



INSTITUTO TECNOLÓGICO DE TUXTLA GUTIÉRREZ



RESIDENCIA PROFESIONAL

“ESTUDIO DE ALGORITMOS PARA EL PRONÓSTICO DEL
RECURSO FOTOVOLTAICO”

EMPRESA

CENTRO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS DEL
INSTITUTO POLITÉCNICO NACIONAL

RESIDENTE

ORTIZ GARCÍA EDUARDO
11270095

INGENIERÍA ELÉCTRICA
FUENTES ALTERNAS DE ENERGÍA

ASESOR INTERNO

M.C. KARLOS VELÁZQUEZ MORENO

ASESOR EXTERNO

DR. JUAN MANUEL RAMÍREZ ARREDONDO

TUXTLA GUTIÉRREZ, CHIAPAS JUNIO 2015

Índice

1. Introducción.....	3
1.1 Antecedentes	3
1.2 Estado del arte	3
1.3 Justificación	4
1.4 Objetivo	5
1.5 Metodología; diagrama a bloques	5
2. Fundamento teórico	8
2.1 El recurso solar	8
2.2 Predicción de irradiancia solar	9
2.3 Métodos de pronóstico	13
Series de tiempo	13
Método de predicción bayesiana	13
Redes neuronales.....	14
Filtros de kalman	15
3. Desarrollo	16
3.1 Enfoque series de tiempo	16
3.2 Metodología Box-Jenkins.....	17
3.3 Modelos de estudio	21
4. Resultados y Conclusiones	34
4.1 Resultados	34
4.2 Conclusiones.....	42
Referencias bibliográficas	43
Anexos	45
Anexo A	45
Anexo B	46
Anexo C	47
Anexo D	48

1. Introducción

1.1 Antecedentes

En el contexto energético actual, los beneficios económicos de las energías renovables han adquirido creciente relevancia, pues éstas contribuyen a reducir el impacto ambiental. Las principales ventajas del uso de las energías renovables son la diversificación del suministro energético, la reducción de la dependencia energética exterior, el incremento del nivel de autoabastecimiento y autosuficiencia energética.

El uso en gran escala de la energía renovable, es importante para el futuro por varias razones; para eliminar la dependencia con respecto a los combustibles fósiles, para combatir el calentamiento global y para elevar el nivel de vida de las poblaciones en los países en desarrollo. Gran parte de la energía renovable es hoy en día un campo de investigación y tecnología para el desarrollo de una nueva industria.

Entre las energías renovables, la energía solar presenta la mayor cantidad absoluta de potencial, la energía solar que llega a la superficie es más de tres órdenes de magnitud superior a la demanda energética mundial actual, pero existen factores que disminuyen su potencial como; La variación temporal de la radiación solar debido a los ciclos diarios, y la variación geográfica de la radiación solar dependiendo de la latitud.

El estudio, la profundización y el perfeccionamiento de los modelos de predicción de la irradiación solar se convierte en el principal pilar fundamental del desarrollo y puesta definitiva por las tecnologías renovables solares para la producción de electricidad. Cuanto más precisas son las predicciones de irradiancia solar, las fuentes productivas de electricidad a partir de este recurso son más fiables en la producción de energía eléctrica.

Existen en la actualidad métodos para el pronóstico de irradiancia solar como; los modelos físicos y modelos estadísticos, los cuales son una herramienta que permiten pronosticar y conocer con anticipación cuanta irradiancia se tendrá en un periodo de tiempo. Los modelos estadísticos se basan principalmente en predicciones confiables a partir del análisis de series de tiempo, para el pronóstico de la radiación a partir de datos históricos.

1.2 Estado del arte

Henzhou Peng, Shinjae Yoo, Dantong Yu, Dong Huang.- Este trabajo se centra en predecir la radiación solar de 30 minutos a 5 horas utilizando satélites geoestacionarios. Se compone de dos partes: el cálculo de movimiento de las nubes y la predicción de la radiación solar utilizando las imágenes de satélite además se propone el uso de estimación de movimiento óptico, dando mejoras significativas, tanto en la estimación de movimiento y predicción de irradiancia. [1]

Gonzalo Casaravilla, Rubén Chaer, Pablo Alfaro.- Exponen un generador de series temporales de radiación global diaria en plano horizontal sobre la superficie terrestre. Se presenta el ejemplo de identificación de los parámetros del modelo para una serie de 8 años. El modelo tiene la propiedad de generar series con igual histograma de amplitud de radiación solar que la del conjunto de datos y la misma función de auto correlación que la serie de datos en un espacio transformado. [2]

Mellit A, Pavan AM.- En este trabajo se abordó un método práctico para la irradiación solar mediante redes neuronales se presenta, el perceptrón multicapa que permite predecir la irradiancia solar sobre una base de 24h, utilizando los valores actuales y la temperatura diaria del aire, para comprobar la capacidad de generalización y la del pronosticador. Los resultados indican que el coeficiente de correlación es de 98% para los días soleados y 94% para los días nublados. [3]

Pérez R, Kivalov S, Schlemmer J.- Este artículo presenta una validación de las previsiones de irradiancia globales a corto y medio plazo que se producen como parte de los EE.UU. Los algoritmos de predicción están validados mediante mediciones en tierra por siete lugares climáticamente distintas en los Estados Unidos durante 1 año. También se presenta un primer análisis de los resultados regionales utilizando irradiancias obtenidas por satélite. [4]

Eugen Diacosnescu.- Este documento le da un enfoque a las redes neuronales recurrentes para la previsión de series de tiempo y en general para el control del sistema dinámico. Usa un modelo de neuronas autorregresivo lineal y el modelo de medias móviles. Además, este trabajo ha intentado identificar una manera de utilizar los métodos estadísticos para obtener métodos y mejorar la eficiencia del proceso de la predicción de series de tiempo con redes neuronales. [5]

Londhe, Vijay Panchang.- En este documento se realizan pruebas de pronóstico utilizando regresiones, modelos autorregresivos integrados de media móvil (ARIMA), funciones de transferencia, redes neuronales y modelos híbridos. Los modelos se estiman sobre la historia antes del inicio del horizonte de proyección, los valores predichos se comparan con los datos reales, los mejores resultados se obtuvieron utilizando el modelo ARIMA con coeficientes variables en el tiempo. [6]

Lo que aquí se propone como proyecto es desarrollar un algoritmo de predicción de irradiancia solar, a partir de datos históricos reportados en el laboratorio de potencia del CINVESTAV, los cuales son requeridos para la identificación del modelado y validación, con ello se pronostica y conoce cuanta irradiancia se tendrá en la entrada del sistema fotovoltaico en un periodo establecido.

1.3 Justificación

El presente proyecto es importante porque con él se desarrollara un algoritmo para el pronóstico de disponibilidad de irradiancia solar que será entregada al sistema fotovoltaico, permitiendo saber cuánta irradiancia es capaz de obtenerse en un periodo de tiempo establecido. Así mismo es necesario que el sistema otorgue

datos confiables de predicción para tener la estimación de potencia que será entregada a la red de CFE.

1.4 Objetivo

Desarrollo e implementación de un algoritmo de pronóstico de irradiancia solar, para predicciones puntuales en la operación de un sistema solar fotovoltaico.

1.5 Metodología; diagrama a bloques

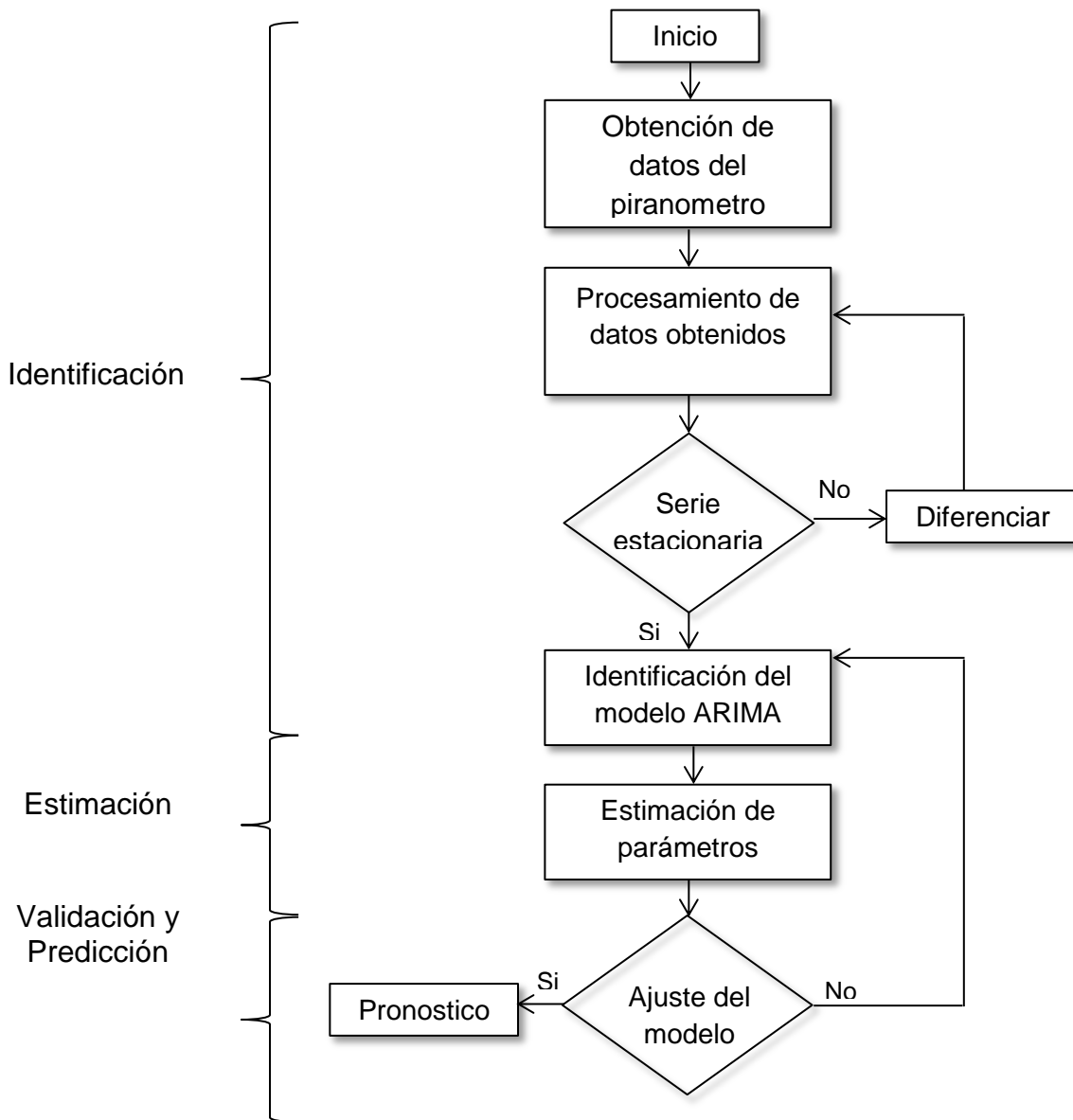


Fig. 1.1 Algoritmo de programación, basado en la Metodología Box-Jenkins ARIMA.

Identificación.- En la etapa de identificación el objetivo es definir si la serie que se quiere analizar contiene estacionariedad, de lo contrario es necesario aplicar algún tipo de transformación y finalmente deducir cuál de los modelos es el indicado, así mismo definir los parámetros que requiera. Para determinar el modelo, utilizamos principalmente las funciones de autocorrelacion y de autocorrelacion parcial.

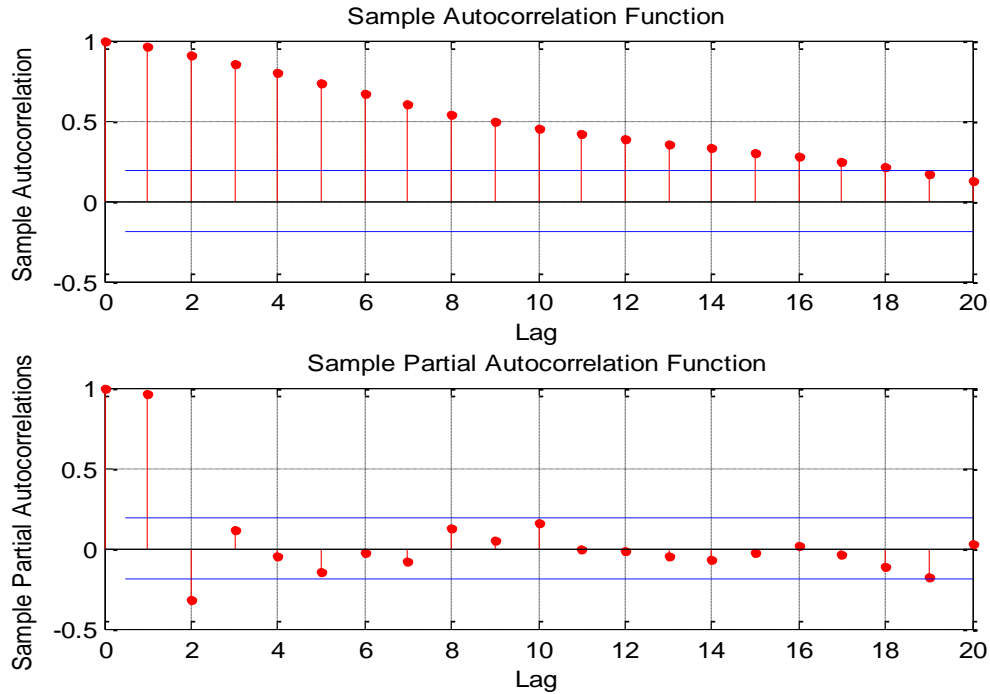


Fig. 1.2 Funciones, ACF y PACF de una serie que no presenta estacionariedad.

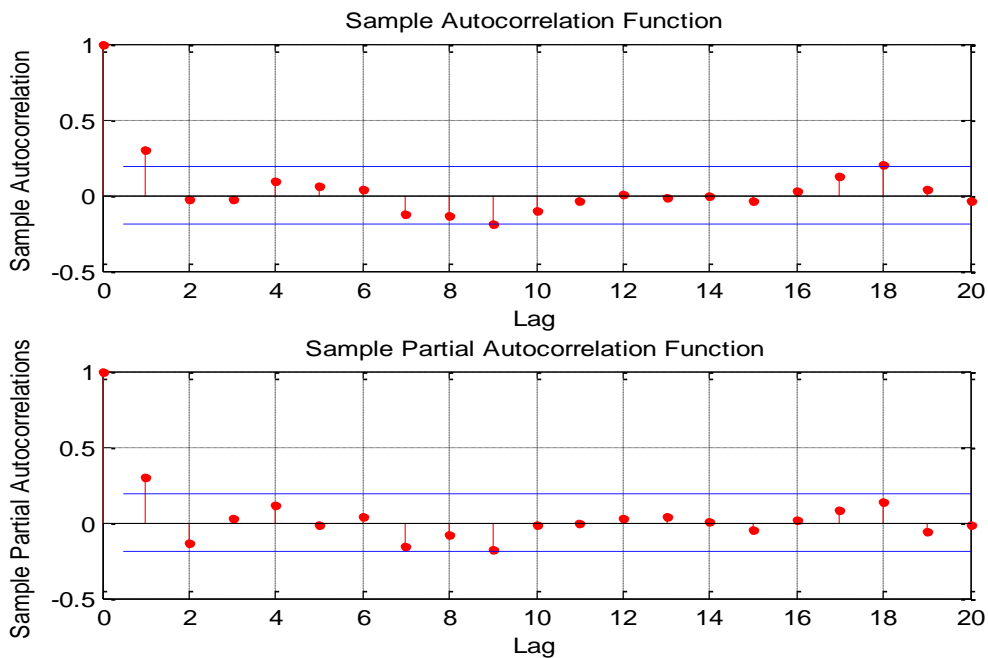


Fig. 1.3 Funciones, ACF y PACF de una serie con estacionariedad.

Estimación.- En la estimación de los modelos ARIMA se recurre a utilizar el método de mínimos cuadrados, el cual nos sirve para determinar los parámetros de cada proceso autorregresivo y de medias móviles, este proceso es importante para para definir el modelo de la ecuación que determine el proceso de validación y predicción.

Validación y Predicción.- Después de estimar los parámetros, será necesario realizar la validación del modelo. Esto consiste principalmente en determinar si el modelo se ajusta adecuadamente de lo contrario se identifica un nuevo modelo, una vez encontrado el modelo adecuado, se pueden llevar a cabo los pronósticos para un periodo, o varios, en el futuro. También pueden construirse intervalos de confianza para las predicciones con base en los pronósticos obtenidos.

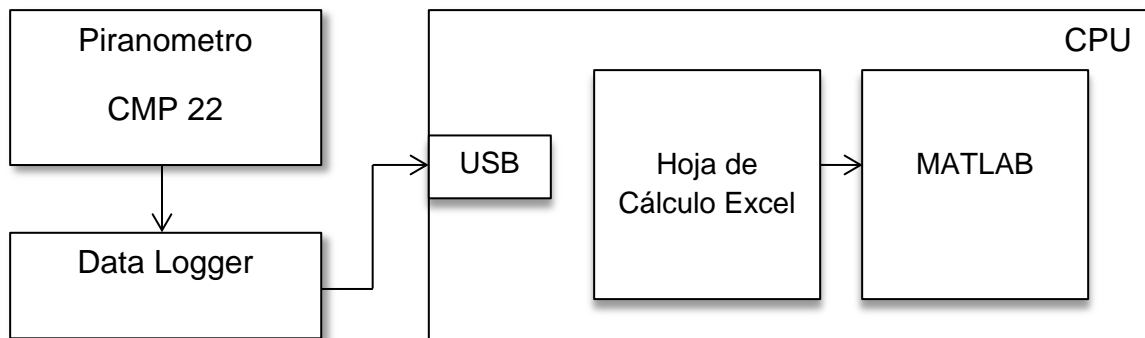


Fig. 1.4 Diagrama de bloques del software.

Piranometro CMP-22.- Es un instrumento meteorológico utilizado para medir la radiación solar incidente sobre la superficie de la tierra. Generalmente se utilizan medidas de radiación: semiesférica, difusa y directa. El principio físico utilizado generalmente en la medida es un termopar sobre el que incide la radiación a través de dos cúpulas de vidrio. Las medidas se expresan en W/m^2 .

Data Logger.- Es un dispositivo electrónico que registra datos en el tiempo o en relación a la ubicación por medio de instrumentos y sensores propios o conectados externamente. Casi todos están basados en microcontroladores. Por lo general son pequeños, con pilas, portátiles, y equipados con un microprocesador, memoria interna para almacenamiento de datos y sensores.

CPU.- La CPU se usa para hacer los procesos en los cuales, influyen las mediciones registradas por el piranometro, las cuales sirven de entrada para el sistema propuesto, este proceso se lleva a cabo mediante el software MATLAB donde se programa el algoritmo para dar resultados, a partir de datos medidos y proporcionados por el piranometro.

2. Fundamento teórico

2.1 El recurso solar

El recurso solar está actualmente considerado como el mayor recurso energético disponible en nuestro planeta. Con el sol como fuente energética la Tierra recibe 174 peta Watios de radiación en la parte más externa de la atmósfera. Aproximadamente el 30% es reflejada y devuelta de nuevo al espacio mientras que el resto es absorbida por las nubes, los océanos y la masa terrestre. [7]

La radiación o irradiancia global, es la radiación recibida en un plano horizontal a nivel de suelo, medida en watts por metro cuadrado. Esta medición de radiación será útil en este trabajo, puesto que es usada para poder estimar la predicción a corto plazo del paso de la irradiancia para un conjunto de paneles solares instalados en el Cinvestav unidad Guadalajara.

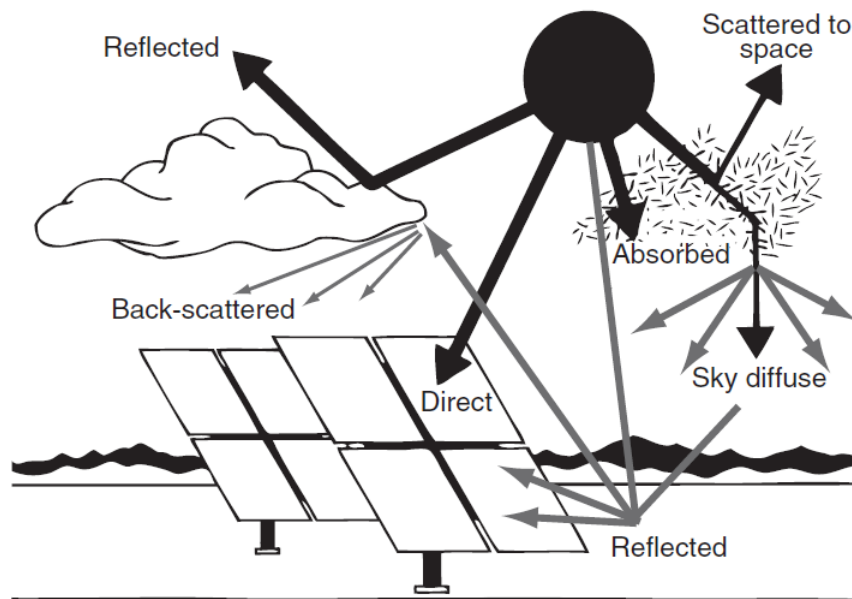


Fig. 2.1 Comportamiento de la radiación al pasar la atmósfera.

La radiación normal directa (DNI), es la radiación por unidad de área recibida por una superficie la cual está siempre perpendicular a la dirección (o en la normal) de los rayos del sol, lo que implica el estar en continuo seguimiento a la posición del sol en todo momento que sea requerido.

Esta radiación es utilizada, al igual que la radiación global horizontal se utiliza en la energía fotovoltaica, en plantas termosolares para calcular la potencia producible. En las plantas termosolares, los heliostatos se orientan para poder captar los rayos perpendiculares del sol y dirigirlos a la torre donde se encuentra la turbina encargada de generar la energía eléctrica.

Distancia Sol-Tierra

La órbita de la tierra alrededor del Sol es elíptica, estando el sol posicionado en uno de los focos. La distancia media del sol a la tierra se ha estimado en $1,496 \cdot 10^8$ km. A esta distancia se le llama Unidad Astronómica. La tierra se encuentra en su posición más próxima al sol el día 3 de Enero. Esta distancia es 0.983 UA. En el punto más lejano se encuentra aproximadamente el día 4 de Julio, siendo la distancia correspondiente de 1.017 UA. [8]

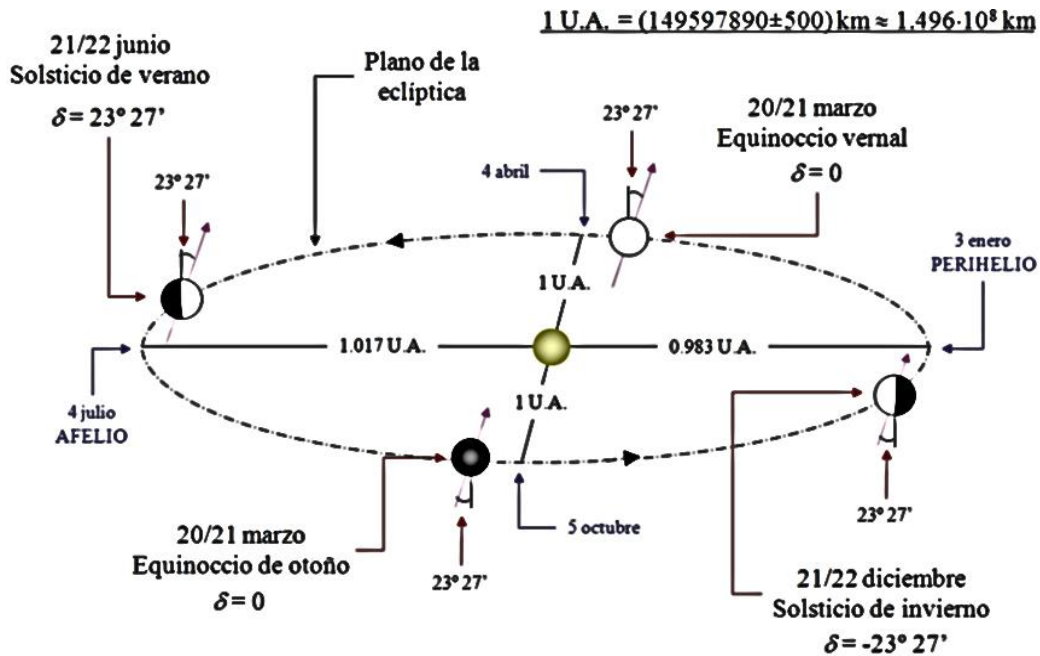


Fig. 2.2 Eclíptica con indicaciones de las posiciones de la Tierra más relevantes a lo largo del año.

2.2 Predicción de irradiancia solar

La predicción solar consiste en anticipar una radiación con distintos horizontes y en un espacio determinado. La técnica del Nowcasting predice a corto plazo, da predicciones entre 0 y 3 horas. El short-term forecasting predice entre 3 y 6 horas. Con predicciones temporales de horas hasta semanas se denomina medium-term forecasting y las predicciones de largo plazo que predicen para años. En la tabla 2.1 se detalla el tipo de horizonte y su paso temporal de predicción de irradiancia.

Dependiendo de la técnica utilizada, es decir, del horizonte de predicción y de la superficie que esta abarca, se utilizan distintos equipos y modelos como por ejemplo las cámaras de cielos claros, celiómetros, radiómetros, imágenes de satélites, métodos numéricos entre otros. Muchos expertos aseguran que para conseguir una predicción más precisa es necesaria la combinación de todas estas técnicas. [9]

Tipo	Horizonte	Paso temporal
Muy corto plazo (Nowcasting)	Una hora	Segundos a minutos
Corto plazo (Short-term forecasting)	Entre tres hora y 6 horas	Minutos, horas
Medio plazo (Medium-term forecasting)	Días, semana	Días
Largo plazo	Un año	Meses

Tabla 2.1 Escalas temporales de predicción.

Dentro de las herramientas utilizadas en la predicción solar se podrían dividir en aquellas utilizadas para medir datos en tiempo real y aquellas basadas en modelos o métodos en las cuales son introducidos los datos medidos para obtener estimaciones. Dentro del primer grupo, también denominado de observación terrestre se encuentran las siguientes herramientas:

Cámaras de cielo.- Se utilizan para localizar, determinar la altura y parametrizar el tipo de nubes y posteriormente aplicando una serie de algoritmos poder determinar cuál será la futura posición de las nubes. La resolución espacial que dan es muy alta, utilizándose casi siempre en plantas solares y pudiendo realizarse predicciones en el plazo de minutos.

El uso de esta herramienta para el muy corto plazo está en la actualidad en investigación y en proceso de optimización por el hecho de determinar cuál es el número de cámaras de cielo optimo a utilizar y de qué manera procesar la información que estas reciben dependiendo de la superficie en la que se va a predecir, hoy en día se están utilizando entre 2 y 3 cámaras por planta aunque nadie asegura que está sea la manera más fiable.



Fig. 2.3 Cámara de cielo

Estaciones meteorológicas.- Las estaciones meteorológicas determinan los valores instantáneos de velocidad de viento, dirección del mismo y valores de radiación la cual puede ser medida en forma de radiación global horizontal (GHI) mediante piranómetros, en forma de radiación difusa horizontal (DHI) utilizando piranómetros con anillo y en forma de radiación directa normal (DNI) con pirheliómetros.



Fig. 2.4 Piranometro



Fig. 2.5 Pirheliómetro

Celiómetros.- Aunque no son siempre utilizados por la mayoría de empresas dedicadas a la predicción solar, los celiómetros que podrían ir incluidos también en el grupo de las estaciones meteorológicas son utilizados para determinar la altura de las nubes cosa que sobre todo es necesaria para días claros en los cuales aparecen nubes altas y muy finas que pasan desapercibidas por las cámaras de cielo.

Radiómetros.- Los radiómetros son utilizados por algunas empresas para determinar la concentración de aerosoles en la atmósfera, este factor se convierte en vital para días claros en los cuales lo que más afecta a la radiación recibida por la superficie terrestre es la cantidad de aerosoles que la atmósfera contiene. Conociendo este dato con precisión y aplicando un modelo de cielos claros se puede reducir la incertidumbre en la predicción de radiación solar.

Imágenes satélite.- Las imágenes de satélites se utilizan para determinar el tipo, la forma y la dirección que llevan las nubes, y a partir de ello determinar una cobertura nubosa. Con el uso de satélites también se es capaz de determinar la cantidad de aerosoles contenida en la atmósfera.

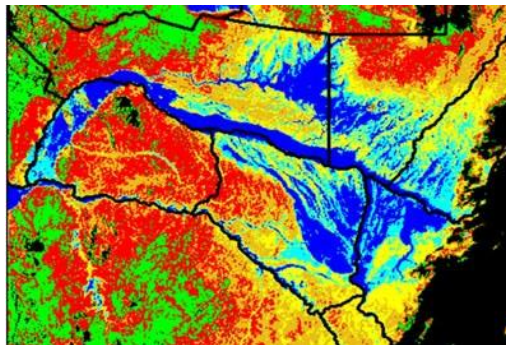


Fig. 2.6 Imagen de satélite

Cuando no se dispone de datos medidos, una de las metodologías para estimar ciertas variables es la utilización de modelos numéricos de predicción. Estos modelos están concebidos no solo para la predicción meteorológica a futuro sino también para la simulación de eventos meteorológicos pasados. El segundo grupo de herramientas lo constituyen todos aquellos métodos y modelos los cuales apoyándose en las variables medidas sacan un resultado.

Métodos Numéricos.- Estos métodos matemáticos funcionan de tal forma que en ellos se introducen las variables medidas mediante las herramientas de observación terrestre y la cobertura nubosa la cual se determina aplicando diferentes algoritmos al movimiento y dirección de las nubes, medido este mediante imágenes de satélites o con cámaras de cielo.

Métodos Meteorológicos Regionales.- Debido a que cada lugar tiene unas características meteorológicas determinadas, mayor cantidad de polvo, niebla por las mañanas. Se crean modelos de cielos claros para lugares específicos en los cuales a partir de estas características, se introducen las aproximaciones de predicción obtenidas en los métodos numéricos y se obtienen las predicciones de radiación definitivas.

Técnicas de predicción de radiación solar.

Nowcasting.- La tecnología más novedosa y por lo tanto la más complicada de llevar a cabo es el Nowcasting. Esta consiste en hacer predicciones a corto plazo con un horizonte de unas 3 o 4 horas. La resolución con la que se hacen las predicciones es muy alta. La obtención de ya sean valores de radiación horizontal o directa a tan pequeño plazo reside en la importancia de detectar el tipo de nubes y el movimiento que estas realizarán en caso de días con cielos cubiertos.

Si se trata de días claros, esta predicción será mucho más sencilla de realizar, encontrando como mayor inconveniente la concentración de aerosoles en la atmósfera donde algunas empresas están introduciendo los celiómetros como forma de detección de los mismos en lugar de imágenes de satélites. La mayoría de empresas aseguran que una combinación del conjunto de herramientas de predicción solar es clave para conseguir predicciones de manera más optimizada.

Forecasting.- Dentro de la tecnología de predicción solar del Forecasting, tecnología más madura y desarrollada que el Nowcasting, existen distintos niveles dependiendo del horizonte de predicción, que en función de cada empresa son asignados de diferente manera. El Forecasting que va a ser más interesante es el que abarca horizontes desde unas pocas horas hasta unas 24 horas.

El forecasting realiza predicciones horarias en cada intervalo del día. En estos casos la resolución con la que se realizan las predicciones se sitúa en torno a los 15 kilómetros cuadrados y en ellas las principales herramientas de predicción que intervienen son las imágenes de satélites junto con los métodos numéricos de predicción donde se podría decir que las imágenes de satélite se utilizan en un horizonte de hasta 6 horas.

2.3 Métodos de pronóstico

Series de tiempo

Los métodos de análisis de series de tiempo consideran el hecho que los datos tomados en diversos periodos de tiempo pueden tener algunas características de autocorrelacion, tendencia o estacionalidad que se debe tomar en cuenta. La aplicación de este método tiene dos propósitos: comprender las fuerzas de influencia en los datos y descubrir la estructura que produjo los datos observados, ajustar el modelo y proceder a realizar pronósticos. [10]

En muchas técnicas de series de tiempo los datos que se obtienen son estacionarios, donde su media, variancia y autocorrelacion no cambia en el tiempo, tampoco se presentan patrones de estacionalidad, sin embargo en la práctica si se presentan estos patrones de tendencia y de estacionalidad, es necesario contar con modelos que las consideren.

Los métodos de series de tiempo incluyen métodos de pronóstico, de suavizamiento simples, métodos de análisis de correlación y métodos de Box-Jenkins ARIMA.

Los métodos de pronóstico y suavizamiento simple se basan en la idea de que hay patrones visibles en una gráfica de series de tiempo que pueden ser extrapolados al futuro. El método se selecciona dependiendo si los patrones son estáticos o dinámicos, la naturaleza de los componentes de tendencia, estacionalidad y que tan lejos se quiera pronosticar, son métodos rápidos de aplicar.

Los Métodos de pronóstico ARIMA también usan patrones de datos, sin embargo puede que no sean fácilmente visibles en la serie de tiempo. El modelo usa funciones de diferencias, autocorrelacion y autocorrelacion parcial para ayudar a identificar un modelo aceptable. El modelo ARIMA representa una serie de pasos de filtraje hasta que solo queda ruido aleatorio. Es un proceso iterativo que consume tiempo de ejecución.

El Modelo ARIMA puede utilizarse para modelar series de tiempo con o sin componentes de tendencia o estacionalidad y proporcionar pronósticos. El perfil de pronóstico depende del modelo de ajuste. Tiene la ventaja de ser más flexible que los métodos de suavizamiento para el ajuste de los datos, sin embargo la identificación del modelo adecuado no puede ser fácilmente automatizado.

Método de predicción bayesiana

Los modelos bayesianos son una alternativa para elaborar pronósticos con pocos datos históricos. La estadística bayesiana no exige numerosos datos históricos; se puede basar en conocimientos anteriores facilitando la predicción, estos métodos constituyen alternativas de pronóstico a corto plazo, que utiliza información muy reciente y permiten actualizar de forma certera dichas predicciones.

Las bases bayesianas parten de la definición de una distribución anterior, otra de los datos, que lleva a la construcción de una distribución posterior para el modelo; según ésta, los pronósticos están condicionados a los mismos datos y los demás parámetros, pero hay otras herramientas que llevan a la actualización recursiva del pronóstico para cada tiempo en un modelo.

El modelo bayesiano se basa en información de los parámetros de la ecuación de predicción del sistema; pero, además la elección de un modelo se basada en la definición de una distribución, que explica una incertidumbre acerca de dicha elección. El propósito es predecir cantidades inciertas con base en el modelo adecuado, que es el que provee un medio para pronosticar a través de la distribución condicional.

La teoría bayesiana aparece cuando se postula el teorema de bayes, el cual toma como base el cálculo de probabilidad de un evento basado en un conocimiento anterior. La inferencia bayesiana hace poco tenía obstáculos técnicos que de alguna manera retardaron su desarrollo, pues generalmente este análisis conlleva a problemas de integración y complejidad en la formulación del problema. [11]

Redes neuronales

Las redes neuronales han experimentado un gran desarrollo gracias a su comportamiento en problemas de reconocimiento de patrones, predicción y clasificación. Las redes neuronales artificiales son mecanismos matemáticos que aprenden a reconocer o clasificar, dicho aprendizaje no descansa sobre un modelo preconcebido sino que busca las correlaciones existentes entre las variables del problema que se propone estudiar.

Las redes neuronales se basan en una estructura de neuronas unidas por enlaces que transmiten información a otras neuronas, las cuales entregan un resultado mediante funciones matemáticas de esta manera aprenden de la información histórica a través de un entrenamiento, mediante el cual se ajustan los parámetros de la red, a fin de entregar la respuesta deseada, adquiriendo entonces la capacidad de predecir el mismo fenómeno.

El funcionamiento de una red consiste en el aprendizaje de un conjunto predefinido de entradas dadas como ejemplo, empleando un ciclo de propagación y adaptación de dos fases: primero, al aplicar un primer patrón como estímulo para la capa de entrada de la red, éste se va propagando a través de las capas siguientes para generar la salida, la cual proporciona el valor del error al compararse con la que se desea obtener.

La importancia de la red consiste en su capacidad de auto adaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe entre un conjunto de patrones dados como ejemplo y sus salidas correspondientes. Dependiendo del tipo de aplicación, se han desarrollado distintos tipos de redes neuronales, que se aplican en sistemas de control, procesamiento, clasificación de patrones, determinación de variables climáticas, procesos químicos.

Filtros de kalman

El filtro de Kalman es un procedimiento matemático que opera por medio de un mecanismo de predicción y corrección. En esencia este algoritmo pronostica el nuevo estado a partir de su estimación previa añadiendo un término de corrección proporcional al error de predicción, de tal forma que este último es minimizado estadísticamente. [12]

El filtro de Kalman consiste en un conjunto de ecuaciones matemáticas que proveen una solución recursiva óptima, por el método de mínimos cuadrados. El filtro se desempeña suponiendo que el sistema puede ser descrito a través de un modelo estocástico lineal, en donde el error asociado tanto al sistema como a la información adicional que se incorpora en el mismo, tiene una distribución normal con media cero y varianza determinada.

El filtro estima el proceso anterior utilizando una especie de control de retroalimentación, estima el proceso a algún momento en el tiempo y obtiene la retroalimentación por medio de los datos observados. Desde este punto de vista las ecuaciones que se utilizan para derivar el filtro de Kalman se pueden dividir en dos grupos: las que actualizan el tiempo o ecuaciones de predicción y las que actualizan los datos observados o ecuaciones de corrección.

Metodología	Ventajas	Desventajas
Box-Jenkins ARIMA.	<ul style="list-style-type: none"> ✓ Los conceptos se derivan de solidas teorías de probabilidad clásica y estadística ✓ Proporcionan intervalos de confianza para las predicciones ✓ Produce optimas predicciones 	<ul style="list-style-type: none"> ✓ Requiere gran número de datos históricos ✓ La identificación del modelo adecuado es proporcionado por el usuario ✓ Requiere experiencia en teoría de probabilidad clásica y estadística
Predicción bayesiana	<ul style="list-style-type: none"> ✓ Pueden manejar bases de datos incompletos 	<ul style="list-style-type: none"> ✓ Requiere experiencia del usuario en teoría previa al modelo
Redes neuronales	<ul style="list-style-type: none"> ✓ Habilidad de aprender y adaptarse al problema 	<ul style="list-style-type: none"> ✓ No tienen fundamento teórico que ayude a validar y mejorar el entrenamiento
Filtro de kalman	<ul style="list-style-type: none"> ✓ Habilidad para predecir el estado de un modelo en el pasado, presente y futuro 	<ul style="list-style-type: none"> ✓ Requiere condiciones iniciales para iniciar el algoritmo recursivo ✓ Requiere conocimientos sobre el sistema

Tabla 2.2 Comparación de los métodos de pronostico

En el estudio se opta por trabajar con la metodología de Box-Jenkins para predicciones a corto plazo, debido a las ventajas que presenta para óptimas predicciones, e intervalos de confianza esta última característica limitante en otras técnicas de predicción. La metodología se basa en conceptos que se derivan de sólidas teorías de probabilidad clásica y estadística.

3. Desarrollo

3.1 Enfoque series de tiempo

Una serie de tiempo es una colección de observaciones tomadas a lo largo del tiempo cuyo objetivo principal es describir, explicar, predecir y controlar algún proceso. Las observaciones están ordenadas respecto al tiempo y sucesivas observaciones son generalmente dependientes. La siguiente ecuación muestra el modelo de una serie de tiempo. [13]

$$Y = \{Y_1, Y_2, Y_3 \dots \dots Y_N\} \quad (3.1)$$

Hoy en día el análisis de datos ordenados en el tiempo es fundamental en muchos campos como la ingeniería, la medicina o el mundo de negocios. Estudiar cómo se ha comportado una variable en el pasado puede ser de gran valor a la hora de pronosticar su comportamiento futuro. Del mismo modo, determinar que otros valores han tenido un comportamiento similar para ayudar a decidir las acciones que se deberán llevar a cabo.

Evidentemente, aunque el valor futuro de una serie de tiempo no sea predecible con total exactitud, para que tenga interés su estudio, el resultado tampoco puede ser completamente aleatorio, existiendo alguna regularidad en cuanto a su comportamiento en el tiempo haciendo posible su modelado, y por consiguiente su predicción.

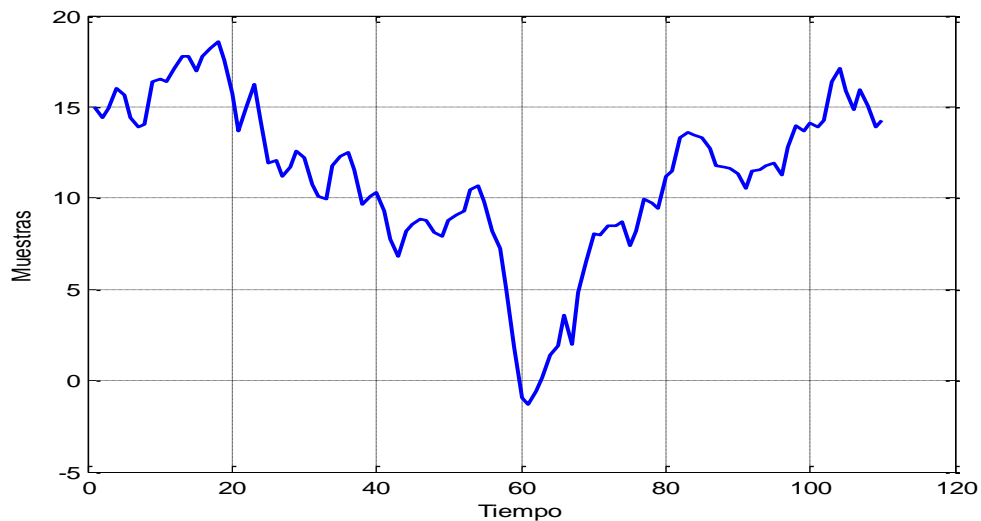


Fig. 3.1 Gráfico de una serie de tiempo.

3.2 Metodología Box-Jenkins

El enfoque de Box-Jenkins es una de las metodologías de uso más amplio para el modelamiento de series de tiempo. Box y Jenkins han desarrollado modelos estadísticos que tienen en cuenta la dependencia existente entre los datos. Cada observación en un momento dado es modelada en función de los valores anteriores. [14]

Modelos autorregresivos.

Definimos un modelo autorregresivo si la variable de un período t es explicada por las observaciones de ella misma correspondientes a períodos anteriores añadiéndose un término de error. En el caso de procesos estacionarios con distribución normal, los procesos estocásticos dice que, bajo determinadas condiciones previas, toda Y_t puede expresarse como una combinación lineal de sus valores pasados más un término de error.

Los modelos autorregresivos se abrevian con la palabra AR tras la que indica el orden del modelo: AR (1), AR (2),...AR (p). El orden del modelo expresa el número de observaciones retrasadas de la serie de tiempo analizada que interviene en la ecuación. Así, por ejemplo, un modelo de orden AR (1) tendría la siguiente expresión:

$$Y_t = \delta + \phi_1 Y_{t-1} + \varepsilon_t \quad (3.2)$$

La expresión general de un modelo autorregresivo AR (p) será la siguiente:

$$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (3.3)$$

Modelo de medias móviles

Un modelo de los denominados de medias móviles es aquel que explica el valor de una determinada variable en un período t en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente. Estos modelos se denotan normalmente con las siglas MA, seguidos, como en el caso de los modelos autorregresivos, del orden. Así, un modelo con q términos de error MA (q) respondería a la siguiente expresión:

$$Y_t = \delta + \varepsilon_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (3.4)$$

Modelo Autorregresivos - Media Móvil, ARMA (p,q)

$$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (3.5)$$

En estos modelos, el proceso se representa en función de observaciones pasadas de la variable y de los valores actuales y rezagados del error. El número de rezagos de la variable de interés (p) y el número de rezagos del error (q) determinan el orden del modelo mixto.

Modelos Autorregresivos Integrados de Media Móvil, ARIMA (p,d,q)

Muchas series de tiempo no son estacionarias, un tipo especial de series de tiempo no estacionarias, son las no estacionarias que se caracterizan porque, al ser diferenciadas, se vuelven estacionarias. Podemos transformar una serie temporal no estacionaria, en valores estacionarios obteniendo las primeras diferencias de los valores.

$$Y_t = y_t - y_{t-1} \tag{3.6}$$

Después de haber diferenciado la serie Y_t se consigue una serie estacionaria y_t , y dicha serie obedece a un proceso ARMA (p,q), se dice que Y_t responde a un proceso ARIMA (p,d,q), donde p se determina por el orden AR, q por el orden MA, y d por el orden de diferenciaciones establecidas:

$$y_t = \delta + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \tag{3.7}$$

Funciones de Autocorrelacion Y Autocorrelacion Parcial

La autocorrelación es la correlación entre observaciones de una serie de tiempo separadas por K unidades de tiempo, su correlograma se denomina función de autocorrelación (ACF), su análisis permite seleccionar los términos a ser incluidos en el modelo ARIMA.

$$r_k = \frac{\sum_{t=b}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=b}^n (y_t - \bar{y})^2} \tag{3.8}$$

Dónde:

$$\bar{y} = \frac{\sum_{t=b}^n y_t}{n - b + 1} \tag{3.9}$$

La autocorrelación parcial es la correlación entre conjuntos de pares ordenados de una serie, mide la fuerza de la relación con otros términos tomados en cuenta. La autocorrelación parcial en una posición K es la correlación entre residuos en tiempo t de un modelo autorregresivo y las observaciones en la posición K con términos para las posiciones que intervienen en el modelo autorregresivo. Su análisis permite seleccionar los términos incluidos en el modelo ARIMA.

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \tag{3.10}$$

Dónde:

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j} \text{ Para } J=1,2,\dots, K=-1 \tag{3.11}$$

Pasos de la metodología de Box-Jenkins:

Identificación

En la etapa de identificación el objetivo es definir nuestra serie de tiempo para ello es necesario conocer si la serie tiene estacionariedad de lo contrario se aplicará la ecuación (3.5) para hacerla estacionaria, finalmente se postulan los modelos matemáticos. Para determinar los modelos, utilizamos principalmente las funciones de autocorrelación y autocorrelación parcial siguiendo el criterio de la tabla 3.1

Modelo	Función de autocorrelación	Función de autocorrelación parcial
MA (q)	Se anula para retardos superiores a q	Decrecimiento rápido sin llegar a anularse
AR (p)	Decrecimiento rápido sin llegar a anularse	Se anula para retardos superiores a p
ARMA (p,q)	Decrecimiento rápido sin llegar a anularse	Decrecimiento rápido sin llegar a anularse

Tabla 3.1 Comportamiento de las funciones ACF y PACF

Estimación

Para estimar los parámetros del modelo matemático se utiliza un algoritmo de mínimos cuadrados de Gauss para minimizar la suma de cuadrados de los residuos. Este algoritmo trata de minimizar la suma de cuadrados de los residuos, comenzando con algún valor de los parámetros del modelo. El algoritmo busca si otro vector de parámetros mejora el valor de la función objetivo y se produce un proceso de iteración hasta que se alcanza su cierto criterio de convergencia.

Validación y pronóstico

Después de estimar los parámetros, será necesario realizar la validación del modelo. Esto consiste principalmente en determinar si el modelo se ajusta adecuadamente al proceso, de lo contrario se identifica un nuevo modelo, después de haber encontrado un modelo adecuado, se pueden llevar a cabo los pronósticos para un periodo, o varios, en el futuro. También pueden construirse intervalos de predicciones con base en los pronósticos.

Lo que se trata con la metodología es identificar el proceso que ha generado los datos, estimar los parámetros que caracterizan dicho proceso, verificar que se cumplan las hipótesis que han permitido la estimación de dichos parámetros. Si dichos supuestos no se cumplieran, la fase de validación sirve como retroalimentación para una nueva fase de identificación. Cuando se satisfagan las condiciones de partida, se puede utilizar el modelo para pronosticar. [15]

Evaluación de Modelos de Series de Tiempo

Para la evaluación del desempeño predictivo de la serie de tiempo se emplean diferentes indicadores que cuantifican qué tan cerca está la variable pronosticada de la serie de datos correspondiente. Una de las medidas más utilizadas es el Promedio del Error Porcentual Absoluto (MAPE).

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left(\frac{|y'_t - y_t|}{y_t} \right) \times 100 \quad (3.12)$$

Otras medidas del error de pronóstico comúnmente empleadas son: Raíz cuadrática media del error (RMSE) y Error absoluto medio (MAE).

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y'_t - y_t)^2} \quad (3.13)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |y'_t - y_t| \quad (3.14)$$

Dónde:

N =Numero de periodos

y_t =Valor real

y'_t = Valor pronosticado

No siempre el modelo que genere pronósticos con un menor MAPE generará los pronósticos con el menor RMSE y viceversa, por lo que en la selección de los mejores modelos de pronóstico se hace necesario establecer la medida de error a utilizar para la elaboración del ranking de desempeño. [16]

3.3 Modelos de estudio

Siguiendo la metodología de Box y Jenkins se abordan modelos de estudio, en el primer modelo los valores manejados fueron proporcionados por el laboratorio de potencia del Cinvestav, la toma de datos proporcionados fue de 6609 muestras que comprenden del 19 al 23 de enero del 2015, los valores tomados para la identificación del modelo fueron del día 19 al 22 de enero con 5169 datos, los datos de predicción son del día 23 con 1440 datos de validación.

En la figura 3.2 se observa la serie de tiempo de los días 19 al 22 de enero del 2015 con 5169 muestras tomadas cada minuto, dichas muestras son graficadas con el software Matlab, en el eje de las abscisas se encuentra el periodo de tiempo y en el eje de las ordenadas se encuentra el valor de la irradiancia, de igual manera se grafican las funciones de autocorrelacion y autocorrelacion parcial basándose en las ecuaciones (3.8 y 3.10).

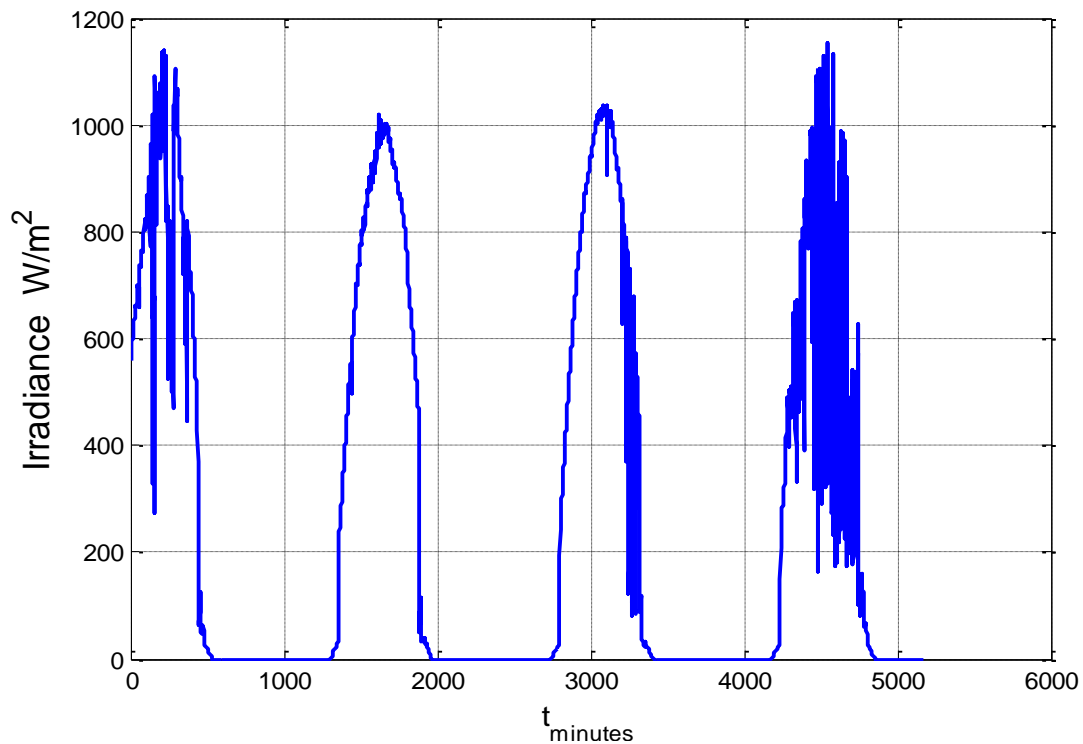


Fig. 3.2 Serie de tiempo del 19 al 22 de enero del 2015

En la figura 3.3 se observan las funciones de autocorrelacion y autocorrelacion parcial de la serie de tiempo original, en el proceso se puede observar que la función de autocorrelacion cae de forma lenta, lo que indica que la serie de tiempo no contiene estacionariedad, y en el caso requiere aplicarle una diferenciación aplicando la ecuación (3.6) para hacer el proceso estacionario.

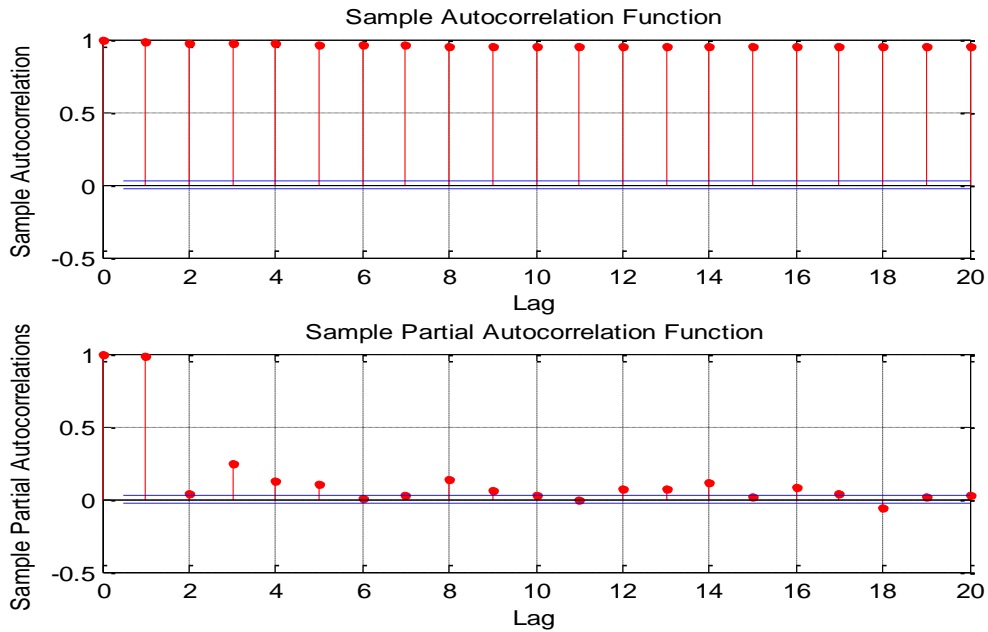


Fig. 3.3 Graficas de Autocorrelacion y Autocorrelacion Parcial de la serie de tiempo original

Al aplicar la diferenciación correspondiente a la serie de tiempo original se grafican nuevamente las funciones de autocorrelacion y autocorrelacion parcial para determinar si la serie presenta estacionariedad. En la figura 3.4 se presentan los datos de la serie original diferenciados y en la figura 3.5 se presentan los correlogramas de las funciones de autocorrelacion y autocorrelacion parcial de un proceso con estacionariedad.

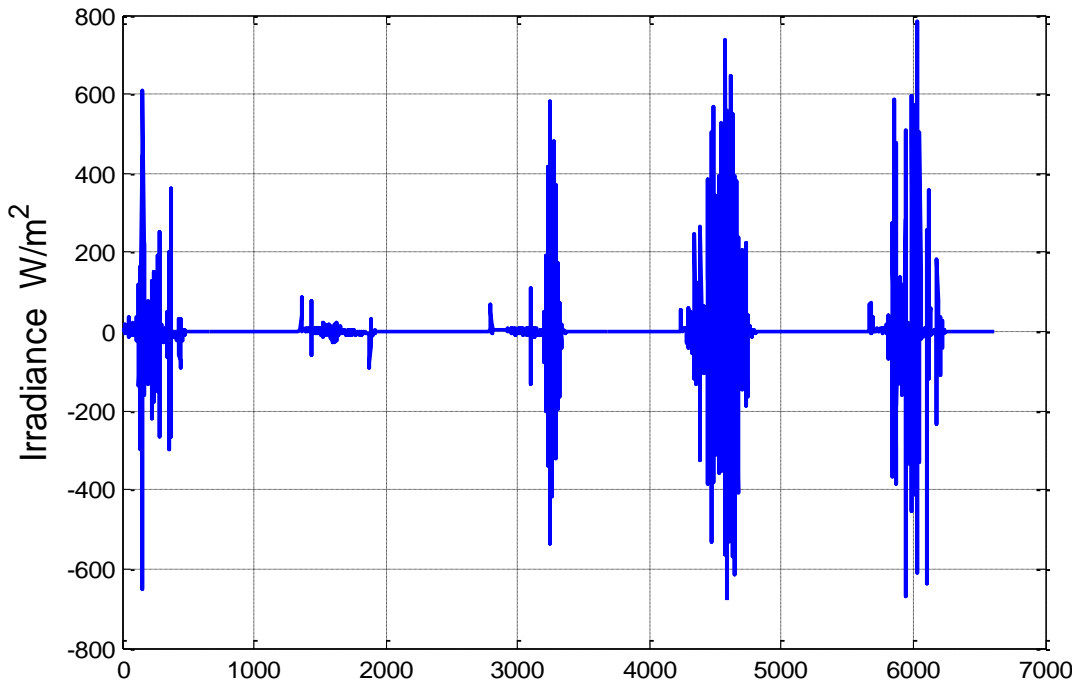


Fig. 3.4 Serie de tiempo diferenciada

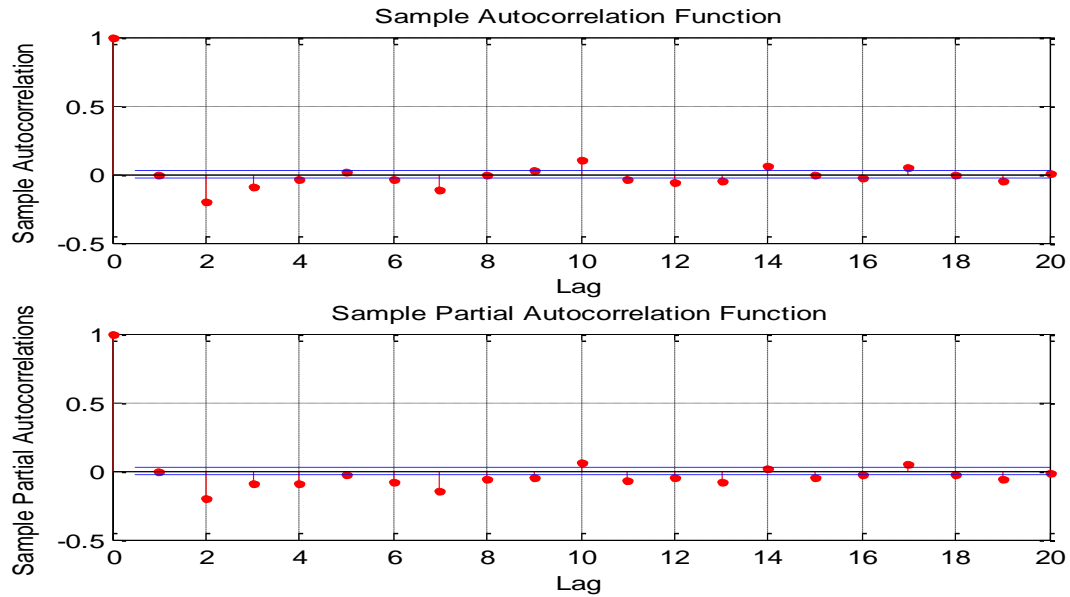


Fig. 3.5 Graficas de Autocorrelacion y Autocorrelacion Parcial de la serie de tiempo diferenciada

En la figura 3.5 se observa que la función de autocorrelación y la función de autocorrelación parcial decaen después del primer rezago, lo que indica que la serie contiene estacionariedad, basándose en los correlogramas y en la tabla 3.1 se siguen los criterios para determinar el grado de los modelos autorregresivos y de medias móviles que son incluidos en los posibles modelos matemáticos ARIMA.

Finalmente se postulan todos los posibles modelos matemáticos que pueden ser evaluados y estimados por el proceso de mínimos cuadrados de Gauss, esto con el fin de minimizar la suma de cuadrados de los residuos, de igual manera mejora el valor de la función objetivo y se produce un proceso de iteración hasta que se alcanza su cierto criterio de convergencia.

$$\text{ARIMA (2,1,2): } y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad (3.15)$$

ARIMA(2,1,2) Model:

 Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.0415752	0.139927	-0.29712
AR{1}	0.42192	0.0218255	19.3315
AR{2}	0.108434	0.0187829	5.77302
MA{1}	-0.485996	0.0212575	-22.8623
MA{2}	-0.330296	0.0204333	-16.1646
Variance	3682.04	15.1188	243.54

$$\text{ARIMA (2,1,3): } y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} \quad (3.16)$$

ARIMA(2,1,3) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.0327601	0.110064	-0.297647
AR{1}	0.640899	0.198168	3.23411
AR{2}	-0.0116746	0.108356	-0.107743
MA{1}	-0.705108	0.198306	-3.55566
MA{2}	-0.197173	0.120741	-1.63302
MA{3}	0.0468842	0.0413929	1.13266
Variance	3682.12	15.134	243.301

$$\text{ARIMA (2,1,4): } y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} + \theta_4 a_{t-4} \quad (3.17)$$

ARIMA(2,1,4) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.0580587	0.198885	-0.291921
AR{1}	-0.283091	0.0223121	-12.6878
AR{2}	0.612373	0.0153524	39.8878
MA{1}	0.223289	0.022837	9.77754
MA{2}	-0.888486	0.0150179	-59.1618
MA{3}	-0.151933	0.00929442	-16.3467
MA{4}	0.0778229	0.00781543	9.9576
Variance	3672.18	15.4125	238.261

$$\text{ARIMA (3,1,2): } y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad (3.18)$$

ARIMA(3,1,2) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.0382799	0.13491	-0.283743
AR{1}	0.467399	0.0725949	6.43846
AR{2}	0.0729691	0.0552203	1.32142
AR{3}	0.00919344	0.0135734	0.677311
MA{1}	-0.531574	0.0726365	-7.31828
MA{2}	-0.292524	0.0596239	-4.90615
Variance	3682.1	15.123	243.478

$$\text{ARIMA (3,1,3): } y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} \quad (3.19)$$

ARIMA(3,1,3) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.00475414	0.0200343	-0.2373
AR{1}	1.30078	0.0350334	37.1298
AR{2}	-0.164841	0.0455239	-3.62097
AR{3}	-0.191258	0.0200806	-9.52451
MA{1}	-1.3647	0.0337789	-40.4009
MA{2}	0.00599175	0.0469295	0.127675
MA{3}	0.384565	0.0207993	18.4893
Variance	3676.28	15.3382	239.682

ARIMA (3,1,4):

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} + \theta_4 a_{t-4} \quad (3.20)$$

ARIMA(3,1,4) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.108727	0.431566	-0.251937
AR{1}	-1.1136	0.0234141	-47.5611
AR{2}	0.250945	0.0198642	12.6331
AR{3}	0.456508	0.010867	42.0086
MA{1}	1.05668	0.0237818	44.4323
MA{2}	-0.572465	0.0181621	-31.5198
MA{3}	-0.829013	0.0161447	-51.3489
MA{4}	-0.102898	0.00876832	-11.7352
Variance	3653.33	15.1049	241.865

$$\text{ARIMA (4,1,2): } y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \quad (3.21)$$

ARIMA(4,1,2) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.0428967	0.146057	-0.2937
AR{1}	0.33433	0.230581	1.44995
AR{2}	0.183343	0.189903	0.965453
AR{3}	-0.0185118	0.0459667	-0.402721
AR{4}	0.0143742	0.0181524	0.79186
MA{1}	-0.398479	0.230242	-1.73069
MA{2}	-0.412037	0.204663	-2.01325
Variance	3681.87	15.1866	242.442

ARIMA (4,1,3):

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} \quad (3.22)$$

ARIMA(4,1,3) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.117879	0.386556	-0.304947
AR{1}	-0.966395	0.0140302	-68.8795
AR{2}	0.384267	0.0111398	34.495
AR{3}	0.393239	0.010804	36.3975
AR{4}	-0.0659032	0.00571771	-11.5261
MA{1}	0.909503	0.0139003	65.4304
MA{2}	-0.694936	0.00977036	-71.127
MA{3}	-0.719537	0.0108325	-66.424
Variance	3654.89	15.1018	242.017

ARIMA (4,1,4):

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \varepsilon_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \theta_3 a_{t-3} + \theta_4 a_{t-4} \quad (3.23)$$

ARIMA(4,1,4) Model:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-0.112754	0.440583	-0.255921
AR{1}	-0.606554	0.0213121	-28.4605
AR{2}	-0.462905	0.00859485	-53.8584
AR{3}	0.493748	0.00898691	54.9408
AR{4}	0.124518	0.0182872	6.80901
MA{1}	0.546716	0.0207247	26.3799
MA{2}	0.183782	0.00666583	27.5708
MA{3}	-0.794887	0.00689841	-115.228
MA{4}	-0.364299	0.0200777	-18.1445
Variance	3655.98	15.2678	239.458

Para el segundo modelo de estudio los valores que se toman fueron proporcionados por The University of Texas at Austin, Webber Energy Group, se proponen 4 modelos con diferentes intervalos de días para la identificación del sistema, la teoría de la metodología Box-Jenkins describe que a mayor cantidad de datos tomados para el modelo el desempeño será mejor en el resultado de la predicción.

La toma de datos proporcionados corresponde a los meses Mayo y Junio del 2013, para la identificación del sistema se toman diferentes intervalos de días en el modelado, validación y predicción, en la tabla 3.2 se observan los intervalos de días y la cantidad de datos totales usados para cada modelo correspondiente estudiado.

Intervalo de días, identificación	Datos del modelo	Datos a validar	Datos totales
5 días	7199	1440	8639
10 días	14399	1440	15839
20 días	28799	1440	30239
35 días	50399	1440	51839

Tabla 3.2 Datos de identificación del segundo modelo de estudio

Para la identificación del modelo se toma la mayor cantidad de días, para el caso 35 días correspondientes del 01/05/13 al 06/06/13, en la figura 3.6 se observa la serie de tiempo con muestras tomadas cada minuto, en el eje de las abscisas se encuentra el periodo de tiempo y en el eje de las ordenadas se encuentra el valor de la irradiancia, de igual manera se grafican las funciones de autocorrelacion y autocorrelacion parcial basándose en las ecuaciones (3.8 y 3.10).

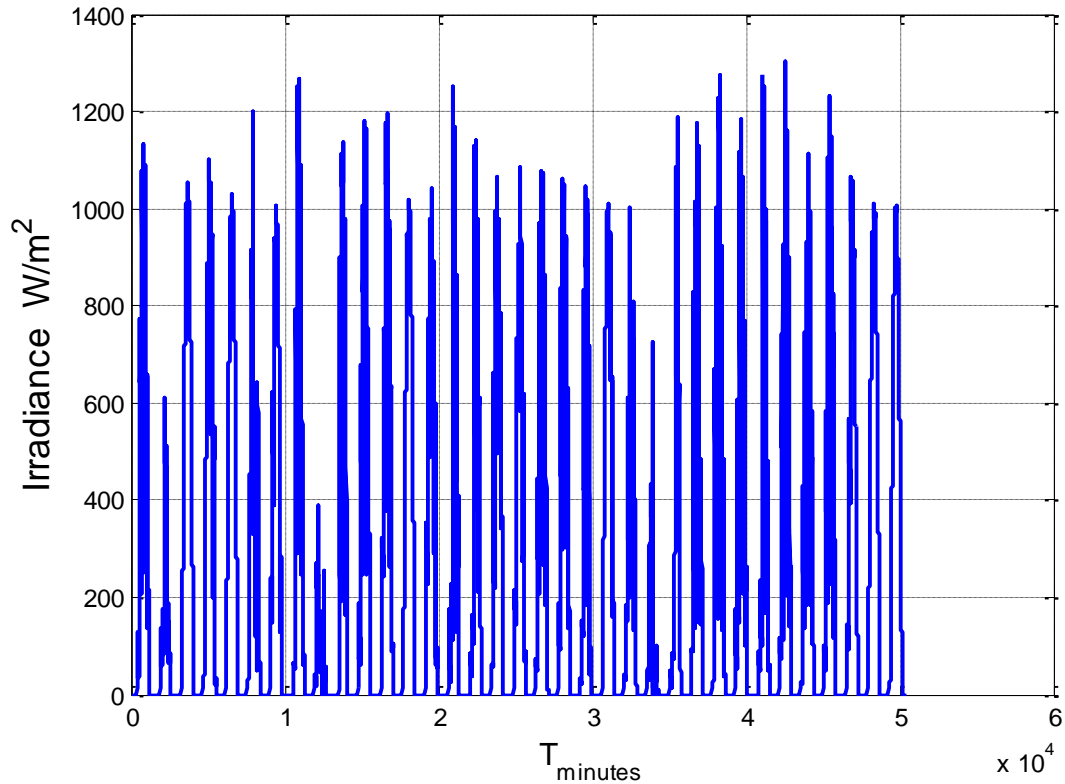


Fig. 3.6 Serie de tiempo original

En la figura 3.7 se observan las funciones de autocorrelacion y autocorrelacion parcial de la serie de tiempo original, en el proceso se puede observar que la función de autocorrelacion decae de forma lenta, lo que indica que la serie de tiempo no contiene estacionariedad, y en el caso requiere aplicarle una diferenciación aplicando la ecuación (3.6) para hacer que el proceso contenga estacionariedad.

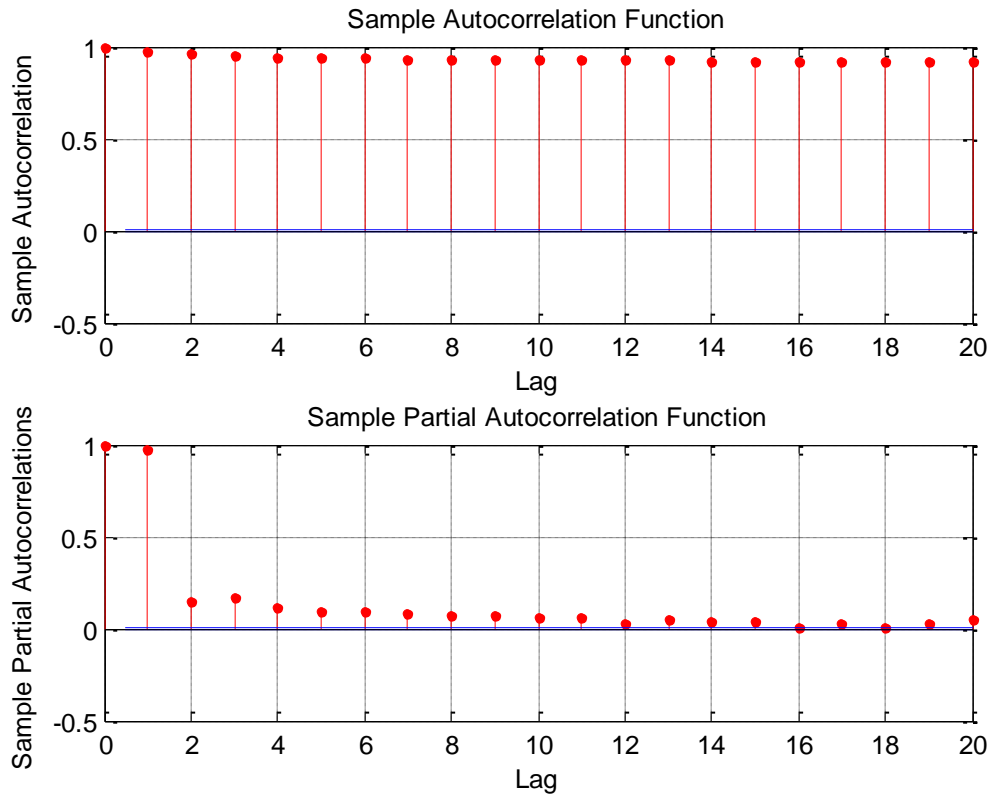


Fig. 3.7 Graficas de Autocorrelacion y Autocorrelacion Parcial de la serie de tiempo original

Al aplicar la diferenciación correspondiente a la serie de tiempo original se grafican nuevamente los correlogramas de autocorrelacion y autocorrelacion parcial para determinar si la serie presenta estacionariedad. En la figura 3.8 se presentan los datos de la serie original diferenciados y en la figura 3.9 se presentan las funciones de autocorrelacion y autocorrelacion parcial de un proceso con estacionariedad.

En la figura 3.9 se observa que la función de autocorrelación y la función de autocorrelación parcial decaen después del primer rezago, lo que indica que la serie contiene estacionariedad, basándose en los correlogramas y en la tabla 3.1 se siguen los criterios para determinar el grado de los modelos autorregresivos y de medias móviles que son incluidos en los posibles modelos matemáticos ARIMA.

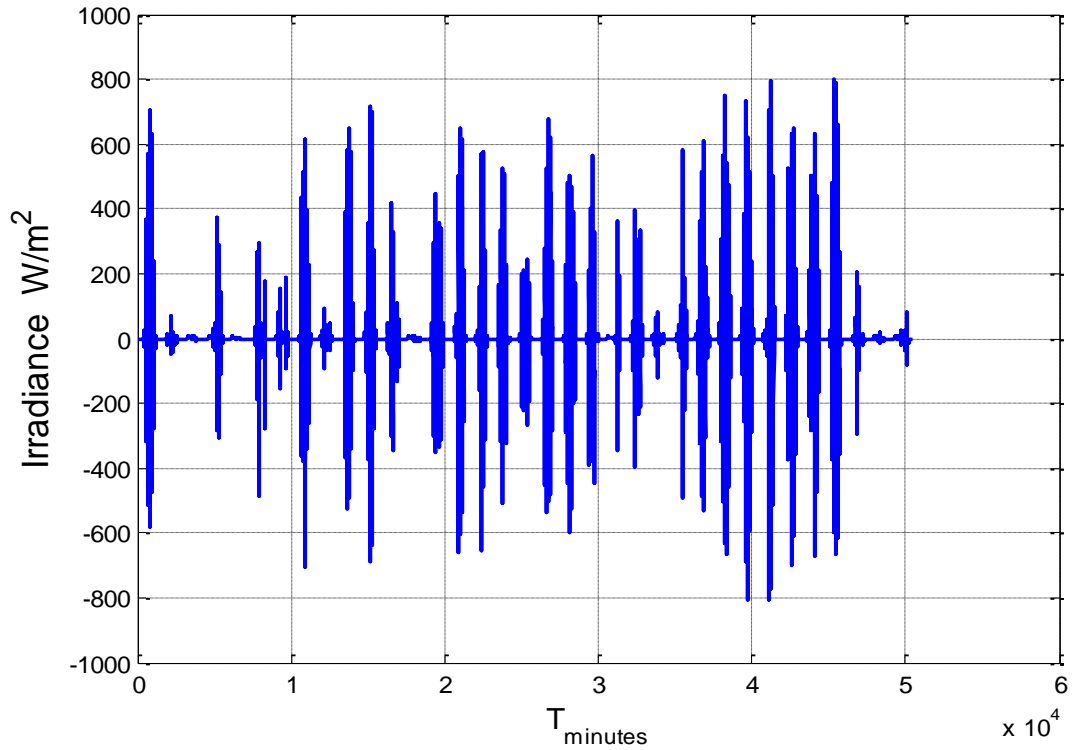


Fig. 3.8 Serie de tiempo diferenciada

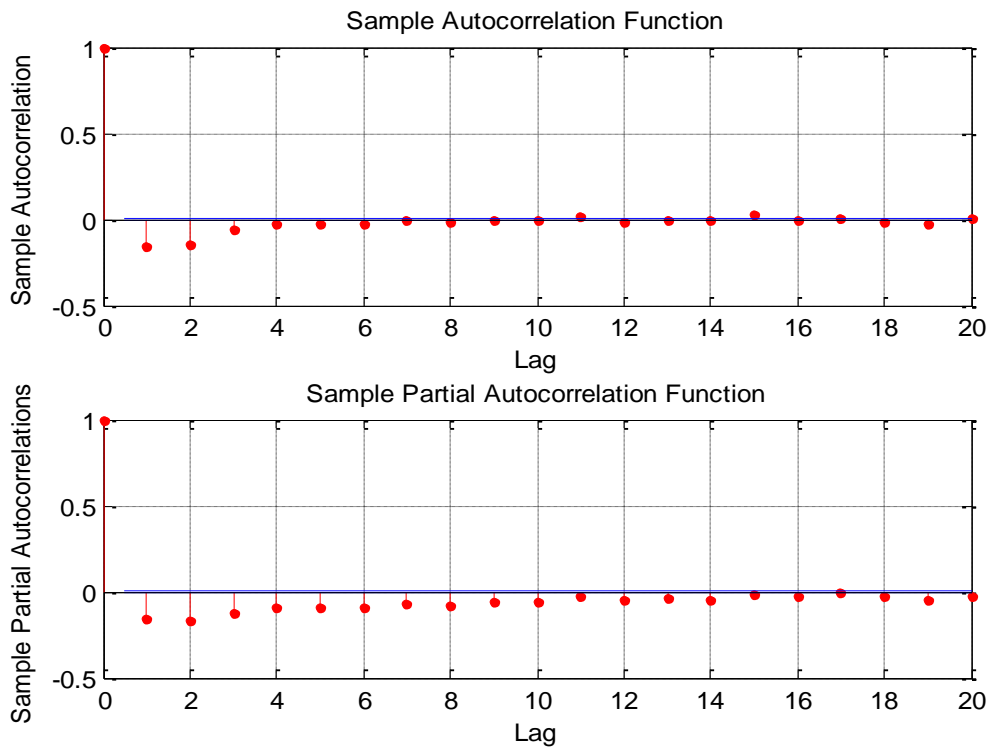


Fig. 3.9 Graficas de Autocorrelacion y Autocorrelacion Parcial de la serie de tiempo diferenciada

Finalmente se postulan todos los posibles modelos matemáticos que pueden ser evaluados y estimados por el proceso de mínimos cuadrados de Gauss, esto con el fin de minimizar la suma de cuadrados de los residuos, de igual manera mejora el valor de la función objetivo y se produce un proceso de iteración hasta que se alcanza su criterio de convergencia.

$$\text{ARIMA 5 días } (1,1,1): y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.24)$$

ARIMA(1,1,1) Model 5 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	4.24565e-06	0.0646909	6.56298e-05
AR{1}	0.601999	0.00537348	112.032
MA{1}	-0.92107	0.00335694	-274.378
Variance	4654.99	24.6483	188.857

$$\text{ARIMA 10 días } (1,1,1): y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.25)$$

ARIMA(1,1,1) Model 10 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-1.48917e-05	0.0512782	-0.00029041
AR{1}	0.632441	0.00421475	150.054
MA{1}	-0.932657	0.00238106	-391.698
Variance	7859.69	35.1693	223.482

$$\text{ARIMA 20 días } (1,1,1): y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.26)$$

ARIMA(1,1,1) Model 20 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-4.18503e-05	0.0372215	-0.00112436
AR{1}	0.609354	0.00282023	216.066
MA{1}	-0.918103	0.0017151	-535.305
Variance	5747.13	16.2938	352.72

$$\text{ARIMA } 35 \text{ días } (1,1,1): y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.27)$$

ARIMA(1,1,1) Model 35 días:

 Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	7.91609e-05	0.0279321	0.00283405
AR{1}	0.621986	0.0020265	306.927
MA{1}	-0.908733	0.00133947	-678.425
Variance	4639.11	8.9339	519.27

Para un tercer caso de estudio los valores que se toman fueron proporcionados por The University of Texas at Austin, Webber Energy Group, se proponen varios modelos con diferentes intervalos de días para la identificación del sistema, La toma de datos proporcionados corresponde a los meses Enero, Febrero y Marzo del 2014, en la tabla 3.3 se observan los intervalos de días y la cantidad de datos totales usados para cada modelo correspondiente.

Intervalo de días, identificación	Datos del modelo	Datos a validar	Datos totales
6 días	8,614	1,440	10,054
11 días	15,637	1,440	17,077
16 días	22,837	1,440	24,277
31 días	42,937	1,440	44,377
59 días	83,181	1,440	84,621
85 días	119,502	1,440	120,941

Tabla 3.3 Datos para la identificación del tercer caso de estudio

Para este caso de estudio los modelos propuestos en la tabla 3.3 no presenta en ninguno de los casos estacionariedad, siguiendo la metodología de Box-Jenkins y para cada modelo requiere la aplicación de una diferenciación para hacer cada modelo estacionario, de igual manera se grafican las funciones de autocorrelacion y autocorrelacion parcial para determinar el grado de los modelos autorregresivos y de medias móviles que serán incluidos en los posibles modelos ARIMA.

Finalmente se postulan todos los posibles modelos matemáticos que pueden ser evaluados y estimados por el proceso de mínimos cuadrados de Gauss, esto con el fin de minimizar la suma de cuadrados de los residuos, de igual manera mejora el valor de la función objetivo y se produce un proceso de iteración hasta que se alcanza cada criterio de convergencia.

$$\text{ARIMA 6 días (1,1,1): } y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.28)$$

ARIMA(1,1,1) Model 6 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	1.43691e-05	0.0106042	0.00135504
AR{1}	0.694781	0.00114176	608.516
MA{1}	-0.862361	0.000958903	-899.321
Variance	644.65	0.405461	1589.92

$$\text{ARIMA 11 días (1,1,1): } y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.29)$$

ARIMA(1,1,1) Model 11 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	-8.46183e-06	0.0342035	-0.000247396
AR{1}	0.712681	0.00351309	202.864
MA{1}	-0.901472	0.00263092	-342.645

$$\text{ARIMA 16 días (1,1,1): } y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.30)$$

ARIMA(1,1,1) Model 16 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	4.98211e-06	0.0202694	0.000245795
AR{1}	0.703569	0.00211624	332.462
MA{1}	-0.883946	0.00167585	-527.46
Variance	1159.89	1.54905	748.776

$$\text{ARIMA 31 días (1,1,1): } y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.31)$$

ARIMA(1,1,1) Model 31 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	4.98211e-06	0.0202694	0.000245795
AR{1}	0.703569	0.00211624	332.462
MA{1}	-0.883946	0.00167585	-527.46
Variance	1159.89	1.54905	748.776

$$\text{ARIMA 59 días (1,1,1): } y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.32)$$

ARIMA(1,1,1) Model 59 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	8.66993e-06	0.0132427	0.000654694
AR{1}	0.704102	0.00141469	497.707
MA{1}	-0.876837	0.00114808	-763.742
Variance	861.055	0.735753	1170.3

$$\text{ARIMA 85 días (1,1,1): } y_t = \delta + \phi_1 y_{t-1} + \varepsilon_t + \theta_1 a_{t-1} \quad (3.33)$$

ARIMA(1,1,1) Model 85 dias:

Conditional Probability Distribution: Gaussian

Parameter	Value	Standard Error	t Statistic
Constant	1.43691e-05	0.0106042	0.00135504
AR{1}	0.694781	0.00114176	608.516
MA{1}	-0.862361	0.000958903	-899.321
Variance	644.65	0.405461	1589.92

Para casos posteriores de estudio, se proponen varios modelos con diferentes intervalos de días para la identificación del sistema, los datos que fueron tomados para la evaluación de cada modelo, fueron tomados a diferentes intervalos de tiempo 1 y 5 minutos, proporcionados por el Cinvestav & The University of Texas at Austin. Los resultados obtenidos se asemejan a los resultados estudiados en dicho documento y anexados en el apartado D del documento.

4. Resultados y Conclusiones

4.1 Resultados

Para el primer modelo de estudio, se comparan los 9 modelos ARIMA propuestos con 1440 datos pronosticados y 1440 datos disponibles de validación del día 23 de enero del 2015, se observa en la tabla 4.1 los índices de errores que son calculados con las ecuaciones (3.12, 3.13, 3.14), con el fin de definir el modelo que se aproxime al valor real medido.

Evaluación Modelo (ARIMA)	MAPE (%)	MAE ($\frac{m}{s}$)	RMSE ($\frac{m}{s}$)
2,1,2	16.5750	22.4392	73.9042
2,1,3	16.5906	22.4408	73.9323
2,1,4	16.5776	22.3366	73.7844
3,1,2	16.5817	22.4341	73.9184
3,1,3	15.6368	22.0537	73.8503
3,1,4	16.5405	22.3250	73.2549
4,1,2	16.6003	22.4452	73.9460
4,1,3	16.5276	22.3498	73.2171
4,1,4	17.9440	22.6685	73.5420

Tabla 4.1 Índices de errores de los modelos ARIMA

Para la evaluación del desempeño predictivo de la serie de tiempo, se comparan los valores pronosticados y los valores medidos, se observa que el modelo que presenta un menor índice de error es el modelo ARIMA (3,1,3) en el Anexo A, se observa la comparación de datos entre la medición real y la medición pronosticada con un error del 15.6368% del error absoluto medio. En las figuras 4.1 y 4.2 se muestran las gráficas de los resultados de predicción del modelo seleccionado.

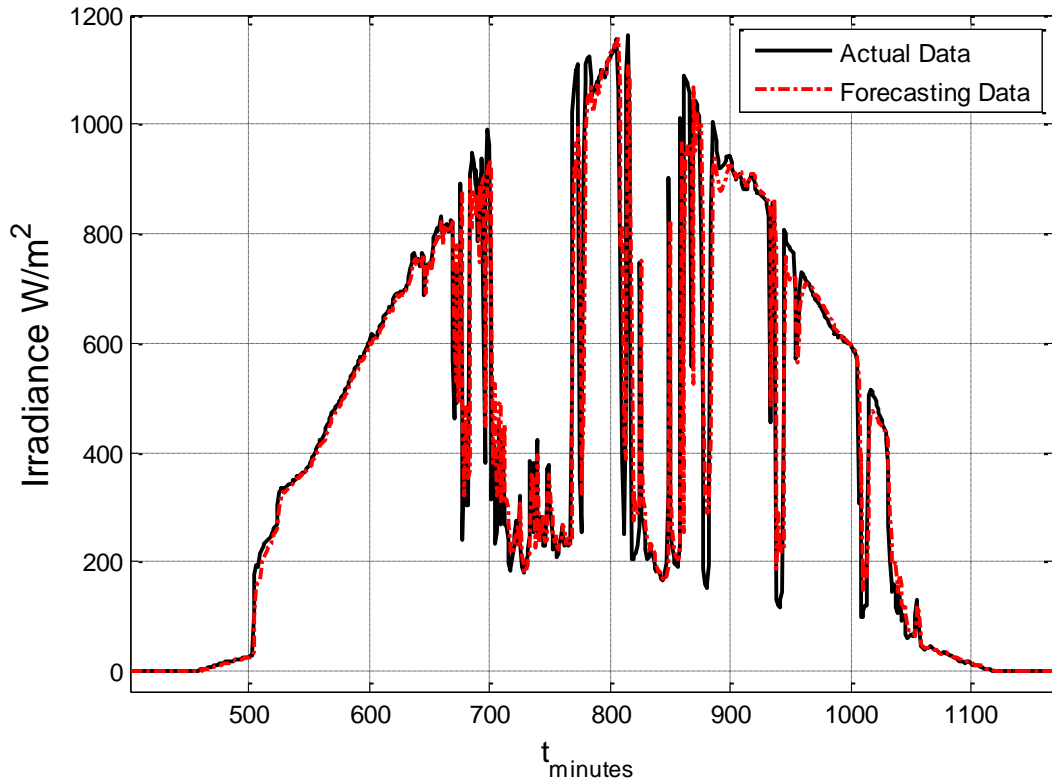


Figura 4.1 Modelo ARIMA (3,1,3), comparación de resultados

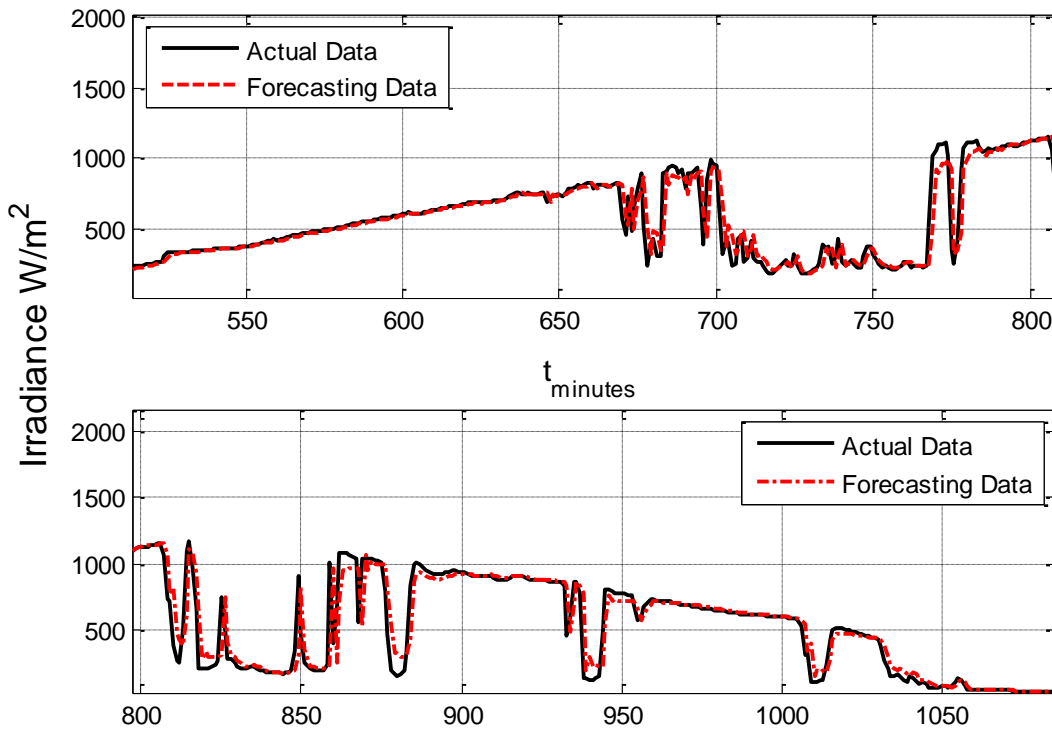


Figura 4.2 Modelo ARIMA(3,1,3), validación de resultados

La figura 4.3 (a,b) nos muestra el modelo ARIMA (3,1,3) con una modificación de un intervalo de confianza del valor pronosticado, de la cual se observa que a partir del valor pronosticado se calculan los valores máximos y mínimos de predicción del 95% de confianza, lo que permite validar que la medición pronosticada se encuentra dentro de los parámetros adecuados del valor real medido.

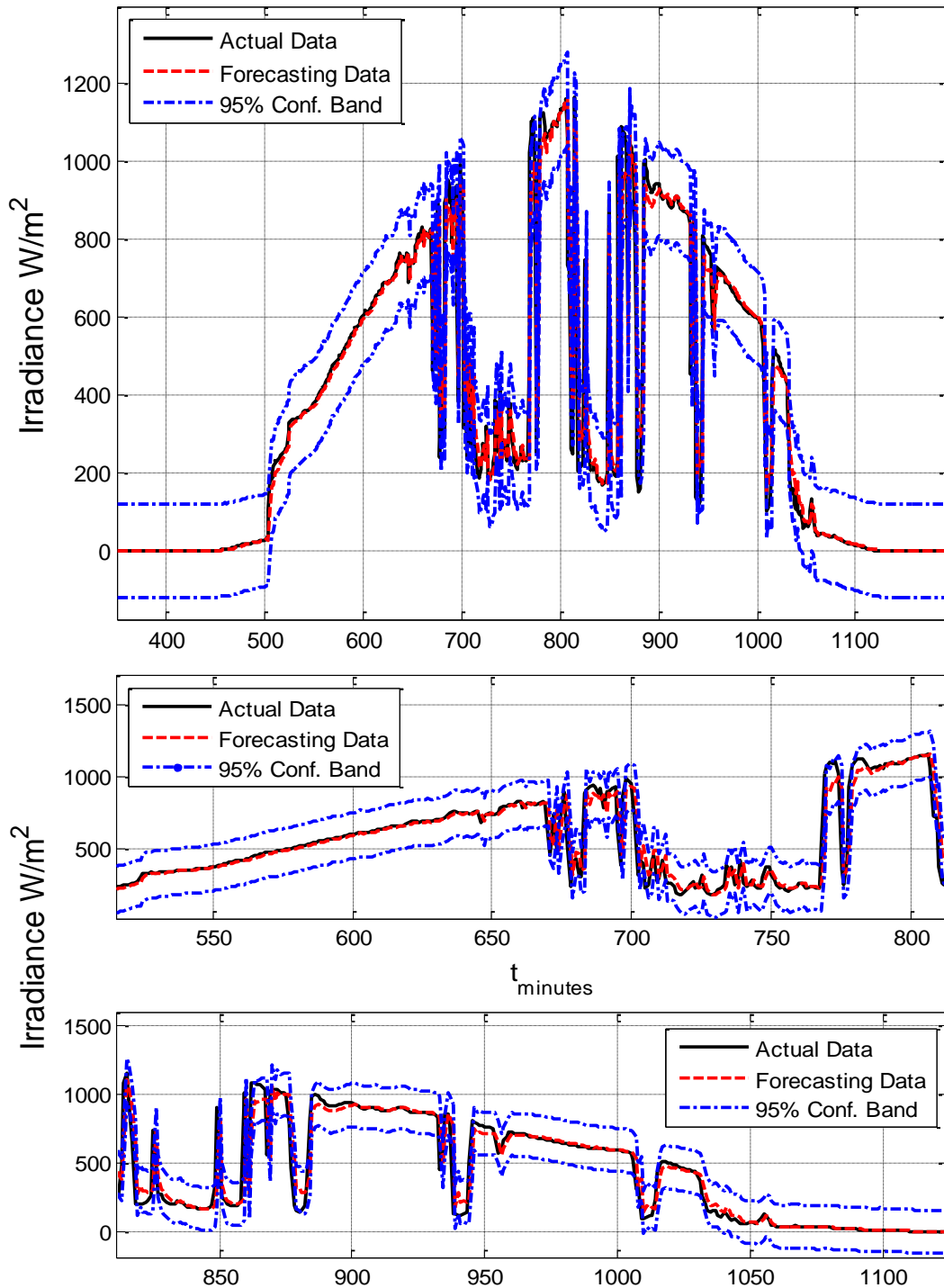


Figura 4.3 (a,b) Modelo ARIMA(3,1,3), Con intervalo del 95% de confianza

Para el segundo modelo de estudio, se comparan los modelos ARIMA propuestos en la tabla 3.2, con 1440 datos pronosticados y 1440 datos disponibles de validación del día 05 de junio del 2013. En la tabla 4.2 se muestran los principales índices de error para cada modelo los cuales son calculados con las ecuaciones (3.12, 3.13, 3.14), con el fin de definir el modelo adecuado que se aproxime al valor real medido.

Evaluación Modelo	MAPE (%)	MAE ($\frac{m}{s}$)	RMSE ($\frac{m}{s}$)
5 días	8.3795	10.2400	21.2944
10 días	8.0636	10.7386	21.7629
20 días	7.8269	9.9229	21.0819
35 días	7.2810	9.1945	20.6393

Tabla 4.2 Índices de errores de los modelos relevantes ARIMA

Para la evaluación del desempeño predictivo de la serie de tiempo, se comparan los valores pronosticados y medidos, se observa que el modelo que presenta un menor índice de error es el modelo de 35 días. En las figuras 4.4 y 4.5 se muestran las gráficas de los resultados de predicción del modelo seleccionado. En el anexo B, se observa la comparación de datos entre la medición real y la medición pronosticada con un error del 7.2810% del error absoluto medio.

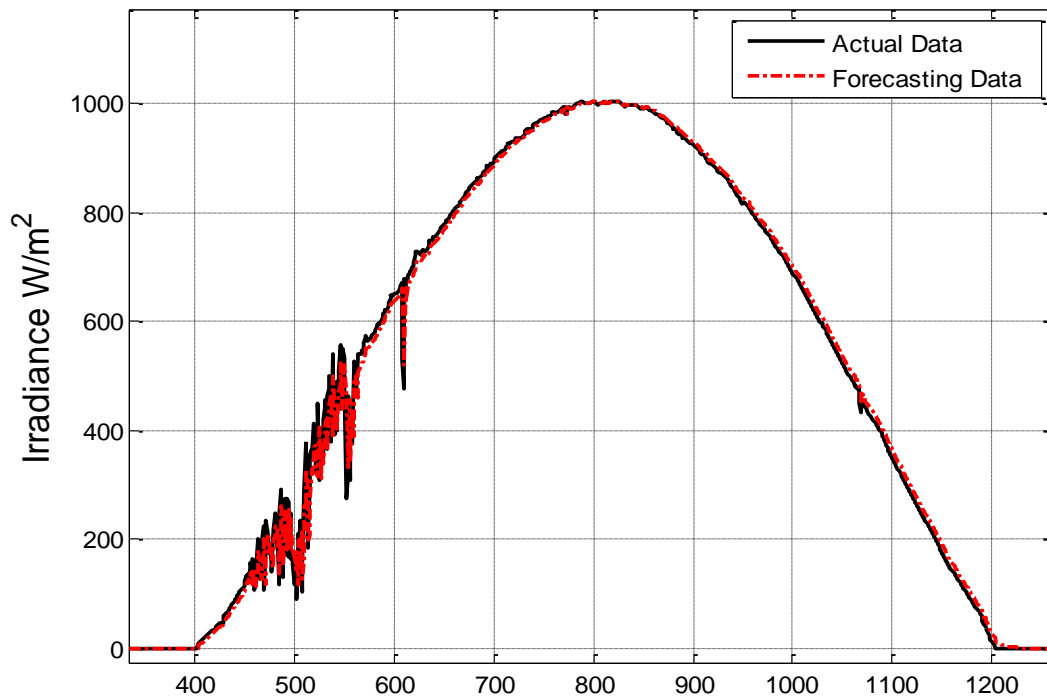


Figura 4.4 Modelo 62 días ARIMA (1,1,1), comparación de resultados

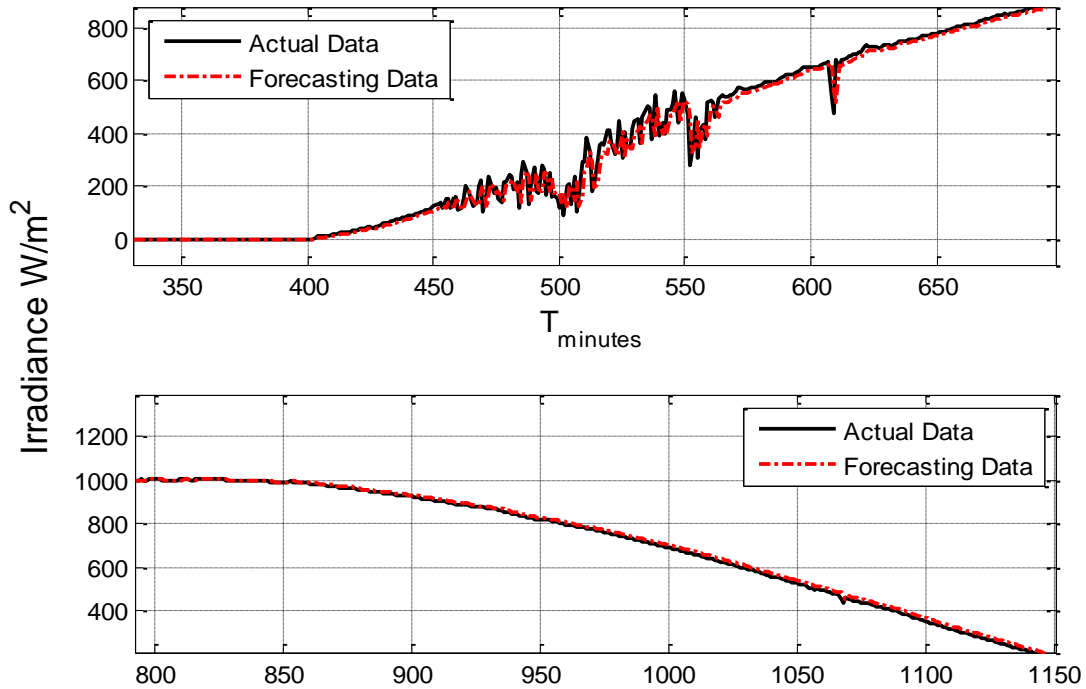
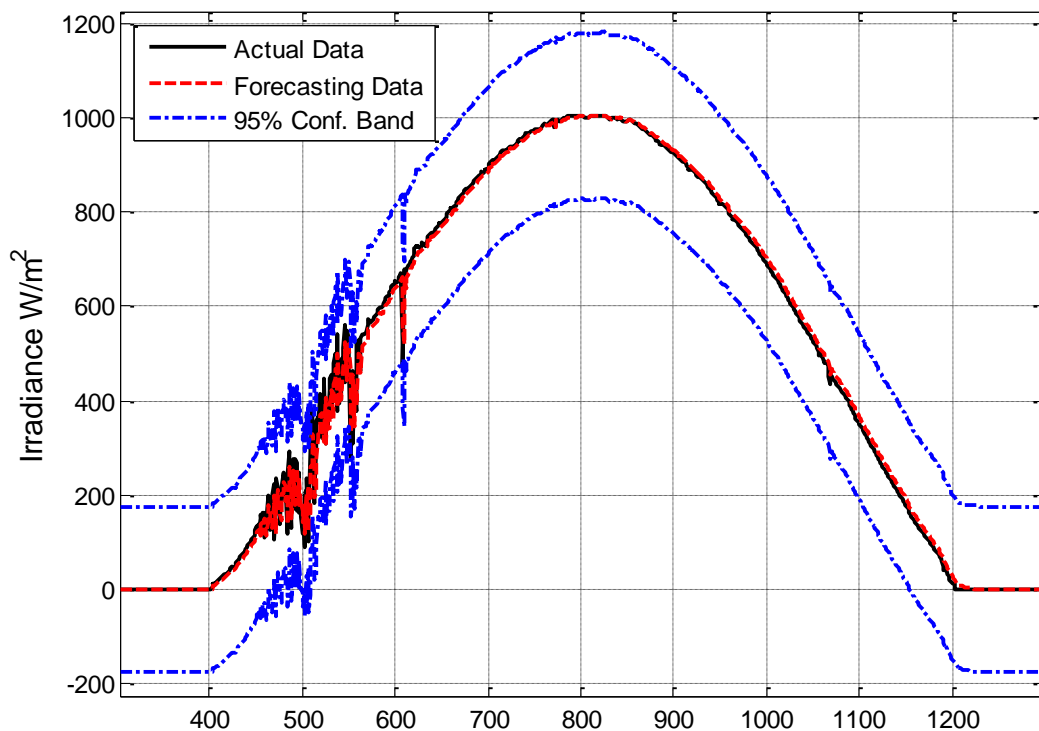


Figura 4.5 Modelo 62 días ARIMA (1,1,1), validación de resultados

La figura 4.6 (a,b) nos muestra el modelo de 35 días con una modificación de un intervalo de confianza del valor pronosticado, de la cual se observa que a partir del valor pronosticado se calculan los valores máximos y mínimos de predicción del 95% de confianza, lo que permite validar que la medición pronosticada se encuentra dentro de los parámetros adecuados del valor real medido.



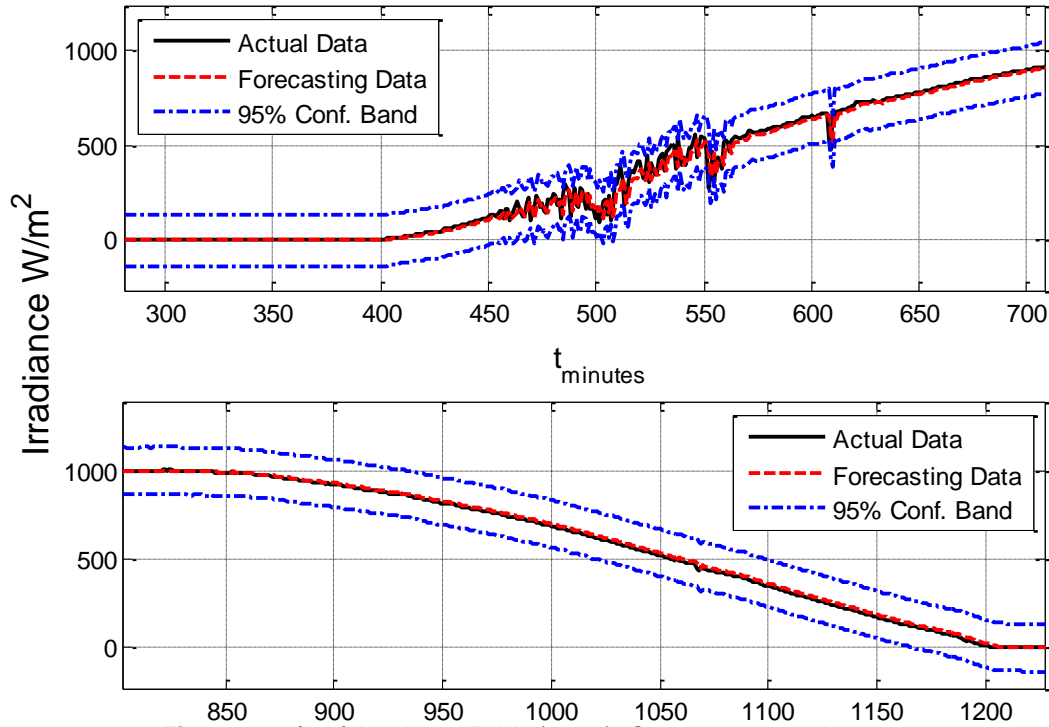


Figura 4.6 (a,b) Modelo ARIMA(1,1,1), Con intervalo del 95%

Para el tercer modelo de estudio, se comparan los modelos ARIMA propuestos en la tabla 3.3, con 1440 datos pronosticados y 1440 datos disponibles de validación del día 31 de marzo del 2014. En la tabla 4.3 se muestran los principales índices de error para cada modelo calculados con las ecuaciones (3.12, 3.13, 3.14), con el fin de definir el modelo adecuado que se aproxime al valor real medido.

Evaluación		MAPE (%)	MAE ($\frac{m}{s}$)	RMSE ($\frac{m}{s}$)
Modelo				
6 Días	2,1,2	25.2831	44.6987	101.9912
11 Días	1,1,1	24.9431	44.4890	102.6145
16 Días	1,1,1	24.7778	44.4265	102.9587
31 Días	1,1,1	24.6887	44.3813	102.9421
59 Días	1,1,1	24.5742	44.3393	103.1792
85 Días	1,1,1	24.0467	44.3115	103.4388

Tabla 4.2 Índices de errores de los modelos relevantes ARIMA

Para la evaluación del desempeño predictivo de la serie de tiempo, se comparan los valores pronosticados y medidos, se observa que el modelo que presenta un menor índice de error es el modelo de 85 días. En las figuras 4.7 y 4.8 se muestran las gráficas del modelo seleccionado. En el anexo C, se observa la comparación de datos entre la medición real y la medición pronosticada con un error del 24.04% del error absoluto medio.

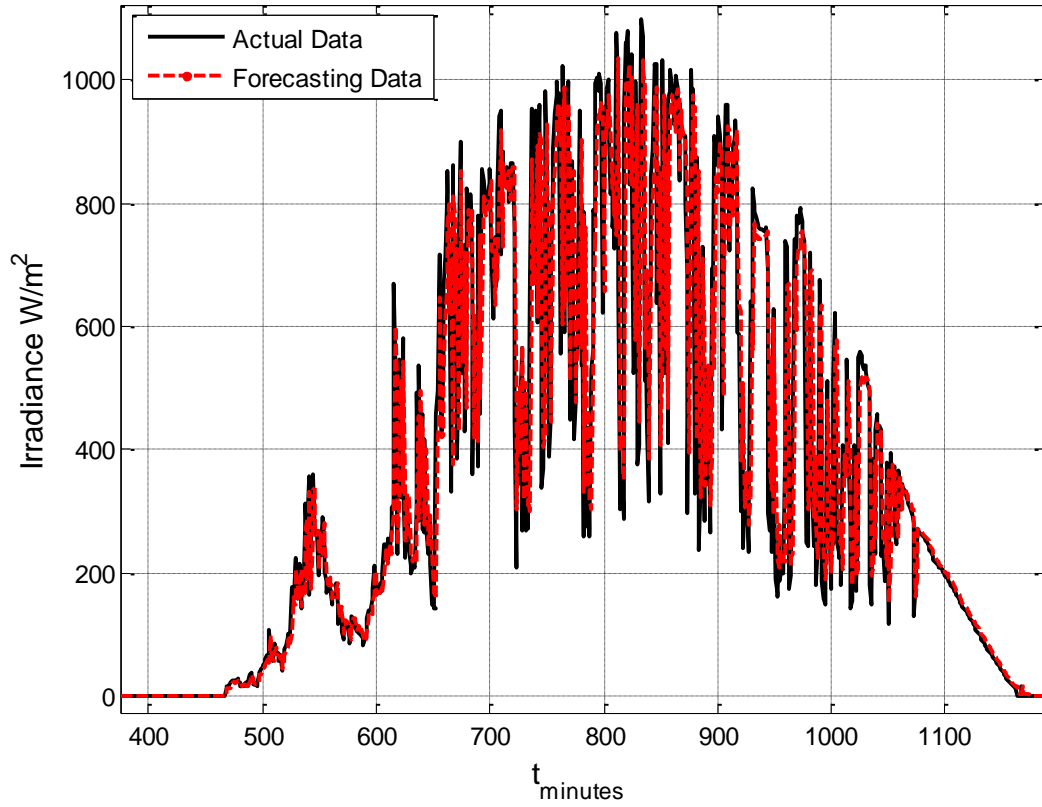


Figura 4.7 Modelo 85 días ARIMA (1,1,1), comparación de resultados

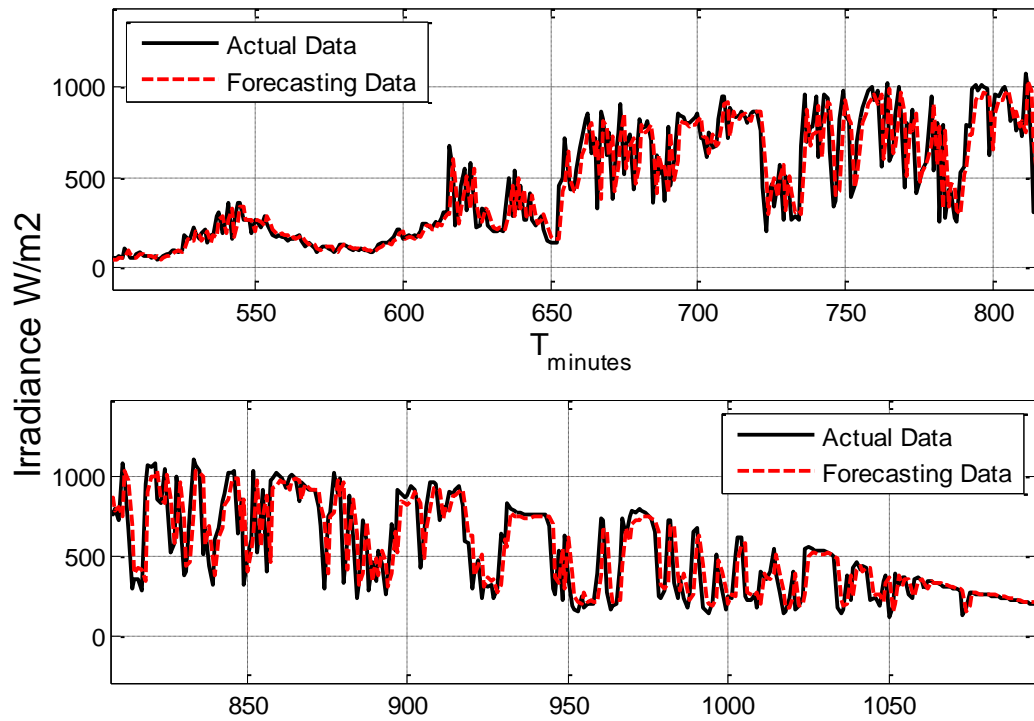


Figura 4.8 Modelo 85 días ARIMA (1,1,1), validación de resultados

La figura 4.9 nos muestra el modelo de 85 días con una modificación con una modificación de un intervalo de confianza del valor pronosticado, de la cual se observa que a partir del valor pronosticado se calculan los valores máximos y mínimos de predicción del 95% de confianza, lo que permite validar que la medición pronosticada se encuentra dentro de los parámetros adecuados del valor real medido.

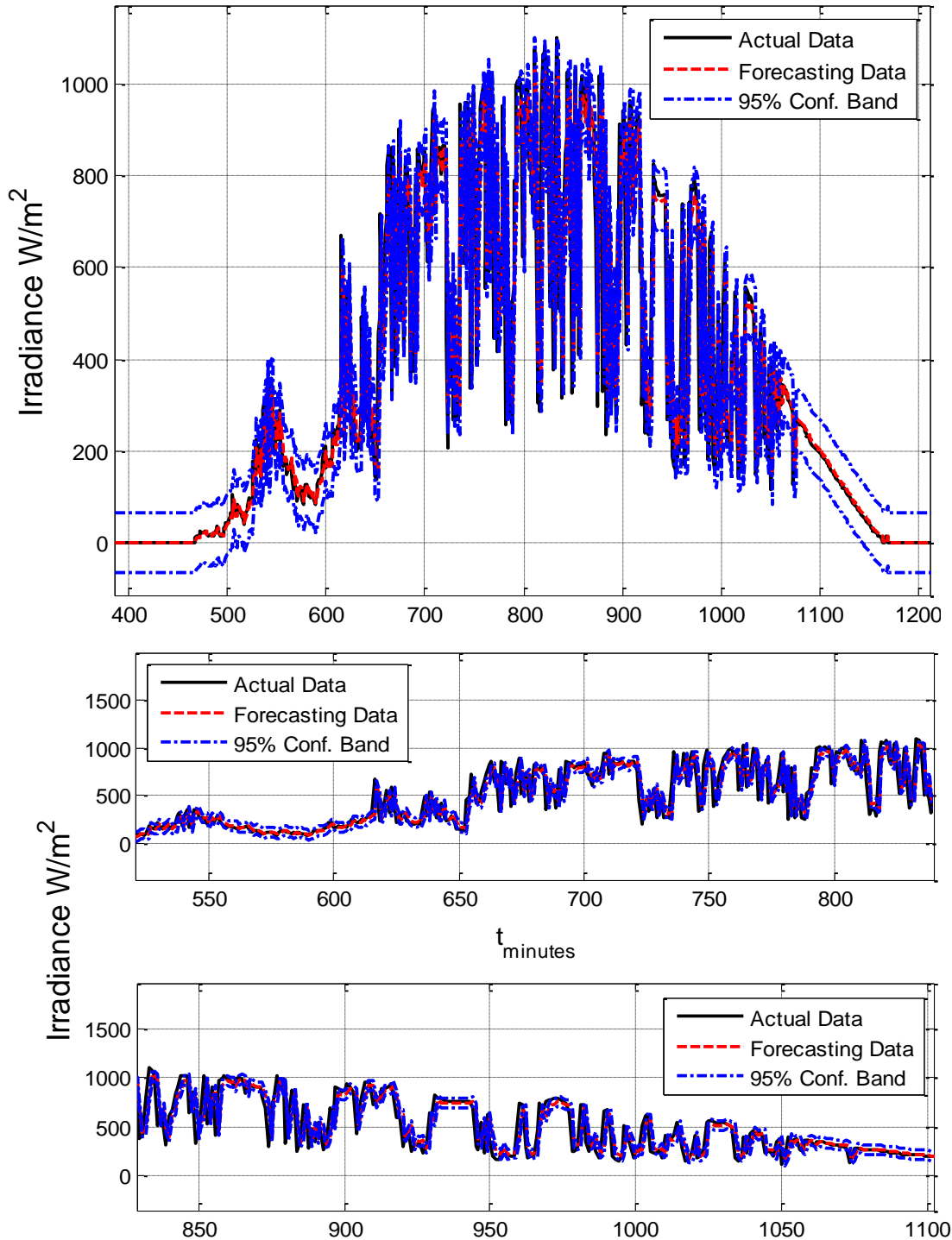


Figura 4.9 (a,b) Modelo 85 días ARIMA(1,1,1), Con intervalo del 95% de confianza

4.2 Conclusiones

Con base en los resultados de la investigación realizada y al diseño de la metodología propuesta para el pronóstico del recurso fotovoltaico, se logró generar un modelo adecuado que describe el comportamiento de las predicciones al igual que su paso temporal, particularmente de las series de tiempo estudiadas, de esta manera se pueden utilizar los datos de predicción y estimación para valores futuros.

La teoría es basada en la metodología de Box-Jenkins que permite describir la relación entre distintas variables en el pasado y adaptarse para las predicciones futuras. En la investigación se realizaron experimentaciones con diferentes muestras de irradiancia para el modelado, entre ellas 6, 15, 30, 60 y 80 días, la métrica usada es la del error porcentual promedio para cada caso de estudio.

Los resultados se analizan conforme se da el proceso de predicción y al mismo tiempo se comparan con los valores reales para observar el paso temporal que se va generando, se observa que las mediciones pronosticadas tienden a no alcanzar los valores mínimos y máximos de los valores reales medidos, con lo que se va generando el error del pronóstico para cada modelo.

Los modelos propuestos demostraron un resultado variante de acuerdo al tiempo y en la toma de las mediciones dependiendo del tamaño de muestras seleccionadas, los óptimos resultados se obtuvieron con muestras mayores a los 30 días en la selección del modelo y con intervalos de muestreo de uno a cinco minutos, de igual manera se calculan los intervalos de confiabilidad, esto con el fin de otorgar un resultado más aproximado al valor real medido del 95%, esta característica es fundamental en la teoría de la metodología Box-Jenkins y limitante para otras técnicas de predicción.

La investigación realizada abre posibilidades para trabajos futuros, entre ellos se propone un modelo híbrido que combine la metodología de Box-Jenkins junto a otras técnicas descritas en el trabajo (filtros de kalman y Predicción bayesiana), esto con el fin de obtener mejores resultados y un modelamiento que describa el óptimo comportamiento de las predicciones a futuro.

Así mismo la obtención de mediciones reales de potencia extraídas del inversor para observar y comparar el comportamiento de la producción de energía, desde dos perspectivas la predicha a partir de irradiancia y la predicha a partir de la misma potencia generada.

Referencias bibliográficas

- [1]. ZHenzhou Peng, Shinjae Yoo, Dantong Yu, Dong Huang, “Solar irradiance forecast system based on geostationary satellite”, IEEE Smart Grid Comm Symposium, p.p. 708-713, Oct. 2013
- [2]. Gonzalo Casaravilla, Rubén Chaer, “Modelo de Series de Tiempo de Radiación Solar para Simulador de Instalaciones Fotovoltaicas”, IEEE, 7º encuentro de Energía, Potencia, Instrumentación y Medidas, Montevideo Uruguay, pp. 88 – 90, Oct. 2008
- [3]. Mellit A, Pavan AM. A “24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at trieste, Italy.” Solar Energy, Volume 84, Issue 5, Pages 807–821, May 2010.
- [4]. Pérez R, Kivalov S, Schlemmer J, “Evaluation of numerical weather prediction solar irradiance forecast in the US”, American solar energy society, Proc. ASES Annual Conference, March 2011.
- [5]. Eugen Diacosnescu, “The use of NARX Neural Networks to predict Chaotic Time Series”, WSEAS transactions on computer research, Volume 3, Issue 3, pages 182-190, March 2008
- [6]. Londhe, Vijay Panchang, “One-Day Wave Forecasts Based on Artificial Neural Networks”, Department of Maritime Systems Engineering, Texas A&M University, Volume 23, Issue 11, Galveston, Texas, pages, 1593-1603, November 2006
- [7]. Iosu Zamarbide Ducun, “Predicción de radiación solar a corto y medio plazo”, Universidad Pública de Navarra, Departamento de Ingeniería Eléctrica, Tesis Maestría, Pamplona, Junio de 2014
- [8]. Roberto Román Diez, “Efectos de los componentes atmosféricos sobre la radiación eritemática y modelo propuesto para la reconstrucción de series”, Laboratorio de Atmósfera y Energía, Departamento de Física Aplicada, Universidad de Valladolid, Julio 2010
- [9]. Iosu Zamarbide Ducun, “Predicción de radiación solar a corto y medio plazo”, Universidad Pública de Navarra, Departamento de Ingeniería Eléctrica, Tesis Maestría, Pamplona, Junio de 2014
- [10]. David K. Hildebrand, Lyman Ott. “Estadística aplicada a la administración y a la economía. Serie de probabilidad y estadística”, Pearson Educación 3ra. Edición, 1998

- [11]. Juan Carlos Martínez, Stephen G. Walker, “Modelación de Series de Tiempo, Estacionariedad y Métodos Bayesianos No Paramétricos”, Documentos de Investigación, Banco de México, Working Papers, Septiembre 2011
- [12]. Blake, Andrew, “State-Space Models and the Kalman Filter: Application, Formulation and Estimation”, HKMA/CCBS May 2004
- [13]. Ramos C, Daniel J, Karelys Del V.- “Aplicaciones estadísticas en la administración y la economía”, Universidad de oriente, Departamento de contaduría, Tesis licenciatura, Cumaná, Mayo de 2009
- [14]. Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. “Time Series Analysis: Forecasting and Control”. 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1994.
- [15]. Box, G. E. and G. M. Jenkins. “Time Series Analysis: Forecasting and Control”. San Fransisco: Holden-Day. Revised Edition
- [16]. Garrick R. Small, Raymond Wong, “The Validity of Forecasting”, University of Technology, Sydney, International Conference Christchurch, August 2002
- [17]. Norman Giraldo Gómez, “Curso de Procesos Estocásticos Trabajo sobre Identificación y Estimación de Señales ARMA”, Escuela de Estadística. Universidad Nacional de Colombia, Octubre 2007
- [18]. George e. P. Box, Gwilym m. Jenkins, Gregory c. Reinsel, “Time Series Analysis, Forecasting and Control”, Fourth edition
- [19]. Tsan-Ming Choi, Chi-Leung Hui, Yong Yu, “Intelligent Fashion Forecasting Systems: Models and Applications”, Springer-Verlag Berlin Heidelberg 2014
- [20]. Ajoy K. Palit and Dobrivoje Popovic, “Computational Intelligence in Time Series Forecasting”, Theory and Engineering Applications, Springer-Verlag London Limited 2005
- [21]. Viorel Badescu, “Modeling Solar Radiation at the Earth’s Surface”, Recent Advances, 2008 Springer-Verlag Berlin Heidelberg

Anexos

Anexo A

Primer modelo de estudio propuesto con 1440 datos pronosticados y 1440 datos disponibles de validación del día 23 de enero del 2015. En la tabla A se observan los valores relevantes de la medición real vs la medición pronosticada.

Modelo 1		Medición Real Irradiancia ($\frac{W}{m^2}$)	Medición Pronosticada Irradiancia ($\frac{W}{m^2}$)
Fecha y Hora			
800	23/01/15 13:19:59	1129	1132.64177712705
801	23/01/15 13:20:59	1132	1134.02661002063
802	23/01/15 13:21:59	1131	1133.94677397996
803	23/01/15 13:22:59	1142	1144.63973118628
804	23/01/15 13:23:59	1145	1145.66826278465
805	23/01/15 13:24:59	1148	1149.21625813495
806	23/01/15 13:25:59	1158	1158.10011300047
807	23/01/15 13:26:59	1065	1069.80339979741
808	23/01/15 13:27:59	733	777.008904329374
809	23/01/15 13:28:59	717	812.769775736556
810	23/01/15 13:29:59	381	469.421196602901
811	23/01/15 13:30:59	207	421.179744873434
812	23/01/15 13:31:59	228	387.506074433260
813	23/01/15 13:32:59	497	612.090615939749
814	23/01/15 13:33:59	1091	1106.11511387453
815	23/01/15 13:34:59	1163	1078.84171718236
MAPE: 15.6368			

Tabla A. Medición real vs la medición pronosticada

Anexo B

Segundo modelo de estudio propuesto con 240 datos pronosticados y 240 datos disponibles de validación del día 18 de mayo del 2013. En la tabla B se observan los valores relevantes de la medición real vs la medición pronosticada.

Modelo 2		Medición Real	Medición Pronosticada
Fecha y Hora		Irradiancia ($\frac{W}{m^2}$)	Irradiancia ($\frac{W}{m^2}$)
510	06/05/2013 08:30	296.2	264.843013520018
511	06/05/2013 08:31	381.3	328.402799927246
512	06/05/2013 08:32	327.2	294.643653669662
513	06/05/2013 08:33	183.6	195.191896264311
514	06/05/2013 08:34	203.5	208.327755610882
515	06/05/2013 08:35	297.5	274.933020063773
516	06/05/2013 08:36	357	319.431286515017
517	06/05/2013 08:37	364.8	328.423527013991
518	06/05/2013 08:38	354.1	324.111771093931
519	06/05/2013 08:39	412.5	368.502774680553
520	06/05/2013 08:40	413	372.874978291313
521	06/05/2013 08:41	363.9	341.516418713176
522	06/05/2013 08:42	323	314.387326051859
523	06/05/2013 08:43	351.1	335.215873531729
524	06/05/2013 08:44	449.1	406.564463563443
525	06/05/2013 08:45	309.4	310.805165065580
MAPE: 7.2810			

Tabla B. Medición real vs la medición pronosticada

Anexo C

Tercer modelo de estudio propuesto con 1440 datos pronosticados y 1440 datos disponibles de validación del día 31 de marzo del 2014. En la tabla C se observan los valores relevantes de la medición real vs la medición pronosticada.

Modelo 3		Medición Real	Medición Pronosticada
Fecha y Hora		Irradiancia ($\frac{W}{m^2}$)	Irradiancia ($\frac{W}{m^2}$)
750	3/31/14 12:29:54	981.98	928.022523324763
751	3/31/14 12:30:54	842.27	819.151849847403
752	3/31/14 12:31:54	810.47	795.862875486687
753	3/31/14 12:32:54	387.42	445.718358353897
754	3/31/14 12:33:54	456.89	495.522447930122
755	3/31/14 12:34:54	696.71	689.835972946002
756	3/31/14 12:35:54	832.43	803.758081949024
757	3/31/14 12:36:54	886.97	853.104612772629
758	3/31/14 12:37:54	947.83	908.426855913886
759	3/31/14 12:38:54	977.15	938.256799446109
760	3/31/14 12:39:54	999.1	961.881623972853
761	3/31/14 12:40:54	921.31	902.250415267219
762	3/31/14 12:41:54	979.29	953.137441223178
763	3/31/14 12:42:54	554.47	603.108638607633
764	3/31/14 12:43:54	796.19	797.626522893538
765	3/31/14 12:44:54	1024.28	987.295361501267
MAPE: 24.0467			

Tabla C. Medición real vs la medición pronosticada

Anexo D

Casos posteriores de estudio, se proponen varios modelos con diferentes intervalos de días para la identificación del sistema, los datos que fueron tomados para la evaluación de cada modelo, fueron tomados a diferentes intervalos de tiempo y proporcionados por el Cinvestav & The University of Texas at Austin.

Modelo 01		Días Modelo: 4 días		Tiempo de muestreo: 1 minuto	
19 /01/15 al 23/01/15		Modelo ARIMA 1,1,1			
Datos del Modelo:	5169	Mape (%)	15.6360		
Datos a Validar	1440	Mae $\left(\frac{m}{s}\right)$	22.0537		
Datos totales	6609	Rmse $\left(\frac{m}{s}\right)$	73.8503		

Tabla D I. Representación de Resultados obtenidos, Modelo 4 días, Datos Cinvestav

Modelo 02		Días Modelo: 35 días		Tiempo de muestreo: 1 minuto	
15 /03/14 al 31/03/14		Modelo ARIMA 1,1,1			
Datos del Modelo:	50399	Mape (%)	7.2810		
Datos a Validar	1440	Mae $\left(\frac{m}{s}\right)$	9.1945		
Datos totales	51838	Rmse $\left(\frac{m}{s}\right)$	20.6393		

Tabla D II. Representación de Resultados obtenidos, Modelo 35 días, Datos University of Texas at Austin

Modelo		Días Modelo: 85 días		Tiempo de muestreo: 1 minutos	
11 /03/13 al 28/05/13		Modelo ARIMA 1,1,1			
Datos del Modelo:	119502	Mape (%)	24.0467		
Datos a Validar	1440	Mae $\left(\frac{m}{s}\right)$	44.3115		
Datos totales	120941	Rmse $\left(\frac{m}{s}\right)$	103.4388		

Tabla D III. Representación de Resultados obtenidos, Modelo 4 días, Datos Cinvestav

Modelo 04		Días Modelo: 29 días	Tiempo de muestreo: 5 minutos	
01/05/13 al 30/05/15		Modelo ARIMA 1,1,1		
Datos del Modelo:	6961	Mape (%)	10.2390	
Datos a Validar	240	Mae ($\frac{m}{s}$)	7.3388	
Datos totales	7200	Rmse ($\frac{m}{s}$)	25.4566	

Tabla D IV. Representación de Resultados obtenidos, Modelo 29 días, Datos Cinvestav

Modelo 05		Días Modelo: 96 días	Tiempo de muestreo: 5 minutos	
01/06/13 al 04/09/13		Modelo ARIMA 1,1,1		
Datos del Modelo:	22800	Mape (%)	12.1037	
Datos a Validar	240	Mae ($\frac{m}{s}$)	10.0535	
Datos totales	23040	Rmse ($\frac{m}{s}$)	19.3996	

Tabla D V. Representación de Resultados obtenidos, Modelo 96 días, Datos Cinvestav

Modelo 06		Días Modelo: 60 días	Tiempo de muestreo: 1 minuto	
01 /06/13 al 31/07/13		Modelo ARIMA 1,1,1		
Datos del Modelo:	86399	Mape (%)	4.4993	
Datos a Validar	1440	Mae ($\frac{m}{s}$)	4.1630	
Datos totales	87838	Rmse ($\frac{m}{s}$)	6.5011	

Tabla D VI. Representación de Resultados obtenidos, Modelo 60 días, Datos University of Texas at Austin

Modelo 07		Días Modelo: 88 días	Tiempo de muestreo: 5 minuto	
01/03/13 al 38/05/13		Modelo ARIMA 1,1,1		
Datos del Modelo:	13680	Mape (%)	14.0727	
Datos a Validar	240	Mae $\left(\frac{m}{s}\right)$	18.6669	
Datos totales	13920	Rmse $\left(\frac{m}{s}\right)$	57.1360	

Tabla D VII. Representación de Resultados obtenidos, Modelo 88 días, Datos Cinvestav

Modelo 08		Días Modelo: 62 días	Tiempo de muestreo: 5 minutos	
01/04/13 al 18/05/13		Modelo ARIMA 1,1,1		
Datos del Modelo:	14639	Mape (%)	20.0107	
Datos a Validar	240	Mae $\left(\frac{m}{s}\right)$	29.8097	
Datos totales	14879	Rmse $\left(\frac{m}{s}\right)$	76.3706	

Tabla D VIII. Representación de Resultados obtenidos, Modelo 62 días, Datos Cinvestav