



---

---

**UNIVERSIDAD AUTÓNOMA DE YUCATÁN**  
**FACULTAD DE INGENIERÍA**  
**UNIDAD DE POSGRADO E INVESTIGACIÓN**

**“ESTIMACIÓN DEL POTENCIAL EÓLICO LOCAL A  
CORTO PLAZO, MEDIANTE SERIES TEMPORALES,  
EN LA ZONA NORTE DE LA CIUDAD DE MÉRIDA,  
YUCATÁN”**

**TESIS**

**PRESENTADA POR:**

**I. F. IRVING SANTIAGO PEÑA QUINTAL**

**EN OPCIÓN AL GRADO DE**

**MAESTRO EN INGENIERÍA**

**OPCIÓN ENERGÍAS RENOVABLES**

**MÉRIDA, YUCATÁN, MÉXICO**

**2016**

## ÍNDICE.

Prologo.....	6
Introducción.....	7
Objetivos.....	8
Capítulo 1: Potencial eólico.....	9
1.1 México, demanda energética y el viento.....	9
1.2 Alcance geográfico del estudio.....	13
1.3 Métodos convencionales para la determinación del potencial eólico.....	15
1.4 Estudios estadísticos y de dispersión empleados en la determinación del potencial eólico.....	19
Capítulo 2: Modelos meteorológicos y computacionales.....	23
2.1 Clasificación de los modelos meteorológicos.....	23
2.1.1 Modelos unidimensionales.....	25
2.1.2 Modelos bidimensionales.....	26
2.1.3 Modelos tridimensionales.....	27
2.2 Métodos computacionales para la implementación de modelos en determinación del clima.....	29
2.3 Técnicas de predicción (Redes neuronales).....	34
2.4 Relación entre los modelos de clima y la determinación del potencial eólico.....	37
Capítulo 3: Predicción de series de tiempo de velocidad del viento.....	39
3.1 Series de tiempo.....	39
3.2 Modelos autorregresivos y de media móvil.....	43
3.3 Algoritmo de predicción.....	49
Capítulo 4: Confrontación de resultados, estimaciones versus mediciones.....	53
4.1 Datos de estaciones meteorológicas.....	53
4.2 Primeras estimaciones.....	56
4.3 Análisis estadístico.....	62
Capítulo 5: Análisis de estimaciones, predicciones y resultados.....	65
5.1 Retroalimentación de modelo y corrección.....	65
5.2 Análisis de datos.....	66
5.3 Análisis de resultados finales.....	69
Conclusiones.....	72
Referencias.....	73
Anexos.....	75

## ÍNDICE DE FIGURAS.

Figura 1.1.- Gráfica de cantidad de usuarios de energía eléctrica, demanda energética.....	10
Figura 1.2.- Capacidad efectiva de generación.....	11
Figura 1.3.- Mapa velocidades de viento de México.....	12
Figura 1.4.- Zonas potenciales para el desarrollo de proyectos de energía eólica....	12
Figura 1.5.- Localización geográfica del estudio.....	14
Figura 1.6.- Estación climatológica en emplazamiento.....	16
Figura 1.7.- Distribución de Weibull, típica de los estudios de potencial eólico.....	17
Figura 2.1.- Sistema climatológico típico y sus componentes.....	24
Figura 2.2.- Ejemplo de rejilla empleada en la parametrización y modelación de un sistema climático en una región.....	28
Figura 2.3.- Diagrama de operación de NOAAPORT.....	31
Figura 2.4.- Diagrama de un perceptrón.....	35
Figura 2.5.- Entradas y salidas de datos en una red neuronal artificial.....	36
Figura 3.1.- Ejemplo de una serie de tiempo, velocidad y dirección del viento a través un periodo.....	39
Figura 3.2.- Esquema de metodología.....	49
Figura 3.3.- Proceso iterativo para la obtención de los coeficientes del modelo ARIMA, en el ejemplo ARIMA (1, 0, 1).....	50
Figura 3.4.- Diagrama de operación para un modelo específico, en este ejemplo ARIMA.....	51
Figura 3.5.- Diagrama de flujo para la aplicación del modelo.....	52
Figura 4.1.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2011.....	53
Figura 4.2.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2012.....	54
Figura 4.3.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2013.....	54
Figura 4.4.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2014.....	55
Figura 4.5.- Distribución de Weibull, Mérida 2011-2014.....	56
Figura 4.6.- Gráfica de predicción para Mérida, empleando los tres modelos con $p=1$ , $d=1$ y $q=1$ .....	58
Figura 4.7.- Gráfica de predicción para Mérida, empleando los tres modelos con $p=1$ , $d=1$ y $q=2$ .....	58
Figura 4.8.- Gráfica de predicción para Mérida, empleando los tres modelos con $p=2$ , $d=1$ y $q=1$ .....	60

Figura 4.9.- Gráfica de predicción para Mérida, empleando los tres modelos con $p=2$ , $d=1$ y $q=2$ .....	60
Figura 4.10.- Gráfica de predicción para Mérida, empleando los tres modelos con $p=1$ , $d=2$ y $q=1$ .....	61
Figura 4.11.- Gráfica de regresión lineal para Mérida, datos medidos vs estimación del modelo ARIMA con coeficientes $q=1$ , $d=1$ y $q=1$ con cuatro años de entrada de datos.....	60
Figura 5.1.- Gráfica de predicción para Mérida, modelo retroalimentado y calibrado ARIMA (1, 1, 1).....	65
Figura 5.2.- Gráfica de predicción para Mérida (1 Enero 2015), modelo retroalimentado y calibrado (factor de corrección 1.09).....	66
Figura 5.3.- Gráfica de predicción para Mérida (2 Enero 2015), modelo retroalimentado y calibrado (factor de corrección 0.935).....	67
Figura 5.4.- Gráfica de predicción para Mérida (3 Enero de 2015), modelo retroalimentado y calibrado (factor de corrección 1.05).....	67
Figura 5.5.- Gráfica de predicción para Mérida (1 Julio 2015), modelo retroalimentado y calibrado (factor corrección final 0.997).....	68
Figura 5.6.- Gráfica de residuos para el análisis estadístico.....	70

## ÍNDICE DE TABLAS.

Tabla 1.1.- Índice de producción de electricidad en Megawatts, por fuente de energía del año 2014.....	11
Tabla 2.1.- Cantidades típicas empleadas de las componentes del sistema atmosférico.....	27
Tabla 3.1.- Características y ejemplos de las componentes de una serie de tiempo...41	
Tabla 4.1.- Supuestos para la aplicación de los modelos autorregresivos y de media móvil (Hernández J. et al, 2013).....	57
Tabla 4.2.- Resultados de análisis estadístico, para la estimación empleando cuatro años como entradas.....	62
Tabla 4.2.- Resultados de la diferencia mínima significativa y grupos mixtos.....	62
Tabla 4.3.- Resultados del análisis de diferencia de medias.....	63
Tabla 4.4.- Tabla de resultados de la regresión lineal de la figura 4.8.....	63
Tabla 5.1.- Resultados de análisis estadístico, para la estimación con el factor de corrección empleando cuatro años como entradas.....	69
Tabla 5.2.- Tabla de modelo del análisis estadístico (promedios de datos).....	69
Tabla 5.3.- Resultados de la diferencia mínima significativa y grupos mixtos.....	70

Tabla 5.4.- Resultados del análisis de diferencia de medias.....	71
--	----

## ÍNDICE DE ECUACIONES.

Ecuación 1.1.- Función de densidad de probabilidad (distribución de Weibull).....	16
Ecuación 1.2.- Potencia teórica viento.....	18
Ecuación 3.1.- Media móvil.....	42
Ecuación 3.2.- Definición de pronóstico.....	42
Ecuación 3.3.- Modelo básico de suavizamiento exponencial.....	43
Ecuación 3.4.- Definición de modelo AR.....	44
Ecuación 3.5.- Polinomio operador de retardos del modelo AR.....	44
Ecuación 3.6.- Definición de modelo AR, alterna.....	45
Ecuación 3.7.- Definición de modelo MA.....	45
Ecuación 3.8.- Polinomio operador de retardos del modelo MA.....	45
Ecuación 3.9.- Definición modelo ARIMA.....	46
Ecuación 3.10.- Polinomio operador de retardos del modelo ARIMA (p, d, q).....	46
Ecuación 3.11.- Modelo ARIMA (p, d, q).....	47
Ecuación 3.12.- Comportamiento en series no estacionarias SARIMA.....	48
Ecuación 3.13.- Definición de Wind Run.....	50

## **PROLOGO.**

El consumo energético es uno de los temas clave en nuestra vida diaria, cada día se emplean millones de litros de combustible fósil para satisfacer una demanda de energía que cada día es más grande. Por ello, recientemente, se han desarrollado una gran cantidad de proyectos e investigaciones relacionadas con las energías renovables. Sin embargo, es bien conocido que las energías renovables, en especial la eólica (obtenida de viento) son de una naturaleza casi caprichosa, errática y poco fiable en términos generales.

Una de las herramientas que mayor desarrollo han tenido son los modelos matemáticos basados en series de tiempo, los cuales son ampliamente utilizados en economía y finanzas para analizar una gran cantidad de información monetaria. Estas herramientas han obtenido mejoras debidas al rápido avance de las herramientas computacionales y el incremento exponencial de la capacidad de cálculo de las computadoras modernas.

Tras muchos años de investigación se sabe que es posible predeterminar, con cierto grado de precisión, la velocidad y dirección del viento de un lugar determinado, esto es gracias a distintas ecuaciones y modelos desarrollados empírica y experimentalmente.

Por lo anterior en esta tesis se propone el empleo de modelos matemáticos basados en series de tiempo para la determinación de potencial eólico, ya que facilitan la toma de decisiones y reducen la inversión monetaria, no es necesario montar instrumentos de medición especializados (anemómetros) y provee de herramientas confiables y probadas.

## **INTRODUCCIÓN.**

En el presente trabajo se describe la implementación y ejecución de un método de predicción de velocidades del viento empleando modelos autorregresivos integrados de medias móviles sustentadas en herramientas computacionales (Mathematica®). Estos modelos son ampliamente utilizados en distintos ámbitos de la ciencia como por ejemplo Economía y Medicina, se propone, además, determinar un factor de corrección para mejorar su resolución local con los datos empleados.

En el Capítulo I, se contextualiza lo relacionado a la demanda energética en México, y el potencial eólico en las diferentes zonas geográficas del País. También resume la metodología para la determinación del potencial eólico en una zona geográfica, mediante la distribución de Weibull, su análisis e interpretación de resultados.

En el Capítulo II, se describen a grandes rasgos, los modelos climatológicos que son empleados normalmente por las distintas instituciones en la predicción del clima de las zonas, se definen sus términos, alcances y limitaciones. Además, se describen las nuevas formas de aplicación con metodología de Box y Jenkins, así como también el uso de redes neuronales artificiales.

En el Capítulo III, se definen y limitan las bases teóricas de las series de tiempo, la metodología para aplicar los modelos, así como las descripciones de los datos de velocidad de viento de los distintos emplazamientos, y los métodos estadísticos mediante los cuales se pueden determinar las posibles correcciones o relaciones que mejoren el rendimiento de la estimación.

Seguidamente, en el Capítulo IV, se compara con la metodología de Weibull y se realiza el análisis pertinente de los datos y estimaciones generadas en la realización de este proyecto, se comparan con mediciones de campo y se generan correcciones para retroalimentar el modelo.

Finalmente, en el Capítulo V se presentan los resultados finales producto de las correcciones, se calculan y reportan los errores asociados a la estimación final y se generan conclusiones.

**OBJETIVO.**

Estimar el potencial eólico local inmediato de la zona norte de la ciudad de Mérida, Yucatán, mediante el uso de modelos matemáticos basado en series de tiempo, para periodos cortos de tiempo.

**OBJETIVOS ESPECÍFICOS.**

- Ingresar al modelo las bases de datos de estaciones climatológicas de FIUADY, CINVESTAV.
- Generar código fuente del modelo con nuestra variable a estimar mediante la herramienta Mathematica® y compilar el modelo.
- Realizar las primeras estimaciones para obtener datos contrastables con los datos de las estaciones climatológicas y generar correcciones estadísticas.
- Retroalimentar el modelo con las correcciones estadísticas y generar nuevas estimaciones.

# CAPÍTULO I

## POTENCIAL EÓLICO Y SU EVALUACIÓN.

### 1.1 México, demanda energética y el viento.

La demanda energética es una de las principales preocupaciones en la actualidad, cada día se debe proveer de una gran cantidad de gigawatts, debido a esto las fuentes de energía renovable juegan un papel muy importante. Una de estas fuentes, altamente utilizada en la última década, es la que se obtiene del viento y es denominada energía eólica.

Los modelos climatológicos son ampliamente utilizados (normalmente para la prevención de desastres naturales) y son útiles para determinar una gran cantidad de datos sobre las variables del clima, generalmente son empleados para determinar temperaturas, radiación, corrientes y precipitación entre otros, pero también son usados en otras áreas, como por ejemplo la agricultura, la aeronáutica y la hidrología. Además de las variables mencionadas anteriormente, los mismos modelos permiten predecir la velocidad y dirección del viento, ya que los modelos ingresan datos de sistemas de medición que se encuentran interconectados a una red de información.

La naturaleza estocástica del viento lo define como un sistema errático y poco fiable, por lo tanto, para determinar la factibilidad de un proyecto de energía renovable (empleando generadores eólicos) en alguna localización geográfica concreta es necesario realizar estudios de potencial eólico; esto conlleva un problema que pocas entidades e instituciones afrontan eficientemente, ya que requiere de periodos de medición muy amplios (duraciones de años), es por esto que se propone el empleo de modelos matemáticos para la determinación de potencial eólico, ya que facilitan enormemente el acceso a estos datos, reducen la inversión monetaria en este rubro debido a que no es necesario montar instrumentos de medición especializados (anemómetros) y proveen de herramientas confiables y probadas (SENER, 2014).

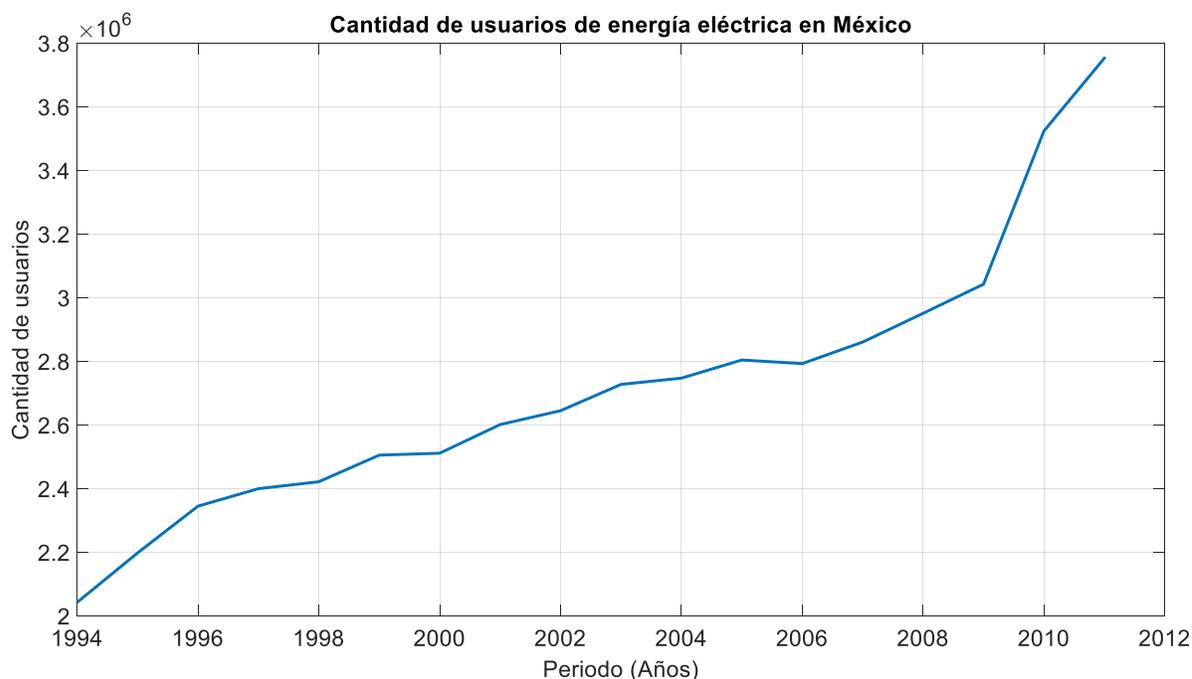


Figura 1.1.- Gráfica de cantidad de usuarios de energía eléctrica, demanda energética (INEGI, 2012).

México es un país con una demanda energética creciente como puede observarse en la figura 1.1, en general, en los estados prácticamente todas las casas habitación cuentan con el servicio de energía eléctrica.

El gran desarrollo urbano de México implica que año con año habrá una mayor demanda de energía. Esta energía es producida en su mayoría por fuentes de energía como la hidroeléctrica, termoeléctrica y ciclo combinado (tabla 1.1). Esto representa importantes repercusiones al medio ambiente, ya que las plantas hidroeléctricas impactan significativamente en la topografía del lugar en el que se construyen, y durante todo su ciclo de vida producen grandes cambios en los ciclos de regeneración hídrica de la región donde se construyen. Por otro lado, tenemos las fuentes que emplean la combustión y la evaporación del agua para generar electricidad, ya que estas se les suele considerar mucho más dañinas al contribuir al efecto invernadero (SENER, 2014).

Tabla 1.1.- Índice de la producción de electricidad en Megawatts, por fuente de energía del año 2014 (CFE, 2014).

Año	Hidroeléctrica	Termoeléctrica	Ciclo combinado (CFE)	Ciclo combinado (PEE)	Duales	Carboeléctrica	Nucleo eléctrica	Geo termoeléctrica	Eoloeléctrica (CFE)	Eoloeléctrica (PEE)	Fotovoltaica	TOTAL
2003	9,615	17,316	3,848	6,756	2,100	2,600	1,365	960	2	0	0	44,562
2004	10,530	16,954	4,776	7,265	2,100	2,600	1,365	960	2	0	0	46,552
2005	10,536	15,715	5,005	8,251	2,100	2,600	1,365	960	2	0	0	46,534
2006	10,566	15,586	5,203	10,387	2,100	2,600	1,365	960	2	0	0	48,769
2007	11,343	15,702	5,416	11,457	2,100	2,600	1,365	960	85	0	0	51,028
2008	11,343	15,734	5,456	11,457	2,100	2,600	1,365	965	85	0	0	51,105
2009	11,383	15,616	6,115	11,457	2,100	2,600	1,365	965	85	0	0	51,686
2010	11,503	15,627	6,115	11,907	2,778	2,600	1,365	965	85	0	0	52,945
2011	11,453	14,732	6,122	11,907	2,778	2,600	1,365	887	87	0	0	51,931
2012	11,498	14,608	6,122	11,907	2,778	2,600	1,610	812	87	511	1	52,534
2013	11,509	14,022	7,420	12,340	2,778	2,600	1,410	823	87	511	6	53,506
2014	12,269	14,002	7,567	12,340	2,778	2,600	1,410	813	86	511	6	54,382

En la tabla 1.1, se ilustra mejor la contribución de cada una de las fuentes de energía y es útil al mismo tiempo para comprobar que al igual que la demanda, la generación se ha incrementado en 10 años, ha pasado de 44 561 a 54 372 Megawatts, es decir un 22.01 % más, esto nos habla de las necesidades de la población, pero es de suma importancia remarcar que estas contribuciones no son proporcionadas por fuentes de energía renovables.

En la figura 1.2, se muestra la distribución de las principales fuentes de energía en México, las principales son la hidroeléctrica y termoeléctrica.

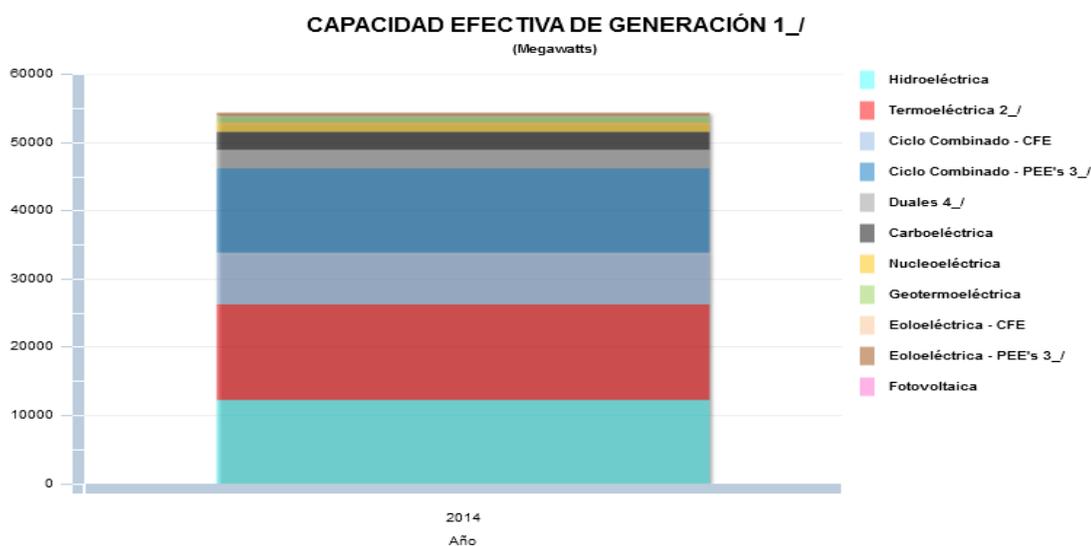


Figura 1.2.- Capacidad efectiva de generación (SENER, 2014).

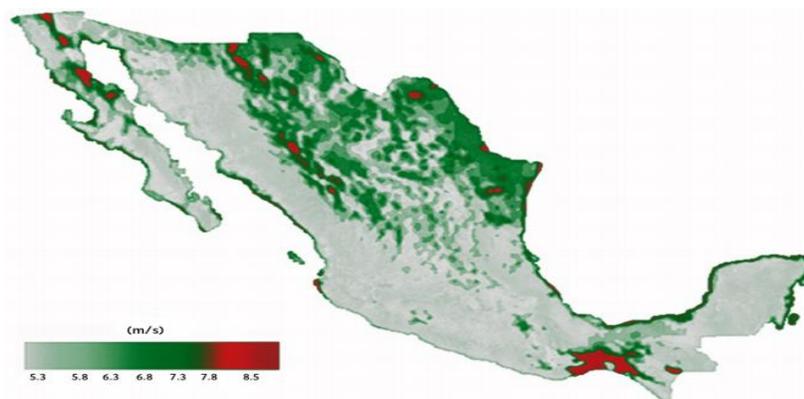


Figura 1.3.- Mapa velocidades de viento de México (SENER, 2014).

En la figura 1.3, se muestra el campo de velocidades de vientos en México. Se estima un potencial aprovechable en torno a los 40,000 MW y se tiene previsto que para finales de 2015 la capacidad instalada de generación mediante energía eólica sea de 358 GW, en Yucatán (figura 1.4) la capacidad instalada se encuentra entre 2 MW. “Yucatán es una de las potenciales zonas en las cuales se tienen como objetivos para llevar a cabo proyectos de generación de energía eólica” (Lozano, 2013).

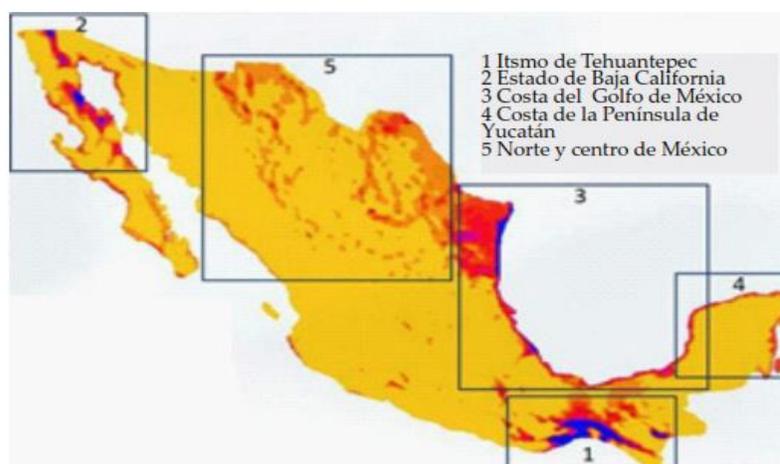


Figura 1.4.- Zonas potenciales para proyectos de energía eólica (AMDEE, 2014).

La Península de Yucatán es una región plana, homogénea y con vegetación baja, lo que la hace atractiva para la implementación de sistemas de conversión de energía eólica.

Al ser una región propensa a los huracanes, el tamaño de las torres debe mantenerse baja (30 a 40 m) para facilitar la protección del equipo bajo condiciones meteorológicas adversas. En los estudios realizados en la Facultad de Ingeniería de la Universidad Autónoma de Yucatán (FIUADY), se reporta una velocidad media de viento del orden de 5 m/s del viento local sobre la región. Sin embargo, otros estudios indican que la velocidad media del viento en la costa norte de la península es del orden de 6.5 m/s. Estos resultados hacen notar que, para emplear la energía eólica en esta región, los dispositivos deben tener como parámetro de funcionamiento de fábrica una velocidad de viento nominal dentro del intervalo 5 a 7 m/s. Esto ha sido de particular importancia para la implementación de turbinas de viento en la gama de 10 kW en nuestra institución educativa (A. Borges et al, 2014).

## **1.2 Alcance geográfico del estudio.**

Este proyecto tiene como límite geográfico la zona norte de la ciudad de Mérida, Yucatán. El estado de Yucatán está situado en el extremo norte de la Península del mismo nombre, en el sureste de la República Mexicana. Colinda al norte con el Golfo de México; al este y sureste con Quintana Roo y al oeste y suroeste con Campeche, es considerada una región en la que predominantemente tiene dos tipos de clima, el cálido subhúmedo que abarca el 90 % de su territorio con temperatura anual fluctuante entre 24.6 y 27.7 °C en promedio y el cálido semiseco que corresponde a la franja costera en el norte del Estado (Yucatán, 2012-2018).

La ciudad de Mérida (figura 1.5) se localiza en las coordenadas 20°58'04"N 89°37'18"O (20.967778, -89.621667), el Campus de Ciencias Exactas e Ingenierías, está ubicado sobre el Periférico Norte de Mérida (Kilómetro 33), con una altitud promedio de 8 metros sobre el nivel medio del mar.

En términos generales, la localidad posee una orografía plana, clasificada como llanura de barrera; sus suelos son generalmente rocosos o cementados. Tiene una situación geográfica privilegiada ya que posee bosques tropicales, cenotes, corrientes subterráneas, cavernas, etc.



### **1.3 Métodos convencionales para la determinación del potencial eólico.**

El método más común para la determinación del potencial eólico utiliza bases datos de mediciones directas de estaciones anemométricas y/o climatológicas, a las cuales se les aplica un tratamiento estadístico no paramétrico como la “distribución de Weibull”, la cual es una función de densidad de probabilidad continua.

El dispositivo encargado de medir la velocidad y dirección del viento se denomina anemómetro; existen diversas formas y principios de funcionamiento para los anemómetros. Sin embargo, es conveniente saber que, a pesar de contar con una amplia gama, algunos de estos poseen ventajas y desventajas, el más usado es el anemómetro de molino o hélice el cual consiste en un pequeño molino o hélice cuyos tres o más aspas poseen cazoletas o copas sobre las cuales actúa la fuerza del viento; el número de vueltas puede ser leído directamente en un contador o registrado por un dispositivo de almacenamiento electrónico.

Una estación meteorológica/anemométrica (figura 1.6) es una instalación destinada a medir y registrar regularmente diversas variables meteorológicas. Estos datos se utilizan tanto para la elaboración de predicciones meteorológicas a partir de modelos numéricos, como para estudios climáticos. Estas estaciones pueden brindar información solamente visual (existen estaciones de las cuales solo se lleva un registro visual y manual), automatizada (permiten almacenar toda la información en registros digitales o impresos) e incluso satélites meteorológicos para ayudar y mejorar la metodología de la predicción del clima y del potencial.

Para realizar cualquier estudio del clima hay que analizar los elementos del tiempo meteorológico: la temperatura, la humedad, la presión, la velocidad promedio del viento y las precipitaciones, los cuales tienen un comportamiento variable a lo largo del tiempo. Generalmente los registros mensuales que se guardan a lo largo de una serie bastante larga de años son los datos más importantes que normalmente aparecen en los gráficos climáticos. Las variaciones a corto plazo de la atmósfera (que generalmente se llama tiempo meteorológico), son las que se relacionan con nuestra vida cotidiana.



Figura 1.6.- Estación climatológica en emplazamiento.

Caracterizar el recurso disponible y determinar los emplazamientos apropiados, se convierten en acciones críticas en la realización de cualquier proyecto de generación de energía eléctrica a través de fuentes renovables. Así, para un aprovechamiento óptimo es importante realizar una valoración energética y de potencial.

Para los estudios sobre generación eólica es muy importante ser capaz de describir estadísticamente la variación de las velocidades del viento. Esta información es necesaria para determinar la viabilidad de un proyecto.

La variación del viento en un emplazamiento típico suele describirse utilizando la llamada distribución de Weibull, representada con la ecuación 1.1, la cual esencialmente emplea dos parámetros, el de forma el cual afecta la *forma o geometría* de la distribución y el de escala el cual asigna mayor o menos amplitud a la distribución.

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(x/\lambda\right)^k} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad 1.1$$

Donde:

f: Función de densidad.

k: Parámetro de forma.

$\lambda$ : Parámetro de escala.

En este tipo de distribuciones el área bajo la curva siempre es igual a uno (figura 1.7), ya que la probabilidad de que el viento sople en cualquiera de las velocidades incluyendo al cero, debe ser del 100 %. Si se miden las velocidades del viento a lo largo de un año se observará que en la mayoría del área de la curva los fuertes vendavales son raros, mientras que los vientos moderados son bastante comunes.

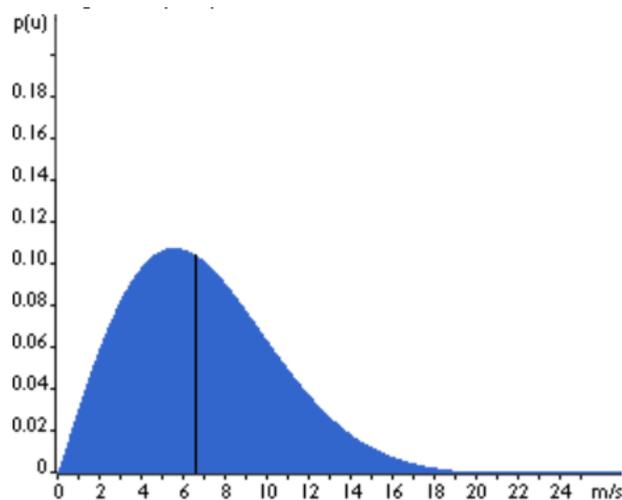


Figura 1.7.- Distribución de Weibull, típica de los estudios de potencial eólico (WindPower, 1998).

La mitad del área azul está a la izquierda de la línea negra vertical a 6.6 m/s. El valor de 6.6 m/s es la mediana de la distribución, esto significa que la mitad del tiempo el viento soplará a menos de 6.6 m/s y la otra mitad soplará a más. La velocidad media del viento, para esta gráfica, es de 7 m/s. Una de las desventajas de la distribución de Weibull, aplicadas a la energía eólica, es que las velocidades de 0 m/s están relacionadas con una probabilidad cero, lo cual no es verdadero ya que existen instantes de tiempo en los que el viento no está presente. La velocidad media del viento realmente es el promedio de las observaciones de la velocidad del viento que se tendrá (WindPower, 1998). La distribución de las velocidades del viento es sesgada, es decir, no es simétrica.

En ocasiones se tendrán velocidades de viento muy altas, pero con poca probabilidad de ocurrencia. Por otro lado, las velocidades del viento menores o iguales a 5.5 m/s son las más comunes. Las velocidades de 5.5 m/s serían el valor modal de la distribución.

Si multiplicamos cada pequeño intervalo de la velocidad del viento por la probabilidad de tener esa velocidad en particular, y los sumamos todos, obtenemos la velocidad del viento media (WindPower, 1998).

La distribución estadística de las velocidades del viento varía de un lugar geográfico a otro, además depende de las condiciones climáticas locales, del paisaje y de su superficie. Por lo tanto, la distribución de Weibull puede variar tanto en la forma como en el valor medio. Si el parámetro de forma es exactamente dos, como en la figura 6, la distribución es conocida como distribución de Rayleigh (WindPower, 1998).

La distribución de Rayleigh es una distribución de probabilidad continua de variables aleatorias de valor positivo. Este tipo de distribución se observa a menudo cuando la magnitud general de un vector está relacionada con sus componentes direccionales.

Un ejemplo en el que surge naturalmente la distribución Rayleigh es cuando la velocidad del viento es analizada en sus componentes vectoriales de dos dimensiones ortogonales. Suponiendo que las magnitudes de cada componente no están correlacionadas, normalmente distribuidos con igualdad de la varianza, y una media de cero, entonces la velocidad del viento en general se caracteriza por una distribución de Rayleigh (WindPower, 1998).

Los fabricantes de aerogeneradores proporcionan gráficas de rendimiento para sus máquinas usando la distribución de Rayleigh. La razón por la que nos interesan las velocidades del viento es por su contenido energético y por lo tanto la potencia que se puede extraer, ecuación 1.2.

$$P_{viento} = \frac{1}{2} \rho v^3 \pi r^2 \quad 1.2$$

Donde:

$P_{viento}$ : Es la potencia del viento (W).

$\rho$ : Densidad del viento (Kg/m<sup>3</sup>).

$v$ : Velocidad del viento (m/s).

$\pi r^2$ : Área del aerogenerador (m<sup>2</sup>).

Las leyes de la física, impiden que se pueda extraer toda la potencia disponible en el viento a su paso por el rotor de un aerogenerador. El viento a su paso se frena, saliendo del mismo con una velocidad menor que con la que ha entrado. La máxima potencia eólica aprovechada la define el Límite de Betz, el cual establece que una turbina eólica puede convertir en energía mecánica como máximo un 59.26 % de la energía cinética del viento que incide sobre ella.

La velocidad del viento está siempre fluctuando, por lo que el contenido energético del viento varía continuamente. La magnitud de esa fluctuación depende tanto de las condiciones climáticas como de las condiciones de superficie locales y de los obstáculos. En la mayoría de localizaciones del planeta el viento tiene velocidades más altas durante el día que durante la noche. Esta variación se debe sobre todo a las diferencias de temperatura, por ejemplo, entre la superficie del mar y la superficie terrestre, el viento presenta turbulencias y tiende a cambiar de dirección más rápidamente durante el día que durante la noche (WindPower, 1998). Aquí es importante mencionar que estos acercamientos estadísticos nos brindan información generalista, es decir, no necesariamente nos darán los valores más puntuales como lo sería la turbulencia.

#### **1.4 Estudios estadísticos y de dispersión empleados en la determinación del potencial eólico.**

En general, los tratamientos estadísticos tienen como objetivo brindar un panorama general sobre la variable que desea evaluar.

Para cualquier base de datos que se desee procesar es necesario analizar y determinar los aspectos básicos como la media, la varianza y el sesgo, además de cumplir los supuestos del tipo de distribución deseada. Existen dos razones básicas por las cuales las distribuciones de probabilidad ocupan un lugar tan prominente en la estadística:

-La primera es que poseen algunas propiedades que la hacen aplicable a un gran número de circunstancias en las cuales es necesario hacer inferencias a través de la toma de muestras.

-La segunda es que es posible hallar una que se ajuste a las distribuciones de frecuencias reales observadas en muchos fenómenos, entre las que se encuentran la velocidad del viento y se representa mediante una función que depende básicamente de la media y la desviación estándar.

La variación de la velocidad del viento es un proceso aleatorio, en la que se debe mediante una determinación estocástica asegurar que su valor se encuentre dentro de ciertos valores límites. Generalmente el primer paso consiste en medir la velocidad y dirección de los vientos a distintas alturas en el lugar o locación deseada, empleando anemómetros (cuyo error sea mucho menor al 1 %).

Se realizan como mínimo, mediciones de corta duración de 1 año. Con mediciones en sitio, de corta duración, y correlaciones que proporcionan información confiable se puede evaluar el recurso eólico y determinar el rendimiento energético esperado. (E. Cadenas et al, 2008). Para ello, se aplica la metodología conocida como MCP (medir, correlacionar y predecir) la cual comprende:

*Medir*, es un proceso básico que consiste en comparar el patrón seleccionado con datos del sitio obtenidos a corto plazo, se mide para ver cuántas veces el patrón está contenido en esa magnitud.

*Correlacionar*, mediante datos de referencia de largo plazo de estaciones de referencia, preferentemente con un régimen de vientos similar al sitio o, en su defecto, empleando datos de referencia sintéticos (generados por modelos de clima), analizarlos y depurarlos aplicando modelos o metodología estadística.

*Predecir*, es el último paso, necesario en la evaluación del recurso eólico en el sitio y a largo plazo. Las mediciones de campo se complementan con la implementación de modelos que calculen el flujo en escala sinóptica (escala de longitud horizontal del orden de los 1,000 km o más), meso-escala (dimensiones horizontales generalmente oscilan de cerca de 9 km a varios cientos de km) y microescala (cerca de 1 km o menos) (Contreras B., 2006).

La predicción meteorológica numérica se refiere a los sistemas que utilizan datos meteorológicos actuales para alimentar complejos modelos fisico-matemáticos de la atmósfera para predecir la evolución meteorológica.

Los primeros intentos por realizar predicciones utilizando este tipo de metodologías se remontan a la década de 1920, sin embargo, no fue hasta la llegada de los equipos computacionales y de la simulación, cuando se pudieron implementar modelos que capaces de realizar cálculos en tiempo real.

Estas herramientas han permitido la manipulación de grandes conjuntos de datos y la realización de cálculos avanzados con una resolución lo suficientemente detallada, y por tanto práctica en las previsiones meteorológicas, requiere del empleo de supercomputadoras. Algunos de los modelos matemáticos de predicción, tanto a escalas global y regional, son utilizados para realizar previsiones para países del mundo entero.

La predicción mediante modelos numéricos se basa en la toma de una muestra del fluido en un instante determinado (conociendo una amplia gama de sus propiedades en diversos puntos, tales como grado de humedad, presión, velocidad y dirección del viento) y la aplicación de ecuaciones de la dinámica de fluidos y termodinámica para calcular el estado del fluido en algún momento del futuro.

Con todo lo anterior es posible también hacer determinaciones en periodos cortos de tiempo, ya que con el sustento que brinda una base de datos, un análisis estadístico y las simulaciones de modelos de escala sinóptica o meso escala es posible determinar los cambios locales en un área geográfica en intervalos de dos a tres horas en el futuro. Esto es una herramienta muy poderosa, ya que da sustento a la toma de decisiones, es decir, decidir qué sistemas de generación emplear, durante cuánto tiempo y en algunos casos reducir la cantidad de sistemas de medición instalados en los emplazamientos. Edward Norton Lorenz en 1963 mostró que hay un límite de dos semanas para predecir el clima, aun teniendo modelos y observaciones perfectas.

Estudió las soluciones de su modelo y se percató que alteraciones mínimas en los valores de las variables iniciales resultaban en soluciones ampliamente divergentes, esta sensible dependencia de las condiciones iniciales fue conocida después como el efecto mariposa. Su investigación dio origen a un renovado interés en la teoría del caos. (Lorenz E., 1963)

Lorenz se dedicó a explorar las matemáticas subyacentes y publicó sus conclusiones en un trabajo titulado “*Flujo determinista no periódico*” (Lorenz E., 1963) en el que describió un sistema relativamente simple de ecuaciones que dieron lugar a un patrón de la complejidad infinita, llamado atractor de Lorenz.

Sin embargo, el análisis numérico y las herramientas computacionales han avanzado muchísimo, y en la actualidad los modelos y pronósticos son capaces de tener una validez de hasta una semana.

Las mejoras del pronóstico numérico se deben a:

- Mejoras del modelo.
- Mejoras de las observaciones.
- Mejoras de la asimilación de datos.

Esto introduce una fuente de predictibilidad más allá del límite de dos semanas. Permiten predecir anomalías del clima, pero no predecir el clima. Por ello se desarrollaron algunas metodologías que permiten ir más allá de este límite. Algunas de estas metodologías son: los pronósticos numéricos del tiempo y los pronósticos de ensamble. Se crean modelos para simular la atmósfera, las inestabilidades aumentan los errores del pronóstico (Caos), los modelos necesitan condiciones iniciales (el análisis estadístico), las condiciones iniciales tienen errores y los errores crecen debido a las inestabilidades y a los errores del modelo, pero son minimizadas al ciclar los modelos.

## CAPÍTULO II

# MODELOS METEOROLÓGICOS Y COMPUTACIONALES.

### 2.1 Clasificación de los modelos meteorológicos.

Los modelos meteorológicos son una serie de sistemas de ecuaciones, las cuales son obtenidas de la investigación y la experimentación de los fenómenos, comprendiendo su análisis desde el punto de vista de la física, la química y la mecánica de fluidos. Se emplean para simular y describir matemáticamente las interacciones que existen entre la superficie terrestre, las diferentes capas que componen la atmósfera y los grandes cuerpos de agua.

El clima es usualmente modelado como la respuesta de la atmósfera, los océanos, los grandes cuerpos de hielo y los continentes al estímulo que representa la radiación solar. La Tierra absorbe la mayor parte de la radiación incidente cerca del ecuador y la redistribuye mediante una serie de procesos de transporte energéticos en los que las nubes, las masas continentales, las corrientes oceánicas y la circulación atmosférica interaccionan de un modo complejo y caótico (Garrido V., 2007).

También se debe considerar la dinámica atmosférica, esto es la suma de las fuerzas que provoca que las superficies de aire se aceleren o desaceleren, es decir crean o alteran el viento. La relación entre las fuerzas y el viento se llama dinámica atmosférica. Las leyes de Newton describen muy bien esta dinámica. La presión, la fricción y la advección son algunas de las fuerzas atmosféricas que actúan en el plano horizontal, existe otras fuerzas como la ocasionada por la rotación de la tierra o por el giro del viento alrededor de una curva, estas son llamadas fuerzas aparentes. La fuerza horizontal más importante para la formación del viento es el gradiente de presión (Garrido V., 2007).

Un elemento en común de todos los modelos es que utilizan la información de esta energía que ingresa al sistema que se desea evaluar (figura 2.1), por lo tanto, términos como la energía de entrada o salida son considerados y empleados por la totalidad de los modelos.

La Meteorología y la Climatología son dos ciencias, aunque a simple vista estas dos disciplinas parezcan ser la misma, pues en el fondo su principal objetivo es predecir el tiempo atmosférico, el principal elemento que las diferencia es la escala de tiempo empleada para llevar a cabo dichas predicciones.

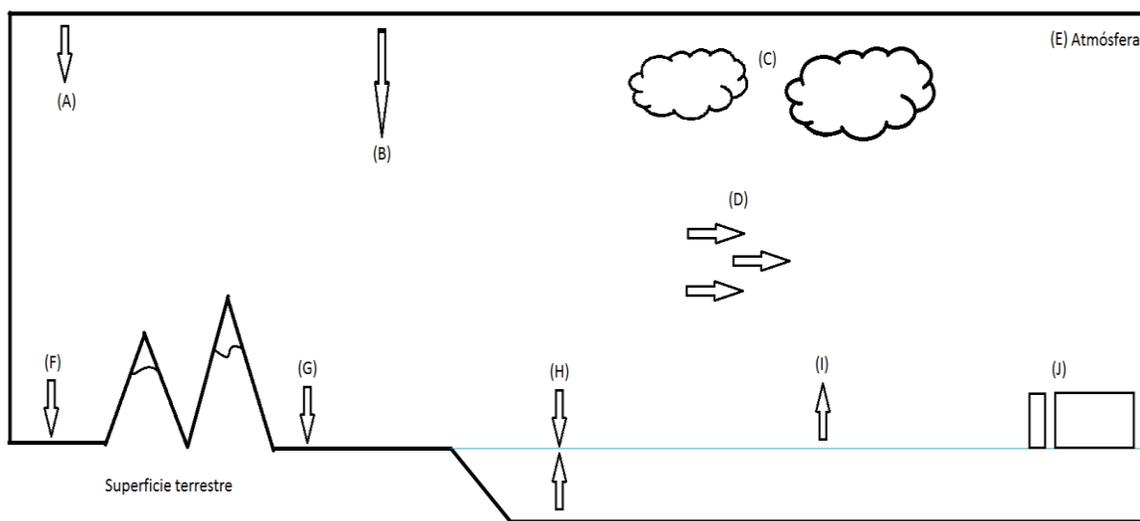


Figura 2.1.- Sistema climatológico típico y sus componentes. (A) Gases componentes de la atmósfera ( $H_2$ ,  $N_2$ ,  $H_2O$ , etc.), (B) Radiación incidente en la atmósfera, (C) Nubosidad, (D) Campos de viento, (E) Estrato atmosférico, (F) Litosfera, (G) Radiación reflejada, (H) Corrientes marinas, (I) Intercambios de calor, (J) Criosfera.

A grandes rasgos, las principales características de estas dos ramas científicas son:

*Climatología*, la escala de tiempo empleada para las predicciones es muy amplia, usa periodos de tiempo muy grandes (décadas, milenios, etc.).

Los métodos utilizados son principalmente cualitativos, de diagnóstico y simplificados que permiten un diagnóstico de la evolución del clima. Entre estos modelos podemos encontrar los Modelos de Balance de Energía.

*Meteorología*, los modelos empleados se caracterizan por proporcionar una predicción del tiempo atmosférico con gran exactitud durante periodos de tiempo pequeños (días, semanas, etc.). Los métodos utilizados para dichas predicciones son principalmente computacionales y predictivos. Utiliza modelos muy realistas que permiten una predicción precisa del tiempo atmosférico (Modelos de Circulación General) (Garrido V., 2007).

Los modelos van desde los relativamente simples, ya sea para cálculos rígidos de la temperatura incidente tratando un punto en la superficie, expandiéndose verticales u horizontalmente hacia la atmósfera hasta modelos climáticos globales acoplados, los cuales involucran la relación atmósfera–océano–hielo continental.

### **2.1.1 Modelos unidimensionales.**

Estos emplean fórmulas y ecuaciones simplificadas, las cuales permiten conocer un valor puntual de la variable que se desea conocer y sólo varían una de las coordenadas espaciales. Para realizar el cálculo emplean valores numéricos precisos de los factores que involucra la variable a evaluar, por lo tanto, son óptimos para una primera aproximación. Estos modelos son empleados, por ejemplo, para determinar la radiación instantánea de un sitio (Garrido V., 2007).

Es claro que son una simplificación muy drástica y reducen de forma importante el valor de los resultados. Sin embargo, existen motivos para aceptar estas soluciones ya que sin duda la simplicidad del modelo y la menor necesidad de potencia de cálculo juegan un papel importante. No obstante, la unidimensionalidad no significa que se deba prescindir de modelar todas las interrelaciones que se producen en el clima tridimensional.

Los *modelos radiativos/convectivos*, estos modelos utilizan solo la dimensión vertical para estudiar las transferencias de radiación a través de la atmósfera. Se emplean para simular la radiación solar terrestre, el efecto de la humedad, de los cambios de composición de la atmósfera, etc. (Garrido V., 2007).

Los *modelos oceánicos de difusión/afloramiento unidimensional*, del mismo modo que los anteriores solamente utilizan la dimensión vertical (Garrido V., 2007).

Para emplearlos es necesario suponer a la atmósfera como una caja homogénea (sin diferencias de composición) y representar al océano como una columna vertical. Para simular los movimientos verticales de agua se considera una segunda columna por donde bajan las aguas frías polares cuya densidad aumenta al absorber la sal que expulsan los hielos al formarse (UCAR, 2004).

Los *modelos de balance de energía unidimensionales* utilizan solo la latitud y solo modelan las transferencias en dirección norte-sur. Los Modelos Climáticos de Balance de Energía descritos vienen formulados mediante un balance energético entre la energía solar absorbida y la emitida por la Tierra. En este balance han de tenerse también en cuenta los procesos de difusión térmica que se producen sobre la superficie terrestre (Garrido V., 2007).

### **2.1.2 Modelos bidimensionales.**

Son aquellos en los que sólo varían dos de las coordenadas espaciales, se consideran diferencias de longitud y latitud o latitud y altura, comúnmente simulan el transporte de masa y energía de manera más eficiente que los modelos unidimensionales.

Esto significa, por ejemplo, que la temperatura asociada a un punto geográfico deberá representar la de toda la masa de tierra y aire sobre el mismo plano vertical o si se trata de un punto sobre el mar deberá representar la temperatura del agua en toda su profundidad y la de la atmósfera que tiene encima (tabla 2.1). Indudablemente, se trata de un avance respecto a los modelos unidimensionales, pero sigue siendo una simplificación drástica que limita el valor de los resultados obtenidos (UCAR, 2004).

“La precisión de los resultados obtenidos dependerá en gran medida de los parámetros a modelar y de los procesos que no es posible simular, para ello se eligen ecuaciones adecuadas para representar los procesos fundamentales que se desarrollan realmente en el entorno, como el intercambio de calor y masa entre la hidrosfera (mares), atmósfera, criosfera (hielos) y litosfera (tierra). A esto se le llama *parametrización*, consiste en incluir ecuaciones para los parámetros climáticos más importantes, intentando modelar las variaciones que se producen a distintas alturas, aunque el modelo trabaje en un solo plano” (UCAR, 2004).

Tabla 2.1.- Cantidades típicas empleadas de las componentes del sistema atmosférico (Garrido V., 2007).

Componente	Superficie (10 <sup>6</sup> km <sup>2</sup> )/%	Masa (10 <sup>18</sup> ) kg	Densidad (kgm <sup>-3</sup> )
Atmósfera	510/100	5	1.3
Océano	361/70.8	1350	1000
Criosfera			
1) Hielo marino	26/5.1	0.04	800
2) Hielo terrestre	14.5/2.8	28	900
Biosfera	103/20.2	0.002	100-800
Superficie	149/29.2	ND	2000

Los *modelos atmosféricos y oceánicos bidimensionales*, emplean la longitud y latitud o latitud y profundidad, se puede combinar un modelo atmosférico y uno oceánico introduciendo interacciones entre ellos, y obteniendo un modelo atmosférico/oceánico bidimensional. El modelo resultante suele ser más sencillo que desarrollar un modelo tridimensional y la potencia de cálculo necesaria no incrementa demasiado (Garrido V., 2007).

### 2.1.3 Modelos tridimensionales.

Los modelos tridimensionales representan una ventaja muy importante respecto a los anteriores. Sin embargo, la confiabilidad de sus resultados dependerá de la capacidad de incluir los procesos más relevantes que intervienen en el clima y en gran medida de la resolución a emplear. Los más sencillos solo aplican la tridimensionalidad a la atmósfera o a la hidrosfera, simplificando las interacciones con los demás elementos del modelo. Cada proceso (viento, corrientes marinas, evaporación, biosfera, etc.) que se incluya al modelo lo complica, aumenta el número de variables a tratar y de parámetros a determinar, pero no necesariamente incrementa las necesidades de cálculo.

Por otro lado, incrementar la resolución del modelo (ver figura 2.2), es decir, disminuir la rejilla de trabajo representa un incremento del número de celdas a evaluar por un factor de hasta ocho veces, lo cual exige una potencia de cálculo mucho mayor.

*Los modelos atmosféricos de circulación general (MCGA)* simulan directamente los vientos y la interacción con la superficie sólida y líquida del planeta, los datos empleados son predeterminados y no se evalúa su evolución (UCAR, 2004).

*Los modelos oceánicos de circulación general (MCGO)* simulan, principalmente, las corrientes oceánicas, los cambios de salinidad, los flujos de calor y la interacción con la litosfera y la atmósfera, los datos empleados son predeterminados y no se evalúa su evolución (UCAR, 2004).

*Los modelos atmosféricos y oceánicos de circulación general (MCGAO)* consisten en el acoplamiento de dos modelos de los tipos anteriores, a los que se añaden módulos que simulan las interacciones entre ellos, como los intercambios de humedad, calor y momento entre ellos (UCAR, 2004). “Actualmente los mejores modelos de circulación general usan celdas de entre 250 y 600 km y entre 10 y 20 capas verticales para la atmósfera y hasta 30 para los océanos. Esto representa, para el caso más sencillo 60 x 30 x 25 celdas, o sea 45000 celdas interactuando entre sí a través de decenas de variables. En el caso de mayor resolución serían 160 x 80 x 50 = 640 000 celdas” (UCAR, 2004).

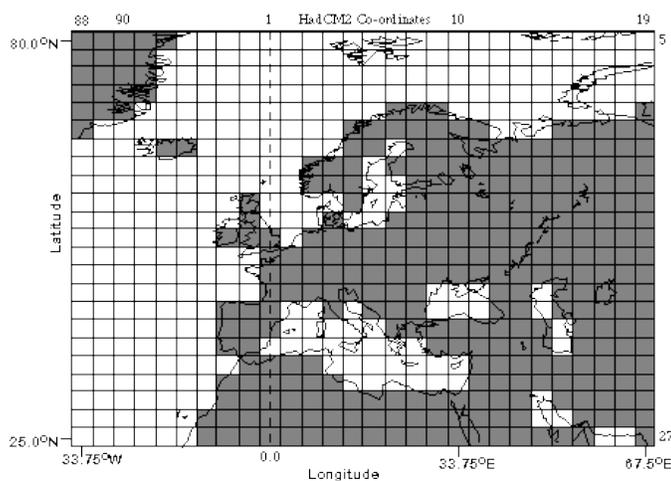


Figura 2.2.- Ejemplo de rejilla empleada en la parametrización y modelación de un sistema climático en una región (UCAR, 2004).

Así, estos modelos tridimensionales son mucho más complejos que los unidimensionales o bidimensionales y por lo tanto exigen capacidades de cálculo grandísimas, tanto por lo que respecta a la cantidad de memoria como a la velocidad de cálculo del proceso, las simulaciones meteorológicas suelen hacerse en los superordenadores más potentes del momento.

La consecuencia de esta exigencia computacional es que las simulaciones son muy caras y los creadores intentan recuperar parte de su inversión cobrando por ofrecer sus servicios, por lo cual no publican sus programas ni facilitan todos sus resultados al dominio público (UCAR, 2004).

Es importante mencionar que la actividad geológica terrestre, los procesos químicos (combustión, emisiones de CO<sub>2</sub>, etc.) y biológicos (deforestación, cambios estacionales de la vegetación) también introducen intercambios energéticos muy importantes entre el interior de la Tierra y la superficie, además son difíciles de modelar ya que además varían en tiempos relativamente cortos.

## **2.2 Métodos computacionales para la implementación de modelos en determinación del clima.**

La aparición de los modelos tridimensionales trajo como consecuencia la necesidad de nuevas y mejores herramientas para resolución de estos. Mediante el uso de computadoras modernas, las cuales actualmente poseen una gran capacidad de procesamiento con costos relativamente admisibles, así como la aplicación de diversos métodos de cálculo numérico, es posible obtener soluciones numéricas aproximadas del sistema de ecuaciones de un MCGA. De este modo, se obtiene la evolución temporal y espacial (tridimensional) del sistema climático que se desea evaluar, en función de las condiciones iniciales. Su solución procede al elegir los valores de ciertos parámetros climáticos, la solución numérica de un MCGA en condiciones preestablecidas se denomina experimento climático (UCAR, 2004).

Los términos que más comúnmente delimitan la eficiencia y el área de trabajo de un modelo meteorológico se definen a continuación.

La *escala sinóptica*, también llamada escala ciclónica, es una escala de longitud horizontal del orden de los 1 000 km o más, corresponde a una escala horizontal típica de depresiones de latitudes medias. Muchas de las zonas de presiones altas y bajas se leen en cartas meteorológicas como sistemas de escala sinóptica (Garrido V., 2007).

La palabra sinóptico deriva del griego σύν y ὀπτικός, significando visión de conjunto. En el análisis sinóptico se emplean las ecuaciones de Navier-Stokes aplicadas al movimiento atmosférico, ya que estas pueden simplificarse en el análisis de escala sinóptica.

Los términos principales en las ecuaciones horizontales quedan expresados en términos de las fuerzas de Coriolis y de gradiente de presión por lo tanto se puede usar una aproximación geostrófica (aproximación física al viento real), aplicado a coordenadas verticales, la ecuación de momento simplifica la ecuación de equilibrio hidrostático (Garrido V., 2007).

La *mesoescala*, en Meteorología es el estudio de sistemas del tiempo atmosférico más pequeños que la escala sinóptica, pero más grandes que la microescala. Sus dimensiones horizontales generalmente oscilan de cerca de 9 km a varios centenares de km. Un ejemplo de sistema de mesoescala meteorológica es la brisa de mar.

La *microescala*, es el estudio de los fenómenos atmosféricos de corto tiempo a una menor extensión que los de mesoescala, cercanos a 1 km o menos, se estudian fenómenos menores que la escala sinóptica; tal que esos estudios generalmente son demasiado pequeños como para colocarlos en una carta sinóptica (Garrido V., 2007).

La meteorología en microescala controla los procesos de más importancia en la mezcla y dilución atmosférica. Algunos de los fenómenos importantes en la meteorología de microescala incluyen transferencias térmicas e intercambios gaseosos entre suelo, vegetación, agua superficial y la atmósfera los cuales causan turbulencia. La velocidad vertical frecuentemente iguala o excede las velocidades horizontales en los sistemas de mesoescala meteorológicas, debido a procesos no hidrostáticos (Garrido V., 2007).

Normalmente, realizar análisis a escala sinóptica requiere de herramientas y cálculos de alto nivel y mucha complejidad, es por esto que existen agencias y organizaciones gubernamentales que se encargan de realizar estos análisis, una de las más importante es la Administración Nacional Oceánica y Atmosférica (National Oceanic and Atmospheric Administration, NOAA por sus siglas en inglés), la cual es una agencia científica del Departamento de Comercio de los Estados Unidos cuyas actividades se centran en las condiciones de los océanos y la atmósfera. Los resultados de los análisis que la NOAA obtiene se encuentran disponibles en su sitio web (NOAA, 2008).

También existen programas o interfaces que mediante la conexión a internet pueden obtener y desplegar la información proporcionada por NOAA, algunos de estos incluyen el NOAAPORT, el cual proporciona una comunicación de difusión unidireccional de datos ambientales de la NOAA y la información casi en tiempo real a la NOAA y los usuarios externos. Este servicio de difusión se lleva a cabo por un proveedor comercial de comunicaciones por satélite que utiliza la banda C (NOAA, 2008). El flujo general de los principales componentes operativos que se alimentan de datos para NOAAPORT se muestra en la figura 2.3 (NOAA, 2008).

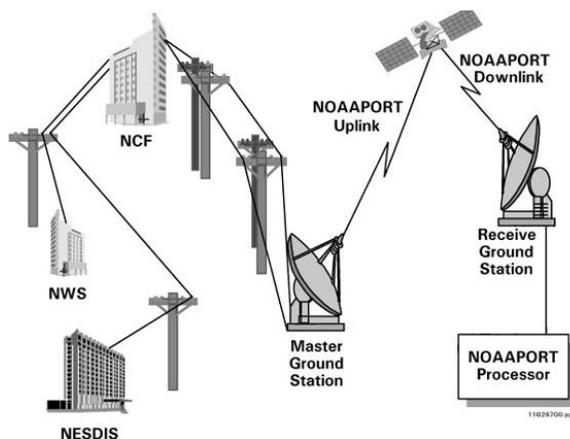


Figura 2.3.- Diagrama de operación de NOAAPORT. (NOAA, 2008).

También existen métodos y modelos, disponibles para el usuario académico o general, mantenidos por comunidades u organizaciones y permiten realizar análisis a mesoescala o escala local. A continuación, se enlistan algunos de los de más fácil acceso.

El modelo de mesoescala PSU / NCAR (conocido como MM5) es un modelo de área limitada, no hidrostático y de seguimiento de terreno diseñado para simular o predecir la circulación atmosférica de mesoescala y fue desarrollado en la Universidad Estatal de Pensilvania (MM5 Community Model, 2015).

El software de este sistema de modelización MM5 es proporcionado y apoyado por el Grupo de Predicción de Mesoescala en su División de Meteorología de mesoescala y microescala, y es distribuido libremente (MM5 Community Model, 2015).

El modelo Weather Research and Forecasting (WRF), es un sistema de predicción numérica del clima a mesoescala de próxima generación diseñado para las necesidades de investigación y predicción operativa atmosférica. Cuenta con dos núcleos dinámicos, un sistema de asimilación de datos, y una arquitectura de software facilitando la computación paralela y extensibilidad del sistema (WRF Model, 2015).

El esfuerzo para desarrollar WRF comenzó en la última parte de la década de 1990 y fue una asociación de colaboración principalmente entre el Centro Nacional de Investigación Atmosférica (NCAR) y diversas instituciones más (WRF Model, 2015).

Este modelo puede generar simulaciones atmosféricas utilizando datos reales (observaciones, análisis) o condiciones idealizadas. Ofrece una previsión operativa para una plataforma flexible y computacionalmente eficiente, al tiempo que proporciona los últimos avances en la física, métodos numéricos y de asimilación de datos aportados por los desarrolladores a través de la comunidad de investigación muy amplio (WRF Model, 2015).

El modelo WAWG nace debido a la evaluación de la importancia de la unión entre las regiones atmosféricas. La propagación vertical de las ondas atmosféricas y el transporte de energía en la troposfera se sabe que juega un papel importante en la dinámica y la química de la atmósfera media y alta. También hay una creciente conciencia de que los cambios en las características de propagación de las ondas planetarias en la estratosfera (debido a factores naturales o antropogénicos) pueden desempeñar un papel en la variabilidad del clima troposférico influyendo en fenómenos como la Oscilación del Ártico (WAWG Model, 2013).

El modelo llamado High Resolution Limited Area Model (HIRLAM), es un sistema de predicción numérica del clima, desarrollado por el programa HIRLAM internacional (HIRLAM Org, 2014), modelo de mesoescala de sistema de predicción numérica del tiempo (permite convección) (HIRLAM Org, 2014).

Existen otras herramientas son el conjunto de técnicas de estudio de series de observaciones dependientes ordenadas en el tiempo y se les denomina Análisis de Series Temporales. El instrumento de análisis que se suele utilizar es un modelo que permite reproducir el comportamiento de la variable de interés en particular en series temporales, un modelo autorregresivo integrado de media móvil o ARIMA (acrónimo del inglés autoregressive integrated moving average) es un modelo estadístico que utiliza variaciones y regresiones de datos estadísticos con el fin de encontrar patrones para una predicción hacia el futuro.

Se trata de un modelo dinámico de series temporales (es una secuencia ordenada de observaciones cada una de las cuales está asociada a un momento de tiempo.), es decir, las estimaciones futuras vienen explicadas por los datos del pasado y no por variables independientes. En Meteorología, tenemos series temporales de temperatura, cantidad de lluvia sobre en una región, velocidad del viento, etc.

El modelo ARIMA necesita identificar los coeficientes y número de regresiones que se utilizarán. Este modelo es muy sensible a la precisión con que se determinen sus coeficientes. Fue desarrollado a finales de los sesenta, Box y Jenkins lo sistematizaron en 1976 (Gonzalez M., 2009).

En el caso de las series temporales de datos climatológicos, hemos de tener en cuenta, que el orden es fundamental ya que tenemos un conjunto de datos consecutivos ordenados, el supuesto de independencia no se sostiene ya que en general, las observaciones son dependientes entre sí y la naturaleza de su dependencia es de interés en sí misma.

Los modelos de series temporales pueden ser:

*Univariantes*, solo se analiza una serie temporal en función de su propio pasado.

*Multivalentes*, se analizan varias series temporales a la vez. Se sabe que existen distintas relaciones entre ambas variables, lo que supone que afecta a la dinámica de ambas series. La forma de reflejar estas interacciones dinámicas entre ambas series es construir un modelo multivariante (Gonzalez M., 2009).

Cuando se construye un modelo multivariante, se establece que existe cierta dependencia o relación entre los pasados de las diversas series.

Cuando las observaciones se recogen solo en momentos determinados de tiempo, generalmente a intervalos iguales, nos referimos a una *serie temporal discreta*. Puede darse el caso de que los datos se generen de forma continua y se observen de forma continua, como es el caso de la temperatura, que se observa de forma continua en el tiempo por medio de aparatos o instrumentos físicos y electrónicos, en este caso contamos con un número muy grande de observaciones y nos referiremos a una *serie temporal continua* (Gonzalez M., 2009).

### **2.3 Técnicas de predicción (Redes neuronales).**

Las redes de neuronas artificiales (llamadas habitualmente RNA) son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en la que funciona el sistema nervioso de los animales. Así, se trata de un sistema de interconexión de neuronas que colaboran entre sí para producir un estímulo de salida. Los primeros modelos de redes neuronales surgen en 1943 gracias a los neurólogos Warren McCulloch y Walter Pitts. En 1949, Donald Hebb desarrolló sus ideas sobre el aprendizaje neuronal, y estableciendo la "regla de Hebb". En 1958, Rosenblatt desarrolló el perceptrón simple (Pino R. et al, 2004).

En inteligencia artificial es frecuente referirse a ellas como redes de neuronas o redes neuronales. El nombre, viene de la idea de imitar el funcionamiento de las redes neuronales de los organismos vivos, es decir, un conjunto de neuronas conectadas entre sí y que trabajan en conjunto, sin que haya una tarea concreta para cada una. Con la experiencia, las neuronas van creando y reforzando ciertas conexiones para "aprender" algo que se queda fijo en el tejido.

Una red neuronal artificial se compone de unidades denominadas neuronas, cada neurona recibe una serie de entradas a través de interconexiones y emite una salida. Esta salida viene dada por tres funciones:

Una *función de propagación (función de excitación)*, que en general consiste en el sumatorio de cada entrada multiplicada por su valor neto. Si el valor es positivo, la conexión se denomina excitatoria; si es negativo, se denomina inhibitoria.

Una *función de activación*, la cual modifica a la anterior. Puede no estar presente, en este caso la salida será la misma función de propagación.

Una *función de transferencia*, que se aplica al valor de la función de activación. Se utiliza para acotar la salida de la neurona y generalmente viene dada por la interpretación que queramos darles a dichas salidas. Algunas de las más utilizadas son la función sigmoidea para obtener valores en el intervalo (0,1) y la tangente hiperbólica para obtener valores en el intervalo (-1,1).

La unidad fundamental de la red neuronal se denomina perceptrón (ver figura 2.4). Un perceptrón es un elemento que tiene varias entradas con un cierto peso cada una. Si la suma de esas entradas por cada peso es mayor que un determinado número, la salida del perceptrón es un uno. Si es menor, la salida es un cero.

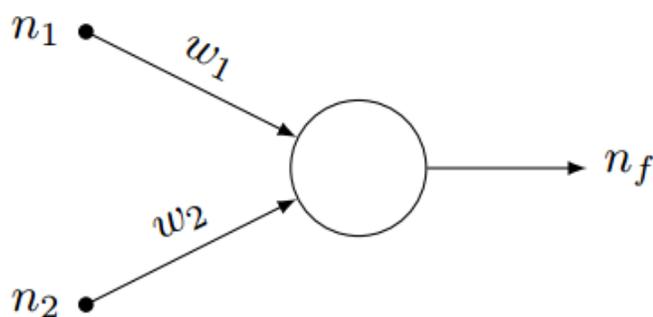


Figura 2.4.- Diagrama de un perceptrón (Pino R. et al, 2004).

Las aplicaciones de las redes neuronales son diversas, se aplican desde el control de sistemas, al análisis de datos y en algunas ocasiones en la predicción de cierta información, la dificultad de estas predicciones está en el grado de detalle y la cantidad de variables a analizar, y por supuesto de la necesidad de contar con amplios rangos de tiempo de previsión.

La literatura sugiere fundamentalmente dos características de las RNAs que las hacen especialmente interesantes: la capacidad de las RN de aproximar prácticamente cualquier función (incluso las no lineales), y la posibilidad de hacer aproximaciones “piece-wise” o por trozos, de las funciones (Pino R. et al, 2004).

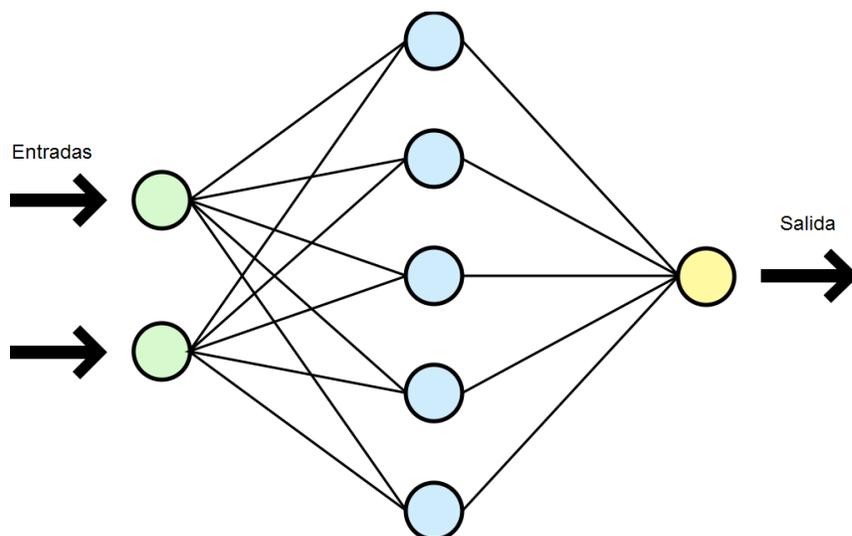


Figura 2.5.- Esquema de una red neuronal artificial.

En la figura 2.5, los círculos representan neuronas y las líneas representan conexiones entre las neuronas. Los círculos verdes representan la capa de entradas, los círculos azules las capas ocultas y el círculo amarillo la capa de salidas.

La mayor ventaja al realizar modelación de las redes neuronales es la capacidad que éstas poseen para aproximar problemas con comportamientos y funciones no lineales, gracias a la forma en que sus datos de entrada (ver figura 2.5) sean tratados, la forma misma de operar de la red (es decir, la forma en que realizan todas las operaciones y cálculos las capas de la red) y sus datos de salida, algo que los modelos ARIMA logran hacer, pero aplicando otras medidas (Collantes J., 2001).

Un punto en contra de las redes neuronales es la existencia de un número no determinado de modelos que se ajustan con la precisión deseada a los datos de entrenamiento, pero pocos de ellos logran generalizar bien estos datos fuera del conjunto de datos de entrenamiento, y se tiene también el efecto de los fuertes ruidos y el hecho de que unos conjuntos de datos de entrenamiento pequeño hacen que el modelo no sea preciso (Collantes J., 2001).

Los modelos cimentados en redes neuronales artificiales no reconocen la relación temporal explícita entre los datos de entrada y los de salida, es decir, no reconocen entre las correlaciones que ocurren en orden temporal y las que no (Collantes J., 2001).

## **2.4 Relación entre los modelos de clima y la determinación del potencial eólico.**

Con todos los elementos descritos anteriormente y la bibliografía existente, es posible discernir que existen muchos estudios y metodologías para la determinación de los diversos parámetros atmosféricos, los cuales sustentan que todo proyecto que se desee realizar para la evaluación del recurso eólico mediante modelos es viable.

Existe una convergencia entre los modelos y los estudios estadísticos, es decir, que siempre serán necesarias las mediciones en campo, pero los modelos de predicción son una gran herramienta en la toma de decisiones. Muchos modelos complejos son capaces incluso de calcular la exposición de cada aerogenerador en el modelo a vientos que varían en el espacio y el tiempo.

Los modelos de clima poseen grandísimas ventajas, como obtener un pronóstico acotado u ofrecer parámetros de comparación, sin embargo, es conveniente recordar que la gran mayoría solo son aplicables a mesoescala, que si bien es cierto brinda una información muy amplia, no es del todo útil en la toma de decisiones de peso.

Las principales ventajas de utilizar la metodología de Box-Jenkins para la predicción de series de tiempo radican en su esencia heurística intrínseca. Es decir, no es necesario que la serie de tiempo cumpla con los supuestos teóricos como la normalidad, la independencia del ruido, la media cero, etc. Tampoco es necesario reconocer o asumir estacionalidad dentro de la serie. Su modelamiento suele ser más sencillo y general que el de otros modelos estadísticos. Por norma general se tiene mayor libertad en cuanto al manejo de parámetros y el mejor ajuste de la serie depende en mayor medida de la habilidad y raciocinio de quien la planteó (Pino R. et al, 2004).

Es importante mencionar brevemente que, en el análisis de series de tiempo, la metodología de Box-Jenkins, nombrada así en honor a los estadísticos George Box y Gwilym Jenkins, se aplica a los modelos autorregresivos de media móvil ARMA o a los

modelos autorregresivos integrados de media móvil (ARIMA) para encontrar el mejor ajuste de una serie temporal de valores, a fin de que los pronósticos sean más acertados. La metodología de Box y Jenkins aplica métodos autorregresivos, promedio móviles y autorregresivo integrados de promedio móvil. (Collantes J., 2001).

Esencialmente, esta metodología se compone de tres etapas de modelado iterativo de los datos y son:

- Identificación y selección del modelo, asegurarse de que las variables son estacionarias (diferenciando en temporada si es necesario).
- Estimación de parámetros usando algoritmos de cálculo para tener coeficientes que mejor se ajustan al modelo ARIMA seleccionado.
- Comprobar el modelo mediante el ensayo, si el modelo estimado se ajusta a las especificaciones de un proceso univariado estacionario. En particular, los residuos deben ser independientes el uno del otro y la media y la varianza constante en el tiempo. Si la estimación es suficiente, tenemos que volver al paso uno y el intento de construir un modelo mejor.

Finalmente, se puede afirmar que la predicción sustentada en series de tiempo resulta de gran interés, ya que permite conocer, con cierto valor de error asignado, los valores futuros de una serie conociendo sus valores pasados, esto resultaría muy útil sobre todo en la toma de decisiones, tanto en el ámbito de la evaluación de viabilidad de los parques eólicos, así como en la integración de sistemas híbridos de generación de energía eléctrica.

## CAPÍTULO III

# PREDICCIÓN DE SERIES DE TIEMPO DE VELOCIDAD DEL VIENTO.

### 3.1 Series de tiempo.

El termino serie de tiempo, hace referencia a un conjunto de datos estadísticos que se recopilan, observan o registran en intervalos de tiempo regulares (diario, semanal, semestral, anual, entre otros). El término serie de tiempo se aplica a datos registrados en forma periódica que muestran (de Arce R. et al, 2001). Algunos ejemplos de series de tiempo son datos de temperatura, precipitación, velocidad del viento, etc. (ver figura 3.1)

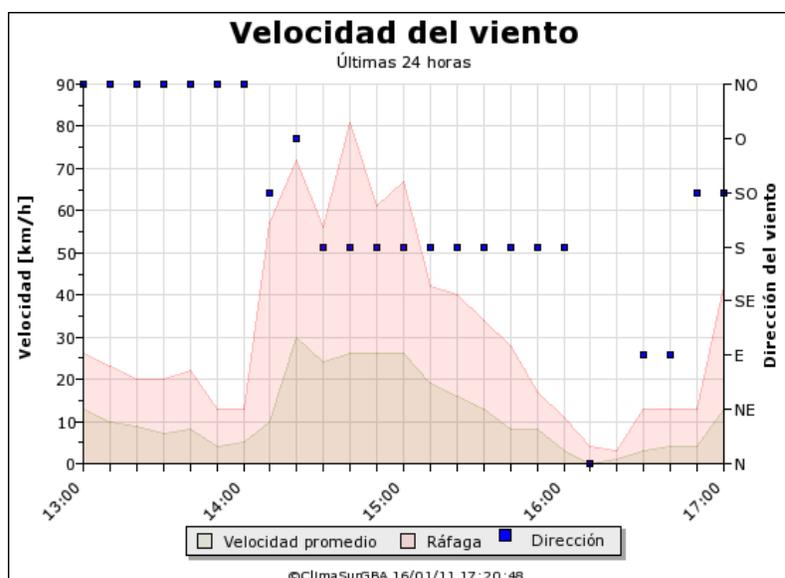


Figura 3.1.- Ejemplo de una serie de tiempo, velocidad y dirección del viento a través un periodo (ClimaSurGBA, 2016).

En términos generales, para las series de tiempo suponemos que existen cuatro tipos básicos de variación, los cuales puede sobreponerse o actuar en conjunto, y contribuyen a los cambios observados en un período de tiempo y dan a la serie su aspecto errático. Estas **cuatro componentes** (ver tabla 3.1) son: *tendencia secular*, *variación estacional*, *variación cíclica* y *variación irregular*.

Es bastante común suponer que además existe una relación multiplicativa entre estas cuatro componentes; es decir, cualquier valor de una serie es el producto de factores que se pueden atribuir a las cuatro componentes (de Arce R. et al, 2001).

*Tendencia secular*, la tendencia secular o tendencia a largo plazo de una serie es por lo común el resultado de factores a largo plazo (ver tabla 3.1). En términos intuitivos, la tendencia de una serie de tiempo caracteriza el patrón gradual y la consistencia de las variaciones de la propia serie, esta se considera una “fuerza” persistente que afecta el crecimiento o la reducción de la misma serie, tales como: cambios en la población, en las características demográficas de la misma, cambios en los ingresos, en la salud, en el nivel de educación y tecnología. Las tendencias a largo plazo se ajustan a diversos esquemas. Algunas se mueven continuamente hacia arriba, otras declinan, y otras más permanecen igual en un cierto período o intervalo de tiempo (de Arce R. et al, 2001).

*Variación estacional*, la componente de la serie de tiempo que representa la variabilidad en los datos debida a influencias de las estaciones, se llama componente estacional. Esta variación corresponde a los movimientos de la serie que recurren año tras año en los mismos meses (o en los mismos trimestres) del año poco más o menos con la misma intensidad (de Arce R. et al, 2001). Por ejemplo: Las más altas temperaturas que se observan en verano.

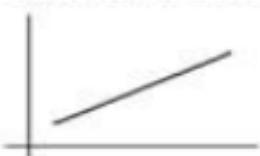
*Variación cíclica*, con frecuencia las series de tiempo presentan secuencias alternas de puntos abajo y arriba de la línea de tendencia que duran más de un año, esta variación se mantiene después de que se han eliminado las variaciones o tendencias estacional e irregular (de Arce R. et al, 2001). Un ejemplo es el clima ya que es una variación son los ciclos cuyos períodos no dependen de otros factores.

*Variación Irregular*, esta se debe a factores a corto plazo, imprevisibles y no recurrentes que afectan a la serie de tiempo. Como esta componente explica la variabilidad aleatoria de la serie, es impredecible, es decir, no se puede esperar predecir su impacto sobre la serie de tiempo (de Arce R. et al, 2001). Existen dos tipos de variación irregular:

a) Las variaciones que son provocadas por acontecimientos especiales, fácilmente identificables, como las inundaciones, tifones, terremotos.

b) Variaciones aleatorias o por casualidad, cuyas causas no se pueden señalar en forma exacta, pero que tienden a equilibrarse en periodos largos.

Tabla 3.1.- Características y ejemplos de las componentes de una serie de tiempo.

Nombre y forma	Ejemplo	Características	Técnica a usar
Tendencia (T). 	Ventas a largo plazo. Precios. Empleos	Se mide en años.	Mínimos cuadrados.
Fluctuaciones cíclicas (C). 	Empleo. Producción. Precio de acciones.	Se mide en años. Ascenso y descenso en periodos mayores. Períodos de prosperidad.	Valores cíclicos relativos.
Variaciones estacionales (E). 	Consumo relacionado con estaciones del año.	Solo se aprecia en datos trimestrales, mensuales. Patrones de cambio.	Promedios móviles.
Variaciones irregulares (I) 	Guerras. Huelgas. Desastres naturales.	No se pueden predecir o medir.	No existe.

Además, las series de tiempo **poseen tendencias**, las cuales pueden ser:

*Tendencia lineal*, la tendencia de una serie viene dada por el movimiento general a largo plazo de la serie. La tendencia a largo plazo de muchas series de negocios (industriales y comerciales), como ventas, exportaciones y producción, con frecuencia se aproxima a una línea recta.

Esta línea de tendencia muestra que algo aumenta o disminuye a un ritmo constante. El método que se utiliza para obtener la línea recta de mejor ajuste es el Método de Mínimos Cuadrados (de Arce R. et al, 2001).

*Tendencia no lineal*, cuando la serie de tiempo presenta un comportamiento curvilíneo se dice que este comportamiento es no lineal. Dentro de las tendencias no lineales que pueden presentarse en una serie se encuentran, la polinomial, logarítmica, exponencial y potencial, entre otras (de Arce R. et al, 2001). Finalmente, para llevar a cabo un análisis de una serie de tiempo es necesario aplicar un método de suavizado, el cual nos permite obtener la información más relevante relacionada con las características intrínsecas de cada serie.

*Promedio móvil*, un promedio móvil se construye sustituyendo cada valor de una serie por la media obtenida con esa observación y algunos de los valores inmediatamente anteriores y posteriores (de Arce R. et al, 2001). Se considerará el promedio móvil (ecuación 3.1) a partir, por ejemplo, de las tres observaciones más recientes. En este caso se utilizará la siguiente ecuación:

$$\overline{X_{movil}} = \frac{\Sigma(n \text{ valores más recientes de datos})}{n} \quad 3.1$$

Los promedios móviles también se pueden construir tomando en cuenta valores adyacentes de las observaciones.

*Promedios móviles ponderados*, consiste en asignar un factor de ponderación distinto para cada dato. Generalmente, a la observación o dato más reciente a partir del que se quiere hacer el pronóstico (ecuación 3.2), se le asigna el mayor peso, y este peso disminuye en los valores de datos más antiguos (de Arce R. et al, 2001).

$$Pronóstico = \frac{x}{y}(a) + \frac{2x}{y}(b) + \frac{3x}{y}(c) \quad 3.2$$

Donde:

a, b y c: Parámetros con valores diferentes, representarán los pesos asignados.

x: Observación (1).

y: Observación siguiente (2).

Puede observarse que el dato más alejado (correspondiente al primer valor) deberá tener el factor de ponderación más pequeño, el siguiente tiene un factor de ponderación del doble que el primero y el dato más reciente (que corresponde al tercer término) tiene un factor de ponderación del triple del primero.

*Suavizamiento exponencial*, emplea un promedio ponderado de la serie de tiempo pasada como pronóstico; es un caso especial del método de promedios móviles ponderados en el cual sólo se selecciona un peso o factor de ponderación: el de la observación más reciente. En la práctica comienza haciendo que  $F_1$ , el primer valor de la serie de valores uniformados, sea igual a  $Y_1$ , que es el primer valor real de la serie (de Arce R. et al, 2001). El modelo básico de suavizamiento exponencial (ecuación 3.3) es el siguiente:

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t \quad 3.3$$

Donde:

$F_{t+1}$ : Pronóstico de la serie de tiempo para el periodo t+1.

$Y_t$ : Valor real de la serie de tiempo en el periodo t.

$F_t$ : Pronóstico de la serie de tiempo para el periodo t.

$\alpha$ : Constante de suavizamiento ( $0 \leq \alpha \leq 1$ ).

### **3.2 Modelos autorregresivos y de media móvil.**

En 1970, Box y Jenkins desarrollaron un cuerpo metodológico destinado a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo juega un papel fundamental. Una parte importante de esta metodología está pensada para liberar al investigador de la tarea de especificación de los modelos dejando que los propios datos temporales de la variable a estudiar nos indiquen las características de la estructura probabilística subyacente. En parte, los procedimientos que vamos a analizar se contraponen a la "forma tradicional" de identificar y especificar un modelo apoyándonos en las teorías subyacentes al fenómeno analizado, aunque, convenientemente utilizados, los conceptos y procedimientos que examinaremos

constituyen una herramienta útil para ampliar y complementar los conocimientos econométricos básicos (de Arce R. et al, 2001).

Para describir los modelos más complejos, primero es necesario describir las partes que lo componen.

*Proceso autoregresivo (AR)*, los modelos autoregresivos se sustentan en el hecho de que el valor actual de la serie  $X_t$ , puede explicarse en función de  $p$  valores pasados  $X_{t-1}, X_{t-2}, \dots, X_{t-p}$ , donde  $p$  determina el número de rezagos necesarios para pronosticar un valor actual.

El modelo autoregresivo de orden  $p$  (ecuación 3.4) está dado por:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad 3.4$$

Donde:

$X_{t-n}$ : Variable a analizar y sucesiones.

$\varepsilon_t$ : Proceso de ruido blanco con media cero y varianza constante.

$\phi$ : Parámetros del modelo AR.

Lo anterior puede ser expresado en términos del operador de retardos (ecuación 3.5).

$$(1 - \theta_1 L - \theta_1 L^2 - \dots - \theta_p L^p) X_t = \varepsilon_t \quad 3.5$$

$$\theta_p(L) X_t = \varepsilon_t$$

Donde:

$X_{t-n}$ : Variable a analizar y sucesiones.

$\varepsilon_t$ : Proceso de ruido blanco con media cero y varianza constante.

$\phi$ : Parámetros del modelo AR.

$L$  : La condición de estacionariedad del modelo AR, es  $|L| = \left| \frac{1}{\phi} \right| > 0$ , si esta condición se cumple el proceso es estacionario, la función de autocovarianza de un proceso AR

estacionario es  $p_k = \begin{cases} 1 & k = 0 \\ \phi p_{k-1} & k > 0 \end{cases}$

Una forma alterna de escribir el modelo AR (ecuación 3.6) es:

$$X_t = \phi X_{t-1} + \varepsilon_t \quad 3.6$$

Donde:

$X_t$ : Variable a analizar.

$\varepsilon_t$ : Proceso de ruido blanco con media cero y varianza constante.

$\phi$ : Parámetro del modelo AR.

*Proceso de media móvil (MA)*, los modelos que incluyen medias móviles determinan el valor de  $X_t$  en función de la innovación actual y su primer retardo.

$$X_t = \varepsilon_t - \theta \varepsilon_{t-1} \quad 3.7$$

Donde:

$X_t$ : Variable a analizar.

$\varepsilon_t$ : Proceso de ruido blanco con media cero y varianza constante.

$\theta$ : Parámetro del modelo MA.

Lo ecuación anterior 3.5, puede ser expresada en función del polinomio de retardos que es:

$$X_t = (1 - \theta)\varepsilon_t \quad 3.8$$

$$X_t = \theta_1(L)\varepsilon_t$$

Donde:

$X_t$ : Variable a analizar.

$\varepsilon_t$ : Proceso de ruido blanco con media cero y varianza constante.

$\theta$ : Parámetro del modelo MA.

$L$  : La condición de invertibilidad del modelo MA, es  $|L| = \left| \frac{1}{\theta} \right| > 0$ .

Los modelos de series de tiempo, se basan en el supuesto de estacionalidad, esto es, la media y la varianza para una determinada serie de tiempo son constantes en el tiempo y la covarianza es invariante en el tiempo.

Dentro de estos modelos univariantes existen muchos que se han desarrollado a lo largo de los años, sin embargo, entre los más populares se desarrollaron suficiente los conocidos con el nombre de ARIMA (p, q, d) donde “p” denota el número de términos autoregresivos, “d” el número de veces que la serie debe ser diferenciada para hacerla estacionaria y “q” el número de términos de la media móvil invertible. Posteriormente se complementará esta perspectiva univariante añadiéndose a la especificación una o más variables exógenas al modelo "tradicional" aproximándonos al estudio de los conocidos como modelos de transferencia.

Se sabe que muchas series de tiempo y en especial las series económicas no son estacionarias, porque pueden ir cambiando de nivel en el tiempo o sencillamente la varianza no es constante en el tiempo, a este tipo de proceso se les considera procesos integrados. Por lo tanto, se debe diferenciar una serie de tiempo “d” veces para hacerla estacionaria y luego aplicar a esta nueva serie diferenciada un modelo (Villavicencio J., 2010). La expresión algebraica queda:

$$X_t^d = c + \underbrace{\phi_1 X_{t-1}^d + \dots + \phi_p X_{t-p}^d}_{\text{AR (p)}} + \underbrace{\theta_1 \varepsilon_{t-1}^d + \theta_2 \varepsilon_{t-2}^d + \dots + \theta_q \varepsilon_{t-q}^d}_{\text{MA (q)}} + \varepsilon_t^d \quad 3.9$$

Expresado en forma del polinomio operador de retardos del modelo ARIMA (p, q, d) queda:

$$\phi(L)(1 - L)^d X_t = c + \Theta(L)\varepsilon_t \quad 3.10$$

Donde:

$X_t^d$ : es la serie de tiempo de las diferencias de orden “d”.

$\varepsilon_t^d$ : es un proceso de ruido blanco.

$c, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ : son los parámetros del modelo.

El modelo ARIMA (p, d, q) se puede representar:

$$X_t = -(\Delta^d X_t - Y_t) + \phi_0 + \sum_{i=1}^p \phi_i \Delta^d X_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad 3.11$$

con  $\Delta X_t = X_t - X_{t-1}$

Donde:

$X_t$ : Serie de la variable.

$d$ : corresponde a las  $d$  diferencias que son necesarias para convertir la serie original en estacionaria.

$\phi_1, \dots, \phi_p$ : Parámetros pertenecientes a la parte autoregresiva.

$\theta_1, \dots, \theta_q$ : Parámetros pertenecientes a la parte de medias móviles.

$\phi_0$ : Constante

$\varepsilon_t$ : Término de error, llamado innovación o perturbación estocástica.

Así, la construcción de los modelos ARIMA (p, q, d) se lleva a cabo de manera iterativa mediante un proceso en el que se puede distinguir cuatro etapas:

*Identificación*, utilizando los datos ordenados cronológicamente se intentará sugerir un modelo ARIMA (p, q, d) que merezca la pena ser investigado. El objetivo es determinar los valores que sean apropiados para reproducir la serie de tiempo. En esta etapa es posible identificar más de un modelo candidato que pueda describir la serie (Villavicencio J., 2010).

*Estimación*, considerando el modelo apropiado para la serie de tiempo se realiza inferencia sobre los parámetros (Villavicencio J., 2010).

*Validación*, se realizan contrastes de diagnóstico para validar si el modelo seleccionado se ajusta a los datos, si no es así, escoger el próximo modelo candidato y repetir los pasos anteriores (Villavicencio J., 2010).

*Predicción*, una vez seleccionado el mejor modelo candidato ARIMA (p, q, d) se pueden hacer pronósticos en términos probabilísticos de los valores futuros (Villavicencio J., 2010).

En ocasiones se tiene series de tiempo que poseen una componente estacional que se repite cada “s” observaciones, por ejemplo, s=12 (12 en 1 año), para poder realizar un análisis de este tipo de series se emplea una modificación al modelo ARIMA (p, q, d) y se le denomina Seasonal ARIMA (p, q, d) o SARIMA (p, q, d). Este comportamiento se aprecia comúnmente en series no estacionarias. La formulación completa de este modelo es:

$$\Phi(B)\Delta^d X_t = \theta(B)\alpha_t \quad 3.12$$

Donde:

$\alpha_t$ : puede analizarse como  ${}_s\Phi(B^s)\Delta_s^D\alpha_t = {}_s\Theta(B^s)\alpha_t$  Y, por lo tanto  $\Phi(B)_s\Phi(B^s)\Delta_s^D\Delta^d X_t = \theta(B)_s\Theta(B^s)\alpha_t$  con lo cual se puede denotar a estos modelos como  $X_t \sim \text{ARIMA}(p, q, d) \times (P, D, Q)_s$ .

Para simplificar la idea detrás de todo esto, los modelos SARIMA son modelos ARIMA (p, d, q) cuyos residuos  $\alpha_t$  se expresan en un nuevo modelo ARIMA (P, D, Q), pero que para la notación se simplifican empleando  $B^s$ .

En general, se dice que los modelos SARIMA y las sucesivas recursiones se emplean en zonas inadmisibles para los modelos ARIMA (p, d, q) pero que solo son expresadas con la componente  $B^s$  y todas sus potencias posibles (de Arce R. et al, 2001).

A continuación, se detallará la notación en la que los algoritmos y/o los softwares computacionales suelen expresar tanto los resultados como las operaciones para todos estos modelos.

1.  $X_t = \alpha_t - \theta_1 \alpha_{t-12}$  esto es la representación de un ARIMA (0, 0, 0) X (0, 0, 1)<sub>12</sub>. En esta solo existe una componente estacional MA, la cual esta señala por s=12, Q=1. Por lo tanto, los pasos k=12 señalan la extensión, Finalmente el decaimiento esta expresado por  $\theta_1$ .
2.  $X_t = {}_s\Phi_1 X_{t-12} + \alpha_t$  esto es un modelo ARIMA (0, 0, 0) X (1, 0, 0)<sub>12</sub>. Esto representa una componente AR estacional definida por s=12 y P=1. Con los pasos k=12.

### 3.3 Algoritmo de predicción.

En esta sección se describe el algoritmo a emplear en este estudio, el cual es un experimento puro, explicativo y con enfoque cuantitativo ya que mediante datos recolectados se contrastan datos y resultados de una simulación.

Es importante notar que se trata de una metodología experimental, debido a que se pueden manipular las variables y que los datos serán no estadísticos.

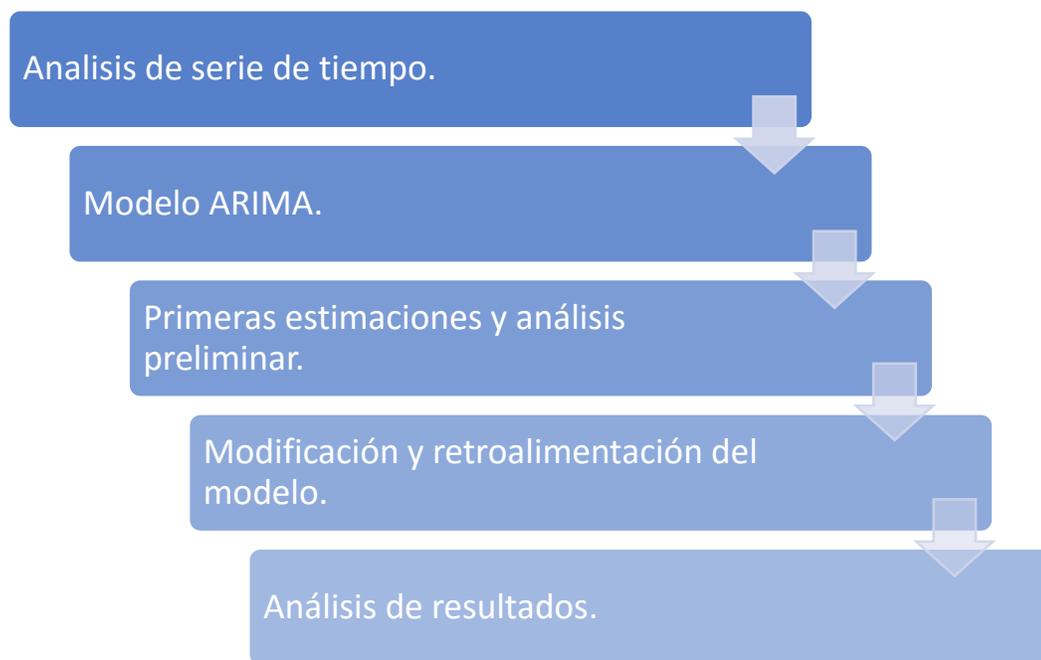


Figura 3.2.- Esquema de metodología.

Para la realización del análisis fue necesario evaluar el comportamiento de la serie de tiempo, calcular los coeficientes para la aplicación del modelo ARIMA, corroborar el correcto funcionamiento del modelo, además de considerar todos los factores que se pueden calcular con un mismo modelo, también es necesario verificar la correcta disposición de las bases de datos, es decir, que contengan datos correctos y no errores de medición relacionados al instrumento con el que se mide (ruido o errores), ya que en caso de tener estos errores es necesario filtrarlos.

Para el análisis se empleará el término de velocidad “Wind Run” que se proporciona en la base de datos, esto, debido a que este término (Wind Run) en m/s está relacionado con el valor promedio de la velocidad del viento sobre el tiempo de

muestreo que emplea la estación, es decir, es un valor que queda expresado como:

$$WR = \frac{\mu WS}{\text{tiempo de muestreo}} \quad 3.13$$

Donde:

$\mu WS$ : Promedio de la velocidad del viento en el emplazamiento (m/s).

A continuación, mediante iteraciones (ver figura 3.3) se procede a obtener los coeficientes necesarios para la aplicación del modelo ARIMA.

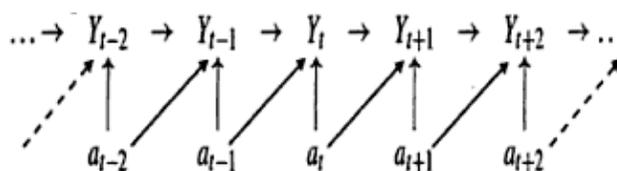


Figura 3.3.- Proceso iterativo para la obtención de los coeficientes del modelo ARIMA, en el ejemplo ARIMA (1, 0, 1).

Seguidamente, se procede a realizar unas primeras estimaciones (ver figura 3.4), empleando periodos de tiempo de entre una a dos horas considerando también la resolución de las velocidades medidas (que tengan órdenes de magnitud correctos y se encuentren expresadas en m/s), esta será definida como el “área efectiva” en la cual se puede aplicar el modelo.

El siguiente paso es comparar los resultados obtenidos mediante la estimación con los históricos medidos (base de datos con periodos de medición de 10 minutos) por las estaciones climatológicas (FIUADY y CINVESTAV) en el mismo periodo que se estimó, con el fin de validar la estimación.

Determinar los errores, la desviación y la dispersión de los datos estimados al contrastarlos con los datos de las estaciones, esto para mejorar la operación intrínseca del modelo y adecuarlo a las condiciones locales. Este paso será realizado mediante el software R Studio (el cual es un lenguaje y entorno de programación para análisis estadístico y gráfico), el cual es una herramienta que será útil en la determinación y análisis de la varianza en una serie de datos.

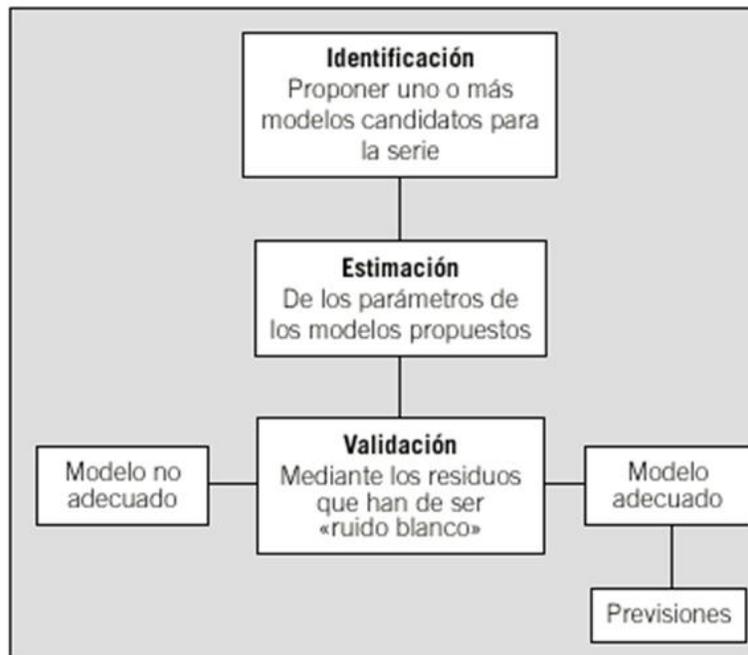


Figura 3.4.- Diagrama de operación para un modelo específico, en este ejemplo ARIMA.

Retroalimentar el modelo con los datos obtenidos de la comparación directa con las estaciones climatológicas y resultado del análisis estadístico.

Efectuar pruebas con las correcciones aplicadas.

Realizar nuevas estimaciones finales con los errores, correcciones aplicadas y utilizarlas para el análisis de datos.

Es importante mencionar que el programa Mathematica® el cual es un programa utilizado en áreas científicas, de ingeniería, matemáticas y áreas computacionales. Originalmente fue concebido por Stephen Wolfram, quien continúa siendo el líder del grupo de matemáticos y programadores que desarrollan el producto en Wolfram Research, la versión 10 del software introduce soporte para series temporales (figura 3.5), en todo el sistema, ya sea que se originen en Finanzas, Economía, Procesamiento de señales, Acondicionamiento físico u otras áreas. La lectura de datos desde dispositivos conectados y el acceso a conocimiento desde datos son ricas fuentes de datos de series temporales.

Así, este software es capaz de analizar mediante cualquier modelo de predicción, sin embargo, para la realización de este proyecto se empleará el modelo ARIMA, por lo descrito en la sección anterior.

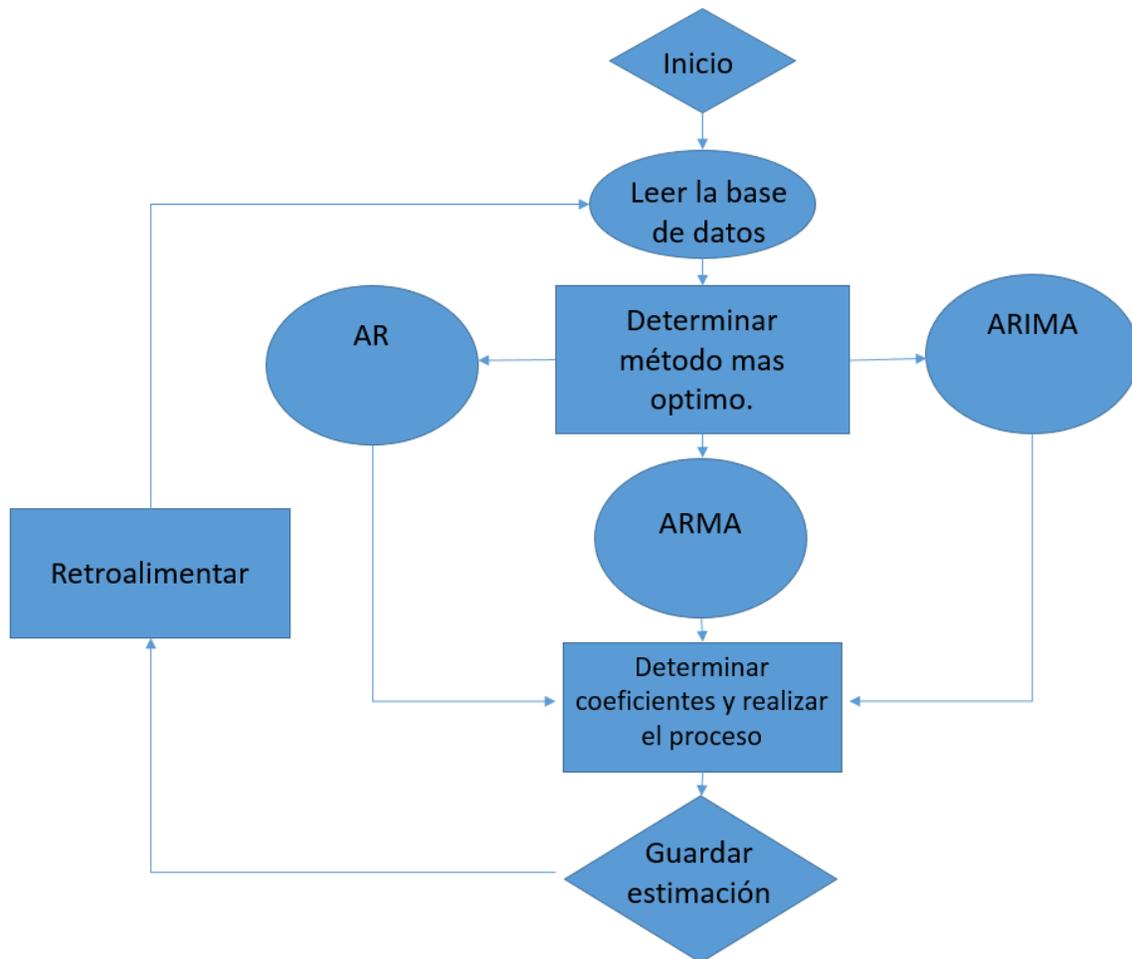


Figura 3.5.- Diagrama de flujo para la aplicación del modelo.

## CAPÍTULO IV

# CONFRONTACIÓN DE RESULTADOS, ESTIMACIONES VERSUS MEDICIONES.

### 4.1 Datos de estaciones meteorológicas.

Las mediciones realizadas por las estaciones meteorológicas se emplean como datos de entrada para prácticamente todas las estimaciones de potencial de las energías renovables, como se mencionó antes; para el caso del potencial eólico normalmente se emplea la distribución de Weibull.

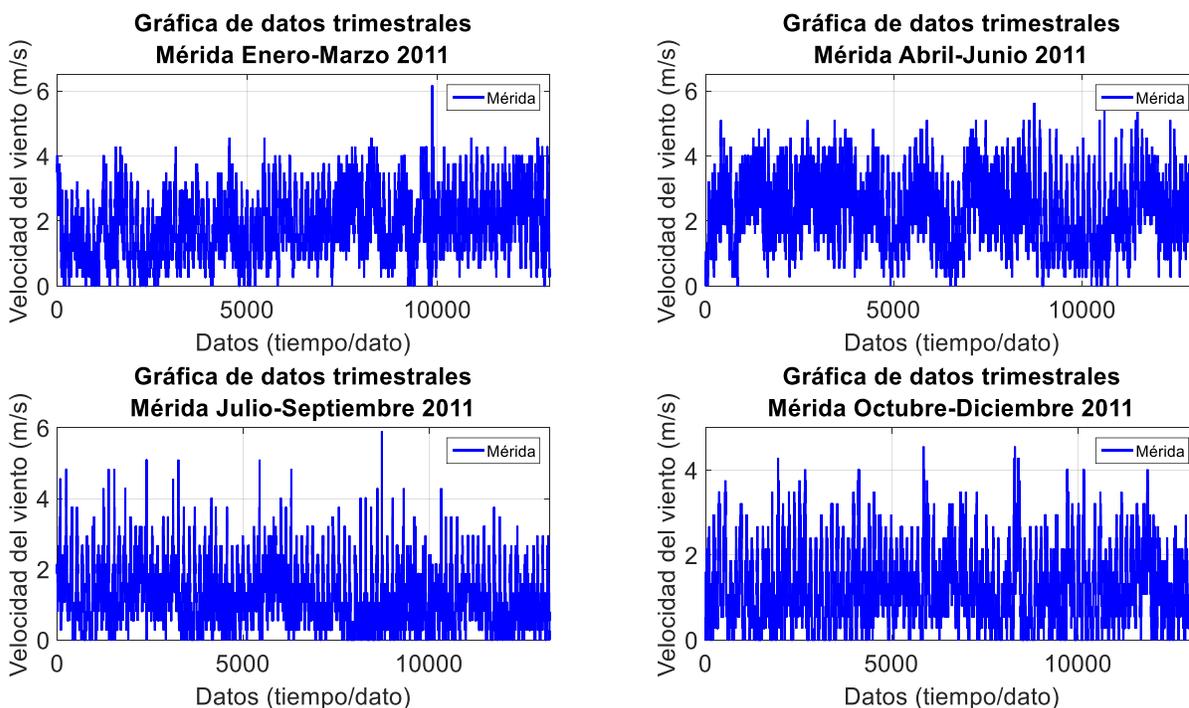


Figura 4.1.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2011.

En esta primera sección se aplica esta distribución, como primer paso, se emplearon como entradas la misma cantidad de datos (figura 4.1 a 4.4), se emplearon todos los datos de cuatro años consecutivos, para datos de solo primavera, solo verano, solo otoño y finalmente solo invierno. A continuación, se presentan los resultados de este primer análisis.

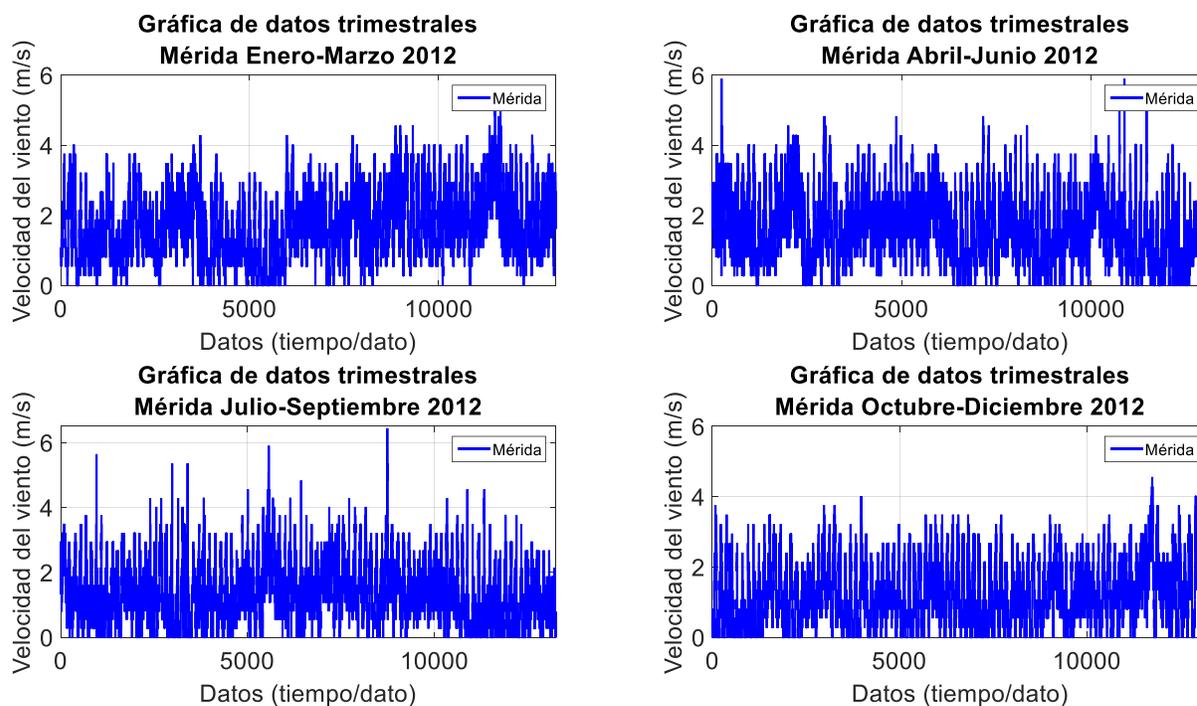


Figura 4.2.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2012.

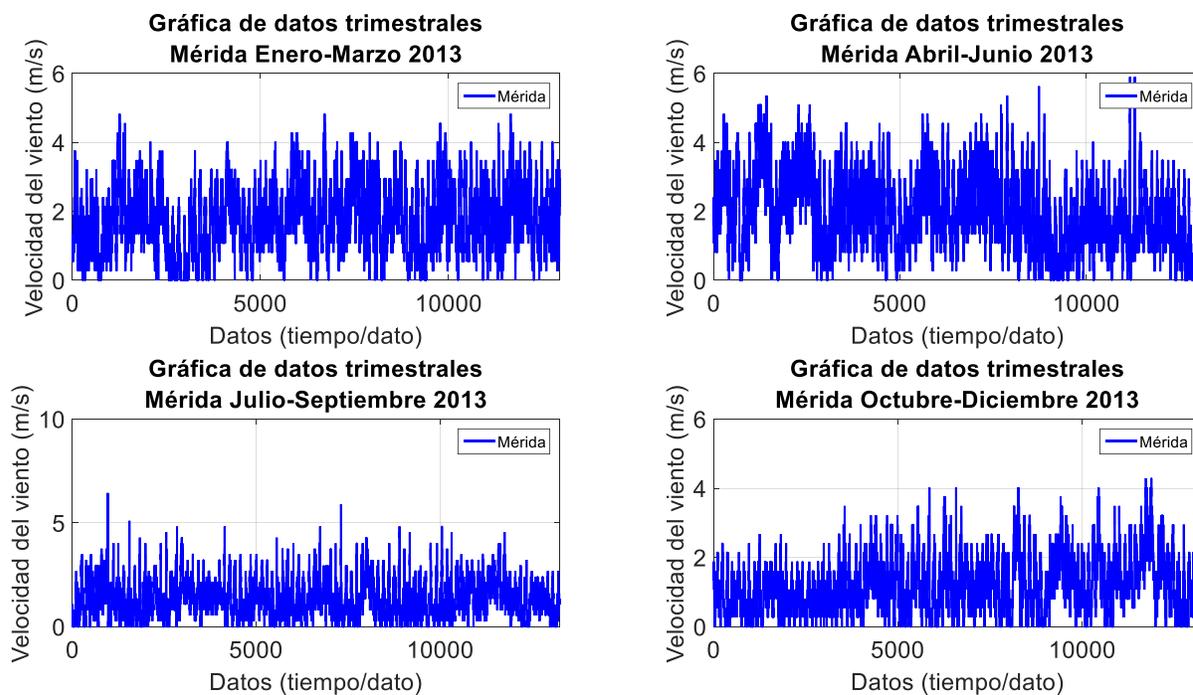


Figura 4.3.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2013.

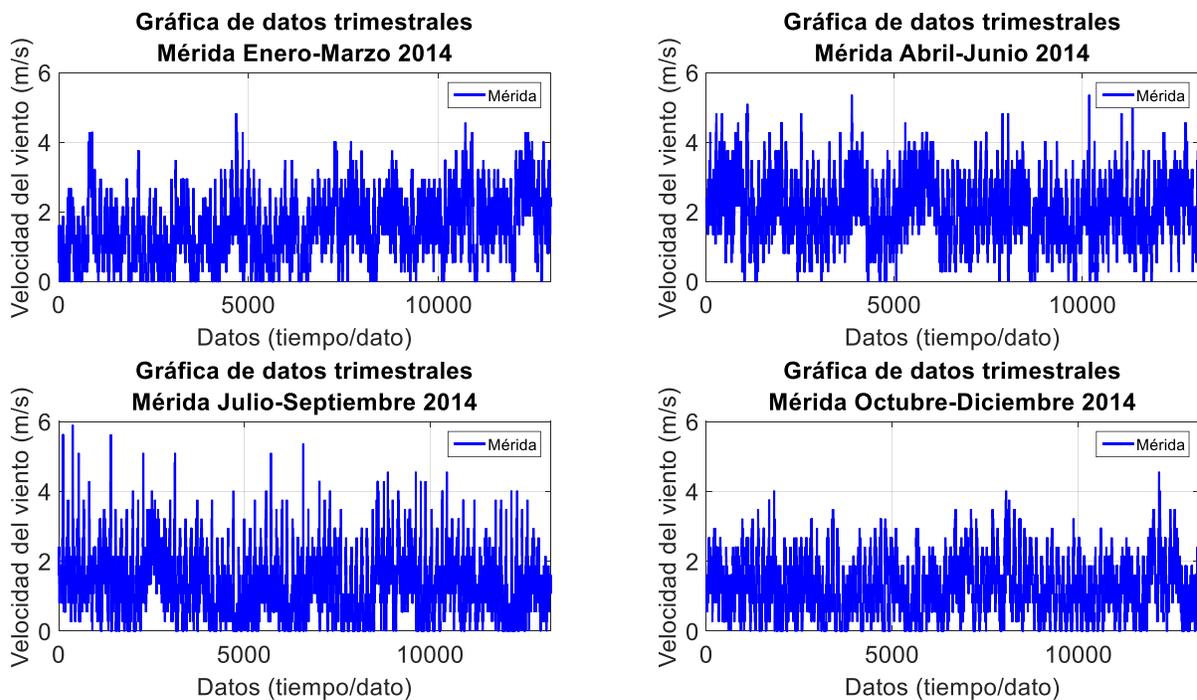


Figura 4.4.- Gráfica de datos de velocidad de viento (m/s) para Mérida 2014.

En la figura 4.5, se muestra la distribución de Weibull para los distintos casos planteados. Lo primero que llama la atención es la variación que ocurre en primavera ya que está presente un incremento del 25.16 % en la velocidad promedio de esa estación, además podemos observar un incremento tanto en la probabilidad como en la velocidad promedio; sin embargo, en términos generales en las estaciones restantes está presente un comportamiento similar.

La grafica además presenta los valores para los factores de forma y escala para cada ajuste, donde vemos que también varía muy poco con excepción de la primavera que tiene valores de  $K=2.05$  y  $C=2.59$  u.

Se observa que con estos datos la velocidad más probable es menor a 1 m/s con mediciones a 10 metros (proporcionada por la estación) y con tiempos de registros cortos (cada 10 minutos), además la distribución de las velocidades del viento es sesgada. En ocasiones se tendrán velocidades de viento “altas”, pero son muy raras.

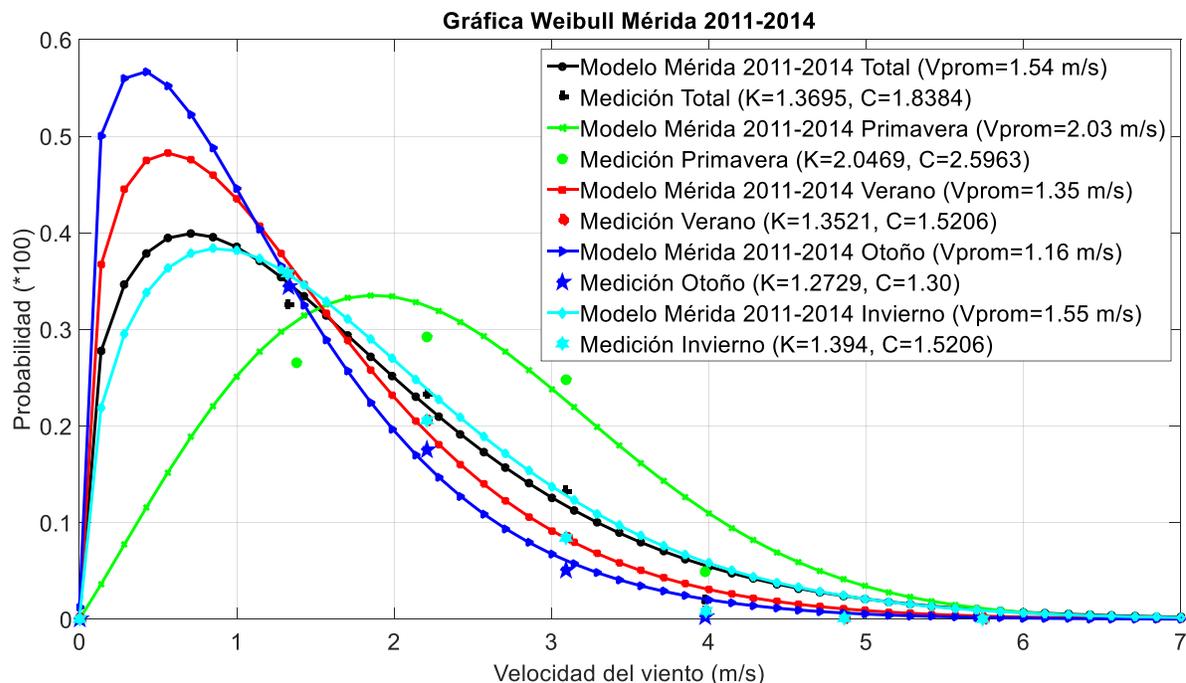


Figura 4.5.- Distribución de Weibull, Mérida 2011-2014.

Por otro lado, las velocidades del viento comprendidas entre 1 y 2 m/s son las más comunes. Se observa que con estos datos la velocidad más probable es menor a 1 m/s con mediciones a 10 metros (proporcionada por la estación) y con tiempos de registros cortos (cada 10 minutos), además la distribución de las velocidades del viento es sesgada. En ocasiones se tendrán velocidades de viento “altas”, pero son muy raras. Por otro lado, las velocidades del viento comprendidas entre 1 y 2 m/s son las más comunes. En términos generales, se puede decir que el comportamiento de la distribución de velocidades durante todo el año es similar al de cada estación del año, y solo cambia significativamente durante la primavera.

## 4.2 Primeras estimaciones.

En esta sección, se describen los resultados de las primeras estimaciones realizadas de la aplicación de los modelos, así como la comparación entre sus efectos. Se sometieron a prueba los tres modelos de predicción más utilizados AR, ARMA y ARIMA, realizando estimaciones cada diez minutos y teniendo como límite superior hasta dos horas de datos.

Para la correcta aplicación de estos modelos es necesario plantear y cumplir los siguientes supuestos:

Tabla 4.1.- Supuestos para la aplicación de los modelos autorregresivos y de media móvil (Hernández J. et al, 2013).

$Y_t = \delta + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \alpha_t$		<b>Hipótesis auxiliares (at)</b>	
<b>Hipótesis auxiliares (at)</b>		Estacionariedad	> Yt, media y varianza constante
Normalidad	> N	Linealidad	> Yt, función lineal
Media nula	> 0	Muestra suficiente	> Datos suficientes en cantidad
Homocedasticidad	> $\sigma^2$	Sin errores en los datos	> y calidad
No autocorrelacion	> independientes	Constancia en los parámetros	> Guarden la misma relación

Como se muestra en la tabla 4.1, la distribución de las velocidades tiene una distribución de errores normal, una media nula, la varianza del error de la variable se mantiene a lo largo de las observaciones (homocedasticidad). Todas las bases de datos presentan estas características y por lo tanto solo se debe tener especial cuidado con la cantidad y calidad de los datos almacenados.

Así, la rutina de operación del código realizado en Mathematica® fue modificada y restringida para no efectuar una estimación automática (valor asignado mecánicamente por el mismo programa) relacionada con cada serie de tiempo, se procedió ingresando manualmente los valores de los coeficientes p, d y q en intervalos en los que el modelo pudiera calcular la serie de tiempo sin incurrir en errores.

Lo anterior está directamente relacionado con la independencia y complejidad de la estimación de la velocidad de viento, por ello es de suma importancia hallar los valores óptimos de los coeficientes ya que impactan directamente en el tiempo que el software requiere para realizar la estimación. Para aplicar estos modelos, se ingresaron bases de datos de Mérida, sin retroalimentar al modelo, en extensiones de años completos, del 1 de enero de 2011 al 31 de diciembre de 2014. Iniciando en un año (solo 2011) de entrada en incrementos de un año hasta llegar a cuatro años (de 2011 hasta 2014), con el objetivo de identificar la serie que mejor resuelve el modelo y contrastarlo con los resultados de los modelos restantes.

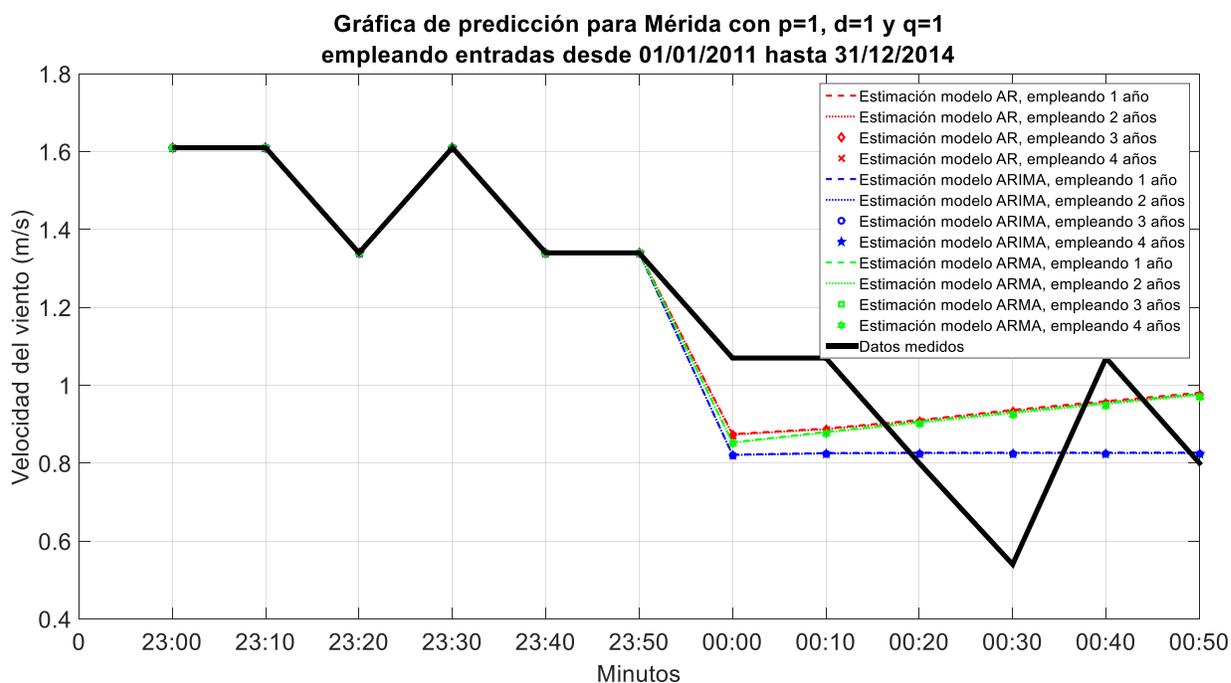


Figura 4.6.- Gráfica de predicción para Mérida, empleando los tres modelos con  $p=1$ ,  $d=1$  y  $q=1$ .

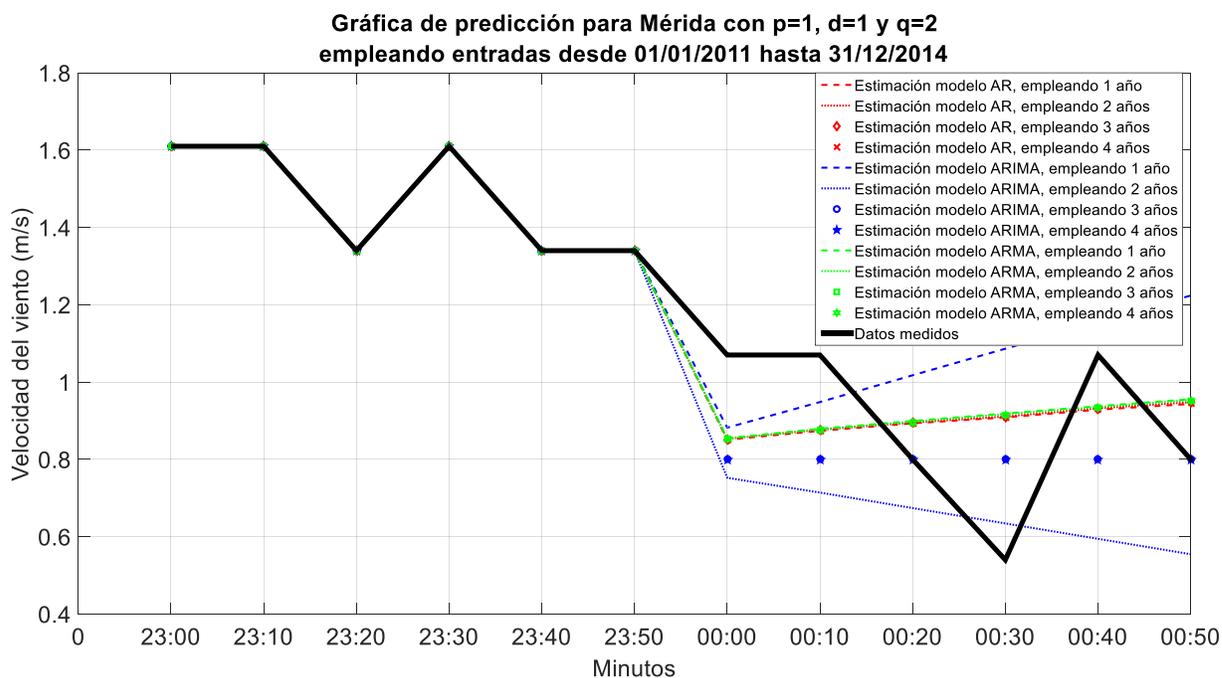


Figura 4.7.- Gráfica de predicción para Mérida, empleando los tres modelos con  $p=1$ ,  $d=1$  y  $q=2$ .

En la figura 4.6, se puede observar el comportamiento en la estimación empleando valores  $q=1$ ,  $d=1$  y  $q=1$ . Los modelos AR y ARMA tienen un comportamiento similar al emplear estos coeficientes, la estimación mediante estos modelos siempre se localiza sobre la media, es decir, lo sobrestima y se incrementa al avanzar en la gráfica en el tiempo. En el caso del modelo ARIMA, este tiene un comportamiento con tendencia hacia a la media de la serie de tiempo, lo cual corrobora los supuestos que se plantean en la tabla 4.1 y en primera instancia indica que estos valores para los coeficientes serían los óptimos.

En la figura 4.7, se despliega el comportamiento en la estimación empleando valores  $q=1$ ,  $d=1$  y  $q=2$ . Los modelos AR y ARMA exhiben un comportamiento muy similar; en este caso llama la atención el comportamiento del modelo ARIMA ya que se presenta una variación tanto hacia arriba como hacia abajo al modificar el número de términos de la media móvil invertible, es decir, incrementar la distancia de los términos de la predicción con respecto a la unidad medida.

En la figura 4.8, el efecto de cambiar el valor de la componente de términos autorregresivos del modelo. Se presenta un comportamiento similar en los modelos AR y ARMA, en el que sobrestiman el valor de predicción cuantos más datos a futuro se requieren para la estimación.

Sin embargo, en este caso es importante recordar que el coeficiente “p” solo será adecuado para representar un modelo autorregresivo siempre que se encuentre comprendido entre cero y uno, en valor absoluto, cumpliendo con ello el criterio de hipótesis de invertibilidad de la media móvil. Al cumplir con este criterio se establece un decrecimiento continuado en los coeficientes del correspondiente modelo autorregresivo y por lo tanto se cumple la estacionariedad (Hernández J. et al, 2013). Por lo tanto, el valor de  $p=2$  en esta gráfica, falla al momento de cumplir con el criterio de estacionariedad de la serie, si bien, en términos generales parecería lo contrario siempre deben cumplirse los criterios y supuestos establecidos.

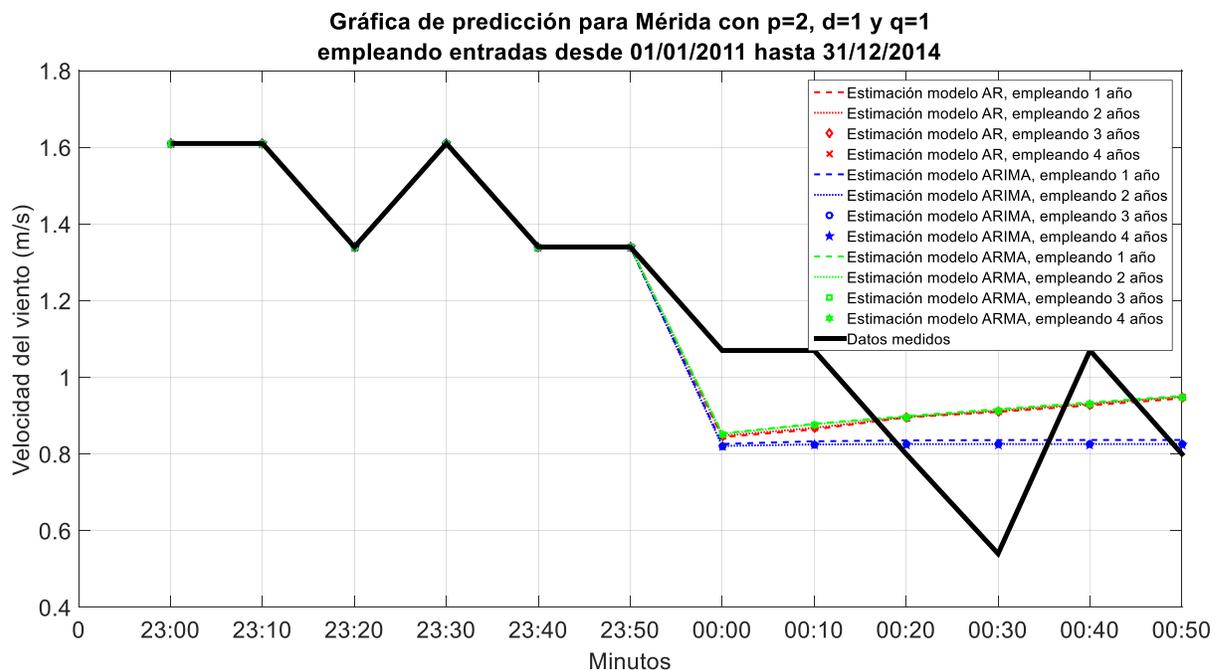


Figura 4.8.- Gráfica de predicción para Mérida, empleando los tres modelos con  $p=2$ ,  $d=1$  y  $q=1$ .

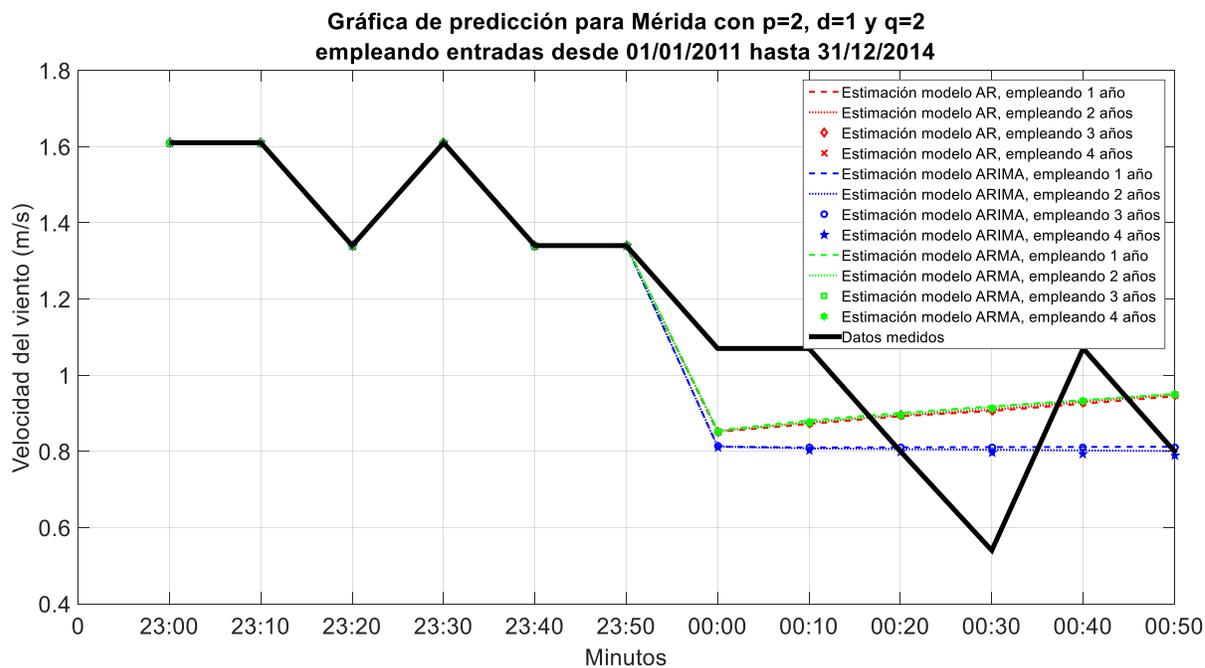


Figura 4.9.- Gráfica de predicción para Mérida, empleando los tres modelos con  $p=2$ ,  $d=1$  y  $q=2$ .

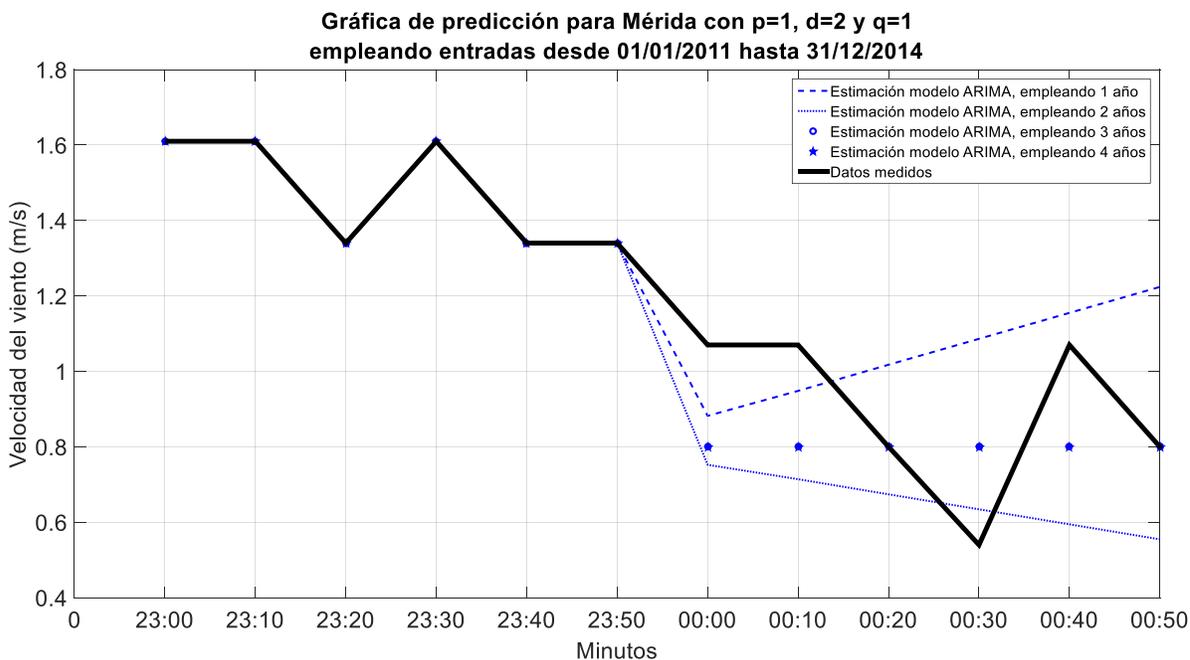


Figura 4.10.- Gráfica de predicción para Mérida, empleando los tres modelos con  $p=1$ ,  $d=2$  y  $q=1$ .

En la figura 4.9, empleando los tres modelos con coeficientes  $p=2$ ,  $d=1$  y  $q=2$ , se presenta un comportamiento similar al de la figura 4.3 y 4.5, es decir, el modelo ARIMA tiene una tendencia hacia la media y los restantes comienza a sobreestimar el valor conforme se pide una mayor cantidad de pasos hacia el futuro.

En la figura 4.10 se presenta el comportamiento del modelo ARIMA empleando valores de  $p=1$ ,  $d=2$  y  $q=1$ . En este caso solo se presenta este modelo, ya los modelos AR y ARMA no pueden realizar dos veces la diferenciación en la serie de tiempo para hacerla estacionaria, es decir, un coeficiente con valor  $d=2$  no es óptimo para la realización de estimación a través de estos modelos. El modelo ARIMA presenta también un comportamiento errático al emplear estos coeficientes.

De las gráficas anteriores, se puede decir que los mejores resultados son los presentados por las figuras 4.6 y 4.9, sin embargo, tomando en consideración los supuestos que debe cumplir la serie al realizar la estimación y que esta sea considerada correcta, el coeficiente "p" expresa el número de términos autoregresivos de la serie, por lo tanto, está relacionado con la sensibilidad de la estimación, así que un valor de  $p=2$  reduce la sensibilidad de la estimación y, en el caso de la figura 4.9 el

coeficiente “q” con valor dos, falla al momento de cumplir el criterio de estacionariedad. Esto deja al resultado presentado en la figura 4.6, es decir, estimación empleando los coeficientes  $p=1$ ,  $d=1$  y  $q=1$  como la opción viable y que cumple con todos los supuestos para realizar la estimación.

### 4.3 Análisis estadístico.

El objetivo de esta sección es determinar y evaluar el modelo, para ello se realizó un análisis estadístico para verificar la cantidad optima de datos mínimos que son necesarios para realizar las estimaciones. A continuación, se presentan los resultados del análisis.

Tabla 4.2.- Resultados de análisis estadístico, para la estimación empleando cuatro años como entradas.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Modelo	1	0.0077	0.0077	0.061	0.807
Residuals	22	2.7625	0.1256		

En la tabla 4.2, se presenta el resultado del análisis estadístico de los datos de la estimación contra el valor medido. En este punto es importante mencionar que si bien el valor de  $Pr=0.807$  se encuentra en la región de aceptación y se considera valida la hipótesis alterna, sin embargo, se trata de la primera aproximación sin retroalimentación del modelo y por lo tanto es altamente sensible a las mejoras mediante el cálculo del factor de calibración.

Tabla 4.5.- Resultados de la diferencia mínima significativa y grupos mixtos.

```
> LSD.test(Resultados_Anova,"Modelo",p.adj="none",console=TRUE)
Study: Resultados_Anova ~ "Modelo"
LSD t Test for Velocidad
Mean Square Error: 0.1255678
Modelo, means and individual ( 95 % ) CI
      Velocidad      std  r      LCL      UCL  Min  Max
ARIMA_4_111  1.147500 0.3562972 12 0.9353560 1.359644 0.82 1.61
Medicion    1.183333 0.3524030 12 0.9711894 1.395477 0.54 1.61
alpha: 0.05 ; Df Error: 22
Critical value of t: 2.073873
Least Significant Difference 0.3000169
Means with the same letter are not significantly different.
Groups, Treatments and means
a      Medicion      1.183
a      ARIMA_4_111   1.148
```

En la tabla 4.3, se plasman los resultados de la diferencia mínima significativa (con un nivel de confianza del 95 %) con un valor de 0.3, un valor crítico de t de Student de 2.073. Además, se crea solo un grupo mixto, marcado con la letra “a”, esto significa que existe una relación estrecha entre el valor medido por la estación y el valor estimado por el modelo. Con esto podemos decir que el modelo genera una estimación cuyo promedio no posee diferencia estadísticamente significativa con el valor medido.

Tabla 4.6.- Resultados del análisis de diferencia de medias.

```
> TukeyHSD(Resultados_Anova)
Tukey multiple comparisons of means
 95% family-wise confidence level

Fit: aov(formula = velocidad ~ Modelo, data = arima_111)

$Modelo
              diff          lwr          upr          p adj
Medicion-ARIMA_4_111 0.03583333 -0.2641835 0.3358502 0.8066647
```

En la tabla 4.4, se despliegan los resultados de la comparación de medias, en la cual se puede observar el valor de máxima diferencia observada durante el análisis estadístico. En este resultado existe una máxima diferencia individual de 0.035 u, hecho que se corrobora con los resultados obtenidos en la tabla 4.3, ya que no se presenta más de un grupo mixto de datos.

Tabla 4.7.- Tabla de resultados de la regresión lineal de la figura 4.8.

```
Linear model Poly1:
f(x) = p1*x + p2
Coefficients (with 95% confidence bounds):
p1 = 0.9147 (0.6183, 1.211)
p2 = 0.06652 (-0.2982, 0.4313)

Goodness of fit:
SSE: 0.2418
R-square: 0.8254
Adjusted R-square: 0.8079
RMSE: 0.1555
```

Para finalizar esta sección, se realizó un análisis de regresión lineal para contrastar los resultados del análisis estadístico.

En la figura 4.11 y tabla 4.5, se muestran los resultados de la regresión lineal, la cual se realizó en el mismo intervalo de confianza del análisis estadístico, de esta primera estimación se puede observar la relación que existe entre estas variables.

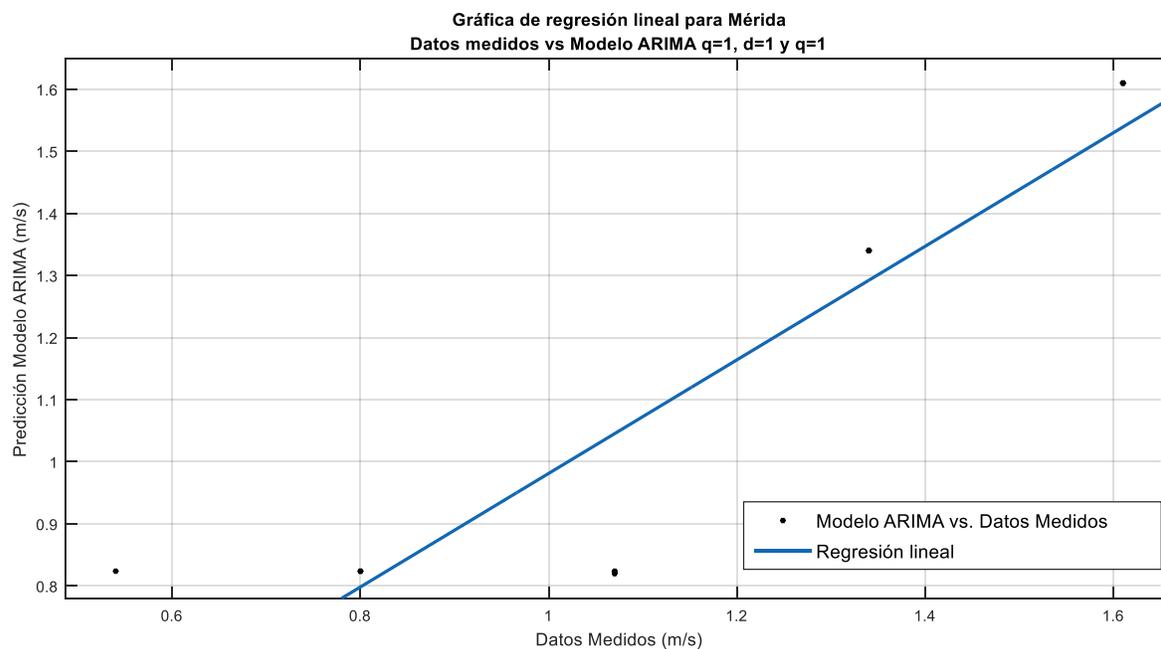


Figura 4.11.- Gráfica de regresión lineal para Mérida, datos medidos vs estimación del modelo ARIMA con coeficientes q=1, d=1 y q=1 con cuatro años de entrada de datos.

En este caso el valor de la  $R^2$  (0.80 u) calculada mediante la regresión lineal, es virtualmente el mismo que el obtenido como probabilidad en el análisis estadístico.

Con los resultados de este análisis se procede a realizar la corrección del modelo con el fin de obtener el mejor ajuste.

## CAPÍTULO V

# ANÁLISIS DE ESTIMACIONES, PREDICCIONES Y RESULTADOS.

### 5.1 Retroalimentación de modelo y corrección.

La retroalimentación del modelo se llevó para obtener una estimación mucho más fina, pero manteniendo los umbrales de estimaciones cada 10 minutos y hasta 2 horas. Esta retroalimentación se realizó a través de una iteración, creando nuevas series de tiempo que ingresan nuevamente al modelo para realizar una nueva estimación.

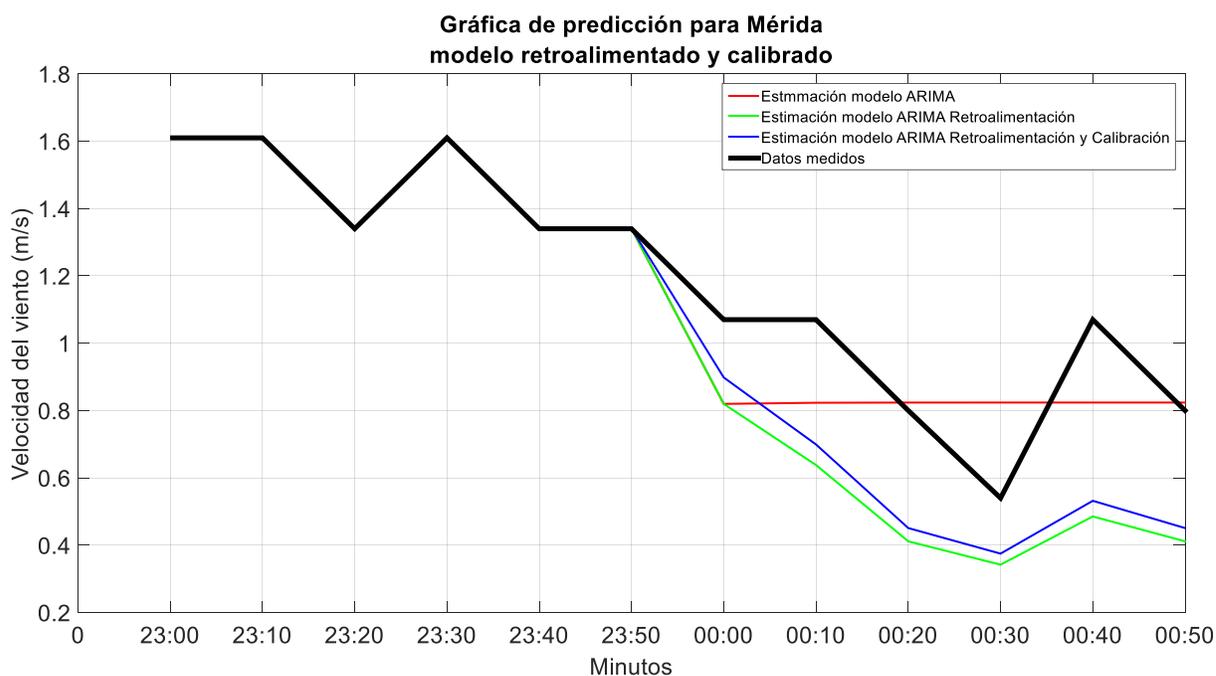


Figura 5.1.- Gráfica de predicción para Mérida, modelo retroalimentado y calibrado ARIMA (1, 1, 1).

En la figura 5.1, se presenta la gráfica del comportamiento de una estimación de una hora mediante el modelo ARIMA (1, 1, 1) retroalimentado, se incluyen los datos de una estimación con factor de corrección, denominado también en este proyecto “factor de calibración” debido a su similitud con los instrumentos de medición típicos (como los anemómetros). Como se sabe, la calibración es el proceso de comparar los valores

obtenidos por un instrumento de medición con la medida correspondiente de un patrón de referencia (o estándar). Así, en la realización de este proyecto se empleó el método de comparación directa, en este método se obtiene un valor en unidades del mensurando, mediante un instrumento, cadena o sistema de medición. El sensor del instrumento es colocado directamente en contacto con el fenómeno que se mide (ISO/IEC, 2005).

## 5.2 Análisis de datos.

En esta sección se presentan los resultados de una serie de estimaciones mediante el modelo ARIMA (1, 1, 1) que han pasado por todo el proceso de retroalimentación y calibración, así como el análisis estadístico pertinente.

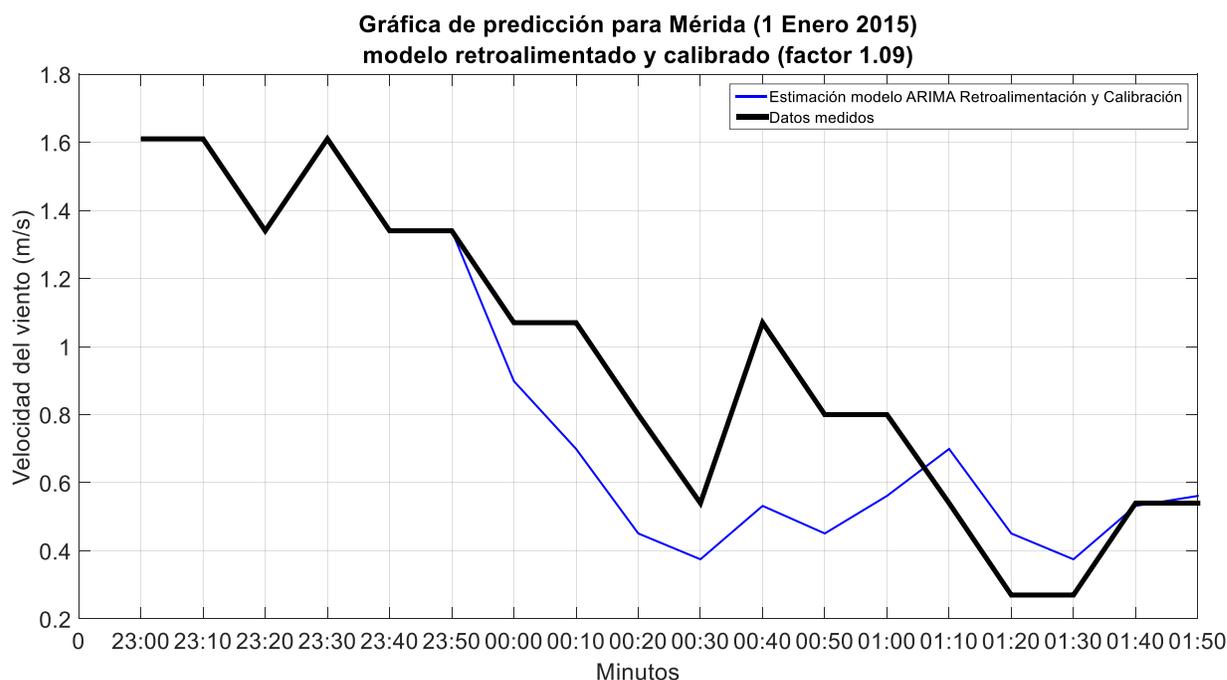


Figura 5.7.- Gráfica de predicción para Mérida (1 Enero 2015), modelo retroalimentado y calibrado (factor de corrección 1.09).

En la figura 5.2, se presentan los resultados finales de la predicción para Mérida del día 1 de enero de 2015, comparando el modelo retroalimentado corregido con el factor de calibración contra los datos medidos, en este caso se presenta una estimación de hasta dos horas. Como se observa en términos generales se trata de una buena estimación empleando un factor de corrección de 1.09 u.

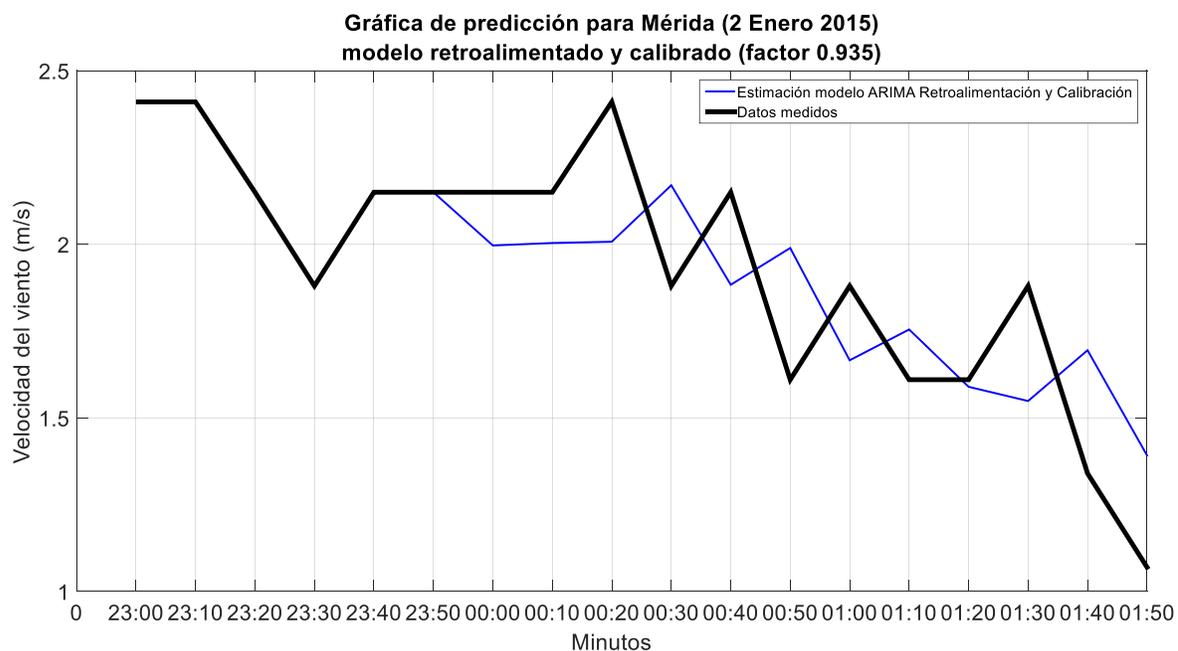


Figura 5.8.- Gráfica de predicción para Mérida (2 Enero 2015), modelo retroalimentado y calibrado (factor de corrección 0.935).

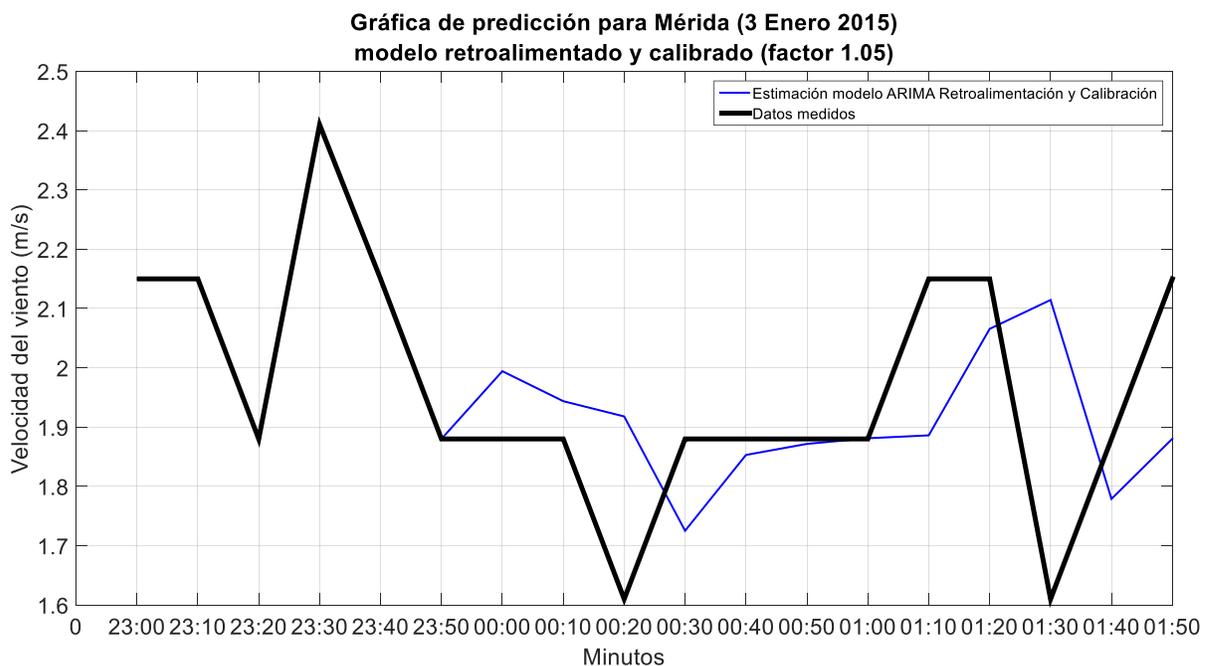


Figura 5.9.- Gráfica de predicción para Mérida (3 Enero de 2015), modelo retroalimentado y calibrado (factor de corrección 1.05).

En la figura 5.3, se muestran los resultados de la predicción para Mérida del día 2 de enero de 2015, comparando el modelo retroalimentado corregido con el factor de calibración contra los datos medidos. En términos generales se trata de una buena estimación empleando un factor de corrección de 0.935 u

Adicionalmente, en la figura 5.4, se exponen los resultados de la predicción para Mérida del día 3 de enero de 2015, comparando el modelo retroalimentado corregido con el factor de calibración contra los datos medidos, en este caso se presenta una estimación de hasta dos horas empleando un factor de corrección de 0.935 u

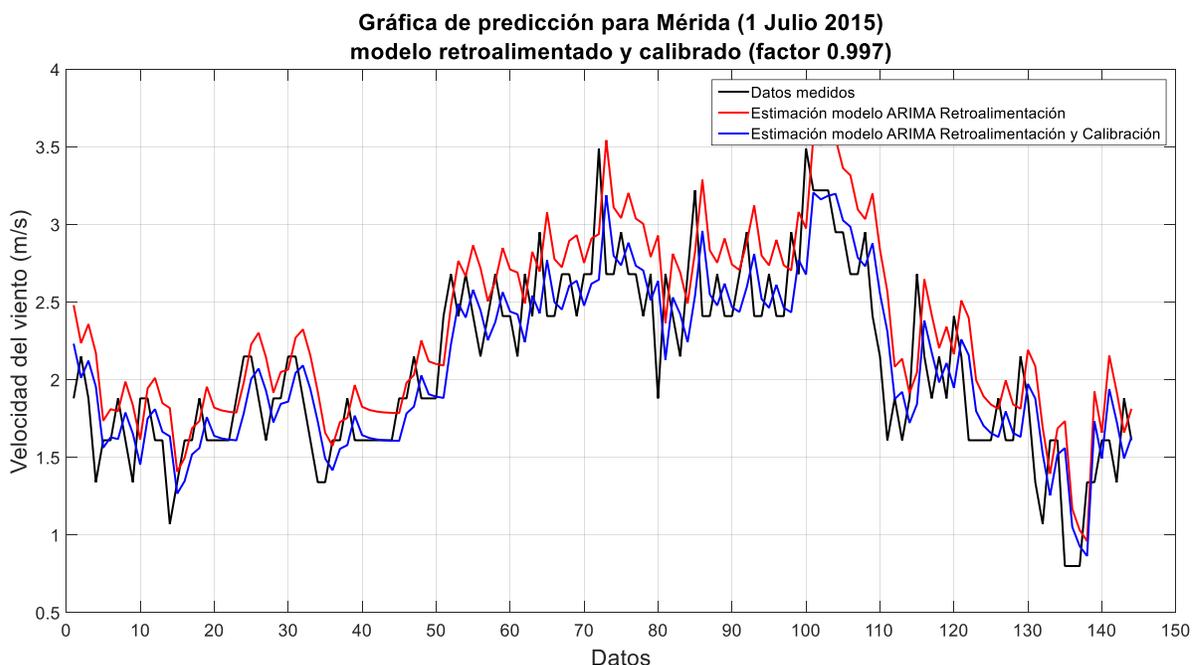


Figura 5.10.- Gráfica de predicción para Mérida (1 Julio 2015), modelo retroalimentado y calibrado (factor corrección final 0.997).

Finalmente, para someter a una prueba mucho más rígida se realizó una predicción final en la cual el modelo calcula los datos de un día completo, empleando la retroalimentación de datos y el factor de corrección final (figura 5.5). Como se puede observar se presenta un pequeño offset o desfase entre los datos de la predicción con respecto a los medidos por la estación, este desfase es inherente al modo de operación del modelo, sin embargo, se puede ver que se trata de una buena estimación y para corroborarlo se procede a realizar el análisis estadístico pertinente.

### 5.3 Análisis de resultados finales.

En esta sección se analizó el modelo final, para ello se realizó un análisis estadístico para determinar las máximas diferencias obtenidas entre el valor estimado y el dato medido. A continuación, se presentan los resultados del análisis.

Tabla 5.1.- Resultados de análisis estadístico, para la estimación con el factor de corrección empleando cuatro años como entradas.

```
> summary(Resultados_Anova)
              Df Sum Sq Mean Sq F value Pr(>F)
v2              1   0.01  0.00923   0.032  0.858
Residuals     286  82.59  0.28879
```

En la tabla 5.1, se muestra el resultado del análisis estadístico de los datos de la estimación contra el valor medido.

En la tabla se observa que el valor de  $Pr=0.858$  se encuentra en la región de aceptación y se considera válida la hipótesis alterna, se corrobora que el modelo es altamente sensible a las mejoras mediante la retroalimentación y el cálculo del factor de calibración. Se tiene además la correcta correspondencia entre los valores de los coeficientes del modelo ARIMA (1, 1, 1) con el valor de grados de libertad del modelo.

Tabla 5.2.- Tabla de modelo del análisis estadístico (promedios de datos).

```
> model.tables(Resultados_Anova, type="mean")
Tables of means
Grand mean

2.095452

v2
v2
Calibración      Medición
      2.1011      2.0898
```

En la tabla 5.2, se presenta en la tabla de modelos, el valor de promedio de los datos, y se observa el valor promedio del modelo retroalimentado y corregido “Calibración” es igual a 2.1011 m/s y el de la medición realizada por la estación meteorológica “Medición” es de 2.0898 m/s, esto nos brinda un panorama más general acerca de la estrecha relación entre el valor medido y el valor estimado para la serie de tiempo.

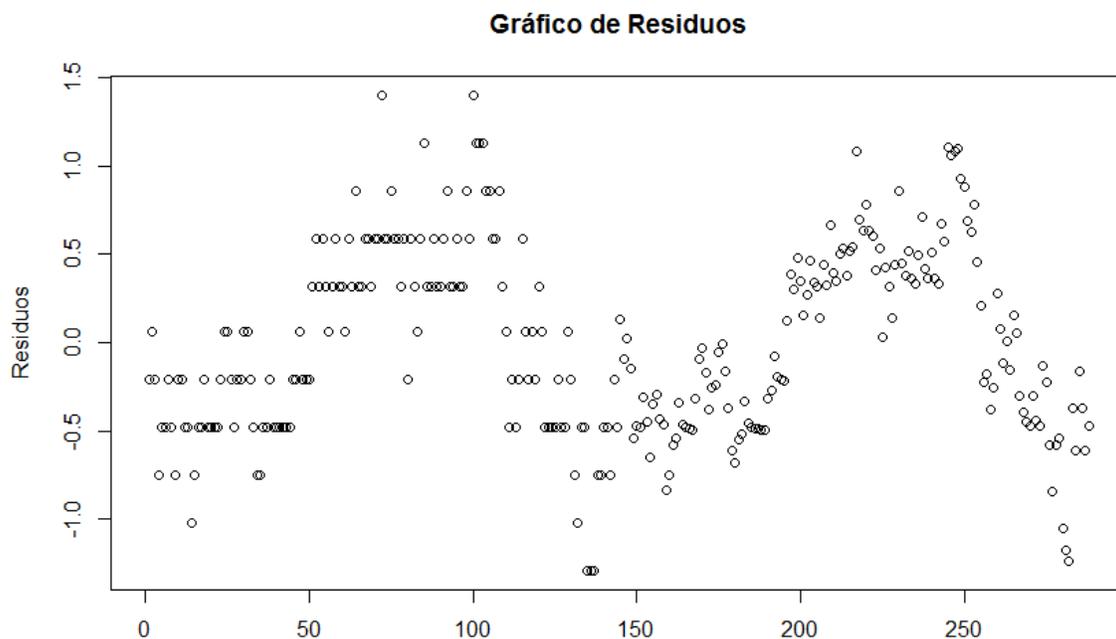


Figura 5.11.- Gráfica de residuos para el análisis estadístico.

En la figura 5.6, se tiene la gráfica de residuos para el análisis estadístico, en la cual se observa el comportamiento de los residuos del análisis. En ella se cumple el criterio de estacionariedad al tener tanto valores positivos como negativos.

Tabla 5.3.- Resultados de la diferencia mínima significativa y grupos mixtos.

```
> LSD.test(Resultados_Anova, "v2", p.adj="none", console=TRUE)
Study: Resultados_Anova ~ "v2"
LSD t Test for v1
Mean Square Error: 0.2887867
v2, means and individual ( 95 %) CI
      v1      std  r      LCL      UCL      Min      Max
Calibracion 2.101113 0.5111012 144 2.012968 2.189258 0.8648902 3.205747
Medicion    2.089792 0.5624490 144 2.001647 2.177937 0.8000000 3.490000
alpha: 0.05 ; Df Error: 286
Critical value of t: 1.968293
Least Significant Difference 0.1246557
Means with the same letter are not significantly different.
Groups, Treatments and means
a      Calibracion      2.101
a      Medicion        2.09
```

Con lo anterior, en la tabla 5.3 se tienen los resultados de la diferencia mínima significativa (con un nivel de confianza del 95 %) con un valor de 0.1246557 m/s, el cual es la diferencia que se aprecia a través de toda la serie de tiempo, con lo cual podemos corroborar lo expuesto en las gráficas 5.2, 5.3, 5.4 y 5.5, las cuales nos presentaban un comportamiento muy similar y corroboran esa diferencia obtenida en el análisis estadístico. Se tiene un valor crítico de t de Student de 1.968.

Además, se crea solo un grupo mixto, marcado con la letra "a", esto significa que existe una relación muy estrecha entre el valor medido por la estación (cuya media se encuentra en 2.09 m/s) y el valor estimado por el modelo (media de 2.101 m/s). Con esto podemos decir que el modelo genera una estimación de velocidad de viento que no tiene diferencia estadísticamente significativa con el valor medido.

Tabla 5.4.- Resultados del análisis de diferencia de medias.

```
> TukeyHSD(Resultados_Anova)
Tukey multiple comparisons of means
 95% family-wise confidence level

Fit: aov(formula = v1 ~ v2, data = final)

$v2
              diff          lwr          upr          p adj
Medicion-Calibracion -0.0113213 -0.135977  0.1133344  0.8582516
```

En la tabla 5.4, se despliegan los resultados de la comparación de medias, en la cual se puede observar el valor de máxima diferencia observada durante el análisis estadístico. En este resultado existe una máxima diferencia individual de 0.0113 m/s, hecho que se corrobora con los resultados obtenidos en la tabla 5.3, ya que no se presenta más de un grupo mixto de datos.

## CONCLUSIONES

- El modelo ARIMA (1, 1, 1) es una herramienta válida para realizar la estimación de velocidad del viento, lo anterior se demuestra en toda la literatura existente ya que se emplea en áreas de la Economía.
- Los sistemas computacionales actuales permiten la realización de estas estimaciones y predicciones a corto plazo, ya que poseen el suficiente poder de cómputo para realizar los cálculos e iteraciones necesarias.
- El software empleado es idóneo al ofrecer calidad, certeza y sobre todo funcionalidad, permitiendo el acceso a cualquier persona.
- Existe una relación muy estrecha entre el valor medido por la estación, con un valor promedio de 2.09 m/s y el valor estimado por el modelo con un promedio de 2.101 m/s.
- El valor máximo observado para la diferencia mínima significativa, empleando un nivel de confianza del 95 %, fue de 0.1246557 m/s entre el valor medido por la estación meteorológica y el calculado por modelo retroalimentado y corregido, además de una máxima diferencia individual de 0.0113 m/s.
- Finalmente, del análisis estadístico se puede concluir que es posible realizar una estimación de la velocidad de viento a corto plazo empleando modelos econométricos.

## REFERENCIAS.

- Administration, N. O. (2008). NOAA. Obtenido de NOAA: <http://www.noaa.gov/>
- Agua, C. N. (2015). CONAGUA. Obtenido de CONAGUA: <http://www.conagua.gob.mx/ocpy/>
- AMDEE. (2014). *Asociación Mexicana de Energía Eólica*. Obtenido de <http://www.amdee.org/>
- Borges A., C. O. (2014). *Predicción del potencial eólico disponible en Mérida, Yucatán, México*. Mérida, Yucatán: Colegio de Ingenieros Civiles de Yucatán.
- consortium, T. H. (2014). *Hirlam NWP in Europe*. Obtenido de Hirlam NWP in Europe: <http://hirlam.org/>
- de Arce R., M. R. (2001). *MODELOS ARIMA*. Madrid: Universidad Autónoma de Madrid.
- Duarte, J. V. (2001). *Predicción con redes neuronales: comparación con las metodologías de Box y Jenkins*. Mérida, Venezuela: Universidad de los Andes.
- Electricidad, C. F. (2014). *Informe Anual 2014*. México: Secretaría de Gobernación.
- Erasmus Cadenas, W. R. (2008). Análisis Estadístico de Mediciones de la Velocidad del Viento Utilizando la Técnica de Valores Desviados. *Simposio de Metrología 2008* (págs. SM2008-M134-1182-7 ). Santiago de Querétaro, México: Centro Nacional de Metrología .
- Garrido, V. J. (2007). *Modelos climáticos de balance de energía y resultados de control para modelos discretizados*.
- Geografía, I. N. (2014). INEGI. Obtenido de INEGI: <http://www.inegi.org.mx/inegi/default.aspx?e=31&s=geo&c=124>
- González, C. M. (2009). *Análisis de series temporales: Modelos ARIMA*. España: Sarriko-On.
- Group, W. A. (2013). *Community Earth System Model*. Obtenido de Community Earth System Model: <https://www2.cesm.ucar.edu/working-groups/wawg>

- Hernández Alonso J, Z. R. (2013). *Modelos econométricos para el análisis económico*. Sevilla: ESIC.
- Intelligence, T. W. (1988). *The Wind Power Wind*. Obtenido de The Wind Power Wind Energy Market Intelligence: <http://www.thewindpower.net/>
- ISO/IEC, 1. (2005). *General requirements for the competence of testing and calibration laboratories* (Second ed.). International Organization for Standardization.
- Lorenz, E. (1963). *Flujo determinista no periódico*. Journal of Atmospheric Sciences. Vol 20: 130-141.
- Lozano, W. (2013). *Plan de energías renovables 2013*. México: Secretaría de Economía.
- Model, T. C. (24 de Diciembre de 2014). *The National Center for Atmospheric Research* . Obtenido de The National Center for Atmospheric Research : <http://www2.cesm.ucar.edu/models>
- NCAR, M. a. (2015). *The Weather Research & Forecasting Model*. Obtenido de The Weather Research & Forecasting Model: <http://www.wrf-model.org/index.php>
- Pino Diez R., P. F. (2004). Pronóstico de la Velocidad y Dirección del Viento mediante Redes Neuronales Artificiales. *VIII Congreso de Ingeniería de Organización Leganés* (págs. 905-9014). Madrid: Dialnet.
- Research, P. S. (2015). *MM5 Community Model*. Obtenido de MM5 Community Model: <http://www2.mmm.ucar.edu/mm5/>
- SENER. (2014). *Prospectiva del Sector Eléctrico 2015-2029*. México: Secretaría de Gobernación .
- Villavicencio, J. (2010). *Introducción a Series de Tiempo*.
- Yucatán, G. d. (2012-2018). *Gobierno del Estado de Yucatán*. Obtenido de Gobierno del Estado de Yucatán: <http://www.yucatan.gob.mx/>

## ANEXOS.

### ANEXO I.

Código fuente del programa en Mathematica® para la realización de la estimación mediante los modelos AR, ARMA y ARIMA.

(\*Condiciones iniciales

-Cargamos el archivo que deseamos leer

-Las fechas de inicio y termino que deseamos leer

-La cantidad de pasos que deseamos que el ARProcess utilice para realizar la prediccion\*)

```
v=Import[NotebookDirectory[]<>"MERIDATOTAL.csv"];
```

```
fechainicial=Table[If[v[[s,1]]=="01/01/2014",1,0],{s,3,Length[v]}];
```

```
fechafinal=Table[If[v[[s,1]]=="31/12/2014",1,0],{s,3,Length[v]}];
```

(\*Numero de estimaciones a realizar\*)

```
Numpredic=6;
```

(\*Resolucion del modelo AR\*)

```
pasosAR=3;
```

(\*Coeficientes para modelos ARIMA y ARMA\*)

```
p=1;
```

```
q=1;
```

```
d=1;
```

```
posfinit=Flatten[Position[fechainicial,1]];posffinal=Flatten[Position[fechafinal,1]];
```

```
vels=Table[v[[s,10]],{s,posfinit[[1]]+3, posffinal[[-1]]+3}];
```

```
velscompa=vels=Table[v[[s,10]],{s,posfinit[[1]]+3, posffinal[[-1]]+3+6}];
```

```
t=Flatten[Table[s,{s,1,Length[vels]}]];
```

```
ts=TimeSeries[vels,{t}];
```

(\*ESTIMACION EMPLEADO EL MODELO ARProcess\*)

```
procAR=EstimatedProcess[ts,ARProcess[pasosAR]];
```

```
forecastAR=TimeSeriesForecast[procAR,ts,{Numpredic}]
```

```
TemporalData[Time: 210321 to 210326 \[SpanFromLeft]
```

```
Data points: 6      Paths: 1
```

```
]
```

```

err=forecastAR["MeanSquaredErrors"];
ListLinePlot[ts]
GAR=ListLinePlot[{velscompa,forecastAR},
                 PlotRange->{{Length[vels]-
30,Length[vels]+10},{0,7}},PlotLabel->"Gráfica de predicción Mérida 2013-2014,
Modelo AR",AxesLabel->{"Fecha","Velocidad"}]
DateListPlot[err, PlotLabel->"Estimación del error",FrameLabel-
>Automatic,TargetUnits->"%"]

```

(\*ARIMA\*)

(\*En este caso es necesario declarar las constantes del tipo de proceso\*)

```

eprocARIMA=EstimatedProcess[ts,ARIMAProcess[p,q,d]]
ARIMAProcess[-1.56711*10^-6,{-0.216372,-0.0512834},2,{-1.},0.142183]
forecastARIMA=TimeSeriesForecast[eprocARIMA,ts,{Numpredic}]
TemporalData[Time: 210321 to 210326 \[SpanFromLeft]
Data points: 6 Paths: 1
]

```

```

GARIMA=ListLinePlot[{velscompa,forecastARIMA},
                    PlotRange->{{Length[vels]-
30,Length[vels]+10},{0,7}},PlotLabel->"Gráfica de predicción Mérida 2013-2014,
Modelo ARIMA",AxesLabel->{"Fecha","Velocidad"}]
errARIMA=forecastARIMA["MeanSquaredErrors"];
DateListPlot[errARIMA, PlotLabel->"Estimación del error",FrameLabel-
>Automatic,TargetUnits->"%"]

```

(\*ARMA\*)

(\*Del mismo modo, es necesario declarar las constantes para realizar el calculo\*)

```

eprocARMA=EstimatedProcess[ts,ARMAProcess[p,q]]
ARMAProcess[0.0381974,{1.10177,-0.12659},{-0.455712,-0.0085562},0.12521]
forecastARMA=TimeSeriesForecast[eprocARMA,ts,{Numpredic}]
TemporalData[Time: 210321 to 210326 \[SpanFromLeft]
Data points: 6 Paths: 1
]

```

```

GARMA=ListLinePlot[{velscompa,forecastARMA},
                   PlotRange->{{Length[vels]-
30,Length[vels]+10},{0,7}},PlotLabel->"Gráfica de predicción Mérida 2013-2014
Modelo ARMA",AxesLabel->{"Fecha","Velocidad"}]

```

```

errARMA=forecastARMA["MeanSquaredErrors"];
DateListPlot[errARMA, PlotLabel->"Estimación del error",FrameLabel->Automatic,TargetUnits->"%"]

(*Exportamos los datos de la prediccion emplean
ARProcess*)SerieTAR=Table[{Length[ts]+k,forecastAR[[2,1,1,k]]},{k,1,Numpredic}];
DatosGAR=Join[GAR[[1,2,1,3,3,1]],SerieTAR];
Export[NotebookDirectory[]<>"Prediccion_AR.csv",DatosGAR];

(*Exportamos los datos de la prediccion emplean
ARIMAProcess*)SerieTARIMA=Table[{Length[ts]+k,forecastARIMA[[2,1,1,k]]},{k,1,Num
mpredic}];
DatosGARIMA=Join[GARIMA[[1,2,1,3,3,1]],SerieTARIMA];
Export[NotebookDirectory[]<>"Prediccion_ARIMA.csv",DatosGARIMA];

(*Exportamos los datos de la prediccion emplean
ARMAProcess*)SerieTARMA=Table[{Length[ts]+k,forecastARMA[[2,1,1,k]]},{k,1,Num
predic}];
DatosGARMA=Join[GARMA[[1,2,1,3,3,1]],SerieTARMA];
Export[NotebookDirectory[]<>"Prediccion_ARMA.csv",DatosGARMA];

```

## ANEXO II.

Código fuente del programa en R Studio® para el análisis estadístico de la estimación mediante los modelos AR, ARMA y ARIMA.

1) Ingresar información (seleccionar archivo en formato .csv).

2) Calcular análisis

```
>Resultados_Anova <- aov(Datos~,Medición=estimación_final)
```

```
> summary(Resultados_Anova)
```

```
> model.tables(Resultados_Anova)
```

```
> model.tables(Resultados_Anova,type="mean")
```

3) Residuos

```
>Residuos <- residuals(Resultados_Anova)
```

4) Gráficos de análisis

```
>plot(Resultados_Anova)
```

```
>plot(Residuos,main="Gráfico de Residuos")
```

5) Diferencia Significativa Mínima

```
>LSD.test(Resultados_Anova,"ind",p.adj="none",console=TRUE)
```

6) Prueba de medias

```
>TukeyHSD(Resultados_Anova)
```

## ANEXO III.

Código fuente del programa en Matlab® para las gráficas de la estimación mediante los modelos AR, ARMA y ARIMA.

```
load('datos.mat')
clc;
figure(1)
%Grafica de Merida 1 año
plot(Merida_2014);
title('Gráfica de datos Mérida 2014','FontSize',18);
legend('Mérida');
set(legend,'FontSize',16);
xlabel('Fecha-Hora','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;
%Grafica de Merida 2 años
figure(2)
plot(Merida_2013_2014);
title('Gráfica de datos Mérida 2013-2014','FontSize',18);
legend('Mérida');
set(legend,'FontSize',16);
xlabel('Fecha-Hora','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;
%Grafica de Merida 3 años
figure(3)
plot(Merida_2012_2014);
title('Gráfica de datos Mérida 2012-2014','FontSize',18);
legend('Mérida');
set(legend,'FontSize',16);
xlabel('Fecha-Hora','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
```

```

grid on;
%Grafica de Merida 4 años
figure(4)
plot(Merida_2011_2014);
title('Gráfica de datos Mérida 2011-2014','FontSize',18);
legend('Mérida');
set(legend,'FontSize',16);
xlabel('Fecha-Hora','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;
%Grafica Merida comparacion modelos p=1,d=1,q=1
figure(5)
plot(MeridaAR1_3,'r --','LineWidth',2);
hold all;
plot(MeridaAR2_3,'r :','LineWidth',2);
plot(MeridaAR3_3,'r d','LineWidth',2);
plot(MeridaAR4_3,'r x','LineWidth',2);
plot(MeridaARIMA1_111,'b --','LineWidth',2);
plot(MeridaARIMA2_111,'b :','LineWidth',2);
plot(MeridaARIMA3_111,'b o','LineWidth',2);
plot(MeridaARIMA4_111,'b p','LineWidth',2);
plot(MeridaARMA1_11,'g --','LineWidth',2);
plot(MeridaARMA2_11,'g :','LineWidth',2);
plot(MeridaARMA3_11,'g s','LineWidth',2);
plot(MeridaARMA4_11,'g h','LineWidth',2);
plot(ComparacionMerida_1,'k','LineWidth',5);
hold off;
title({'Gráfica de predicción para Mérida con p=1, d=1 y q=1','empleando entradas desde 01/01/2011 hasta 31/12/2014'},'FontSize',18);
legend('Estimación modelo AR, empleando 1 año','Estimación modelo AR, empleando 2 años','Estimación modelo AR, empleando 3 años','Estimación modelo AR, empleando 4 años','Estimación modelo ARIMA, empleando 1 año','Estimación modelo ARIMA, empleando 2 años','Estimación modelo ARIMA, empleando 3 años','Estimación modelo ARIMA, empleando 4 años','Estimación modelo ARMA,

```

empleando 1 año','Estimación modelo ARMA, empleando 2 años','Estimación modelo ARMA, empleando 3 años','Estimación modelo ARMA, empleando 4 años','Datos medidos');

```
set(legend,'FontSize',16);
```

```
xlabel('Minutos','FontSize',18);
```

```
ylabel('Velocidad del viento (m/s)','FontSize',18);
```

```
grid on;
```

```
%Grafica Merida comparacion modelos p=2, d=1 y q=1
```

```
figure(6)
```

```
plot(MeridaAR1_5,'r --','LineWidth',2);
```

```
hold all;
```

```
plot(MeridaAR2_5,'r :','LineWidth',2);
```

```
plot(MeridaAR3_5,'r d','LineWidth',2);
```

```
plot(MeridaAR4_5,'r x','LineWidth',2);
```

```
plot(MeridaARIMA1_211,'b --','LineWidth',2);
```

```
plot(MeridaARIMA2_211,'b :','LineWidth',2);
```

```
plot(MeridaARIMA3_211,'b o','LineWidth',2);
```

```
plot(MeridaARIMA4_211,'b p','LineWidth',2);
```

```
plot(MeridaARMA1_21,'g --','LineWidth',2);
```

```
plot(MeridaARMA2_21,'g :','LineWidth',2);
```

```
plot(MeridaARMA3_21,'g s','LineWidth',2);
```

```
plot(MeridaARMA4_21,'g h','LineWidth',2);
```

```
plot(ComparacionMerida_1,'k','LineWidth',5);
```

```
hold off;
```

```
title({'Gráfica de predicción para Mérida con p=2, d=1 y q=1','empleando entradas desde 01/01/2011 hasta 31/12/2014'},'FontSize',18);
```

```
legend('Estimación modelo AR, empleando 1 año','Estimación modelo AR, empleando 2 años','Estimación modelo AR, empleando 3 años','Estimación modelo AR, empleando 4 años','Estimación modelo ARIMA, empleando 1 año','Estimación modelo ARIMA, empleando 2 años','Estimación modelo ARIMA, empleando 3 años','Estimación modelo ARIMA, empleando 4 años','Estimación modelo ARMA, empleando 1 año','Estimación modelo ARMA, empleando 2 años','Estimación modelo ARMA, empleando 3 años','Estimación modelo ARMA, empleando 4 años','Datos medidos');
```

```
set(legend,'FontSize',16);
```

```
xlabel('Minutos','FontSize',18);
```

```

ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;
%Grafica Merida comparacion modelos p=1, d=1 y q=2
figure(7)
plot(MeridaAR1_7,'r --','LineWidth',2);
hold all;
plot(MeridaAR2_7,'r :','LineWidth',2);
plot(MeridaAR3_7,'r d','LineWidth',2);
plot(MeridaAR4_7,'r x','LineWidth',2);
plot(MeridaARIMA1_112,'b --','LineWidth',2);
plot(MeridaARIMA2_112,'b :','LineWidth',2);
plot(MeridaARIMA3_112,'b o','LineWidth',2);
plot(MeridaARIMA4_112,'b p','LineWidth',2);
plot(MeridaARMA1_12,'g --','LineWidth',2);
plot(MeridaARMA2_12,'g :','LineWidth',2);
plot(MeridaARMA3_12,'g s','LineWidth',2);
plot(MeridaARMA4_12,'g h','LineWidth',2);
plot(ComparacionMerida_1,'k','LineWidth',5);
hold off;
title({'Gráfica de predicción para Mérida con p=1, d=1 y q=2','empleando entradas
desde 01/01/2011 hasta 31/12/2014'},'FontSize',18);
legend('Estimación modelo AR, empleando 1 año','Estimación modelo AR, empleando
2 años','Estimación modelo AR, empleando 3 años','Estimación modelo AR,
empleando 4 años','Estimación modelo ARIMA, empleando 1 año','Estimación modelo
ARIMA, empleando 2 años','Estimación modelo ARIMA, empleando 3
años','Estimación modelo ARIMA, empleando 4 años','Estimación modelo ARMA,
empleando 1 año','Estimación modelo ARMA, empleando 2 años','Estimación modelo
ARMA, empleando 3 años','Estimación modelo ARMA, empleando 4 años','Datos
medidos');
set(legend,'FontSize',16);
xlabel('Minutos','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;
%Grafica Merida comparacion modelos p=2, d=1 y q=2
figure(8)

```

```

plot(MeridaAR1_9,'r --','LineWidth',2);
hold all;
plot(MeridaAR2_9,'r :','LineWidth',2);
plot(MeridaAR3_9,'r d','LineWidth',2);
plot(MeridaAR4_9,'r x','LineWidth',2);
plot(MeridaARIMA1_212,'b --','LineWidth',2);
plot(MeridaARIMA2_212,'b :','LineWidth',2);
plot(MeridaARIMA3_212,'b o','LineWidth',2);
plot(MeridaARIMA4_212,'b p','LineWidth',2);
plot(MeridaARMA1_22,'g --','LineWidth',2);
plot(MeridaARMA2_22,'g :','LineWidth',2);
plot(MeridaARMA3_22,'g s','LineWidth',2);
plot(MeridaARMA4_22,'g h','LineWidth',2);
plot(ComparacionMerida_1,'k','LineWidth',5);
hold off;

title({'Gráfica de predicción para Mérida con p=2, d=1 y q=2','empleando entradas
desde 01/01/2011 hasta 31/12/2014'},'FontSize',18);

legend('Estimación modelo AR, empleando 1 año','Estimación modelo AR, empleando
2 años','Estimación modelo AR, empleando 3 años','Estimación modelo AR,
empleando 4 años','Estimación modelo ARIMA, empleando 1 año','Estimación modelo
ARIMA, empleando 2 años','Estimación modelo ARIMA, empleando 3
años','Estimación modelo ARIMA, empleando 4 años','Estimación modelo ARMA,
empleando 1 año','Estimación modelo ARMA, empleando 2 años','Estimación modelo
ARMA, empleando 3 años','Estimación modelo ARMA, empleando 4 años','Datos
medidos');

set(legend,'FontSize',16);
xlabel('Minutos','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;
%Grafica Merida comparacion modelos p=1,d=2,q=1
figure(9)
plot(MeridaARIMA1_121,'b --','LineWidth',2);
hold all;
plot(MeridaARIMA2_121,'b :','LineWidth',2);
plot(MeridaARIMA3_121,'b o','LineWidth',2);

```

```

plot(MeridaARIMA4_121,'b p','LineWidth',2);
plot(ComparacionMerida_1,'k','LineWidth',5);
hold off;
title({'Gráfica de predicción para Mérida con p=1, d=2 y q=1','empleando entradas
desde 01/01/2011 hasta 31/12/2014'},'FontSize',18);
legend('Estimación modelo ARIMA, empleando 1 año','Estimación modelo ARIMA,
empleando 2 años','Estimación modelo ARIMA, empleando 3 años','Estimación
modelo ARIMA, empleando 4 años','Datos medidos');
set(legend,'FontSize',16);
xlabel('Minutos','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;
%Grafica Merida prediccion con retroalimentacion
figure(10)
plot(MeridaARIMA4_111,'r ','LineWidth',2);
hold all;
plot(MeridaARIMA4_111_Retro,'g ','LineWidth',2);
plot(MeridaARIMA4_111_Retro_Error,'b ','LineWidth',2);
plot(ComparacionMerida_1,'k','LineWidth',5);
hold off;
title({'Gráfica de predicción para Mérida', 'modelo retroalimentado y
calibrado'},'FontSize',18);
legend('Estmmación modelo ARIMA','Estimación modelo ARIMA
Retroalimentación','Estimación modelo ARIMA Retroalimentación y Calibración','Datos
medidos');
set(legend,'FontSize',16);
xlabel('Minutos','FontSize',18);
ylabel('Velocidad del viento (m/s)','FontSize',18);
grid on;

```