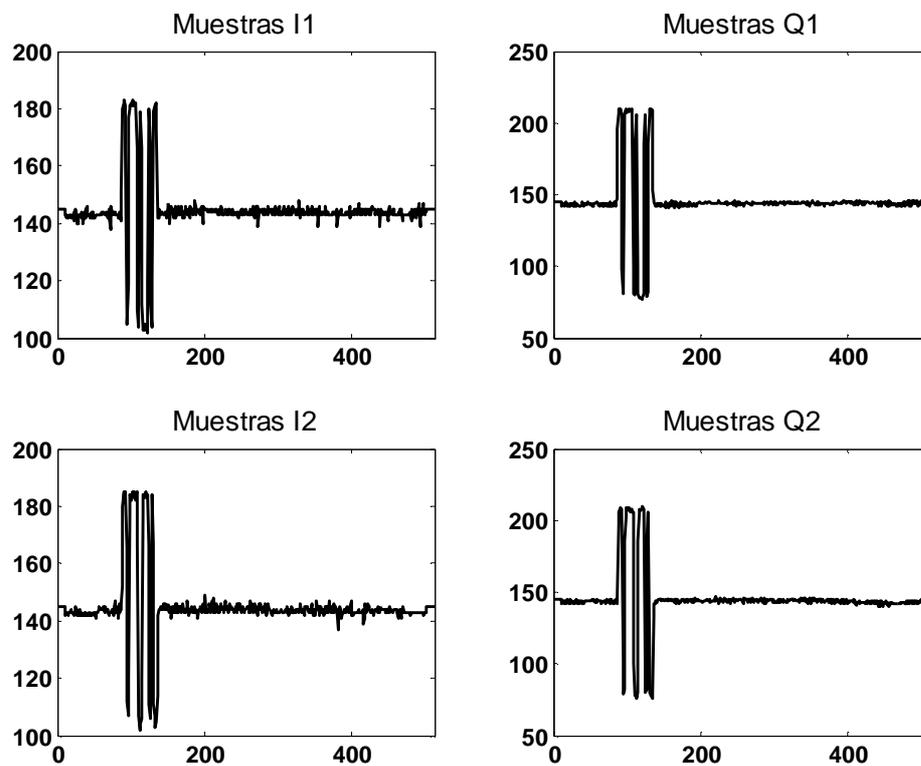


Figura 5.10. Muestreo en cuadratura para un sondaje a lazo cerrado. El eje x representa el número de muestra mientras que el eje y corresponde a las amplitudes.

La frecuencia de muestreo fue de 100kHz y se tomaron 512 muestras para cada canal y cada secuencia de código. Esto se significa que por cada chip de código de 30 μ se muestrea tres veces (se toma una muestra cada 10 μ). Dado que se toman 512 muestras en total, el tiempo de muestreo T_m puede calcularse como:

$$T_m = 10\mu \times 512 = 5120\mu = 5.12 \text{ ms}$$

Lo que equivale a que la altura virtual (h_v) sea de (ecuación 2.11):

$$h_v = \frac{cT_m}{2} \cong 768 \text{ Km}$$

Este es el valor de altura de diseño del AIS-INGV

El objetivo es tener una ventana de tiempo de 5.12 ms antes de disparar otro pulso para evitar error de ambigüedad.

Una vez realizado el muestreo en cuadratura se realizan las correlaciones y el cálculo de la amplitud de la señal final (Figuras 5.11 y 5.12).

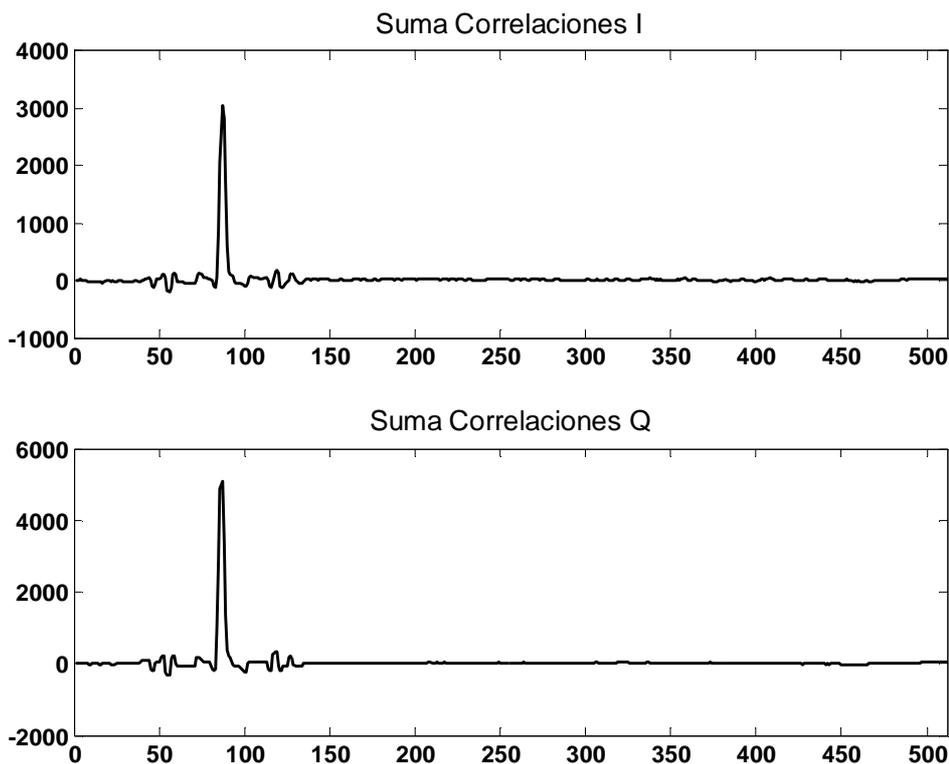


Figura 5.11. Suma de correlaciones para cada canal

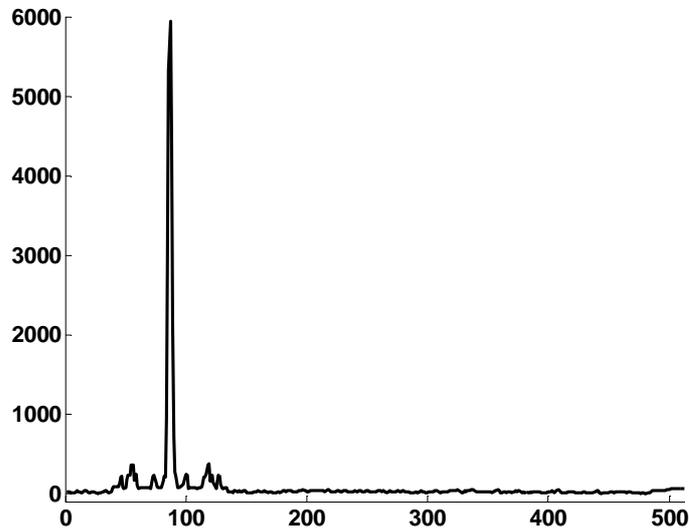


Figura 5.12. Estimación de la amplitud de la señal recepciónada después de ser correlacionada con el código local

Al finalizar se obtiene una señal que posee un pico de energía distintivo que corresponde a un eco. Los bajos niveles de ruido, casi despreciables, hacen de este caso un caso ideal. Es decir, que muy poco probable que se corresponda con un escenario real pero que sirve para estudiar el mayor SNR alcanzable en el equipo. Para ello se considera la amplitud del máximo (eco) y se lo compara con el nivel de ruido que será estimado como el valor promedio de las muestras diferentes al eco.

Sultzer y Woodman (1984) sugirieron como criterio general de diseño en radares de alta atmósfera la utilización de códigos con funciones de autocorrelación con grandes lóbulos principales y lóbulos laterales mínimo donde especifican que, la potencia total en la correlación de los lóbulos laterales debe ser del orden del 20% o menos que el lóbulo principal. Este criterio es teórico y supone que las señales no se encuentran degradadas, inmersas en ruido ni desplazadas en el tiempo. Sin embargo puede servir como referencia para estimar un valor de SNR aceptable al momento de diseño del radar.

Se calculó que el valor de SNR en este caso es de alrededor de 38 dB de acuerdo a la ecuación 5.18.

Si se sigue el criterio de Sultzer y Woodman, se puede pensar entonces que si la SNR de un eco respecto al ruido es menor a 30 dB, ésta no cumple con el criterio. En tal caso, será necesario usar algún tipo de técnica de mejoramiento para la relación señal/ruido. En particular y como se mencionó anteriormente, se utilizaran múltiples pulsos para integrar coherentemente.

b. Datos reales

Se analizaron diferentes casos típicos utilizando sondajes reales para una frecuencia fija adquiridos por el AIS-INGV ubicado en la estación ionosférica Roma (41.8 N, 012.5 E). Para ello se realizaron experimentos de campo en el que la tesista participó obteniendo datos in situ del sistema. Para ello, se realizó una campaña de medición de 3 meses en el año 2014.

Con condiciones reales nos referimos a que estos datos fueron obtenidos a partir de realizar el proceso real de transmitir una onda de radio codificada y luego muestrear en cuadratura en la antena receptora para posteriormente almacenarlas en un archivo de salida.

Las señales muestreadas son almacenadas en archivos binarios de extensión “.raw” y corresponden a una frecuencia fija. Cada archivo corresponde a una fecha, hora y minuto específico y contiene múltiples pulsos. En cada archivo se almacenaron datos correspondientes a un tiempo de sondaje de aproximadamente 100 segundos durante los cuales se transmitió y recibió pulsos de 480 μ s (Arokiasami et al, 2003).

Se realizó una campaña de adquisición de datos entre los meses de Marzo y Mayo de 2014. Se realizaron sondajes durante los minutos 05, 07, 09, 20, 22, 24, 35,37, 39, 50, 52, 54 por cada hora del día obteniéndose un número de muestras suficientes para realizar estudios sobre detección.

El experimento se diseñó utilizando diferentes valores de frecuencia para analizar los diferentes estratos ionosféricos (a diferentes alturas) y las diversas condiciones ionosféricas. Para cada una de estas frecuencias, se realizaron sondajes por un período

que varía entre 2 y 3 días durante los cuales la frecuencia seleccionada permaneció constante. A lo largo de este trabajo de tesis se utilizaran diferentes datos de esta campaña para ilustrar los diferentes temas que se traten.

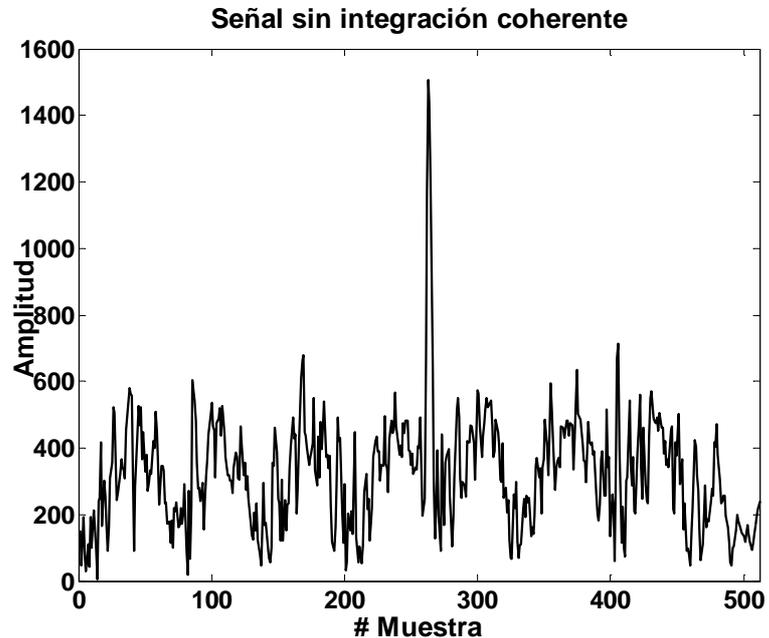
Los casos analizados son:

- i. Señal que contiene un eco de gran energía respecto al ruido.

Este caso es el más cercano al ideal y el que en principio requerirá un menor número de integración de pulsos. Si bien se realizaron cálculos sobre un gran número de archivos se seleccionó a modo de ejemplo el correspondiente al día 01 de Abril de 2014. El sondaje se realizó a las 19:35 UT usando una frecuencia de 8 MHz.

La figura 5.13 compara los resultados obtenidos sin realizar integración coherente (Figura 5.13.a) con el resultado de integrar 30 pulsos (Figura 5.13.b).

Figura 5.13.a



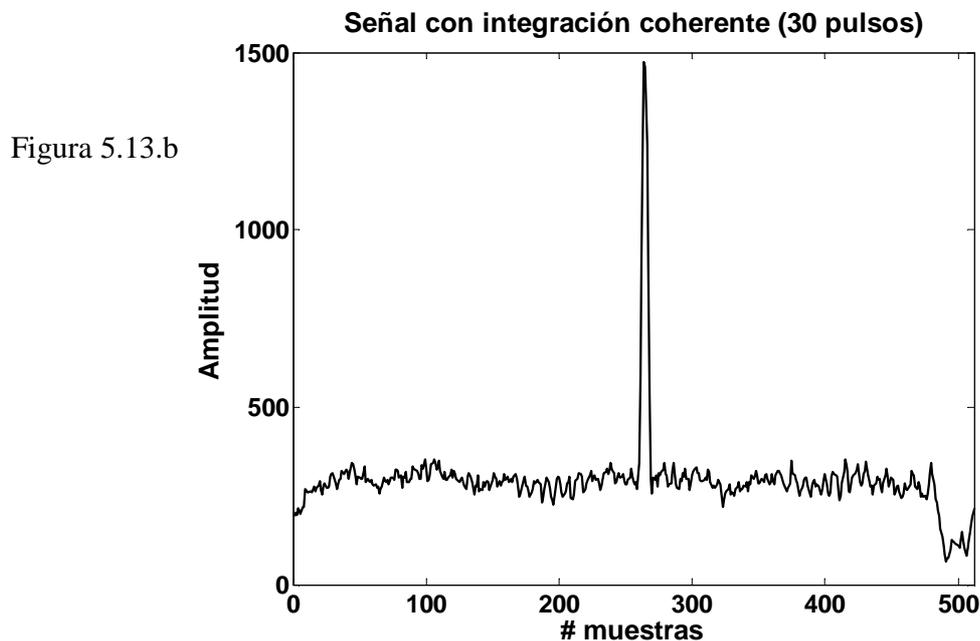


Figura 5.13. Comparación entre señales resultantes utilizando un único pulso y de realizar integración coherente (30 iteraciones)

Es fácilmente observable como al integrar múltiples pulsos el nivel de ruido decrece significativamente debido a las características estocásticas del mismo lo que hace que tienda a cancelarse. Mientras que en los sucesivos pulsos los ecos se encuentran aproximadamente a la misma altura y sus amplitudes se suman coherentemente. Una vez realizada la integración la amplitud de la suma coherente de las n muestras es dividida por este escalar (factor de escala del aumento de la amplitud durante la integración). Los valores estimados para la SNR entre el eco y el ruido es de 29 dB en caso de usar un único pulso y de 33 dB utilizando 30 iteraciones. Esto significa que se cumple el criterio de diseño establecido en el caso patrón ($SNR > 30$ dB) usando integración coherente.

ii. Señal que contiene un eco inmerso en altos niveles de ruido

Un caso similar con un único eco pero inmerso en mayores niveles de ruidos también se analizó. En la Figura 5.14 a y b se muestran los resultados

obtenidos para el caso de solo utilizar un pulso y de utilizar 30 pulsos. Estos datos corresponden al sondaje realizado el día 09 de Abril de 2014 a las 17:22 hs para una frecuencia fija de 3 MHz.

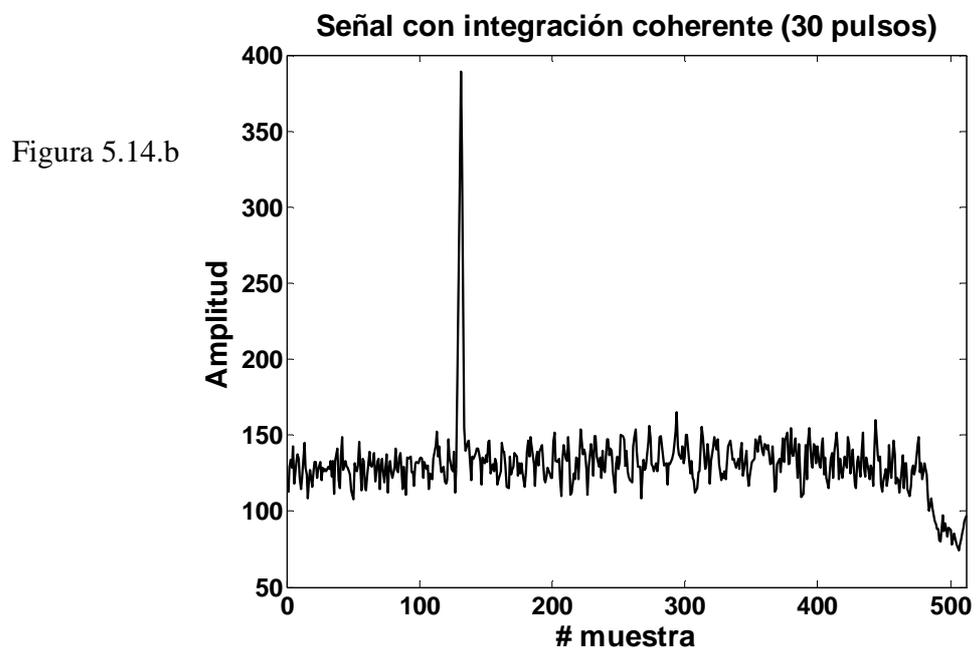
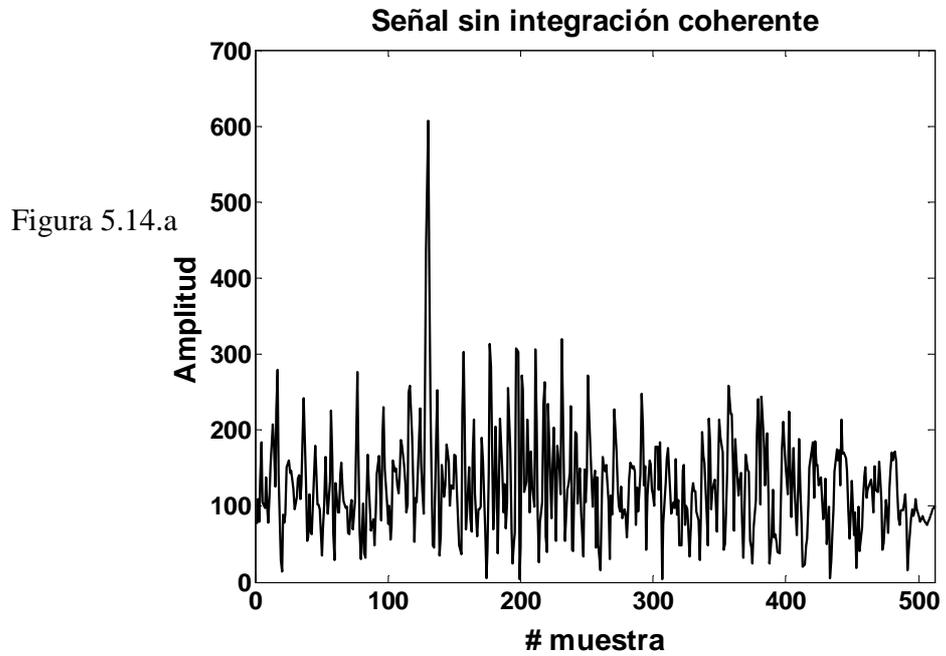


Figura 5.14. Comparación entre señales resultantes utilizando un único pulso y de realizar integración coherente (30 iteraciones)

En este caso también es claro el aporte de múltiples pulsos integrados. El ruido es atenuado mientras que el eco es realzado. Se estimó que el valor de SNR para el caso de un pulso es aproximadamente 29 dB mientras que al cabo de 30 iteraciones se logro un valor de 34 dB.

iii. Señal que contiene un eco inmerso en una señal de interferencia.

Un caso específico de señales que degradan la detección del eco son las señales de interferencia. Estas señales indeseadas son producto de algún otro equipo emisor de señales de radio que afectan a las señales del radar. Se analizaron diversos casos de interferencia.

En particular la Figura 5.15 a y b muestran los resultados obtenidos para el sondaje realizado el día 03 de Abril de 2014 a las 19:37 UT para una frecuencia de 8 MHz. En este caso se puede observar que existe una señal moduada superpuesta a la señal con el eco. Una observación importante es que al cabo de 30 iteraciones disminuyó la amplitud del ruido pero no de la señal de interferencia. Esto sucede debido a que tanto el eco como la interferencia son coherentes. Es observable además que al bajar el nivel de ruido un segundo eco aparece y se puede suponer la presencia de dos blancos o lo que es equivalente en nuestro caso, se observan dos estratos ionosféricos para una misma frecuencia.

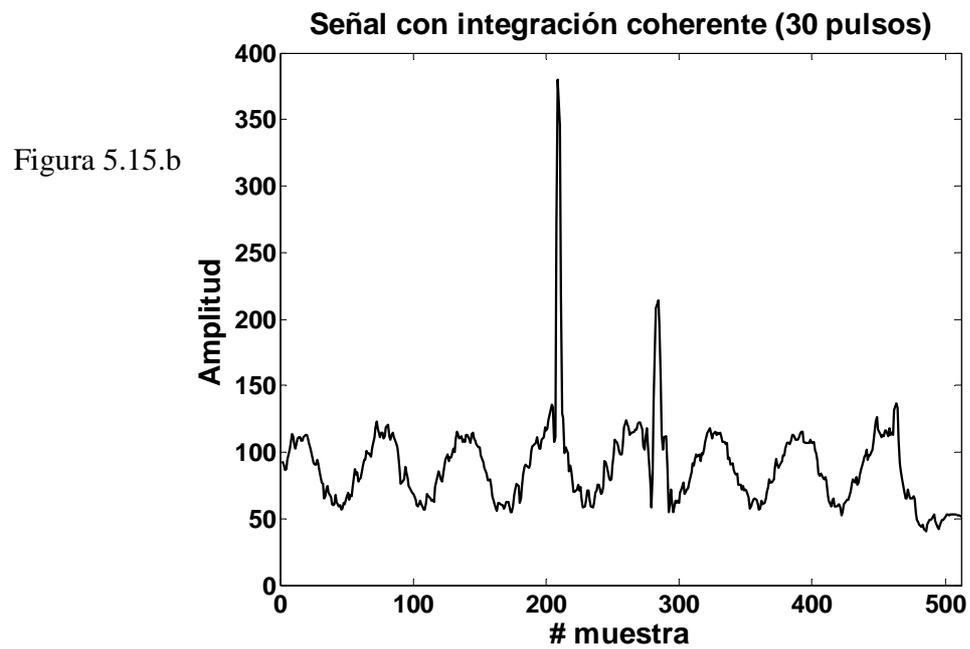
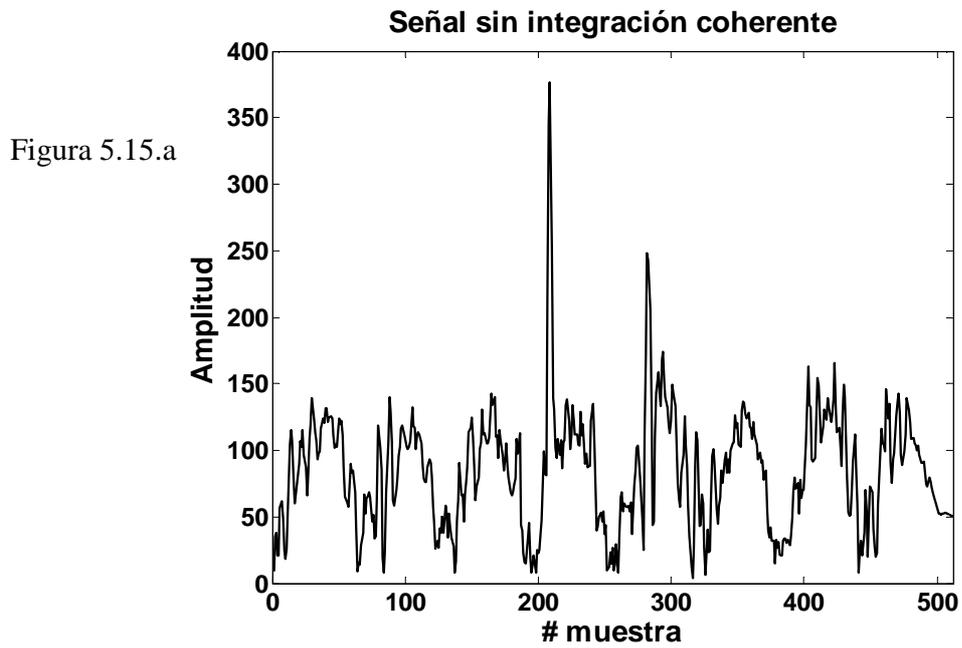


Figura 5.15. Comparación entre señales resultantes utilizando un único pulso y de realizar integración coherente (30 iteraciones)

Se calculó el valor SNR alrededor de 28.4 sin utilizar múltiples pulsos y de 31.8 dB usando 30 iteraciones.

Una observación importante es la relacionada al cálculo del SNR (ecuación 4.18) ya que éste se realiza comparando la amplitud del pico máximo contra el resto de las muestras que representan el ruido y/o interferencia. En este caso dentro el segundo eco no es tratado como un blanco sino que es computado como ruido lo que hace que la SNR sea menor.

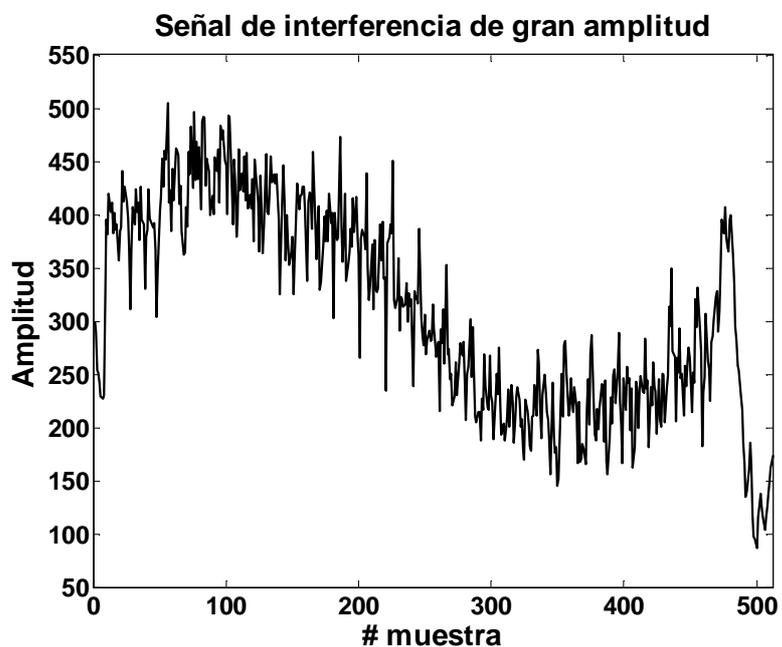


Figura 5.16. Señal de interferencia sensada el 09 de Abril de 2014, 17:36 UT a una frecuencia de 6 MHz.

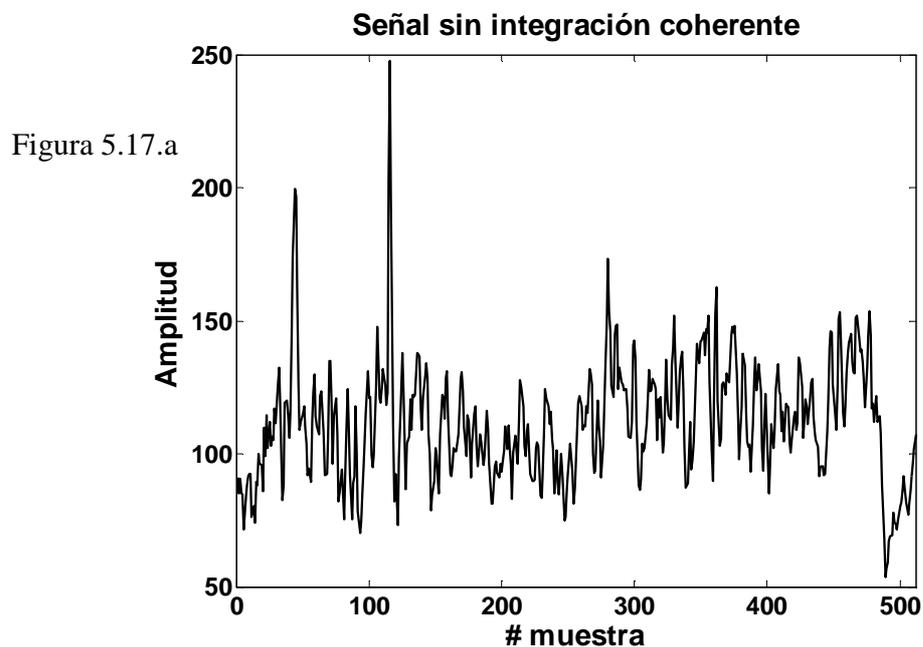
Otra nota importante es que en este caso la amplitud de la interferencia es mucho menor a la del eco lo que no ocurre en todos los casos. La figura 5.16 ilustra este caso. La gráfica corresponde a las mediciones realizadas el 09 de Abril de 2014 a las 17:36 UT para una frecuencia fija de 6 MHz. En algunas circunstancias la señal de interferencia es tan importante que se impone sobre la fase del código haciendo imposible la detección de

cualquier capa ionosférica. En particular el ejemplo se trata de una señal de interferencia que ocurre periódicamente en la zona cercana a la antena del equipo en frecuencias muy cercanas a los 6 MHz.

iv. Señal que contiene más de un eco.

En muchas ocasiones no hay un único eco sino que se pueden encontrar ecos múltiples. En el caso de por ejemplo contar con dos ecos (Figura 5.17), y al considerar un único pico como blanco (valor máximo de amplitud) el cálculo del SNR tiende a ser más bajo. Esto se debe fundamentalmente a que la amplitud del ecos secundario es sustancialmente mayor al ruido circundante y contribuye positivamente en el cálculo de la varianza.

El valor calculado del SNR si se utiliza un solo pulso es de 27.8 dB mientras que después de integrar 30 pulsos se alcanza un SNR de 33.7 dB.



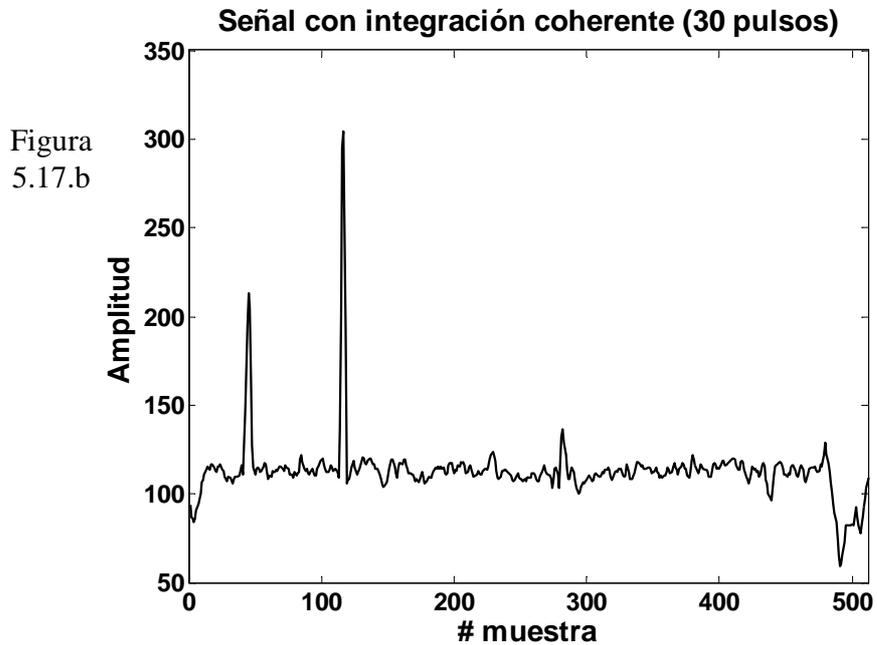


Figura 5.17. Comparación entre señales resultantes utilizando un único pulso y de realizar integración coherente (30 iteraciones)

v. Señal que contiene únicamente ruido.

Ya se estudiaron diversos casos de señales que contienen blancos inmersos en ruido u/o interferencia. Aún hay casos donde solo se recibe ruido proveniente de diferentes fuentes. Esta situación muchas veces suele ser engañosa. Si ocurre que los niveles de ruido son muy altos un eco débil puede ser enmascarado y a simple vista no ser observable (Figura 5.18 a) donde el SNR alcanza solo a 25.8dB. Si en cambio se ejecutan múltiples sumas, el eco débil prevalece (Figura 5.18.b) mejorando el SNR hasta 30 dB. Estas mediciones corresponden al día 01 de Abril de 2014 las 21:20 hs a una frecuencia fija de 8 MHz.

Figura 5.18.a

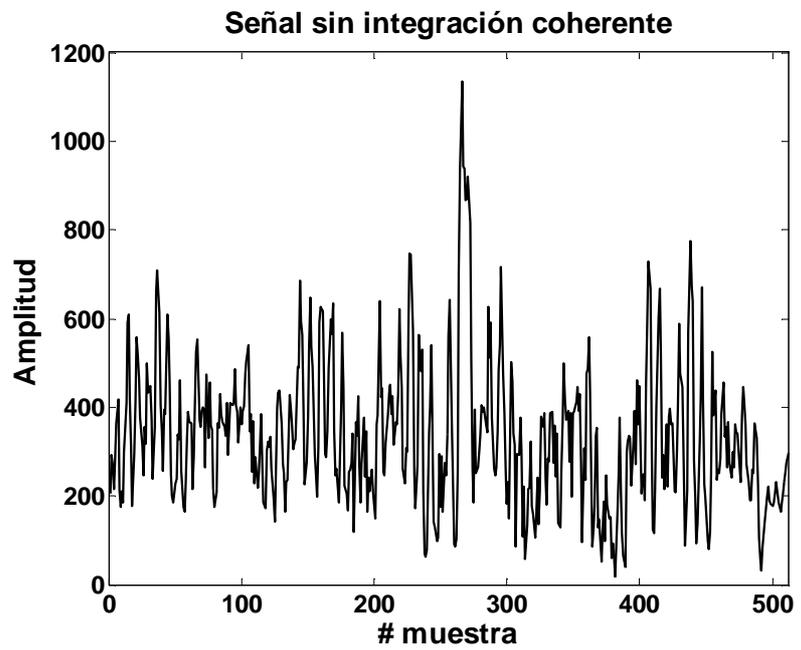


Figura 5.18.b

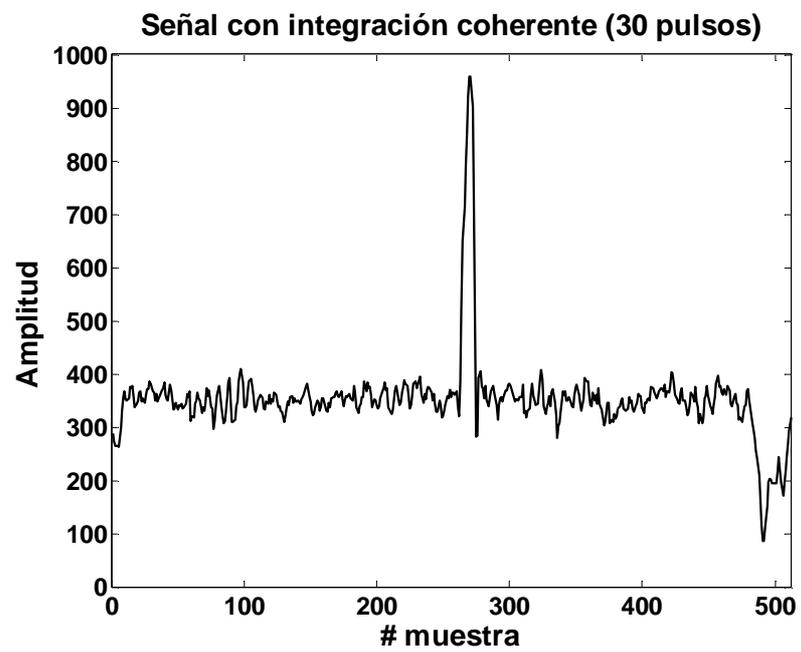


Figura 5.18. Comparación entre señales resultantes utilizando un único pulso y de realizar integración coherente (30 iteraciones)

Sin embargo en algunas ocasiones a pesar de integrar múltiples pulsos no se alcanza a apreciar la presencia de algún eco. Entonces se supone que la señal esta compuesta únicamente por ruido. Esta situación se muestra en las Figuras 5.19 a y b.

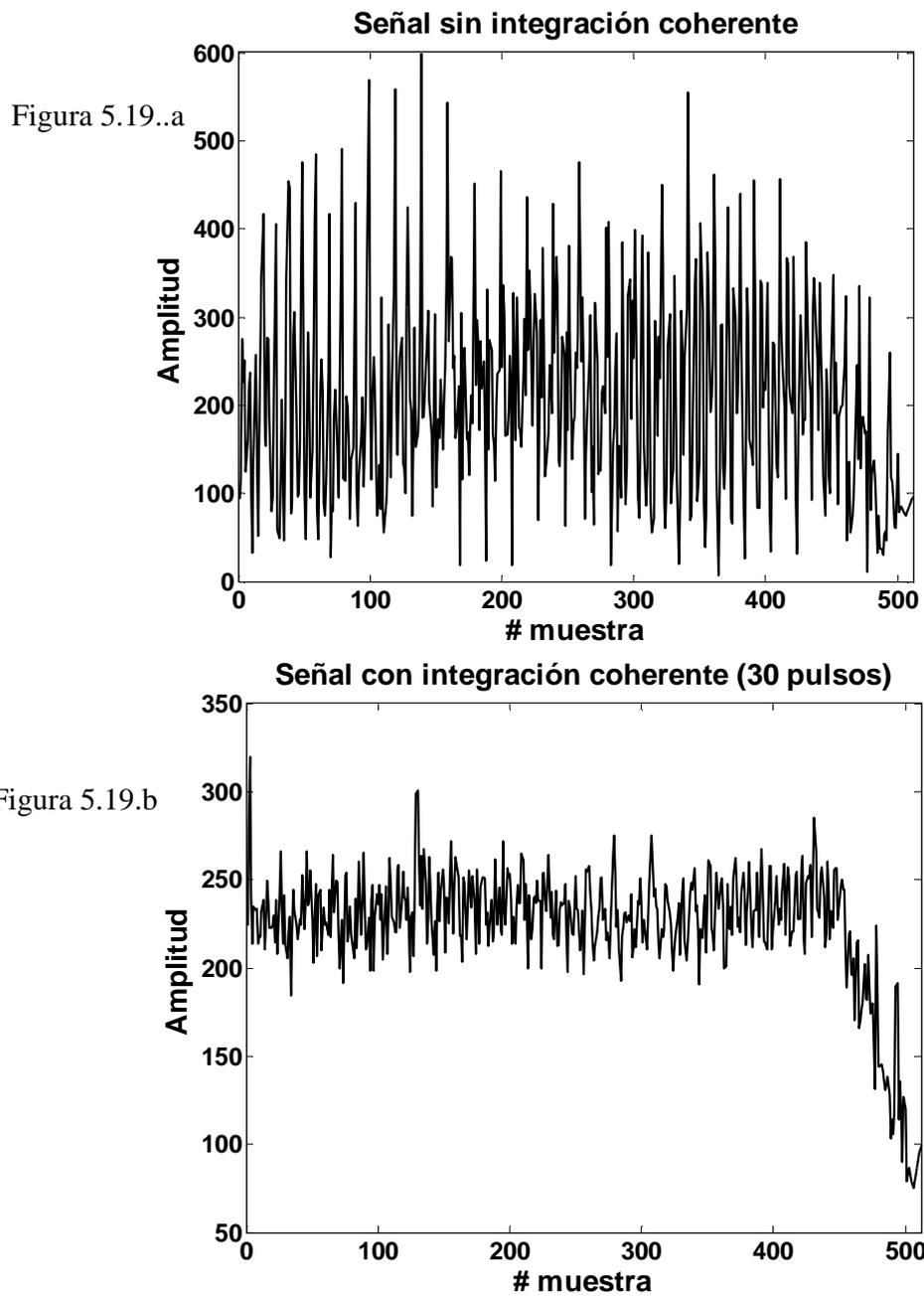


Figura 5.19. Comparación entre señales resultantes utilizando un único pulso y de realizar integración coherente (30 iteraciones)

Se puede observar que a pesar de integrar 30 pulsos no se puede determinar la existencia de algún eco. En este caso los valores de SNR son 22.8 dB y 20.4 dB para un pulso y 30 pulsos integrados respectivamente.

Cabe recordar que el cálculo del SNR se realiza comparando la máxima amplitud respecto al resto de las amplitudes que componen la señal. Esto se debe a la suposición de que la máxima amplitud corresponde a un blanco lo que no se cumple aquí.

En todos los casos estudiados se compararon señales resultantes de utilizar un único pulso con aquellas resultantes de integrar 30 pulsos. Un análisis interesante será estudiar a partir de qué cantidad de pulsos se alcanza el valor deseado de 30 dB. Para ello se calcularon los valores de SNR utilizando diferentes cantidades de pulsos (Figura 5.20).

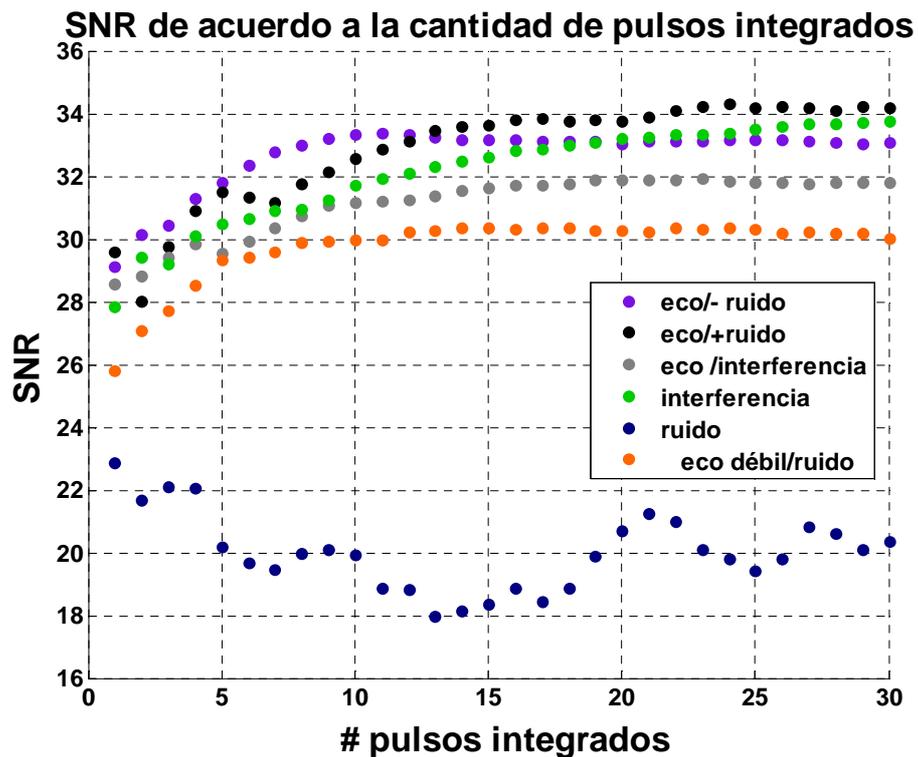


Figura 5.20. Cálculo de la relación señal/ruido como función de los pulsos integrados para diferentes los diferentes casos estudiados

Algunas conclusiones importantes que se pueden derivar de esta gráfica son:

- Como es de esperarse en el caso donde solo se tiene ruido, aún cuando se sumen muchos pulsos la SNR no mejora. De hecho, cada vez que se usa un nuevo pulso se estima un valor máximo diferente (no es coherente) para realizar el cálculo. Por esto el comportamiento de los valores de SNR no tienen una tendencia clara.
- En el resto de los casos se observa que a medida que el número de pulsos integrados aumenta, el valor del SNR tiende a estabilizarse y a partir de algún punto las mejoras no son realmente significativas. Consideremos como objetivo establecer algún valor de referencia para cantidad de pulsos integrados a partir del cual el SNR no mejora considerablemente. Además un aspecto deseable desde el punto de vista computacional es realizar la mínima cantidad de iteraciones que garanticen una mejora de al menos 30 dB (el valor pre-establecido de SNR anteriormente).
- De acuerdo a la gráfica se puede observar que en la mayoría de los casos un valor aceptable es el de 10 iteraciones. En general, y estudiando otros casos no graficados aquí, este es un buen valor para integrar pulsos. No es exigente en cuanto a computación pero garantiza en la mayor parte de los casos una buena relación señal/ruido.
- Es importante resaltar que estos cálculos están sesgados debido a que se uso como criterio que en cada señal se encuentra un único eco que corresponde al máximo valor de amplitud lo cual no siempre se cumple. A pesar de esto, esta es una buena aproximación para tener un criterio de elección de la cantidad de pulsos que se usaran y usar 10 iteraciones, en la práctica, resultó ser un número acorde en la mayor parte de los casos.
- Si bien sería necesario un estudio más exhaustivo utilizando un gran número de casos, en general una buena decisión de diseño en este equipo es utilizar 10 pulsos al momento de integrar. Adicionalmente el sistema permite modificar este número, por ejemplo incrementándolo al doble.

5.9. Última etapa del modelo

En este capítulo se propuso como modelo de detección para un sondador ionosférico que consta de cinco etapas:

1. Primera etapa: Determinación de un esquema de codificación. En esta etapa se seleccionó el esquema de codificación denominada código complementario.
2. Segunda etapa: Recepción y Muestreo. Luego del realizar un muestreo en cuadratura se procedió al análisis de la señal recepcionada.
3. Tercera etapa: Correlación. El método "Quadrature Component Correlation" fue el elegido por su robustez.
4. Cuarta etapa: Mejoramiento de la Relación Señal/Ruido. Finalmente mediante la integración coherente de múltiples se estimó como un valor aceptable el uso de 10 pulsos. La última etapa del modelo propuesto corresponde a la determinación de un algoritmo de detección Automática y que será tratado en el siguiente capítulo.

CAPITULO VI

Etapa de Detección Automática de Ecos de Radar: Algoritmos

6.1. Introducción

Como se planteó en el Capítulo IV, el modelo computacional de detección propuesto está compuesto por diferentes etapas. Cada una de estas etapas se basa en metodologías específicas para cada instante desde la generación y transmisión de las señales, hasta su recepción y procesamiento. Finalmente cada una de estas etapas son acopladas para obtener un único algoritmo general de detección.

La Figura 6.1 esquematiza las cinco etapas en un modelo de detección para un radar pulsado y coherente y sus etapas equivalentes en el diseño del modelo de detección para un sondador ionosférico que es el caso de estudio de este trabajo de tesis. Nótese que la quinta etapa correspondiente al diseño del algoritmo para la detección automática no está aún determinado (zona sobreada).

Este capítulo tratará sobre la quinta y última etapa del modelado correspondiente al diseño de un algoritmo adecuado para el sondador a partir de los algoritmos de detección automática estudiados en el capítulo precedente. Posteriormente el modelo computacional propuesto será validado utilizando señales reales adquiridas por un radar ionosférico. El experimento se diseñó en base a diferentes condiciones ionosféricas, obteniendo así diferentes casos de estudio.

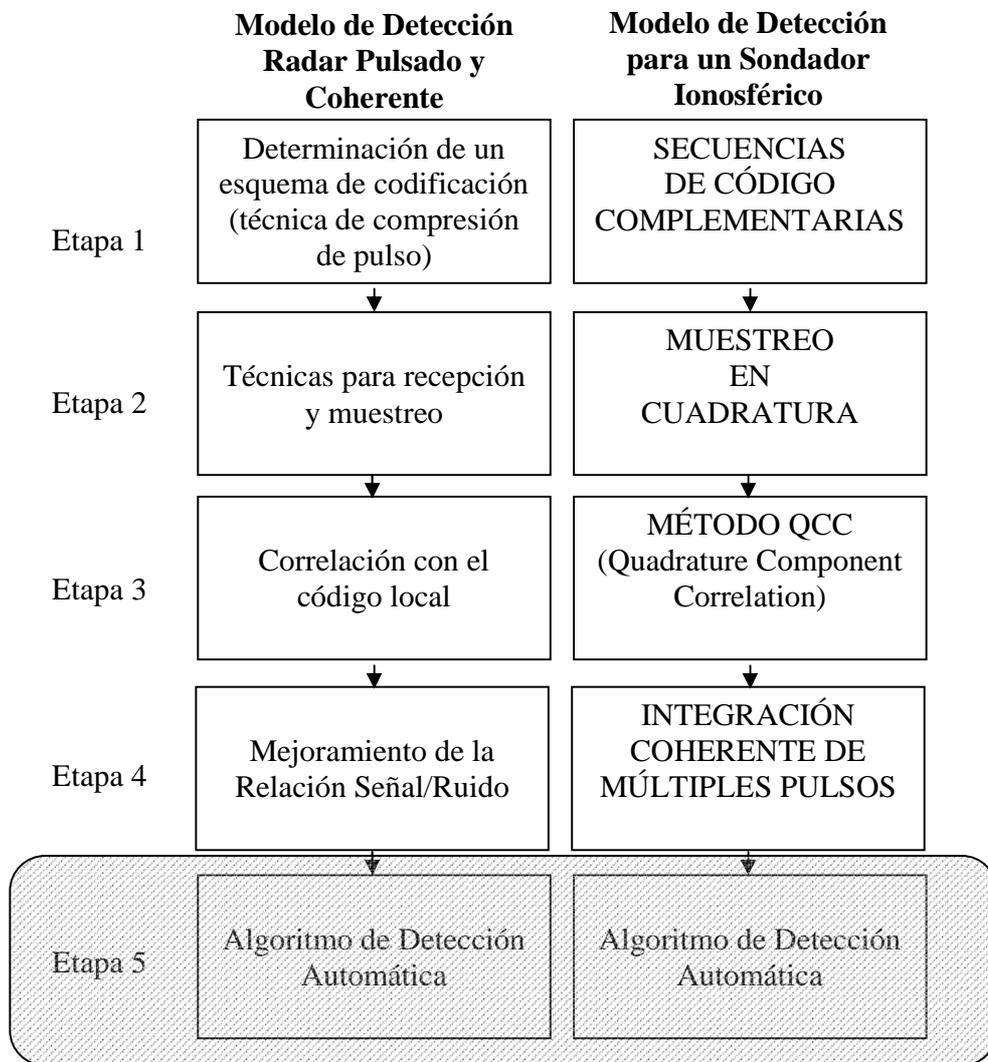


Figura 6.1. Etapas del modelado de detección de ecos para un sondador ionosférico

6.2. Implementación de la Funcionalidad Automática de Detección

Si se analiza la cadena de procesos que se ejecutan después de la recepción de la señal (muestreo, correlación e integración), se puede observar que ésta va sufriendo una serie de modificaciones. Al finalizar la etapa de integración coherente se obtiene una señal discreta sobre la que se aplicará algún algoritmo que permita decidir si alguna de las muestras presentes en la señal corresponde o no a un blanco. Esta última etapa en el modelado es la que implementa la funcionalidad de detección automática.

Como se mencionó en el capítulo III, la función del algoritmo de detección automática es reconocer aquellos ecos presentes y determinar el retardo temporal en el cual ocurren.

Desde el punto de vista geofísico esta tarea de decisión significa determinar un umbral de tal manera que cada muestra que lo supere representa un eco de la señal reflejada en la ionósfera. Por lo tanto se puede decir que dado un valor de umbral T ,

$$Deteccion = \begin{cases} \text{Capa ionosférica} & \text{si } y_i > T \\ \text{No capa ionosférica} & \text{si } y_i \leq T \end{cases}$$

Donde y_i representa la i -ésima muestra de la señal y .

Sabemos que (Capítulo IV) no existe un único algoritmo que se adapte a todos los tipos de blancos ni a todos los sistemas de radares. Por lo tanto, es necesario evaluar el desempeño de los algoritmos tradicionales y sus variantes adaptadas al caso del sondador ionosférico. Para este tipo de radar, la ionósfera es el objeto de interés o blanco.

Si bien como se mencionó anteriormente la ionósfera puede ser considerada cuasi-estacionaria por un período corto de tiempo, esta hipótesis pocas veces se cumple en un escenario real. En realidad ocurre que:

- Esta hipótesis se cumple por un periodo corto en general pero difícil de determinar dada la gran variabilidad de la ionósfera. Su comportamiento es aún más impredecible a priori en caso de existir fenómenos pasajeros.

- Cómo blanco, no es un objeto con una morfología definida, por lo tanto su sección transversal o RCS no es fácil de determinar. Es un blanco muy complejo.

En términos de detección, significa que el análisis respecto a las características del blanco y a las estrategias a seguir por el modelo, deben ser abordados de manera estadística sobre las amplitudes que presentan las señales involucradas.

Se puede pensar en cuatro tipos de estrategias para la detección automática:

1) Si sólo consideramos la hipótesis de la no fluctuación de la ionósfera durante un periodo de tiempo, una primera aproximación para modelar este blanco, sería considerar a la ionósfera como Clutter. Entonces la estrategia sería, diseñar un mapa de Clutter con el objetivo de “extraer” todos aquellos ecos no fluctuantes de la señal original bajo la suposición de que éstos corresponden a la ionosfera.

Sin embargo, existen dos supuestos importantes que no se cumplen en este caso. El Clutter se considera estacionario, sin embargo la ionosfera fluctúa a diferentes velocidades y de diferente manera (no siempre se puede predecir su comportamiento) ante diversos forzantes o eventos. Esto mismo significa que su RCS no siempre es la misma, y no posee las mismas características refractarias. Para generar un mapa de Clutter es preciso caracterizar completamente los objetos que constituyen el Clutter lo cual no se cumple en el caso de la ionósfera. Por esto, los algoritmos de mapas de Clutter no son utilizados para el caso de este blanco y por lo tanto son descartados para el modelado que se propone.

2) Descartada la posibilidad de utilizar mapas de Clutter, la siguiente estrategia está relacionada con la posibilidad de estimar el nivel de ruido presente en toda la señal que se analiza y establecer un único umbral T (Capítulo III):.

3) Otra opción es la de estimar un umbral adaptativo que para cada muestra de la señal determine un umbral particular de acuerdo a los valores estadísticos del ruido adyacente a la misma. En este caso se utilizó el denominado algoritmo CA-CFAR que para estimar

los niveles de ruido, para la construcción del umbral adaptativo, utiliza como valor estadístico el promedio de las amplitudes alrededor de una celda de interés (Capítulo III):

4) Finalmente se propone analizar el algoritmo OS-CFAR que mediante una reordenación de las amplitudes de las muestras se estima un valor de umbral para cada muestra. Este tipo de algoritmo se denomina “detector paramétrico” (Capítulo III).

Estos últimos tres tipos de estrategias llevaron a implementar dichos algoritmos para luego ser contrastados con datos reales obtenidos del sondador AIS-INGV.

6.3. Evaluación de la capacidad de detección

Antes de mostrar los resultados de los diferentes algoritmos es preciso establecer un “evaluador” apropiado para verificar y validar los requerimientos de detección del diseño del sondador ionosférico. Se necesita obtener un criterio para determinar cual es el algoritmo que tiene mejor desempeño en la detección de estratos ionosféricos. A partir del criterio de optimización para obtener la mayor cantidad de detecciones positivas y evitar la mayor cantidad posible de falsas alarmas, es necesario encontrar una herramienta de evaluación adecuada que refleje esta situación.

La manera natural sería evaluar la relación señal/ruido en cada caso. Es decir, evaluar la relación entre la amplitud de los ecos y el nivel de ruido o interferencia que posee la señal.

Con anterioridad en el Capítulo IV se explicaron las dificultades que existen para filtrar un eco del resto de la señal. El problema principal es que cada eco detectado no corresponde únicamente a una reflexión en la ionósfera, sino que es una señal compuesta tanto por el eco como por el ruido o interferencia en el canal de transmisión.

Esto lleva a un problema cíclico: para evaluar la detección con el cálculo de SNR (ecuación 5.18), es necesario poder asegurar una detección positiva (no detectar una falsa alarma), discriminar el eco, estimar su amplitud y además estimar un valor real del nivel de ruido. En la práctica, esto se convierte en un inconveniente.

Supongamos el caso extremo que se ilustra en la Figura 6.2, donde a simple vista se puede observar que: a) El valor bajo del umbral T permite realizar muchas detecciones donde posiblemente se encuentren numerosas falsas alarmas. b) Suponiendo que el nivel de ruido es aproximadamente el estimado con T , entonces todas las detecciones contribuyen a elevar el valor del SNR aún cuando se trate de falsas alarmas. c) No es fácil determinar cuales de los picos corresponden a un eco y cuales solamente a ruido, por lo tanto es difícil estimar el valor numérico del SNR.

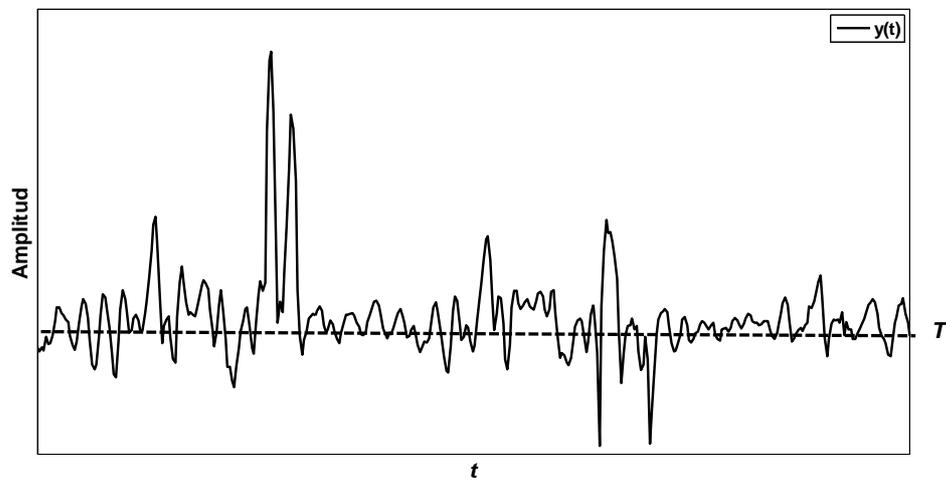


Figura 6.2. Si se cuentan todas las detecciones para el umbral T , la relación señal/ruido es alta aunque es visiblemente evidente que más detecciones no implica mejor calidad en la detección.

Por lo tanto es preciso establecer un criterio por el cual se pueda discriminar una detección positiva de un eco mientras las falsas alarmas sean desechadas. Es decir, tener mayor número de detecciones no significa necesariamente un mejor desempeño de detección. Este criterio debe realizarse con la restricción de no poder separar señal de ruido.

Teniendo en cuenta estos inconvenientes, está claro que se necesita otro criterio. Se propone utilizar un nuevo evaluador para el desempeño de los algoritmos de detección automática. La idea subyacente es poder estimar cuanto aportan en energía las detecciones a la señal. Si las detecciones corresponden a falsas alarmas lo más probable sea que contribuyan menos que un eco real.

En otras palabras, se evalúan las contribuciones en energía de las detecciones versus la energía de la señal original (compuesta por ecos y ruido). A diferencia de SNR que es una proporción entre las amplitudes del eco versus la amplitud promedio del ruido.

Para ello se utiliza la siguiente relación:

$$EC = \frac{Pe}{Psn} \quad (6.1)$$

Donde Pe corresponde a la potencia del eco, Psn es la potencia de la señal más ruido y EC es la contribución en energía de los ecos.

A su vez,

$$Pe = \frac{\sum_{i=1}^N (eco(i) - \bar{x})^2}{M} \quad (6.2)$$

Donde N es la cantidad de ecos detectado $eco(i)$ es la amplitud del i -ésimo eco detectado, \bar{x} corresponde al promedio de las amplitudes de la señal y M es el número total de muestras.

La potencia del ruido se calcula como:

$$Psn = \frac{\sum_{i=1}^N (x(i) - \bar{x})^2}{M} \quad (6.3)$$

Donde $x(i)$ es la i -ésima muestra de la señal (Molina et. al., 2016).

El hecho de que la energía proveniente de los ecos detectados sea dividida por la estimación del ruido global de la señal (utilizando todas las muestras de la señal discreta), asegura que tener mayor cantidad de detecciones no significa mejor detección.

El indicador EC sugiere una relación similar a la relación señal/ruido en cuanto a que si no hay detecciones Pe es nulo y por lo tanto no hay contribución en energía alguna. Esta es la situación donde todas las muestras se encuentran debajo del umbral.

Si en el otro caso extremo, se detectan todos los puntos $Pe=Ps_n$ y por lo tanto $EC=1$ (todas las muestras superan el umbral). La Figura 6.3 ilustra esquemáticamente estos dos

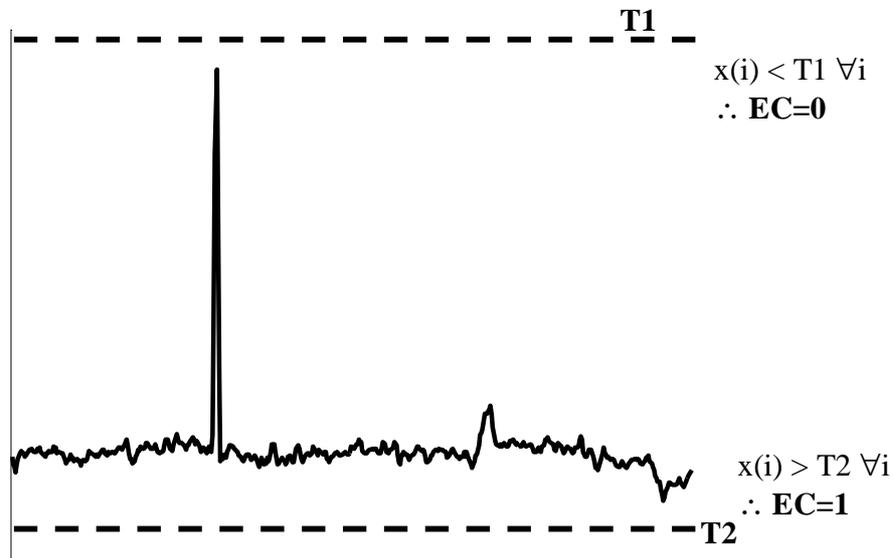


Figura 5.3. Relación de la contribución en energía (EC) en relación a la posición del umbral casos hipotéticos.

En este trabajo se implementaron tres algoritmos para detección automática donde se utilizó este nuevo criterio para evaluar sus capacidades de detección. Para cada uno de estos algoritmos se ejecutaron sobre un mismo set de datos reales adquiridos por el AIS-INGV con el objetivo de que puedan ser comparables. Si bien, durante el proceso de investigación, se utilizó un gran número de sondajes realizados durante la campaña de medición en este trabajo de tesis se mostrarán casos de estudio correspondientes a características ionosféricas típicas.

6.4. Consideraciones respecto a los parámetros usados en los algoritmos

Al momento de realizar la implementación se utilizaron diferentes parámetros en los algoritmos. En todos los casos se trabajó con una señal digital de 512 muestras resultante al final de la etapa de integración coherente.

De este modo, se cuenta con una señal a la que ya se le aplicaron todas las etapas de detección mencionadas anteriormente. Es decir, la señal es el resultado de:

- Codificar una portadora con un par de secuencias de código complementario que es transmitida hacia el medio y posteriormente recepcionada en el equipo.
- La señal es muestreada en cuadratura (en dos canales I y Q).
- Luego es correlacionada con el código local utilizando el método QCC.
- Estos pasos se repiten para múltiples pulsos que son integrados coherentemente. En particular se utilizan 10 pulsos (Capítulo V).

A continuación se implementaron los diferentes algoritmos de detección automática seleccionados a saber:

- Algoritmo de umbral fijo (TFix)
- Algoritmo CA-CAFAR
- Algoritmo OS-CFAR

Actualmente el AIS-INGV implementa el algoritmo de umbral fijo o TFix. Este algoritmo, de acuerdo a lo analizado en el Capítulo IV (ecuación 4.1), considera que el umbral puede ser calculado como un factor del promedio de amplitudes de las muestras dentro de una ventana de las muestras que componen la señal. En particular este algoritmo utiliza todas las muestras disponibles por sondaje. Por lo tanto, al momento de estimar el nivel de ruido se utilizan las amplitudes de las 512 muestras. Este algoritmo se caracteriza porque al tener fijo el umbral, para cada muestra la probabilidad de falsa alarma puede ser diferente, lo cuál se cumple en la mayoría de los casos.

En cuanto a los algoritmos de umbral adaptativo y no paramétrico, CA-CFAR y OS-CFAR, el valor de probabilidad de falsa alarma se mantiene constante (Capítulo IV). Por lo tanto, es preciso determinar un valor de probabilidad de falsa alarma que se considere aceptable de acuerdo al diseño del equipo.

Para ello debemos considerar un valor aceptable de tiempo entre ocurrencias de falsas alarmas. Es decir, cada cuanto tiempo se desea que ocurra una falsa alarma. Debido a que si la probabilidad de detección aumenta, también aumenta la probabilidad de falsa alarma y viceversa, no se puede suponer $Pfa=0$. Es necesario establecer un compromiso entre ambos valores. En particular en el caso del sondador ionosférico se puede utilizó un valor de tiempo entre falsas alarmas (Tfa) alrededor de 10 segundos (Molina et. al., 2013). Este valor es acorde a las características cuasi estacionarias de la ionósfera durante cortos períodos de tiempo. Dado que durante ese tiempo la ionosfera permanece coherente, las falsas alarmas que probablemente ocurran lo harán por los niveles de ruido o interferencia. En función a este criterio se puede estimar al Pfa como:

$$Pfa = \frac{1}{Tfa Bif} \quad (6.4)$$

Donde $Tfa=10$ segundos y $Bif= 60$ kHz es el ancho de banda en frecuencia intermedia en el caso del AIS-INGV (Zucheretti et. al., 2003). Entonces:

$$Pfa = \frac{1}{10s \ 60kHz} \cong 1.6 \cdot 10^{-6}$$

Por lo tanto, en los casos de estudio de este trabajo de investigación se usó $Pfa=10^{-6}$ (Molina et. al., 2013).

Cabe destacar que este valor de Pfa es un valor típico para muchos sistemas de radares, e incluso es utilizado normalmente en la bibliografía para evaluar diferentes variantes de algoritmos CFAR (Garvanov y. Kabakchiev, 2002; Richards et. al, 2010; Tie-Yan, 2003; entre otros).

6.5. Implementación de los tres algoritmos propuestos

Los algoritmos fueron evaluados de acuerdo a diferentes condiciones ionosféricas típicas y a parámetros de detección. Los datos fueron adquiridos en el observatorio ionosférico INGV, Roma de latitud media.

El AIS-INGV es un sondador que se encuentra activo y monitorea continuamente la ionósfera. Durante el sondaje para obtener datos sin procesar fue necesario desconectar la posibilidad de generar ionogramas. Por lo tanto, el experimento se diseñó de tal modo que por cada sondaje completo se almacenaron archivos binarios sin procesamiento y el ionograma con la configuración actual para detección (TFix). Entre 5 y 10 minutos posteriores a la obtención de los archivos binarios se procedió a la obtención del ionograma. De este modo se puede observar cada ionograma generado por el sistema actual antes y después del procesamiento de cada archivo binario.

Con el objetivo de analizar el comportamiento de los algoritmos de detección para diferentes estratos ionosférico se realizaron campañas de sondaje de aproximadamente 3 días a una frecuencia fija.

En todos los casos se procedió a ejecutar los diferentes algoritmos de detección y a calcular las contribuciones en energía de las detecciones realizadas.

1) Caso de estudio: Señal sin eco

El primer caso de estudio es el caso de tener una señal compuesta únicamente por ruido sin tener ningún eco. Se corresponde a la hipótesis nula. De tal manera que si un determinado algoritmo genera alguna detección, ésta corresponde a una falsa alarma.

Dadas las frecuencias f_s de la onda y f_p la frecuencia de oscilación del plasma, en el caso de que $f_s > f_p$ la onda atraviesa el estrato ionosférico. Dada esta condición, si $f_s > f_0F2$ no hay ecos de amplitud suficiente para detectar y solo se lee ruido (Rishbeth and Garriott, 1969). En este caso de estudio para la campaña en latitud media y en horas nocturnas ocurre que $f_0F2 < 8\text{MHz}$.

Bajo estas condiciones cualquier detección corresponde a una falsa alarma. La Figura 6.4 ilustra esta situación.

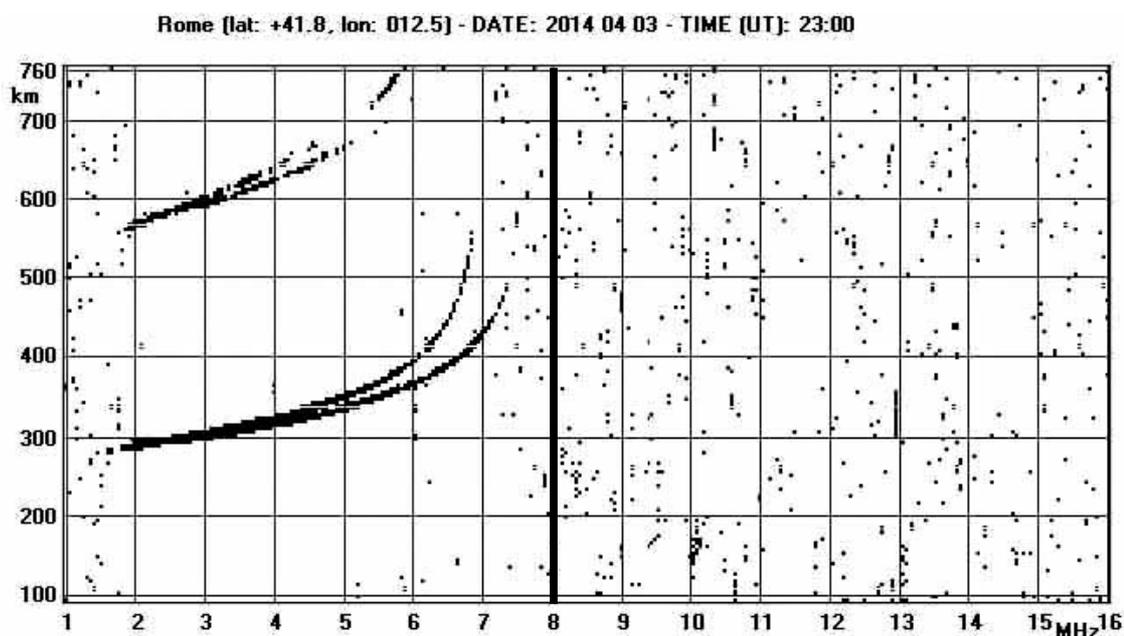


Figura 6.4. Ionograma correspondiente al día 03 de Abril de 2014, 23:00 UT que muestra que a partir de los 8 MHz (remarcado en el gráfico con una línea perpendicular en negrita) solo se receptiona ruido en el sistema.

En la Figura 6.5 se realiza una comparación en performance de los tres algoritmos propuestos. Los datos procesados en este caso corresponden al 03 Abril de 2014 a las 23:07 UT para una frecuencia fija de 8 MHz. En la grafica se puede muestra la amplitud de la señal en función de la altura en km. Se puede observar claramente que no existe un pico distintivo y que en efecto todas las muestras parecen corresponde solamente a ruido. De acuerdo a lo deducido en la Ecuación 5.1 la contribución en energía de las detecciones (EC) debería ser nula.

Se puede ver que, mientras que los algoritmos CA-CFAR y OS-CFAR no detectan ningún eco (ninguna muestra supera el umbral correspondiente), el algoritmo TFix si lo hace. Si se observa nuevamente el ionograma de la Figura 5.3, se pueden observar que para frecuencias superiores a 8 MHz aparecen graficados numerosos puntos. Los mismos

claramente son falsas alarmas. Esto se debe a que la actual configuración del equipo es usando TFix.

Si bien el algoritmo de umbral fijo genera falsas alarmas, la contribución $EC=0.28$ no es demasiado alta. Sin embargo, se puede decir que este algoritmo no tiene un comportamiento satisfactorio en detección.

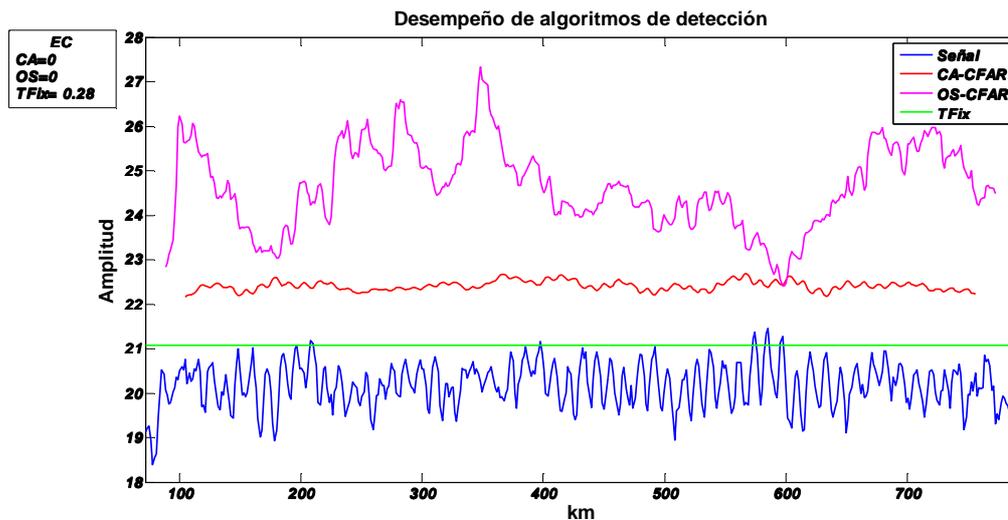


Figura 6.5. Grafica comparativa del desempeño de los diferentes algoritmos de detección automática propuestos. La señal corresponde al sondaje realizado el día 03 de Abril de 2014 a las 23:07 UT a una frecuencia fija $f=8.0$ MHz en la estación ionosférica de Roma por el AIS-INGV

En la mayoría de los sondajes estudiados se observa un comportamiento similar respecto a la performance de los algoritmos. Aunque en algunos casos CA-CFAR y OS-CFAR detecten falsas alarmas las contribuciones en energía de las mismas es siempre inferior al algoritmo de umbral fijo. La Figura 6.6 compara el desempeño de cada algoritmo para el sondaje registrado el día 03 de Abril de 2014 a las 23:50 UT para $f=8.0$ MHz.

Aún cuando OS-CFAR detecta falsas alarmas $EC=0.05$ es muy inferior al caso de TFix (0.36). Por lo tanto, OS-CFAR se comporta de manera más robusta.

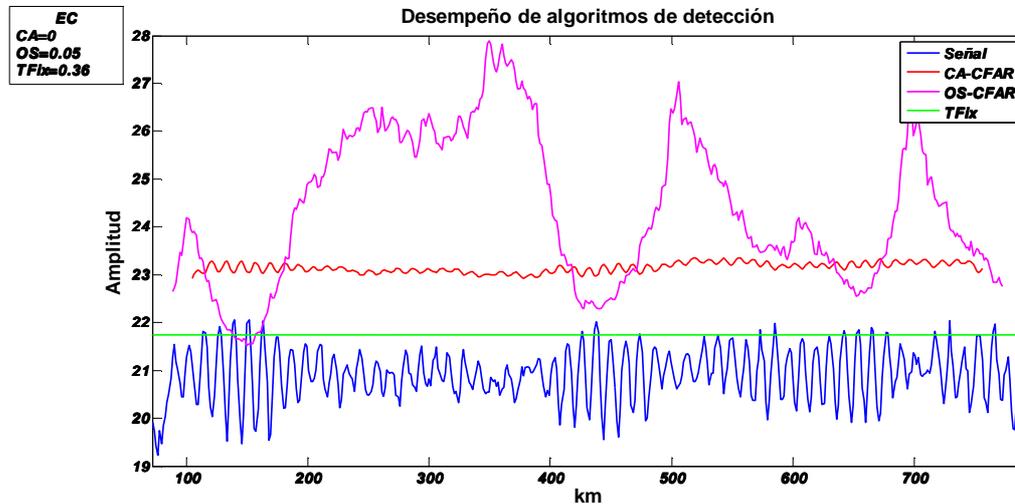


Figura 6.6. Grafica comparativa del desempeño de los diferentes algoritmos de detección automática propuestos. La señal corresponde al sondaje realizado el día 03 de Abril de 2014 a las 23:50 UT a una frecuencia fija $f=8.0$ MHz en la estación ionosférica de Roma por el AIS-INGV

2) Caso de estudio: Señal con un único eco

Si se considera la situación donde se tiene un único pico distintivo correspondiente a un eco que está inmerso en ruido, casos todos los algoritmos mostraron un comportamiento aceptable. Este puede considerarse el mejor escenario para la detección. Esto se ilustra por el caso del sondaje adquirido el 01 Abril de 2014 a las 19:35 UT para una frecuencia fija de 8 MHz (Figura 6.7). En este caso las contribuciones en energía (EC) son: CA-CFAR=0.87, OS-CFAR=0.89 y TFix=0.86. Estos valores próximos a 1, indican que la mayor contribución en energía de la señal proviene del único pico principal (eco) mientras que el nivel de ruido es considerablemente bajo.

Si bien todos los algoritmos detectan apropiadamente, es interesante resaltar que OS-CFAR no solo detecta la muestra correspondiente al eco sino también las muestras adyacentes. Envolviendo así todos los picos de mayor energía alrededor del eco y por

consiguiente describe mejor el blanco (forma y rango del eco). Esta situación se ve reflejada, además, en el leve incremento de la EC.

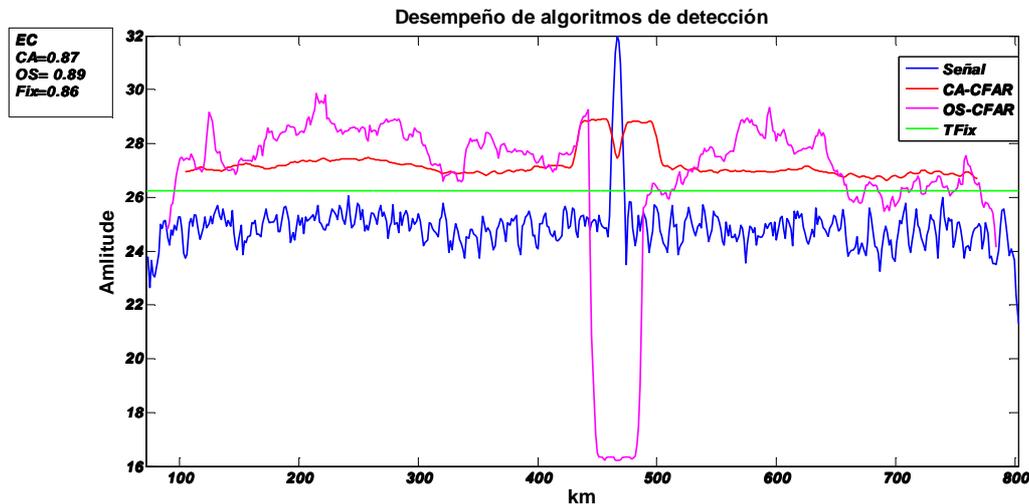


Figura 6.7. Grafica comparativa del desempeño de los diferentes algoritmos de detección automática propuestos. La señal corresponde al sondaje realizado el día 01 de Abril de 2014 a las 19:35 UT a una frecuencia fija $f=8.0$ MHz en la estación ionosférica de Roma por el AIS-INGV

3) Caso de estudio: Señal con múltiples ecos

En muchas situaciones en una misma señal están presentes múltiples ecos o blancos. Desde el punto geofísico esto ocurre cuando ocurren múltiples reflexiones de la onda electromagnética y puede tener diferentes interpretaciones.

Un ejemplo es cuando, para una determinada frecuencia, la onda de radio transmitida es reflejada en diferentes estratos ionosféricos a diferentes alturas. Esto ocurre cuando la capa esporádica Es, la capa F se encuentra presentes para una misma frecuencia de sondaje. Esto sugiere que la onda electromagnética es parcialmente reflejada, refractada y absorbida por la capa Es y finalmente reflejada por la capa F.

Otra situación común es la que se percibe cuando el estrato ionosférico está muy bien definido. De esta manera la onda es reflejada con tanta energía en la ionosfera que, a su

vez, es re-reflejada en la tierra para volver a ser enviada hacia la ionosfera. Si esta situación persiste, se pueden observar múltiples ecos separados por intervalos de tiempo aproximadamente iguales.

Desde el punto de vista del procesamiento de señales, ambos casos se corresponden con la aparición de múltiples picos dentro de la señal. Cada uno de estos picos tendrá una amplitud que se corresponde con la energía de la onda reflejada en cada capa ionosférica. Al momento de realizar la detección tener múltiples ecos puede degradar la capacidad del algoritmo. Esto se complica al tener múltiples ecos de diferentes amplitudes, llegando a tener situaciones donde alguno de los blancos se encuentre enmascarado por la presencia de otro eco próximo de mayor energía.

Los siguientes ejemplos ilustran los diferentes inconvenientes en la detección de ecos múltiples.

Para el sondaje realizado el 04 de Abril de 2014 a las 17:52 UT se seleccionó como frecuencia fija igual a 3.0 MHz (Figura 6.8). En general en horario diurno para esta frecuencia ocurre que se puede encontrar presente más de un estrato ionosférico. En esta situación es observable la presencia de múltiples ecos. En primera instancia el eco cercano a los 150 km (P1) puede deberse a una reflexión en la capa semi transparente Es que refleja parcialmente la energía de la onda mientras que el resto es absorbida. A su vez, a una altura mayor se producen nuevas reflexiones. Los picos cercanos a los 230 (P2) y 245 km (P3) pueden deberse a los rayos ordinarios y extraordinarios reflejados de la capa F. También son observables probables ecos más lejanos de mucha menor energía. El eco cercano a los 380 km (P4) puede ser producto de una reflexión oblicua. A su vez, el pico cercano a 460 km (P5) podría corresponderse con una segunda reflexión en la capa F (la distancia duplica la altura de P2). El eco P6 ocurre alrededor de los 690 km y es posible que sea una tercera reflexión en la capa F (triplica en altura a P2). Estas son posibles explicaciones de acuerdo a la hora del día, la frecuencia, latitud de la estación ionosférica y otras condiciones geofísicas.

Los valores de EC muestran que en general se tiene un buen desempeño, siendo todos superiores a 0.8. Sin embargo el algoritmo de umbral fijo falla en detectar P4 lo que se ve reflejado en un decremento de la contribución de las detecciones en energía siendo ésta

de 0.89. Mientras que en los algoritmos de umbral adaptativo se logran valores de 0.9 (CA-CFAR) y 0.93 (OS-CFAR).

La Figura 6.9 muestra dos ionogramas obtenidos a las 17:45 (Figura 6.9 a) y a las 18:00 (Figura 6.9 b) UT. En los mismos se puede observar claramente, para la frecuencia marcada (3.00 MHz), la presencia los rayos ordinarios y extraordinarios mientras que la capa E está levemente insinuada al igual que un tenue eco doble de la capa F. Cabe recordar que estos ionogramas se obtienen después de utilizar el algoritmo TFix.

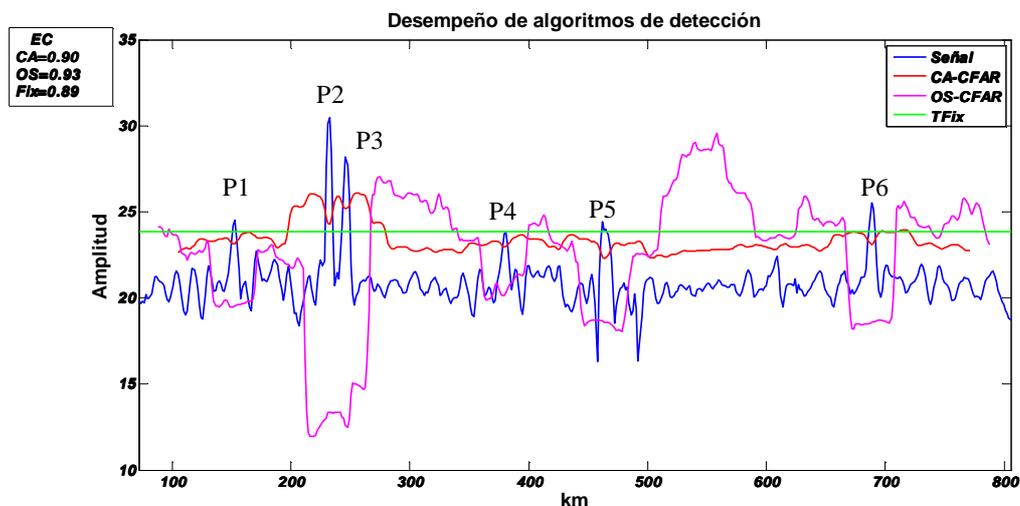


Figura 6.8. Gráfica comparativa del desempeño de los diferentes algoritmos de detección automática propuestos. La señal corresponde al sondaje realizado el día 04 de Abril de 2014 a las 17:52 UT a una frecuencia fija $f=3.0$ MHz en la estación ionosférica de Roma por el AIS-INGV

La falla principal del algoritmo TFix radica en que al utilizar el total de muestras al momento de calcular el nivel de ruido de la señal se utilizan todas las muestras disponibles, los ecos (mayor energía) contribuyen a elevar el umbral. En contraposición las características locales para el cálculo de estadísticas de los otros dos algoritmos (dentro de una ventana de muestras), hace que la energía de ecos más alejados no sea tenida en cuenta.

Figura 6.9 a. Ionograma generado 04 de Abril de 2014 a las 17:45 UT para una frecuencia de 3.0 MHz (línea punteada)

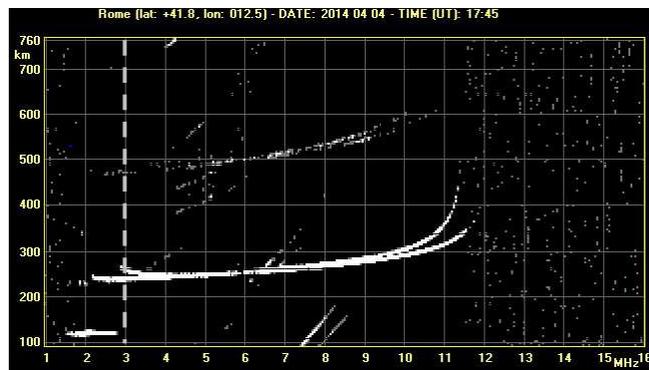


Figura 6.9 b. Ionograma generado el 04 de Abril de 2014 a las 18:00 UT para una frecuencia de 3.0 MHz (línea punteada)

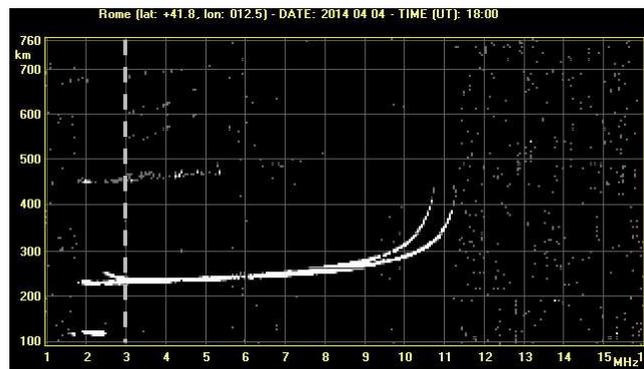


Figura 6.9. Ionograma obtenido mediante el actual sistema de detección del AIS-INGV 15 minutos antes (aproximadamente) y después de obtener las muestras pre-procesadas y que se utilizaron en el análisis de los algoritmos.

Existen circunstancias donde un eco reflejado en la ionósfera incide en la tierra y es reflejado nuevamente hacia la ionósfera, Esto puede ocurrir más de una vez obteniendo así múltiples ecos a diferentes retardos temporales y en consecuencia a diferentes alturas. El tiempo estimado entre ecos es aproximadamente el mismo aunque la energía va disminuyendo con las múltiples reflexiones. Si bien el eco que realmente puede determinar la altura de la ionósfera será la de la primera reflexión, detectar múltiples

reflexiones puede colaborar en el estudio de la morfología de la ionósfera en un determinado momento y bajo ciertas condiciones (Cabrera et. al, 2010).

La Figura 6.10 muestra múltiples ecos correspondientes a múltiples reflexiones en la capa F. Se pueden observar 5 ecos cuyas energías disminuyen con la altura. El algoritmo TFix falla nuevamente al no detectar la dos últimas reflexiones. Las contribuciones en energía para cada algoritmo son, 0.96 para CA-CFAR, 0.97 para OS-CFAR y 0.95 para TFix. Esta claro que estos últimos ecos no contribuyen demasiado pero aún así son ecos válidos.

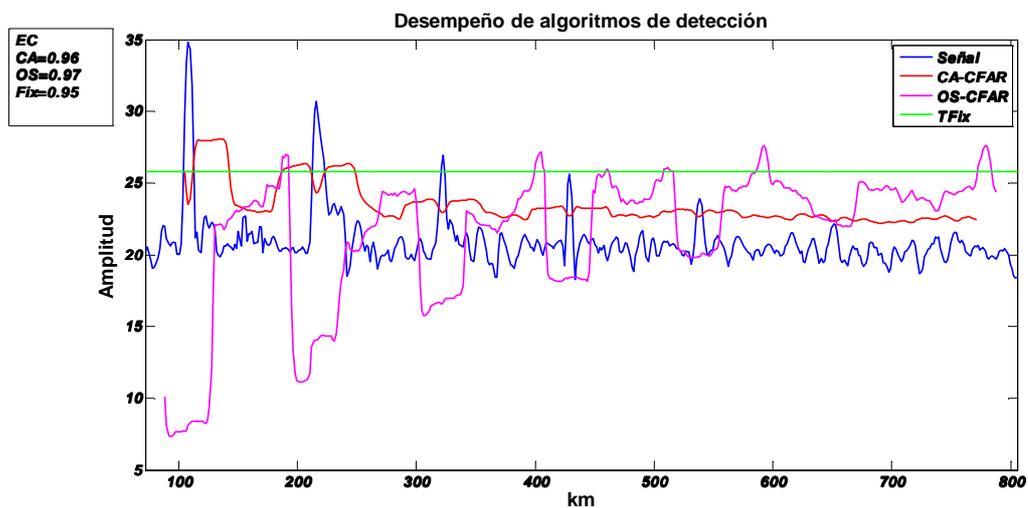


Figura 6.10. Gráfica comparativa del desempeño de los diferentes algoritmos de detección automática propuestos. La señal corresponde al sondaje realizado el día 11 de Abril de 2014 a las 18:22 UT a una frecuencia fija $f=3.0$ MHz en la estación ionosférica de Roma por el

4) Caso de estudio: Señal con interferencia

Un caso extremo para las capacidades de detección de un radar es ante la presencia de señales de interferencia. En la mayoría de los casos la detección resulta ser compleja. A veces es tan severa la interferencia que puede llegar a corromper completamente la codificación haciendo imposible cualquier tipo de detección.

La Figura 6.11 muestra los resultados obtenidos al procesar una señal con interferencia. Los datos usados corresponden al día 03 de Abril de 2014 a las 19:35 UT para una frecuencia de sondaje de 8.0 KHz. A simple vista se puede observar la modulación de la señal de interferencia y además se pueden distinguir dos ecos. Estos ecos corresponden en este caso a los rayos ordinarios y extraordinarios de la ionósfera que se encuentran a un altura aproximada de 325 km y 485 km respectivamente.

Se observa además que todos los algoritmos detectan ambos ecos. Sin embargo OS-CFAR consigue tener una leve mejora respecto a los otros dos. Las contribuciones en energía de las detecciones son 0.49 para CA-CFAR, 0.55 para OS-CFAR y 0.49 para TFix. La mejoría en detección de OS-CFAR se puede explicar porque no solo se detectan los valores máximos del eco sino que el algoritmo detecta la morfología del mismo. En estudios posteriores podría ser de interés tener esta información. Intuitivamente, cuantas más muestras adyacentes al máximo contengan energía del blanco, significa que ese estrato ionosférico se encuentra de algún modo esparcido.

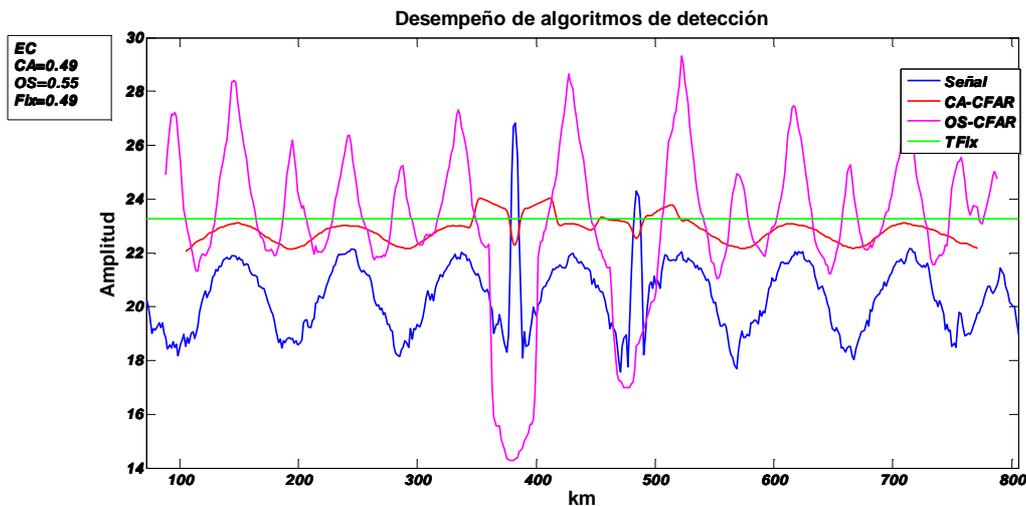


Figura 6.11. Gráfica comparativa del desempeño de los diferentes algoritmos de detección automática propuestos. La señal corresponde al sondaje realizado el día 03 de Abril de 2014 a las 19:35 UT a una frecuencia fija $f=8.0$ MHz en la estación ionosférica de Roma por el AIS-INGV

El principal inconveniente radica en que el nivel de la interferencia es suficientemente alto como para elevar los valores de umbrales y con ello esconder o enmascarar parcialmente al blanco. En particular el segundo eco se encuentra ubicado en la parte ascendente de la pendiente de la onda de interferencia y por lo tanto su energía compite con la misma. El umbral constante resulta tal que detecta apenas al segundo eco. La principal razón es que al calcular el promedio usando todas las muestras de la señal, el alto nivel de interferencia sesga este cálculo.

Esta situación persiste en el caso de usar un valor medio local (dentro de una ventana) ya que las muestras inmediatamente adyacentes a la muestra que contiene el eco, están también afectadas por los altos valores de interferencia.

En cambio OS-CFAR al re-ordenar las muestras dentro de la ventana de manera ascendente asegura que las mayores amplitudes estén alejadas del valor medio o k -ésimo valor usado en el cálculo. Por lo tanto la performance de OS-CFAR es significativamente mayor a los otros dos algoritmos.

No siempre es posible la detección ante la ocurrencia de interferencia. La Figura 6.12 muestra un caso donde la interferencia es tal que supera incluso a la codificación. Este es

el caso del sondaje correspondiente al 09 de Abril de 2014 a las 17:37 UT para una frecuencia de 6.0 MHz.

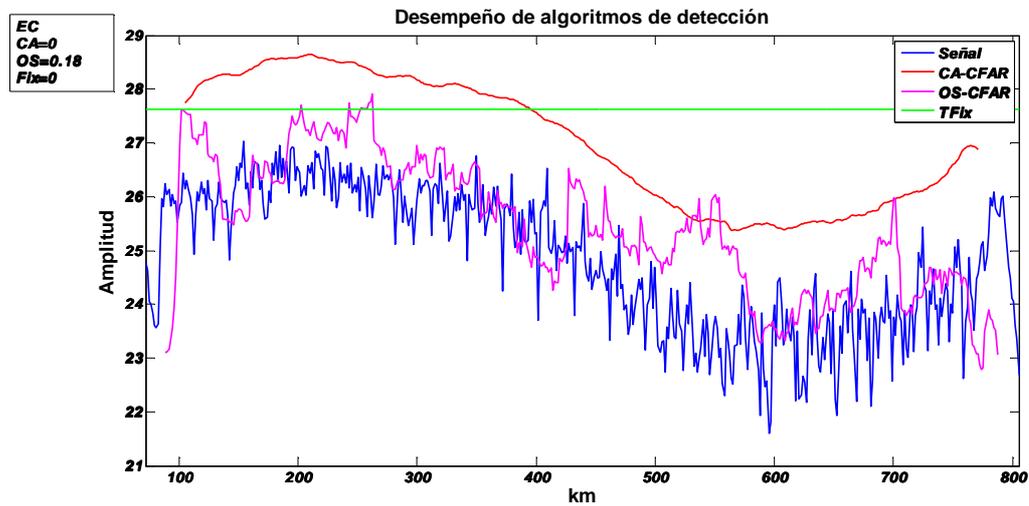


Figura 6.12. Gráfica comparativa del desempeño de los diferentes algoritmos de detección automática propuestos. La señal corresponde al sondaje realizado el día 09 de Abril de 2014 a las 19:37 UT a una frecuencia fija $f=6.0$ MHz en la estación ionosférica de Roma por el AIS-INGV

Los ionogramas obtenidos el 09 de Abril de 2014 a las 17:30 (Figura 6.13 a) y a las 17:45 (Figura 6.13 b) UT muestran la presencia de la interferencia para la frecuencia de 6 MHz.

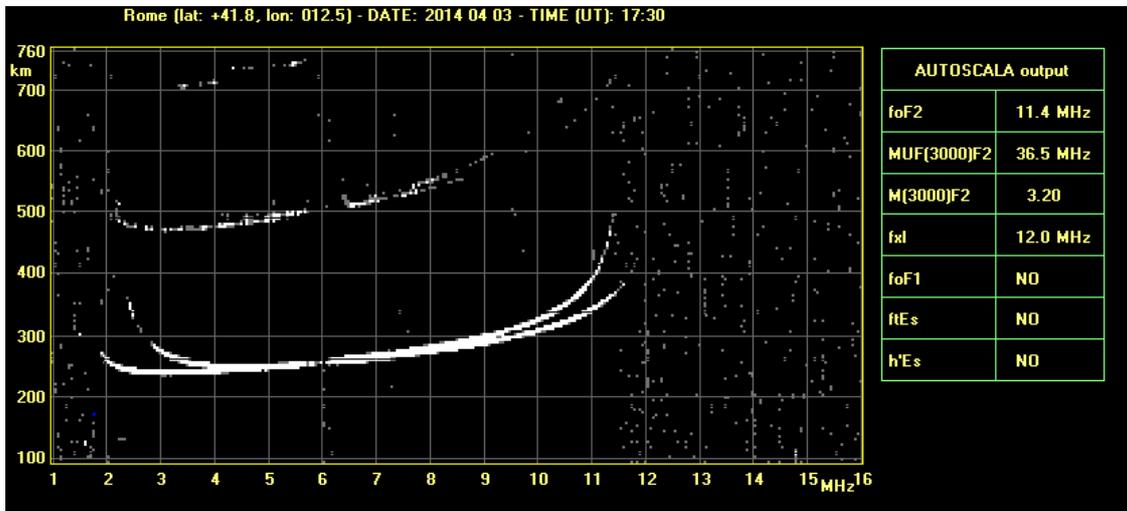


Figura 6.13 a. Ionograma correspondiente al día 09 de Abril de 2014 a las 17:30 UT

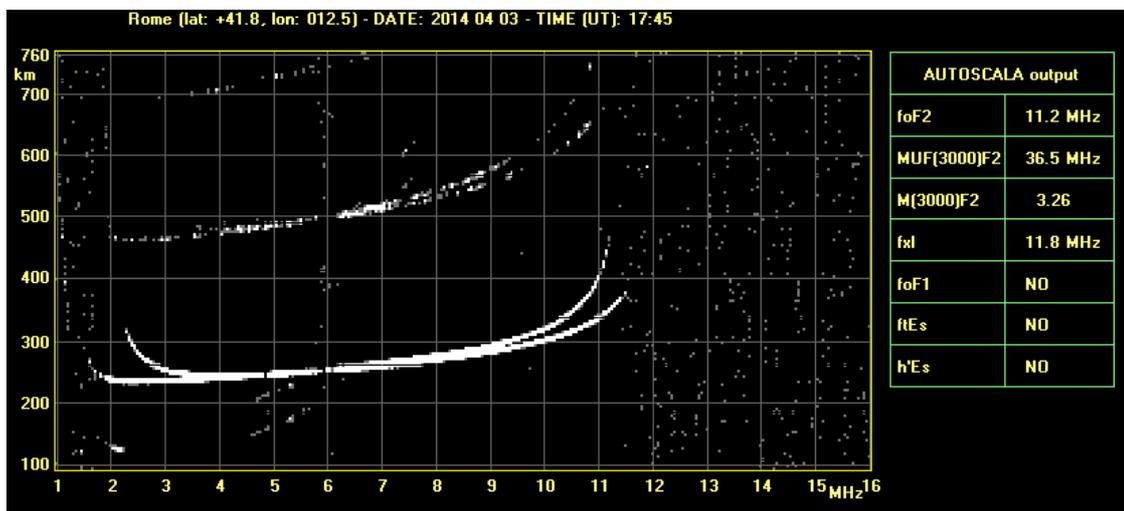


Figura 6.13 a. Ionograma correspondiente al día 09 de Abril de 2014 a las 17:30 UT

Esta claro que no hay ningún eco para ser detectado, sin embargo el algoritmo OS-CFAR detecta numerosas falsas alarmas ($EC=0.18$) pero que en suma no contribuyen demasiado a la energía total de la señal.

6.6. Análisis del desempeño de los algoritmos propuestos

Este análisis se puede realizar mediante dos enfoques:

1) Respecto a la validación del modelo

En general se puede observar que el desempeño de los algoritmos CA-CFAR y OS-CFAR superan ampliamente al desempeño que puede alcanzar el algoritmo TFix y que es el que actualmente se encuentra implementado en el AIS-INGV.

Se puede observar que para el caso de tener una señal sin ecos, que el algoritmo que más falsas alarmas produce es TFix. Los algoritmos basados en CFAR, al mantener acotada la Pfa , logran que el umbral se adapte a las condiciones adyacentes al eco. Esto significa que valores alejados (posiblemente otros ecos de mayor energía) no interfieren y no aportan información para el cálculo del umbral.

Incluso en casos complejos, como ocurre ante la presencia de múltiples ecos, la performance de los algoritmos basados en CFAR es superior a la de TFix.

Se puede decir que los algoritmos basados en CFAR son más robustos. Esto se explica además dado que el algoritmo TFix se basa completamente en valores empíricos mientras que los algoritmos CFAR se basan en teoría de detección.

También se pudo observar que en los casos donde todos los algoritmos tienen un rendimiento similar, los algoritmos CA-CFAR y OS-CFAR pueden agregar mayor información. Por ejemplo, dado que el umbral que calcula OS-CFAR permite envolver todo el eco, es posible reconocer por ejemplo la morfología del mismo.

2) Respecto al desempeño al desempeño computacional

Escribir un programa involucra una serie de pasos para resolver un determinado problema, desde la formulación del problema y su especificación, el diseño de una solución computacional hasta su implementación en un lenguaje determinado y su posterior ejecución en un computador de características particulares. También se debe contar entre estos pasos las pruebas y documentación para finalmente evaluar la solución. Una vez que se puede encontrar un modelo matemático que se ajusta a las necesidades del problema, se procede a encontrar un modelo computacional para obtener una solución. Por lo tanto, el primer paso es el de determinar un modelo computacional en términos de un algoritmo. Podemos definir algoritmo como una secuencia de instrucciones finitas, cada una de cuales no deben poseer ambigüedades y deben poder ejecutarse una cantidad finita de tiempo (Aho, 2001).

La evaluación del modelo computacional no es fácil de realizar. Para un determinado problema podemos encontrar numerosos algoritmos que lo resuelven y lo hacen de manera correcta (los resultados que arroja son válidos). No es fácil determinar la bondad de un algoritmo, y más difícil si este es óptimo. Existen numerosos criterios para determinar la bondad de un algoritmo. En muchos casos estos criterios son subjetivos e involucran características como la simplicidad, claridad, o considerar cuanto se ajustan las soluciones a los datos esperados.

Una manera más objetiva de realizar esta evaluación es evaluar el tiempo de ejecución de programa que implementa el algoritmo y la eficiencia del uso de los diferentes recursos de computación.

En el caso de implementar código científico se valora la robustez del programa pero sobre todo los tiempos de ejecución de mismo.

En un primer análisis se puede contabilizar, mediante código fuente incrustado en el mismo programa, cuanto tiempo demora la ejecución. Como una primera aproximación para evaluar eficiencia, ésta es una buena medida. Sin embargo, si el problema crece en tamaño de su entrada (por ejemplo crece el número de muestras que se deben procesar) es indispensable entender en que proporción aumentará el tiempo de ejecución. Una observación importante es que esta medición del tiempo que demoró en correr el

programa responde a una computadora específica (donde se realizó la medición). Lo deseable es poder establecer la eficiencia de un algoritmo de manera independiente al hardware que se tenga disponible.

Con el objetivo de poder sacar provecho a la tecnología de procesamiento disponible es importante escribir código eficiente que aproveche el hardware subyacente.

Escribir código computacional eficiente requiere del conocimiento de diferentes técnicas de programación científica relacionadas tanto a la optimización del software como al hardware. Esto significa, escribir código en un lenguaje de programación que permita el uso eficiente de recursos aplicados a la ciencia y al cálculo numérico. Además de diseñar algoritmos eficientes que permitan estipular de ante mano el comportamiento del software de acuerdo a cómo crecen las entradas del sistema. Estos algoritmos deben estar adaptados a los tipos de datos requeridos en el cálculo así como también a las estructuras de datos más acordes al problema.

Finalmente, al momento de escribir el programa se debe tener en cuenta la arquitectura subyacente para hacer uso eficiente por ejemplo de todas las unidades de procesamiento disponibles y de utilizar eficientemente la capacidad de almacenamiento del equipo.

En el caso del modelo propuesto, el procesamiento de las señales involucradas se realizó en una computadora de propósitos generales sin la necesidad de recurrir a técnicas de programación en paralelo. Recordemos que el tiempo entre sondajes típicamente es de aproximadamente 15 minutos. Los algoritmos implementados en este trabajo no superan los pocos segundos de tiempo de procesamiento. Esto quiere decir que el tiempo de ejecución es suficientemente bajo para permitir realizar el procesamiento entre sondajes sin mayores dificultades.

Sin embargo cuando los requisitos del sistema son más exigentes, por ejemplo para un determinado estudio es preciso realizar sondajes cada un minuto, es necesario realizar un estudio detallado de la eficiencia del tiempo de ejecución de los algoritmos propuestos.

Estudiar la complejidad algorítmica significa estimar en que medida el tiempo de ejecución aumentará a medida que el tamaño de la entrada crezca.

6.6.1. Análisis de Complejidad Algorítmica

A continuación se realiza el análisis de complejidad algorítmica para los algoritmos propuestos en este trabajo de tesis:

Algoritmo TFix

POCEDIMIENTO TFIX (y:arreglo; var M: numero total de muestras)

- (1) **begin**
- (2) Calcular Promedio de muestras y obtener T
- (3) **PARA** cada muestra y_i de y **HACER** (M veces)
- (4) **SI** $y(i) > T$ **ENTONCES**
- (5) **ESCRIBIR**(Deteccion positiva, i)
- (6) **FIN**
- (7) **FIN**
- (8) **FIN.**

El paso (2) que calcula T utilizando el promedio de todas las muestras requiere de recorrer el arreglo y completamente. Esto significa que cada vez que varíe el valor total de muestras, este paso consume una cantidad de operaciones igual. Es decir, este paso tiene una complejidad de $O(M)$, donde M es la cantidad de muestras torales que posee la señal.

A su vez, de manera independiente, es necesario recorrer nuevamente todo el arreglo para hacer M comparaciones de cada valor de la muestra contra el valor del umbral T. Esto también tiene un orden lineal en complejidad algorítmicamente.

El resto de las operaciones no depende del tamaño del arreglo y por consiguiente se puede considerar como de orden constante.

En conclusión, el algoritmo TFIX tiene un orden de complejidad lineal.

Algoritmos CFAR

Tanto el algoritmo CA-CFAR como OS-CFAR tienen una misma formulación general. Sólo varían en el modo de obtener el valor del umbral variable para cada muestra. En consecuencia se analizará la complejidad algorítmica del algoritmo CFAR.

Para realizar el análisis se utilizará la hipótesis que el tamaño de la ventana siempre es inferior a la cantidad de muestras total de la señal. Esta hipótesis siempre se cumple en el caso CFAR (ver más detalles en Capítulo IV).

El algoritmo presentado en el Capítulo IV es el siguiente:

POCEDIMIENTO CFAR (y:arreglo; var M: numero total de muestras; N: numero de muestras en la ventana)

```
(9) begin
(10)   PARA cada muestra yi de y HACER (M veces)
(11)       Obtener Ti usando N celdas
(12)       SI y(i) >Ti ENTONCES
(13)           ESCRIBIR(Deteccion positiva, i)
(14)       FIN
(15)   FIN
(16)   FIN.
```

Una vez prefijado el tamaño de la ventana, sin importar como crezca el tamaño del arreglo de muestras, el proceso de obtener el umbral particular de cada muestra (T_i) dado en el paso (3), siempre tendrá un costo constante. En este caso, el mayor costo computacional está dado por el lazo de comparación de cada muestra con su umbral. La cantidad de comparaciones realizadas si depende del número total de muestras (M).

En el peor caso, que no sabemos de antemano la cantidad de muestras (N) usadas en la ventana, el algoritmo tendrá un costo computacional total de $O(N*M)$.

Cualquiera de los tres algoritmos analizados la complejidad computacional tiene un costo polinómico, por lo tanto se puede decir que los algoritmos utilizados son eficientes.

CAPITULO VII

Conclusiones Generales y Perspectivas

El objetivo principal de este trabajo de tesis se ha propuesto un modelo de detección automática para de ecos de señales de radar acorde a características específicas del blanco. En particular diseñar un modelo computacional para la detección automática de ecos de un sondador ionosférico. Luego del estudio, el análisis y el diseño de diferentes técnicas para el procesamiento de señales se arribó a un modelo computacional adecuado para un sondador ionosférico. Los objetivos particulares relacionados a las diferentes etapas en el modelo también fueron alcanzados.

Los principales aportes logrados durante el desarrollo de esta tesis se enuncian en los párrafos subsiguientes.

Se logró caracterizar a los estratos ionosféricos como un blanco complejo, su sección eficaz de radar, los tiempos típicos de coherencia para una ionósfera de latitud media y finalmente determinar un criterio aceptable para la ocurrencia de falsas alarmas. Se realizó un estudio sobre la probabilidad de falsa alarma deseable para este sistema y se obtuvo como valor estimado de $P_{fa}=10^{-6}$.

En base a los modernos criterios de diseño de radares, esto es emitir bajos niveles de potencia manteniendo una alta resolución, se utilizó la técnica de compresión de pulso. Para esto se realizó el estudio de esquemas de codificación de la señal portadora y en base a los requisitos planteados se seleccionó la codificación de secuencias de código complementario. En el diseño global del sistema de detección automática, esta fue la primera etapa.

Dado que el procesamiento de las señales recepcionadas y muestreadas en cuadratura necesitan, cada una de las secuencias, ser correlacionadas con el código base correspondiente; se diseñó un método apropiado. Después de evaluar dos métodos propuestos, el más adecuado resultó ser el método QCC. Siendo este un aporte novedoso y cuyo resultado dio lugar a una publicación en una revista científica de nivel internacional.

Las señales recepcionadas en principio son señales donde se encuentra algún eco producido por alguno de los estratos inosféricos pero además es una señal que se encuentra atenuada, retardada en el tiempo e inmersa en ruido e interferencia. Esto impacta en el sistema deteriorando la capacidad de detección al correlacionar un único pulso. A partir de las características del blanco se realizó un estudio para estimar la cantidad de pulsos que podían integrarse coherentemente para obtener un mejoramiento en la relación señal/ruido. Se determinó que durante una cantidad de 10 sondajes la ionósfera se mantiene coherente.

Se estudiaron tres tipos de algoritmos numéricos para cálculo de umbral de detección, se analizaron y fueron comparados en su desempeño mediante simulaciones numéricas utilizando datos reales. fueron publicados en revistas científicas internacionales.

Se realizó la comparación entre el algoritmo actualmente vigente en el sondador AIS-INGV (algoritmo TFix) y los algoritmos basados en CFAR. Se pudo mostrar que debido a las características empíricas del algoritmo TFix, su desempeño es mucho menor respecto a los algoritmos CA-CFAR y OS-CFAR. Estos últimos se basan son mucho más robustos en casos reales, lo que se debe principalmente a la capacidad de mantener constante la P_{fa} . El resultado son algoritmos con umbral variable que se adaptan a las condiciones locales del eco mejorando la capacidad de detección automática del detector. Debido a la imposibilidad de discriminar eficazmente la amplitud de los ecos así como también el nivel de ruido presente en las señales, no fue posible utilizar el cálculo de la SNR. Es por esto que se diseñó un evaluador del rendimiento de la capacidad de detección de los algoritmos para un escenario real. Mediante el cálculo de la contribución de energía de cada detección se pudo evaluar cada algoritmo utilizando diferentes casos de estudio. Los casos de estudio presentes en este trabajo de tesis fueron seleccionados de una campaña de medición utilizando el AIS-INGV. Para ello se diseñó un experimento de acuerdo a la caracterización del blanco, en este caso la ionósfera de latitud media.

El desempeño computacional de cada uno de los algoritmos y sus implementaciones también fueron analizados. Se demostró que los algoritmos son computacionalmente eficientes y que permiten el procesamiento de las señales necesario durante el tiempo entre sondajes.

Se pudo probar que el modelo computacional propuesto (compuesto por diversas etapas) para la detección automática de ecos de un sondador ionosférico es el más adecuado y que permite mejorar sustancialmente la calidad de detección del modelo vigente en la actualidad.

A partir de este trabajo de tesis se pueden sugerir nuevas líneas de investigación y de desarrollo sobre la tecnología de radar. Una de las perspectivas en el estudio de detección es la de poder diseñar un detectores inteligentes que puedan ser capaces de seleccionar automáticamente el tipo de algoritmo de detección de acuerdo al estado del medio.

Referencias bibliográficas

- Aho, Alfred V; Hill, Murray; Hopcroft, John E.; Ullman Jeffrey D. (2001). Data Structures and Algorithms
- Ariokiasamy, James B., C. Bianchi, U. Sciacca, G. Tutone and E. Zuccheretti (2003); “The new AIS-INGV Digital Ionosonde”; Desing Report, Sezione di Geomagnetismo, Aeronomia e Geofísica Ambientale, Istituto Nazionale di Geofísica e Vulcanologia (INGV), Roma, Italia.
- Barry, G. H. (1971), A low vertical incidence ionosonda, IEEE Trans. Geosci. Electron. GE-9(2), pp. 86-89.
- Barton, D.K., Leonov, S.A. (Eds.). Radar Technology Encycloped ia. Artech House, USA, 1998
- Baskaradas Arokiasamy, J., C. Bianchi, U. Sciacca, G. Tutone y E. Zuccheretti (2002). The new INGV digital ionosonde design report, Rapporti Tecnici INGV, 12, 59 .
- Baskaradas, J.A., Bianchi, S., Pietrella, M., Pezzopane, M. (2014). Description of ionospheric disturbances observed by vertical ionospheric sounding at 3MHz. Ann. Geophys. 57 (1), A0187.
- Bianchi Cesidio, Sciacca Umberto, Zirizziotti Achille, Zuccheretti Enrico, Arokiasamy Baskaradas James (2003), Signal Processing Techniques for Phase – Coded HF-VHF Radars, Annals of Geophysics, V46, N°4,pp 697-705.
- Bianchi Cesidio and David Altadill (2005), Ionospheric Doppler measurements by means of HF-radar techniques, Annals Geophysics, V48, N°6, pp. 989-993.
- Bianchi, Cesidio (2006), *comunicación personal*, Sezione di Geomagnetismo, Aeronomia e Geofísica Ambientale, INGV, Roma, Italia.
- Cabrera M. A (2003). Modelado de la distribución de la densidad electrónica en la ionósfera. “Modelado del contenido electrónico total oblicuo”. FACET – UNT, Tucumán, Argentina.
- Cabrera Miguel A., R. G. Ezquer, J. Lopez, C. Bianchi, S. M. Radicella, G. Molina, E. Garay, C. Ivan, M. Valdez, M. Ferreyra, L. A. Zalduendo (2007); Proyecto de desarrollo de un sondador digital para investigación Ionosférica, Revista Ciencia Tecnología y Medio Ambiente, FRT, UTN, N°6, pp 30-33, Argentina.
- Cabrera , M., Zuccheretti, E., Ezquer, R., Sciacca, U., Lopez, J., Molina, M., Baskaradas, J. Some considerations for different time-domain signal processing of pulse compression radar. Ann. Geophys. 53 (5–6), 1–11, <http://dx.doi.org/10.4401/ag-4758>, 2010
- Curry, G. Richard (2005), Radar System Performance Modeling, 2nd. Edit., Ed. Artech House, Norwood, MA.
- Davies, K. (1969). Ionospheric Radio Waves, Blaisdell Publishing Company, London,.
- Dominici P (1998). My first fifty years in ionospheric research. Ann. Geofis 41:857-883.
- Garvanov y Chr. Kabakchiev (2002). Average decision threshold of CA CFAR and excision CFAR detectors in the presence of strong pulse jamming, Proc. of the German Radar Symposium GRS 2002, Bonn,Germany.
- Ghebrebrhan, Ogubazghi; Luce Hubert; M. Yamamoto and S. Fukao (2004); Interference suppression factor characteristics of complementary codes for ST/SMT radar applications, Radio Science, V39, RS3013, doi:10.1029/2003RS002910.

- Golay, M.J.E.(1961), Complementary series, *IRE Trans. Inf.Theory*, 7, 82-87.
- Hines C.O. (1960).Internal gravity waves at ionospheric heights. *Cn J Phys* 38: 1441-1481
- Jursa, A. S. (1985), Handbook of Geophysics and the Space Environment, AFGL, Air Force, USA.
- Kay Steven M (2013).Fundamentals of Statistical Signal Processing,Volume III Practical Algorithm Development. ISBN-13: 978-0-13-280803-3.
- Langtangen, H. P. (2016). A Primer on Scientific Programming with Python. 3ra edicion. 2016.Springer.
- Melvin W. y Scheer J. (2013), Principles of Modern Radar: Radar Applications. ISBN: 9781891121548
- Mahafza, B.R. (2008), Radar signal analysis and signal processing using MATLAB. Chapman and Hall/CRC.
- Mohamed, Nasser J. (1991); Resolution Function of Nonsinusoidal Radar Signals: II – Range – Velocity Resolution with Pulse Compression Techniques, *IEEE Transactions Electromagnetic Compatibility*, V33, N°1.
- Nathanson F.E, Reilly J.P. (1991), “Design Principle. Signal Processing and the Environment”, 2° Edic., Mc Graw Hill.
- Nozaki, Kenro (1993), Application of FM/CW techniques to ionosondes. World Data Center, Report UAG-104, Proceedings of Session G6 at the XXIVth general Assembly of the URSI, Kyoto, Japan.
- Oberkampf, [William L.](#) (2010). Verification and Validation in Scientific Computing. Polytechnic Institute and State University. ISBN: 9780521113601
- Oppenheim, Alan (2005), Tratamiento de señales en tiempo discreto. Pearson Prentice Hall, Madrid.
- Poole, A. W. V. (1985), Advanced Sounding, The FMCW alternative, *Radio Science*, V.20, pp. 1609-1616.
- Pacheco, P. (2011). An Introduction to parallel computing. ISBN-10, 123742609
- Pratt T.W. y Zelkowitz (1998), Lenguajes de Programación, Prentice Hall.
- Pressman, Roger S. (2010). Ingeniería del software: Un enfoque práctico. 7ma Ed. MacGraw-Hill
- Rastogi, Prabhat K. (1990); Signal processing and data analysis in middle Atmosphere radar; *Radio Science*, V25, N°5, pp 1071-1086.
- Rastogi, P. K., and G. Sobolewski(1990); New quasi-complementary code sets for atmospheric radar applications, *Radio Science.*, V25, N°5, pp 1087-1094.
- Reinisch, B. W. (2000), Radio Sounding of Geospace Plasmas, *Física de la Tierra*, pp. 105-126, ISSN: 0214-4557.
- Richards, M.A. (2005). Fundamentals of Radar Signal Processing. McGraw-Hill, New York.
- Richards M. A., Scheer J. A., Holm W. A. (2010). Principles of Modern Radar. SciTech Publishing
- Rishbeth, H., and O. K. Garriot (1969), Introduction to Ionospheric Physics, Academic Press, England.
- Rohling, H. (1983). Radar CFAR thresholding in clutter and multiple target situations. *IEEE Trans. Aerosp. Electron. Syst.* 19, 4.
- Roberts, M. J. (2004), Señales y Sistemas, Ed. Mc. Graw Hill, Mexico.

- Skolnik, M. I.(1980), "Introduction to radar systems, Mc. Graw-Hill, Tokyo.
- Skolnik, M.I. (2001). Introduction to Radar Systems. McGraw-Hill.
- Skolnik, M.I. (2008). Radar Handbook, Third Edition. McGraw Hill Professional.
- Sivaswamy, R.(1978), "Multiphase Complementary Codes", IEEE Trans. on Information Theory, vol. IT-24, No. 5, p.546-552.
- Simth, S. W. (1999). The scientist and engineer's guide to digital signal processing. 2nd Ed. California technical publishing.
- Sultzer, M, P., y R. F. Woodman (1984); Quasi-complementary codes: A new technique for radar sounding, Radio Science., 19, pp 337-344.
- Tie-Yan Liu,Kwok-Tung Lo, Xu-Dong Zhang, y Jian Feng (2003). A new cut detection algorithm with constant false-alarm ratio for video segmentation. Journal of visual communications and image representations.
- Tomasi, Wayne (1996); Sistemas de Comunicaciones Electrónicas, Ed. Prentice Hall, Mexico.
- Williams, David (2004). 50 years of Computing at CERN. Proceedings CHEP2004 Interlaken, 2004.
- Zolesi B. y Cander J. (2014). Ionospheric Prediction and Forecasting, Springer, Berlin.
- Zuccheretti, E., G. Tutone, U. Sciacca, C. Bianchi, J.A. Baskaradas (2003). The new AIS-INGV digital ionosonde, Annals of Geophysics, 46 (4), 647-659.
- Zuccheretti Enrico, Bianchi C. y Cabrera M. (2007), Mediciones realizadas en el Laboratorio de Radiofrecuencia, Sezione di Geomagnetismo, Aeronomia e Geofisica Ambientale; Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italia.