



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

PREDICCIÓN DE CORTO PLAZO DE POTENCIA GENERADA EN UN
AEROGENERADOR USANDO MODELO SARIMA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELECTRICISTA

RAMÓN SIMÓN ANDRÉS NORAMBUENA ORTEGA

PROFESOR GUÍA:

SR. PATRICIO PARADA SALGADO

MIEMBROS DE LA COMISIÓN:

SR. HECTOR AGUSTO ALEGRÍA

SR. JORGE SILVA SÁNCHEZ

SANTIAGO DE CHILE

ENERO 2011

RESUMEN DE LA MEMORIA
PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELECTRICISTA
POR: RAMÓN SIMÓN ANDRÉS NORAMBUENA ORTEGA
FECHA: 11 DE ENERO DE 2011
PROF. GUÍA: SR. PATRICIO PARADA SALGADO

PREDICCIÓN DE CORTO PLAZO DE POTENCIA GENERADA EN UN AEROGENERADOR USANDO MODELO SARIMA

El aumento del aporte energético por parte de las centrales eólicas dentro de la matriz de energías renovables no convencionales de Chile, crea la imperiosa necesidad de desarrollar herramientas que ayuden a gestionar el funcionamiento de parques eólicos, y en particular de los aerogeneradores que lo componen, con el fin de hacer más eficiente la integración y manejo en el sistema interconectado. En esta línea, el propósito de este trabajo es desarrollar un modelo predictivo para la potencia generada en un aerogenerador en base a series de tiempo históricas de variables atmosféricas del lugar donde éste se encuentra.

El trabajo de memoria presenta los resultados de la implementación de un modelo SARIMA (siglas en inglés de *Seasonal Auto Regressive Integrated Moving Average*) y un modelo de persistencia, para predicción de velocidad de viento a horizontes de tiempo de uno y cinco pasos en una escala de tiempo de una hora por cada paso, resultados que luego son transformados a potencia eléctrica por medio de la curva de potencia del aerogenerador considerado.

La investigación conecta los campos de la física, generación de energía y de teoría de estimación. Mientras que el primero aporta las ecuaciones con las cuales se describe el viento en la atmósfera y el segundo aporta la base técnica con la cual se relaciona la velocidad del viento con la potencia generada por un aerogenerador, el tercero entrega las herramientas para poder realizar predicción a distintos horizontes por medio de series de tiempo. Por ello, el reporte comienza por los fundamentos físicos que describen la velocidad del viento en la atmósfera, para seguir con los principios técnicos de un aerogenerador y continúa mencionando técnicas utilizadas en el ámbito de la predicción. Además, se trabaja con datos muestreados durante el año 1990 en la localidad de Punta Lengua de Vaca y que fueron obtenidos por el proyecto EOLO del Departamento de Geofísica de la Universidad de Chile.

Los resultados de este trabajo permitieron conocer las limitaciones, ventajas y desventajas que poseen tanto el modelo de persistencia como los modelos SARIMA en el ámbito de predicción. En la misma línea, se cuantificó por medio de indicadores de desempeño la exactitud en las predicciones realizadas usando ambos modelos, para finalmente compararlos bajo distintos horizontes de predicción y usando datos de distintos lugares.

Finalmente se concluye que el modelo SARIMA puede ser utilizado para predicción de potencia generada en un aerogenerador y que, en comparación con el modelo de persistencia, presenta mejores resultados en predicción a cinco pasos, pero no así en el caso de predicción a un paso, donde la relación se invierte.

A mi familia.

Agradecimientos

En primer lugar agradezco todo el apoyo brindado por mi madre Laura, mi padre Ramón y mis hermanos Rodrigo y Bárbara durante todos estos años de esfuerzo, por haber estado conmigo en todo momento y por haber apoyado cada una de mis decisiones. También agradecer a mi tía Margarita y mi tío Rubén, de quienes sé que nunca me faltó ni faltará apoyo.

En segundo lugar agradezco el apoyo brindado por el profesor Patricio Parada, quien siempre tuvo la mejor disposición a escuchar y aconsejarme en el desarrollo de este trabajo. Un excelente profesor y una gran persona. También agradezco el apoyo del profesor Jorge Silva, cuyos comentarios fueron valiosos para el desarrollo de este trabajo.

Además, mis agradecimientos para mis amigos de la vida, en especial para Gonzalo y Alejandro de quienes recibí el apoyo y consejo en todo momento. También agradecer a mis compañeros de Universidad: Sebastián Campos, Sebastián Cubillos, Juan Pablo M., Juan Pablo K., Lilian, Tomás, Diego, Ignacio, Franz, Alex, Juan Alejandro, Andrés, Andrea, Daniela, Carlos, Mauricio y muchos otros, con los cuales compartí durante estos años de estudio y dejaron recuerdos que espero jamás olvidar. A mis compañeros en el laboratorio: Alex, Andrés, Eduardo y Patricio, con quienes compartí mayormente mi último semestre, llevándome los más agradables recuerdos y de quienes recibí valiosos aportes para la elaboración de este documento.

Finalmente, agradecer al profesor Javier Ruiz del Solar, con quien trabajé durante gran parte de estos años y tuve la oportunidad de desarrollarme como estudiante, y al profesor Felipe Camel, quien contribuyó con información para el desarrollo de este trabajo.

A todos, muchas gracias.

Simón Norambuena.
Santiago, Enero de 2011.

Índice General

Agradecimientos	III
1. Introducción	1
1.1. Motivación	1
1.2. Objetivos Generales	2
1.3. Objetivos Específicos	3
1.4. Alcances	3
1.5. Estructura del Documento	4
2. Antecedentes	5
2.1. Energía Eólica en Chile	6
2.2. Naturaleza y Fenomenología del Viento	6
2.2.1. Ecuaciones de Navier-Stokes	9
2.2.2. Ecuaciones Promediadas de Reynolds	10
2.2.3. Rugosidad	12
2.2.4. Variación del viento con la altura	13
2.3. Potencia Extraíble del Viento	13
2.3.1. El Viento y su Potencia	13
2.3.2. Extracción de la Potencia del Viento	15
2.4. Descripción General de un Aerogenerador	16
2.4.1. Sistema de Soporte	16
2.4.2. Sistema de Transmisión	18
2.4.3. Sistema de Rotor	18
2.4.4. Sistema de Control	19
2.5. Tecnologías de Generación Eólica	19
2.5.1. Generador de Inducción de Velocidad Fija	19
2.5.2. Generador de Inducción de Velocidad Fija con Capacidad FRT	20
2.5.3. Generador de Inducción Doblemente Alimentado	20
2.5.4. Generador Síncrono de Velocidad Variable	20
2.6. Relación Potencia Mecánica - Velocidad del Viento	21

3. Estimación de Velocidad del Viento	24
3.1. Estimación de la Velocidad del Viento	24
3.2. Métodos Basados en Histogramas	25
3.2.1. Histogramas de Viento y Estimación de Energía	25
3.2.2. Distribución de Weibull	26
3.2.3. Distribución de Rayleigh	29
3.2.4. Modelo de Persistencia	30
3.3. Modelos Lineales Estacionarios	30
3.3.1. Proceso Auto Regresivo (AR)	31
3.3.2. Proceso de Media Móvil (MA)	31
3.3.3. Filtro de Kalman	34
3.4. Métodos no Paramétricos	41
3.4.1. Redes Neuronales Artificiales	41
3.4.2. Modelo Gris y otros métodos	42
3.5. Modelos Numéricos de Predicción de Tiempo Climático	42
4. Metodología de Trabajo	45
4.1. Análisis estadístico y de frecuencia	47
4.2. Desarrollo de un modelo SARIMA para predicción	47
4.3. Indicadores de desempeño	51
4.3.1. Error Cuadrático Medio	51
4.3.2. Error Cuadrático Medio con pesos	51
4.3.3. Error Absoluto Medio	51
4.3.4. Error Absoluto Porcentual Medio	51
4.3.5. Error Absoluto Porcentual Efectivo	52
4.3.6. Criterio de Información Bayesiano	52
4.3.7. Criterio de Información de Akaike	53
5. Resultados y Discusión	54
5.1. Series de tiempo	55
5.2. Función de Autocorrelación Empírica	57
5.3. Espectro de Frecuencias	58
5.4. Resultados de modelo SARIMA	59
5.5. Resultados de modelo de persistencia	65
5.6. Comparación entre los modelos implementados	70
6. Conclusiones	73
Referencias	74

Apéndices	81
A . Función de Autocorrelación	81
B . Transformada a Tiempo Discreto de Fourier	82
C . Códigos	82
C .1. simular	82
C .2. busqueda_ordenes_sarima2	83
C .3. boxjenkins3	88

Índice de figuras

2.1.	Regímenes de viento en la Tropósfera	8
2.2.	Masa de aire que mueve las aspas del aerogenerador	14
2.3.	Velocidades del viento en una Turbina Eólica	15
2.4.	Tipos de Torres para aerogeneradores	17
2.5.	Curva de potencia de un aerogenerador	22
3.1.	Esquema del Método de Histogramas	25
3.2.	Algoritmo de Filtrado de Kalman	39
3.3.	Estructura de red neuronal artificial.	41
3.4.	Modelo NWP del servicio metereológico alemán con resolución espacial de $60 \times 60 [km^2]$	43
4.1.	Diagrama de Bloques de las funciones programadas	48
5.1.	Series de tiempo de Punta Lengua de Vaca: (a) Mediciones tomadas en la semana 24 de 1990. (b) Mediciones tomadas en la semana 12 de 1990. (c) Mediciones tomadas en la semana 14 de 1990.	55
5.2.	Series de tiempo en Punta Lengua de Vaca con mayor número de muestras	56
5.3.	Función de Autocorrelación empírica para series de tiempo de Punta Lengua de Vaca para datos tomados en: (a) Semana 24 de 1990. (b) Semana 12 de 1990. (c) Semana 14 de 1990.	57
5.4.	Espectro de frecuencias para series de tiempo de Punta Lengua de Vaca para datos tomados en: (a) Semana 24 de 1990. (b) Semana 12 de 1990. (c) Semana 14 de 1990.	58
5.5.	Error cuadrático medio de modelo SARIMA de ordenes $q = 1$ $d = 0$ $D = 1$ sobre series de tiempo en Punta Lengua de Vaca con datos tomados en: (a) Semana 24 de 1990. (b) Semana 12 de 1990. (c) Semana 14 de 1990.	59
5.6.	Resultados de predicción a un paso con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	60
5.7.	Resultados de predicción a un paso con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	61

5.8.	Resultados de predicción a un paso con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	61
5.9.	Resultados de predicción a cinco pasos con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	63
5.10.	Resultados de predicción a cinco pasos con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	63
5.11.	Resultados de predicción a cinco pasos con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	64
5.12.	Resultados de predicción a un paso con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	66
5.13.	Resultados de predicción a un paso con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	66
5.14.	Resultados de predicción a un paso con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	67
5.15.	Resultados de predicción a cinco pasos con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	68
5.16.	Resultados de predicción a cinco pasos con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	69
5.17.	Resultados de predicción a cinco pasos con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.	69

Capítulo 1

Introducción

En el presente capítulo se realiza una presentación general del trabajo, mostrando la motivación del mismo, los objetivos y los alcances de éste.

1.1. Motivación

En la actualidad, gran cantidad de países se encuentran potenciando el desarrollo de energías renovables no convencionales (ERNC) [1], entre los cuales se encuentra Chile. En el caso de nuestro país, al año 2010 existen cuatro centrales eólicas operando en el Sistema Interconectado Central (SIC), lo cual hace relevante contar con herramientas que permitan optimizar el costo de operación de las plantas, así como también del sistema eléctrico en general. La aplicación más importante que tiene la predicción de potencia generada en la actualidad es reducir la necesidad de equilibrio de energía y reserva de poder, los cuales son requeridos para integrar la energía eólica en un sistema eléctrico de potencia [2, 3]. Esto conlleva menores costos de integración para los parques eólicos, menores emisiones de las plantas eléctricas que son usadas para balancear el sistema que en general son centrales térmicas y, consecuentemente, un aumento del valor de la energía eólica [4].

Los objetivos de predecir la potencia generada dependen de la aplicación. Entre ellos podemos mencionar los siguientes:

1. Programación óptima de las plantas de energía y de equilibrio de potencia. Una predicción certera de la energía generada es requerida por el sistema de control de las plantas.

El horizonte de tiempo depende del marco regulatorio y técnico por el cual se encuentra determinada la planta.

2. Determinación de la reserva de poder que se debe tener lista para balancear la potencia. Ello requiere una predicción de la certeza de la energía eólica a producir.
3. Operación de la malla y su respectivo manejo de la congestión, se debe tener una predicción de las corrientes y de la potencia en cada punto o área de la malla.
4. Correcta integración con un sistema interconectado como el SIC o el Sistema Interconectado del Norte Grande (*SING*). Se debe tener una predicción de la potencia a generar dentro de los próximos 15 días o, al menos, tener un valor mínimo de potencia a generar, de tal manera que la entidad correspondiente realice un despacho económico que optimice el costo del sistema.
5. Programación de mantenimiento de los parques eólicos, plantas de energía convencionales o líneas de transmisión. Se debe tener una predicción de la potencia a generar dentro de los días que duren los trabajos.

De lo anterior, la principal motivación de construir un modelo que sea capaz de predecir la potencia generada por un parque eólico es la gran utilidad que este tiene en el período de planificación y operación del parque, con lo cual se contribuye a mejorar la gestión y eficiencia que ellos tienen.

1.2. Objetivos Generales

El objetivo general del presente trabajo de título es desarrollar un modelo predictivo, de corto plazo, que permita pronosticar la potencia generada por un parque o turbina eólica, por medio de mediciones históricas de variables como el viento, temperatura, presión y la potencia eléctrica generada por un parque o turbina.

1.3. Objetivos Específicos

Los objetivos específicos de este trabajo de título son:

- Determinar si existen modelos físicos del viento que puedan ser empleados en un modelo predictivo.
- Conocer el estado del arte de modelos de predicción de energía eólica y en especial la relación existente entre la potencia mecánica del viento y la potencia eléctrica generada que permite construir un modelo predictor.
- Disponer de modelos de estimación utilizados en el problema de predicción de velocidad de viento.
- Desarrollar un modelo de predicción de corto plazo de la potencia generada en parques eólicos usando modelo SARIMA (por las siglas en inglés de Auto Regresivo Integral de Media Móvil Estacional).
- Evaluar el comportamiento de cada modelo utilizando datos simulados y reales, que permita comparar ambas soluciones del punto de vista del error de predicción.

1.4. Alcances

Para la realización de este trabajo, se presentan los alcances de éste:

1. Desarrollo e implementación de un modelo SARIMA para la predicción de potencia generada en parques eólicos ubicados en ciertas localidades a seleccionar.
2. Selección de lugares para los cuales se implementa el modelo de predicción.
3. Estudio estadístico de las series de tiempo reales de velocidad de viento.
4. Realización de cálculos y fijación de requerimientos a partir del sistema diseñado.
5. Establecimiento de límites de horizonte de predicción para el sistema diseñado, así como también restricciones de ubicación del parque eólico a estudiar.

Cabe destacar que el presente trabajo no pretende desarrollar un modelo universal de predicción de corto plazo de potencia generada parques eólicos, para parques ubicados en cualquier lugar de Chile o el mundo, sino que será específico para cada localidad seleccionada, pudiéndose extender a otros lugares. Además se trabaja solamente con datos de velocidad de viento obtenidos en la realidad, y no con historiales de potencia producida, por lo que esta última se simula en base a la relación que existe entre ambas en un aerogenerador.

1.5. Estructura del Documento

El presente trabajo de título se encuentra organizado como sigue. A lo largo del capítulo 2 se muestran antecedentes y el contexto del problema a resolver. Luego, en el capítulo 3 se muestra el estado del arte en técnicas de predicción de velocidad de viento y de potencia generada en una turbina o en un parque eólico. A continuación, en el capítulo 4 se aborda la metodología con la que se enfrenta el problema, para luego en el capítulo 5 mostrar los resultados obtenidos con su análisis y discusión respectivo. Finalmente en el capítulo 6 se dan a conocer las conclusiones de este trabajo de título y cuales son las potenciales líneas de trabajo que podrían desarrollarse para extender este trabajo.

Capítulo 2

Antecedentes

La energía eólica es un recurso altamente variable, pues depende directamente de las condiciones de viento en área geográfica donde se ubique un aerogenerador. Factores como la temperatura, presión y humedad, época del año, topología del terreno, cercanía con el borde costero, juegan un rol central en el comportamiento de un generador eólico. Los modelos de viento pueden ser de tres tipos:

1. Fenomenológicos, basados en la física del viento, y descrito mediante las ecuaciones de Navier-Stokes.
2. Estadísticos, basados en series de tiempo y de modelos de filtrado estadístico.
3. Mixtos.

En este capítulo se presentan los antecedentes necesarios para comprender el contexto y la importancia del problema de estimación de potencia generada en parques eólicos.

El capítulo se encuentra organizado como sigue: En la sección 2.1 se entregan antecedentes sobre el estado de desarrollo de los parques eólicos en Chile, a continuación en la sección 2.2 se muestran las ecuaciones diferenciales derivadas de la física del viento, para luego mostrar la energía que es posible extraer del viento de manera teórica. El capítulo finaliza entregando antecedentes sobre la composición de aerogeneradores, tecnologías de generación y la relación Potencia Eléctrica - Velocidad del viento.

2.1. Energía Eólica en Chile

Desde el año 2004, el gobierno de Chile ha potenciado la explotación de los recursos naturales que tiene Chile para utilizarlas como fuentes de energías renovables no convencionales (*ERNC*). En especial, en los últimos años el número de proyectos de inversión en parques eólicos ha aumentado, ya sea porque el tema ha alcanzado madurez así como también, porque los costos de los equipos ha disminuido por la mayor cantidad de fabricantes existentes. Este fenómeno ha llevado a que se hayan construido 4 parques eólicos en los últimos 5 años en sectores costeros ubicados desde la IV hasta la VIII región [5]. Según datos del Centro de Despacho Económico del Sistema Interconectado Central (*CDEC-SIC*), existe una potencia instalada de $165,7[MW]$ y con los proyectos aprobados por la Comisión Nacional del Medio Ambiente (*CONAMA*) se alcanzará una capacidad instalada que superará los $200[MW]$ sólo en centrales eólicas en el año 2011. Además, mensualmente los parques eólicos inyectan una energía del orden de $50.000[MWh]$ al Sistema Interconectado Central y Sistema Interconectado del Norte Grande.

Cabe destacar que en los modelos de optimización utilizados por CDEC-SIC, las centrales de ERNC no son incluidas en la programación ya que las empresas generadoras que son propietarias de los parques eólicos no son capaces de entregar un pronóstico de la potencia que pueden generar en el corto plazo. Ello conlleva a que el punto de operación del sistema interconectado central no sea óptimo en el sentido del mínimo costo de la energía del sistema.

Estos antecedentes, dan a conocer la gran importancia que tiene el desarrollar un modelo aceptable para predecir la potencia generada en parques eólicos.

2.2. Naturaleza y Fenomenología del Viento

En teoría, la dinámica de la atmósfera es tan compleja como la de un sistema no lineal con infinitos grados de libertad. Su dinámica puede ser descrita matemáticamente en términos de las ecuaciones de movimiento que son derivadas de los principios de conservación de masa, momentum y calor. A pesar de esto, la estructura no lineal de estas ecuaciones fundamentales

hace que no se tengan soluciones analíticas, por lo que se realizan aproximaciones numéricas que se asemejan a estados realistas de la atmósfera y que en la mayor parte de los casos, solamente pueden ser obtenidas bajo hipótesis simplificadoras.

En el contexto de la generación de energía eólica existen dos enfoques importantes para describir el comportamiento de la atmósfera. El primero se basa en un pronóstico de su comportamiento dinámico de la atmósfera y, en particular, de la velocidad del viento. Ella puede ser obtenida resolviendo numéricamente las ecuaciones de movimiento del viento por medio de modelos numéricos de predicción tiempo (*NWP - Numerical Weather Predictions*). Su resultado es utilizado para generar la predicción de potencia basada en un modelo de transformación mecánico-eléctrico (la curva de potencia del aerogenerador) y la resolución espacial con que fue resuelto el problema. El segundo enfoque busca adaptar el modelo *NWP* por medio de modelos más simples para describir las condiciones locales de viento en un sitio dado.

Para determinar la complejidad de resolver las ecuaciones de movimiento, se observa que la atmósfera se encuentra dividida en diversas capas que separan diferentes regímenes de flujo. Estas capas están definidas por los efectos físicos predominantes que influyen en la dinámica. En el caso de la energía eólica, la tropósfera es la capa considerada, ya que se extiende en los primeros cinco a diez kilómetros de atmósfera sobre la Tierra, cuyos regímenes de viento son mostrados en la figura 2.1.

Alturas mayores de $1[km]$ son el dominio de sistemas sinópticos de presión de gran escala. Su campo de viento es ampliamente dominado por los gradientes horizontales de presión y temperatura, así como también por la fuerza de Coriolis causada por la rotación de la Tierra. Dado que la fuerza de la superficie terrestre es débil, la fuerza de Coriolis y los gradientes de presión se consideran a menudo como balanceadas, llevando a que el viento se mueva paralelo a las isóbaras (región cuyos puntos tienen igual presión).

El campo de viento geostrófico es considerado como la principal fuerza directiva del flujo en la capa atmosférica subyacente denotada como la capa límite planetaria (*PBL - planetary boundary layer*). El viento en la PBL es dominado por la influencia de la fricción ejercida por

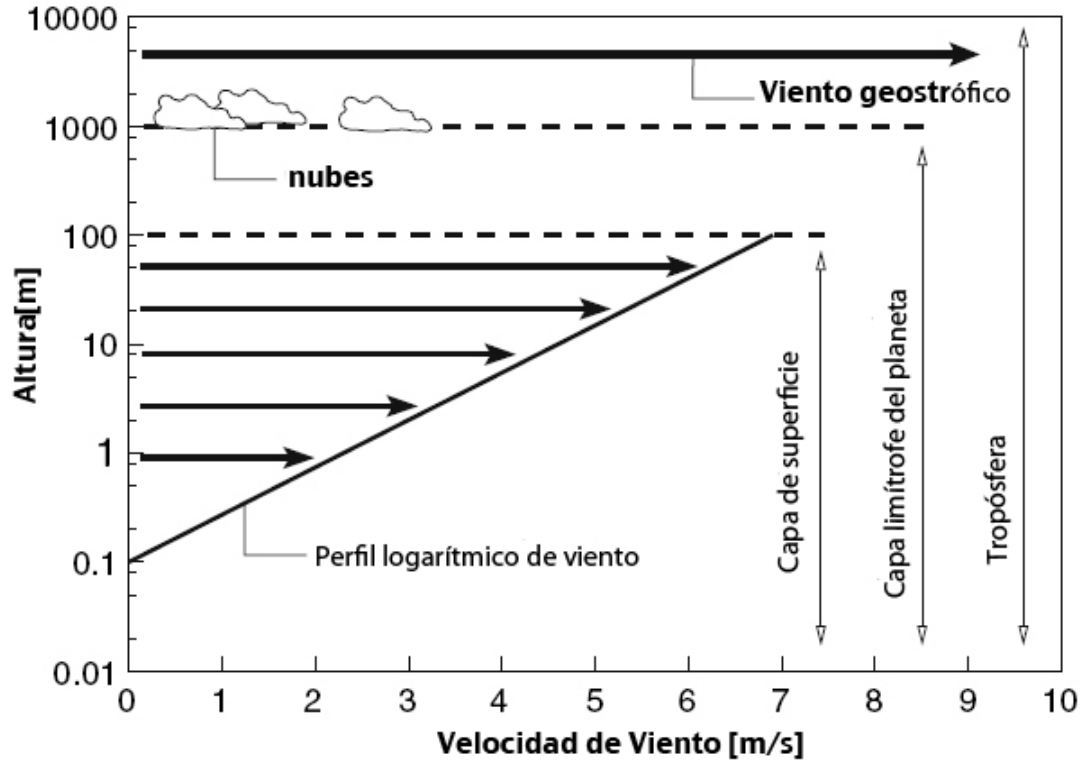


Figura 2.1: Regímenes de viento en la Tropósfera

la superficie de la Tierra. Típicamente, el flujo de aire cercano a la superficie es turbulento, lo cual provee un mecanismo de alto acomplamiento efectivo entre las velocidades del viento a distintas alturas, llevando al transporte vertical del momento horizontal que es dirigido hacia la Tierra donde la velocidad del viento se anula. Este flujo de momentum basado en la mezcla turbulenta es por lejos mayor que aquel basado solamente en la viscosidad molecular, así es que el cambio de superficie del viento con la altura depende fuertemente del grado de turbulencia en la atmósfera.

A continuación se detallan modelos cuantitativos de la dinámica de la atmósfera. Para mayor información acerca de la fenomenología del viento se puede consultar [6].

2.2.1. Ecuaciones de Navier-Stokes

Un fluido en la atmósfera se encuentra expuesto a fuerzas exteriores e interiores. Se consideran fuerzas exteriores a las fuerzas de gradiente de presión, la fuerza de Coriolis y la fuerza de gravedad, mientras que se considerarán fuerzas interiores a aquellas causadas por la fricción entre los elementos del fluido. La diferencia entre altas y bajas presiones, con la correspondiente fuerza dirigida desde alta hacia baja presión, constituye un gradiente de presión. La gravedad es dirigida directamente hacia el centro de la tierra independientemente del estado del flujo, mientras que la fuerza de Coriolis producida por la rotación de la tierra es proporcional a la velocidad del elemento y actúa perpendicularmente a la dirección de movimiento. Dado que el aire es un fluido Newtoniano, se asume que presenta fricción molecular que es proporcional al gradiente local de velocidad, por lo que la viscosidad molecular η es una característica del flujo y no depende del estado de éste.

El principio básico de la conservación del momentum de la dinámica del flujo de aire, puede ser empleado para modelar pues pueden utilizarse para describir las ecuaciones del cambio de momentum que ejercen las fuerzas sobre las partículas del fluido. Esto lleva a la ecuación de Navier-Stokes de la atmósfera, dada por:

$$\underbrace{\rho(\partial_t \vec{u} + \vec{u} \cdot \nabla \vec{u})}_{\text{cambio de momentum}} = \underbrace{-\nabla p}_{\text{gradiente de presión}} + \underbrace{2\rho \vec{u} \times \vec{\Omega}}_{\text{Fuerza de Coriolis}} + \underbrace{\eta \nabla^2 \vec{u}}_{\text{fricción molecular}} - \underbrace{f_g}_{\text{fuerza de gravedad}}, \quad (2.1)$$

Donde Ω es la velocidad rotacional de la tierra, ρ es la densidad del aire y η es la viscosidad dinámica. Aquí se consideran cuatro variables, tres correspondientes al vector de velocidad \vec{u} y una a la presión atmosférica p . Por lo tanto, se tienen tres ecuaciones para cuatro variables. La cuarta ecuación necesaria se deriva a partir del principio de conservación de masa y la asunción que la densidad del aire es constante. Esta es la llamada ecuación de continuidad y está dada por:

$$\nabla \cdot \vec{u} = 0. \quad (2.2)$$

Estas dos ecuaciones toman en cuenta todas las situaciones de flujo, a excepción de los efectos térmicos, y describen regímenes de viento en todas las escalas.

La no linealidad en la ecuación 2.1 permite una abundante variedad de estados de flujo posibles y es, en particular, responsable del comportamiento caótico del viento para algunas condiciones de borde. Desafortunadamente, esta estructura no lineal no permite encontrar una solución analítica a la ecuación de Navier-Stokes. Aunque podemos describir la dinámica de la atmósfera en la mayor parte de los casos, no es posible determinar una expresión analítica o una solución única del campo de velocidades del viento. Adicionalmente, el flujo atmosférico real es mayormente turbulento, lo que dificulta la solución del problema debido a la irregularidad del movimiento del fluido y las fluctuaciones de sus variables

Con el propósito de enfocarse en el flujo promedio y eliminar los detalles del movimiento fluctuante de las partículas del fluido, las ecuaciones de movimiento son simplificadas. Este enfoque es llamado promedio temporal de las ecuaciones de Reynolds (*Reynolds-Averaged Equations*) [6]. Estas ecuaciones son la base para el estudio analítico y modelamiento numérico del flujo atmosférico.

2.2.2. Ecuaciones Promediadas de Reynolds

Las variables en la ecuación de Navier-Stokes junto con la ecuación de continuidad describen la dinámica de las variables en todo instante y posición. El efecto macroscópico de estas fluctuaciones en el comportamiento promedio del flujo es lo que generalmente se desea conocer. Por lo tanto, si seguimos el enfoque de Reynolds [6], las ecuaciones de Navier-Stokes pueden ser reescritas para describir las variables promedio y no instantáneas, y al mismo tiempo, cubrir las propiedades de turbulencia del flujo del aire.

En estas ecuaciones se utiliza el promedio temporal de cada una de las variables de las ecuaciones de Navier-Stokes, y de continuidad, por lo que cada variable es descrita por medio de su valor medio en un intervalo de tiempo dado T y por una parte que varía en el tiempo; Por ejemplo la componente i -ésima de velocidad del viento se puede escribir como:

$$u_i(t) = U_i(t) + u'_i(t) \text{ con } U_i(t) = \overline{u_i(t)}, \quad (2.3)$$

donde el promedio móvil sobre un período T se encuentra dado por:

$$\overline{u_i(t)} = \frac{1}{T} \int_t^{t+T} u_i(t') dt'. \quad (2.4)$$

El largo del intervalo de tiempo es arbitrario, pero fijo. Debe ser elegido de tal manera que la escala de tiempo de las turbulencias se encuentre separado de la escala de tiempo del promedio de la velocidad. A menudo se utiliza $T = 10[min]$. Además, por definición se deriva que:

$$\overline{u'_i(t)} = 0. \quad (2.5)$$

Las variables \vec{u} y p son substituidas por sus descomposiciones análogas a las mostradas anteriormente. El resultado es el promedio temporal de las ecuaciones de Navier-Stokes, las cuales son muy parecidas a las originales, salvo por un término adicional que describe una fuerza adicional en el flujo promedio debido a la turbulencia. Esta ecuación está dada por:

$$\rho(\partial_t \vec{U} + \vec{U} \cdot \nabla \vec{U}) = -\nabla p + 2\rho \vec{U} \times \vec{\Omega} + \eta \nabla^2 \vec{U} - f_g - \underbrace{\tau_t}_{turbulencia}. \quad (2.6)$$

El nuevo término adicional es τ_t un tensor denominado de estrés turbulento:

$$\tau_{t,ij} = \rho \overline{u'_i(t)u'_j(t)}. \quad (2.7)$$

el cual está presente debido a la no linealidad de la ecuación de Navier-Stokes y describe la correlación desconocida *a priori* entre las fluctuaciones de las componentes del viento.

La ventaja que presentan estas ecuaciones es que se encuentran referidas a variables poseen mejor comportamiento y sin fluctuaciones irregulares que disturben seriamente los términos de la ecuación de Navier-Stokes. Por lo tanto, las ecuaciones obtenidas habilitan la identificación de distintos regímenes de flujo en la atmósfera concordante con los efectos físicos dominantes.

Finalmente, para hallar las soluciones de estas ecuaciones se debe describir el efecto de las turbulencias, teniéndose la existencia de distintos métodos para parametrizar su efecto.

2.2.3. Rugosidad

La rugosidad de una superficie se determina en base al tamaño y distribución de los elementos de rugosidad que contiene, y se evalúa mediante un parámetro denominado “longitud de rugosidad”, el cual representa la altura a la cual la velocidad media del viento es cero, considerando una variación logarítmica del viento con la altura. Consideremos una altura dada h , una sección S de cara al viento de los elementos de rugosidad y un obstáculo de área horizontal A . Luego, la rugosidad Z_0 puede expresarse en términos de los parámetros h , S y A de la siguiente forma:

$$Z_0 = \frac{hS}{2A}.$$

Esta expresión entrega valores razonables cuando $A \gg S$, pero si A es del orden de S , el valor obtenido para Z_0 es muy grande. Este escenario es propio de ciudades o bosques, sitios en donde los elementos de rugosidad están muy juntos. En la tabla 2.1 se muestran valores típicos de rugosidad para diversos escenarios:

Z_0 [m]	Lugar
1	Ciudad o bosque
0.5	Suburbios
0.3	Cinturones verdes
0.2	Arbolado abundante
0.1	Campo
0.05	Campo abierto
0.03	Campo sin construcciones ni arboles
0.01	Pistas de aeropuertos, Hierba cortada
0.005	Terreno descubierto
0.001	Superficies nevadas
0.0003	Superficies de arena
0.0001	Superficie acuática

Cuadro 2.1: Rugosidad en diversos escenarios.

2.2.4. Variación del viento con la altura

La rugosidad de la superficie hace disminuir la velocidad del viento a baja altura. Así, la velocidad del viento suele aumentar con la altura. Normalmente los datos de viento se obtienen a una altura estándar de 10 metros. Como en general los rotores de los aerogeneradores se ubican a alturas superiores, se hace necesario extrapolar los datos a la altura deseada. Si se manejan registros de velocidad de viento correspondientes a dos alturas, es posible determinar la rugosidad local que explica la diferencia entre las velocidades mediante la relación:

$$r = A_1^{\frac{V_2}{V_2 - V_1}} \cdot A_2^{\frac{-V_1}{V_2 - V_1}}.$$

donde A_1 es la altura 1, A_2 es la altura 2, V_1 es la velocidad del viento a la altura 1 y V_2 es la velocidad del viento a la altura 2. Con la rugosidad r determinada, es posible emplear un método simple para extrapolar la velocidad a otra altura. Asumiendo que la velocidad del viento varía logarítmicamente con la altura, se cumple la siguiente expresión [6]:

$$V_2 = V_1 \cdot \frac{\ln(A_2) - \ln(r)}{\ln(A_1) - \ln(r)}. \quad (2.8)$$

2.3. Potencia Extraíble del Viento

Dado que el viento se traslada a cierta velocidad, tiene asociada una cantidad de energía cinética dada, la cual puede ser transformada en energía mecánica y ésta en eléctrica. Para esto, se debe cuantificar teóricamente de manera aproximada la energía cinética asociada a una masa de aire.

2.3.1. El Viento y su Potencia

El viento es una masa de aire que se traslada de una zona a otra, producto de gradientes térmicos y de presión. Toda masa m de aire que se mueve a velocidad constante v , tendrá una energía cinética dada por:

$$E = \frac{1}{2}mv^2. \quad (2.9)$$

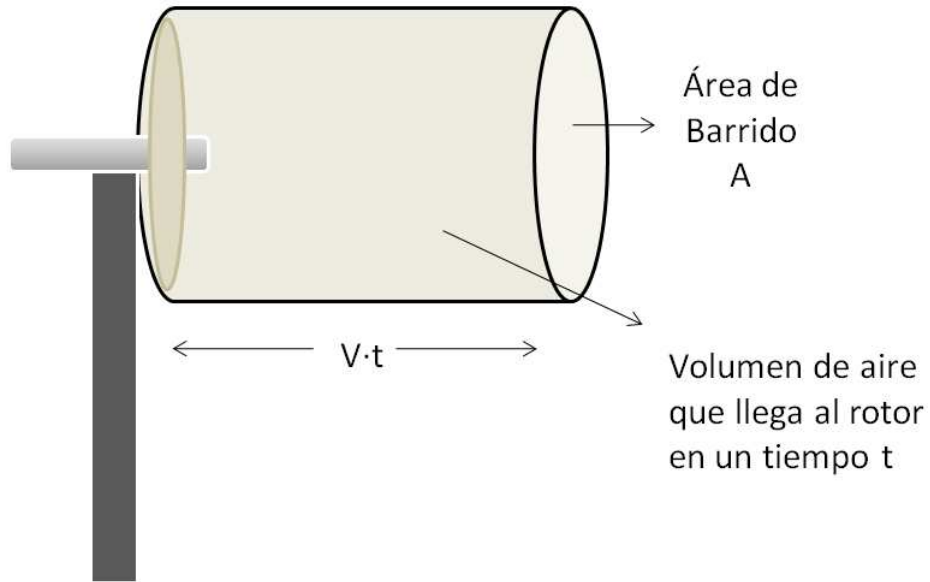


Figura 2.2: Masa de aire que mueve las aspas del aerogenerador

La potencia que se extrae de este bloque de aire, depende de la masa de éste y de la velocidad con que se desplaza en ese instante. Es decir, la potencia del viento es la derivada de la energía, la cual depende de la tasa de cambio de masa de un instante a otro:

$$P_{viento} = \frac{dE}{dt} = \frac{1}{2} \frac{dm}{dt} v^2 + mv \frac{dv}{dt}. \quad (2.10)$$

La tasa de cambio de masa $\frac{dm}{dt}$ del bloque de aire en un instante dado está determinado por su densidad ρ , su área A y su velocidad instantánea $\frac{dx}{dt}$, que supondremos de valor constante v en una sola dirección:

$$\frac{dm}{dt} = \rho A \frac{dx}{dt}. \quad (2.11)$$

De lo anterior, se concluye que la potencia del viento es función de la densidad del aire, el área del bloque de aire, y la velocidad del viento al cubo:

$$P_{viento} = \frac{dE}{dt} = \frac{1}{2} \rho A v^3. \quad (2.12)$$

Por lo tanto, entre mayor sea el área de la masa interceptada por la turbina, mayor resulta ser la potencia extraíble por ella. Por esta razón, los aerogeneradores poseen aspas que son de gran tamaño.

2.3.2. Extracción de la Potencia del Viento

La conversión de la energía cinética del viento a energía mecánica rotatoria del eje, consiste en reducir la velocidad de la masa de aire interceptada por la turbina. Esto se ejemplifica a continuación.

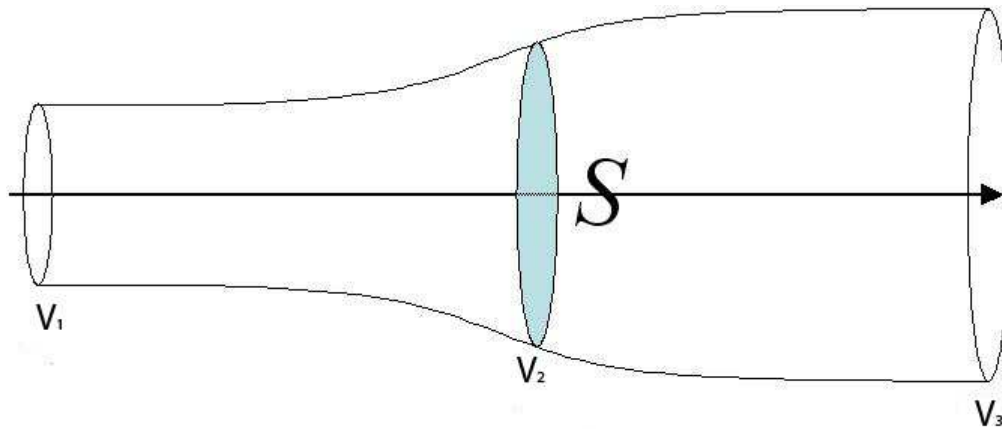


Figura 2.3: Velocidades del viento en una Turbina Eólica

$$E_{\text{extraíble}} = \frac{1}{2}m(v_1^2 - v_3^2) \quad (2.13)$$

con lo cual la potencia extraíble es:

$$P_{\text{extraíble}} = \frac{1}{2}\dot{m}(v_1^2 - v_3^2) \quad (2.14)$$

Sin embargo, aun en el caso ideal, no es posible extraer el 100 % de la potencia del viento. Esto se debe a que en el caso extremo en que $v_3 = 0$ se provocaría una congestión en el flujo de aire, teniendo como consecuencia que $\dot{m} = 0$, lo cual se traduciría en una detención total del bloque de aire, y consecuentemente en una potencia extraída de valor nulo. Por lo tanto la potencia extraída se podrá conocer si se conoce la diferencia de masa del aire incidente en el rotor de la turbina, el cual depende de la velocidad del viento en ese punto.

$$P_{\text{extraíble}} = \frac{1}{2} A \rho v_2 (v_1^2 - v_3^2). \quad (2.15)$$

Lo anterior, permite divisar que debe existir una velocidad v_2 óptima en el eje tal que se pueda maximizar la extracción de potencia del viento incidente. Según Betz y Schmitz [7] la velocidad v_2 se puede expresar en función de la velocidad incidente v_1 y la velocidad posterior a la extracción v_3 . Según la teoría de Betz, esta velocidad v_2 corresponde al promedio entre v_1 y v_3 . De tal manera, la máxima potencia extraída se puede expresar como:

$$P_{\text{textextraíbleMAX}} = \frac{1}{2} A \rho \frac{(v_1 + v_3)}{2} (v_1^2 - v_3^2) \quad (2.16)$$

2.4. Descripción General de un Aerogenerador

Un aerogenerador es una máquina eléctrica rotatoria que es capaz de generar energía eléctrica por medio de la conversión electromecánica de la energía que posee el viento.

Existen distintos tipos de aerogeneradores, clasificándose mayormente entre aquellos de eje horizontal y de eje vertical. Para el caso de los aerogeneradores con hélice de eje horizontal, por lo general, poseen un diseño relativamente estandarizado el cual se detalla a continuación.

2.4.1. Sistema de Soporte

Consiste en la torre de soporte del aerogenerador y su correspondiente fundación de anclaje. Su función es mantener el aerogenerador a la altura correcta de funcionamiento por sobre el nivel del suelo, debiendo ser capaz de tolerar el peso de éste y las diversas exigencias que conllevan la exposición a fuertes vientos y el sostener un cuerpo de gran envergadura rotando.

Hay cuatro tipos de torres: tubulares, de enrejado, de mástil tensado e híbridas, tal como se aprecia en la figura 2.4.

Las torres de tipo enrejado facilitan la circulación de aire a través de la misma, disminuyendo las solicitaciones mecánicas sobre el rotor de la máquina. Estas pueden estar hechas de hierro o madera y usualmente requieren menos material que sus contrapartes tubulares.



Figura 2.4: Tipos de Torres para aerogeneradores

Los diseños tubulares son construidos en metal o concreto, y requieren escaso mantenimiento, menor que el caso de la torre de enrejado. La mayoría de los grandes aerogeneradores se entregan con torres tubulares de acero, fabricadas en secciones de 20-30 metros con bridas en cada uno de los extremos, y son unidas con pernos en el lugar de instalación. Las torres son tronco-cónicas, con el fin de aumentar su resistencia y al mismo tiempo ahorrar material.

Muchos de los aerogeneradores pequeños están construidos con delgadas torres de mástil sostenidas por cables tensores. La ventaja es el ahorro de peso y, por lo tanto, de costo. Su principal desventaja es el difícil acceso a las zonas alrededor de la torre, lo que las hace menos apropiadas para zonas agrícolas.

Las fundaciones se realizan de concreto reforzado, y deben ser acordes con la altura de la torre y la envergadura de la máquina rotatoria que ésta debe soportar.

En la zona superior de la torre existe un cabezal y una plataforma, que permitan anclar

el aerogenerador a la torre y a la vez lo deja libre para rotar, con el fin de orientarse según la dirección del viento.

2.4.2. Sistema de Transmisión

El sistema de transmisión de un aerogenerador está compuesto por el eje de rotación, la caja amplificadora de revoluciones y el generador eléctrico. Su función principal es transmitir la energía cinética de la rotación de las aspas al generador eléctrico, equipo que se encarga de de la conversión electromecánica de la energía.

El generador eléctrico puede ser de diversos tipos, ya sea de corriente alterna o continua. Sin embargo los más usados en la industria hoy en día son los generadores eléctricos de corriente alterna, de preferencia generadores sincrónicos, principalmente para máquinas de baja potencia (menores a 100 kW), y generadores de inducción, principalmente en máquinas de alta potencia [7].

La caja amplificadora de revoluciones es un componente opcional, que sólo se ocupa de modificar la velocidad de giro del eje que conecta con el generador eléctrico. Luego, dependiendo del criterio de diseño de la máquina y la forma de operar de ésta, puede incluirse o no.

2.4.3. Sistema de Rotor

El sistema de rotor está compuesto por el conjunto de aspas y una masa central, y que en conjunto conforman la hélice, encargada de extraer parte de la energía cinética del viento y transformarla en energía cinética de rotación del eje del aerogenerador.

Las aspas pueden seguir diversos perfiles aerodinámicos, diseñados para poder captar la máxima energía cinética. A su vez la hélice puede estar compuesta por una, dos, tres o más aspas, siendo el modelo de tres aspas el que predomina fuertemente en el mercado mundial de aerogeneradores.

Según sea la forma de operación del aerogenerador, se eligen distintos tipos de control de éste.

2.4.4. Sistema de Control

El principal objetivo del sistema de control es mantener constante el nivel del voltaje generado por la máquina, tarea que se ve dificultada por la naturaleza errática y variable de la intensidad y dirección del viento, lo que genera fluctuaciones en los niveles de tensión que se registran en bornes de la máquina. Por lo que la mayor tarea de este sistema es regular la frecuencia y magnitud del voltaje generado dado que la velocidad de giro del eje del generador depende de la velocidad instantánea del viento.

Rectificar el voltaje generado para luego entregar un voltaje controlado en frecuencia y magnitud.

Cabe mencionar que cada generador tendrá sus propios equipos de electrónica de potencia, de esta forma se logra independencia entre las máquinas.

2.5. Tecnologías de Generación Eólica

Los aerogeneradores utilizados actualmente para la transformación de energía eólica en eléctrica no incorporan tecnología uniforme ya que existen varios prototipos, los cuales implementan distintos mecanismos y tecnologías de generación. Por ejemplo, existen aerogeneradores que utilizan desde máquinas sincrónicas de imanes permanentes con inversores, hasta máquinas de inducción con etapa de control dinámico de deslizamiento [8]. Cada tecnología de generación presenta sus ventajas y desventajas en distintos aspectos tales como su costo, mecanismos de control, calidad, etc.

A continuación se discute brevemente distintas tecnologías de generación.

2.5.1. Generador de Inducción de Velocidad Fija

Esta tecnología de generación es la más simple y común, ya que introduce un generador de inducción de rotor de jaula de ardilla, el cual se acopla directamente a la red.

Este mecanismo de generación se denomina entre los fabricantes como el concepto danés, debido a que fue generado por la empresa danesa Vestas. Sus mayores ventajas consisten en el menor tamaño y bajo costo de inversión, operación y mantenimiento de las máquinas

de inducción. Además, permite implementar métodos simples de control reduciendo de esta forma aun más los costos. Sin embargo, esta tecnología presenta desventajas en lo referente a la calidad del suministro ya que necesita de compensación reactiva de potencia, y además el sistema de control que utiliza no ejerce control sobre velocidad ni el deslizamiento lo cual provoca que las inyecciones de energía a la red se vean fuertemente afectadas por variaciones en la velocidad del viento.

2.5.2. Generador de Inducción de Velocidad Fija con Capacidad FRT

Esta tecnología de generación es igual a la anterior a excepción que presenta además equipos que permiten un control dinámico del consumo de reactivos mediante diversos equipos como son los SVCs, STATCOMs y bancos de condensadores electrónicamente controlados.

2.5.3. Generador de Inducción Doblemente Alimentado

Esta tecnología utiliza una máquina de inducción con acceso al rotor mediante una etapa de inversión, lo cual permite inyectar o extraer energía del rotor controlando el nivel de reactivos y, por consiguiente, la velocidad de giro del rotor de la máquina. Se denomina doblemente alimentado debido a que los reactivos necesarios para la excitación de la máquina son suministrados tanto por el estator como por el inversor.

Las ventajas que presenta el generador doblemente alimentado consisten en un control variable de los reactivos consumidos por el rotor, lo cual permite un control de la velocidad frente a variaciones del viento y perturbaciones del sistema eléctrico, de esta manera ofrece una mejor calidad del servicio. No obstante, estas ventajas se obtienen a un costo mayor debido a la necesidad de un inversor. Además, el acceso al rotor introduce mayores costos de mantenimiento de los generadores.

2.5.4. Generador Sincrónico de Velocidad Variable

La utilización de generadores sincrónicos en sistemas eléctricos requiere que el generador gire siempre a una velocidad constante y en sincronismo con las restantes máquinas sincróni-

cas conectadas a la red, lo cual en el caso de la energía eólica es complicado debido a la velocidad variable del viento. Sin embargo, gracias a la electrónica de potencia, es posible conectar estos generadores sincrónicos de velocidad variable a la red eléctrica mediante una etapa de rectificación e inversión. Esta interfaz permite rectificar la señal de corriente alterna desacoplándola de la red eléctrica, para luego ser convertida, mediante una etapa inversora, en una señal de corriente alterna a la misma frecuencia de la red eléctrica.

Las principales ventajas de este tipo de generación recaen en la mayor capacidad de control en la inyección de potencia activa y reactiva a la red, mejorando sustancialmente la calidad del servicio. No obstante, el uso de rectificadores e inversores de potencia igual a la del parque, elevan los costos de inversión en gran medida. Además, el uso de estas etapas de rectificación e inversión conllevan a un aumento de la contaminación armónica de la red, los cuales en el caso de una integración masiva de esta tecnología puede dañar la integridad del sistema interconectado.

2.6. Relación Potencia Mecánica - Velocidad del Viento

La curva de potencia de un aerogenerador es un gráfico (entregado por el fabricante) que caracteriza potencia eléctrica generada por el aerogenerador a diferentes velocidades del viento. Las curvas de potencia se obtienen a partir de medidas realizadas en campo, donde un anemómetro es situado sobre un mástil relativamente cerca del aerogenerador (no sobre el mismo aerogenerador ni demasiado cerca de él, pues el rotor del aerogenerador puede crear turbulencia, y hacer que la medida de la velocidad del viento sea poco fiable).

Si la velocidad del viento no varía bruscamente, pueden usarse las medidas de la velocidad del viento realizadas con el anemómetro y leer la potencia eléctrica disponible directamente del aerogenerador, y dibujar los dos tipos de valores conjuntamente en un gráfico como el que se muestra en la figura 2.5.

En la curva de potencia mostrada, puede verse una nube de puntos esparcidos alrededor de las curvas de potencia. El motivo es que en la práctica la velocidad del viento siempre fluctúa, y no se puede medir exactamente la columna de viento que pasa a través del rotor del

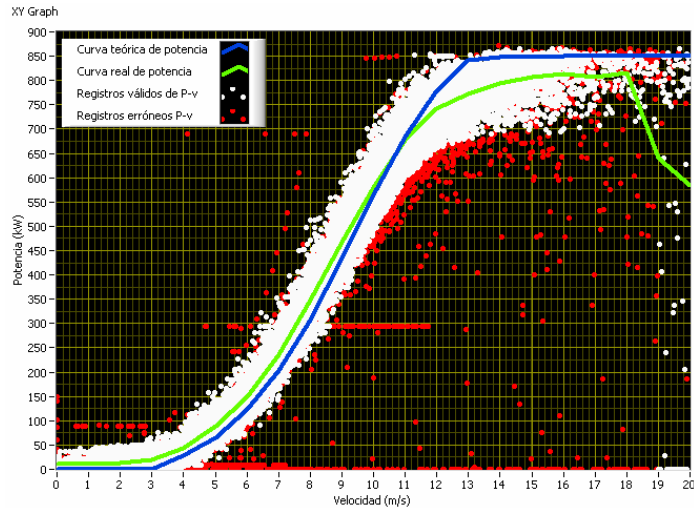


Figura 2.5: Curva de potencia de un aerogenerador

aerogenerador (colocar un anemómetro justo enfrente del aerogenerador no es una solución factible, ya que el aerogenerador también frenará el viento enfrente de él). Por esta razón, en la práctica se toma un promedio de las diferentes medidas para cada velocidad del viento, y se traza el gráfico con esos promedios.

Si se tiene un 3 % de error en las mediciones de la velocidad del viento, entonces el error en las cifras de energía del viento puede contener un 9 % de error. En consecuencia, pueden existir errores hasta de 10 % incluso en curvas certificadas. De igual manera, las curvas de potencia están basadas en medidas realizadas en zonas de baja intensidad de turbulencias, y con el viento viniendo directamente hacia la parte delantera de la turbina. Las turbulencia locales y los terrenos complejos, tal como aerogeneradores situados en una pendiente rugosa, pueden implicar que ráfagas de viento golpeen el rotor desde diversas direcciones. Por lo tanto, en la práctica es difícil reproducir exactamente la curva en una localización cualquiera. Una curva de potencia no indica de manera exacta cuanta potencia produce un aerogenerador a una cierta velocidad del viento media sino más bien el valor esperado de esta, tomando un rol fundamental la desviación estandar de la potencia generada para cada velocidad.

Finalmente, debe tenerse el cuenta el hecho de que la turbina puede no estar girando a la temperatura, rugosidad y presión de aire estándar, y consecuentemente, los cambios en la

densidad del aire.

Capítulo 3

Estimación de Velocidad del Viento

3.1. Estimación de la Velocidad del Viento

La tarea de estimar la velocidad del viento debe tener como primer paso definir si lo que se desea obtener es un único valor de velocidad de viento o bien, un conjunto de valores de velocidad cada uno con una probabilidad asociada, conformando de esta manera una función de densidad de probabilidad [3, 4, 9].

Existen distintos métodos bajo los cuales se aborda el problema de predicción descrito anteriormente, que por su naturaleza, es altamente variable y no-lineal. Los principales enfoques son:

- Enfoque Físico, el cual tiene como objetivo describir con leyes físicas la dinámica de la velocidad del viento así como también el proceso de conversión en energía eólica.
- Enfoque estadístico paramétrico, el cual tiene como objetivo describir la conexión existente entre la energía eólica generada y las condiciones atmosféricas y de época del año, para un instante dado, por medio de relaciones estadísticas.
- Enfoque de inteligencia computacional, el cual usa métodos estadísticos no-paramétricos o heurísticos para “aprender” la relación que existe entre la potencia generada y las condiciones atmosféricas y de época del año, para un instante dado.

A continuación, se detallan distintas técnicas de estimación de velocidad del viento que se encuentran dentro de los enfoques mencionados anteriormente.

3.2. Métodos Basados en Histogramas

3.2.1. Histogramas de Viento y Estimación de Energía

Esta técnica se basa en la toma de medidas de velocidad media del viento, en intervalos de tiempo regulares T (que pueden ser minutos, días, horas, etc.) para luego, usando la muestra de datos obtenida anteriormente, construir un histograma de las velocidades medias. El histograma del viento H_v , se construye agrupando los datos en categorías de viento previamente definidas, como por ejemplo pueden ser 2, 4 ó 6 $\frac{m}{s}$. Luego, usando el histograma de frecuencias de velocidad, se puede calcular el porcentaje de veces en que la velocidad del viento se encontró en un intervalo dado durante el período de observación.

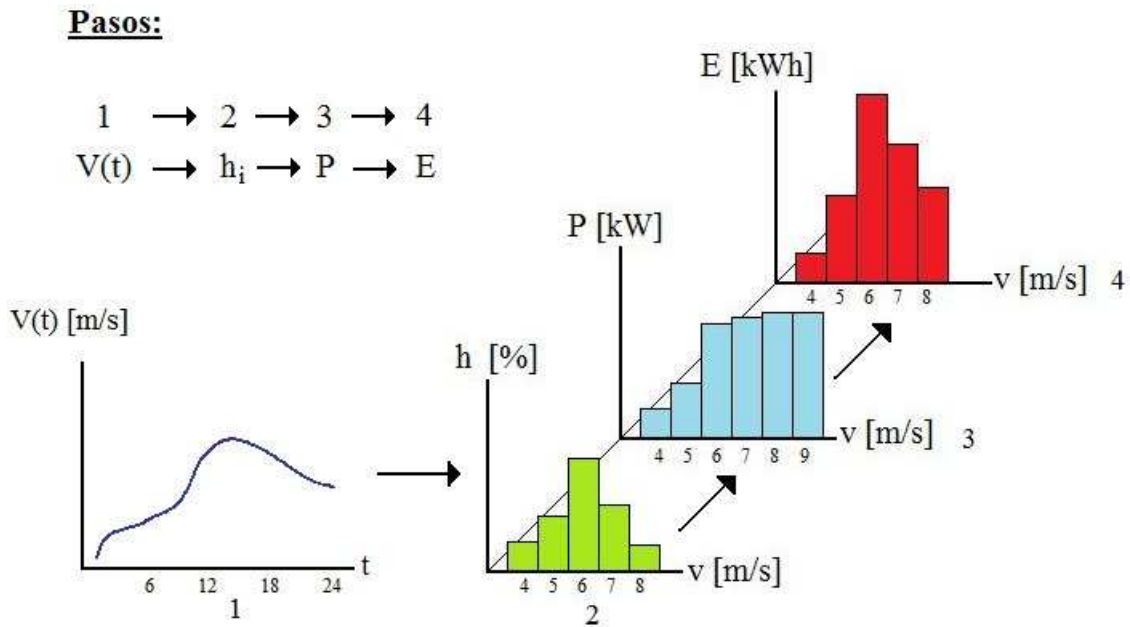


Figura 3.1: Esquema del Método de Histogramas

Una vez calculado el porcentaje de veces que se tuvo velocidades para cada intervalo dado, por medio de la curva de potencia del conjunto turbina-generador, se calculan las potencias asociadas a cada intervalo de velocidades y el histograma con los porcentajes, se calcula el promedio ponderado de potencia obtenida, que corresponde al estimador insesgado de la potencia generada esperada.

3.2.2. Distribución de Weibull

Esta técnica de estimación para la velocidad del viento, corresponde a un refinamiento de la técnica mencionada anteriormente, en donde se asume como hipótesis que la velocidad del viento sigue una distribución de Weibull.

Una variable aleatoria v sigue una distribución de Weibull de parámetros c , k si su función de densidad de probabilidad es:

$$f_W(v) = \left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1} \exp\left[-\left(\frac{v}{c}\right)^k\right], \quad (3.1)$$

cuya función de probabilidad acumulada es:

$$F_W(v) = 1 - \exp\left[-\left(\frac{v}{c}\right)^k\right]. \quad (3.2)$$

El momento de orden n de la distribución de v es:

$$E[v^n] = \int_0^{+\infty} v^n f_W(v) dv = c^n \Gamma\left(1 + \frac{n}{k}\right), \quad (3.3)$$

en donde $\Gamma(\cdot)$ corresponde a la función Gamma que se define por:

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt. \quad (3.4)$$

Esta distribución describe aproximadamente [10, 11] a la distribución que tiene la velocidad del viento para un intervalo de tiempo de un mes aproximadamente, teniéndose que variar los parámetros k y c de escalamiento respectivos para ajustarlos a la época del año en la que se desea estimar la velocidad del viento. De esta manera, es posible estimar la potencia media generada por el parque para cada período de tiempo que se desee.

Para ajustar los valores de los parámetros k y c existen distintos métodos, entre los cuales se destacan mayormente.

Método de la varianza

En este método el parámetro c se estima sólo en base a la velocidad media. Para el factor de forma k se han determinado tres expresiones distintas dependiendo del nivel de la varianza

asignada a los datos, la cual se clasifica como baja, media o alta. En el caso que la velocidad media se encuentre expresada en $\frac{m}{s}$ las expresiones son las siguientes:

$$c = \frac{\bar{v}}{1 + \frac{1}{k}},$$

$$k = s\bar{v}^{0.5}$$

donde el parámetro s es igual a 1.05, 0.94 y 0.73 si el nivel de varianza estimado es bajo, medio o alto, respectivamente.

Método de los Cuartiles

Cuando se dispone del histograma de velocidades, pero se desconoce la serie de valores que lo generaron, se puede utilizar la expresión de la función de distribución de probabilidad acumulada para calcular las velocidades correspondientes a los cuartiles de la distribución de probabilidades. Definimos V_p como el cuartil cuya probabilidad acumulada es p :

$$V_p = c[-\ln(1 - p)]^{\frac{1}{k}}.$$

Si se toman la división de los valores de velocidad de los cuartiles 0.75 y 0.25, se obtiene:

$$\frac{V_{0.75}}{V_{0.25}} = \left[\frac{\ln(0.25)}{\ln(0.75)} \right]^{\frac{1}{k}}.$$

Luego, el valor de k es:

$$k = \frac{1.573}{V_{0.75} - V_{0.25}}.$$

Estimando la mediana V_{me} a partir del histograma, se deduce una expresión para c usando el hecho que $V_{me} = V_{0.5}$ y por lo tanto:

$$c = \frac{V_{me}}{(\ln 2)^{\frac{1}{k}}}.$$

Método de máxima verosimilitud

Este método consiste en encontrar los valores de c y k de modo de maximizar la probabilidad conjunta de obtener los valores observados en la serie. Para ello se define la función

de verosimilitud L como el logaritmo natural del producto de la función de densidad de probabilidad para todas las velocidades observadas.

$$L = \ln \prod_{i=1}^n p(v_i).$$

Utilizando las propiedades del logaritmo natural, L puede ser expresada como:

$$L = \sum_{i=1}^n \ln(p(v_i)).$$

Esta función depende de los parámetros c y k a través de la función $p(\cdot)$ y su máximo se obtiene en los puntos donde las derivadas parciales con respecto a cada variable se anulan:

$$\frac{\partial L}{\partial \Psi} = \frac{\partial}{\partial \Psi} \left[\sum_{i=1}^n \ln(p(v_i)) \right] = 0.$$

En donde Ψ es cualquiera de los dos parámetros y n corresponde al total de observaciones disponibles para realizar el ajuste. Introduciendo la expresión de $p(\cdot)$ en la ecuación anterior y derivando parcialmente con respecto a c y k , se obtienen las siguientes expresiones:

$$\begin{aligned} c &= \frac{1}{n} \sum_{i=1}^n \left(\frac{V_i}{c} \right)^{(k-1)} V_i \\ k &= \frac{n}{\sum_{i=1}^n \left(\frac{V_i}{c} \right)^k \ln\left(\frac{V_i}{c}\right) - \sum_{i=1}^n \frac{V_i}{c}}. \end{aligned} \tag{3.5}$$

Estas ecuaciones se resuelven simultáneamente mediante métodos numéricos.

Método de mínimos cuadrados

Puesto que la función de probabilidad de una distribución Weibull es no lineal, se trabaja con la función de probabilidad acumulativa mencionada al principio de la sección:

$$F_W(v) = 1 - \exp\left[-\left(\frac{v}{c}\right)^k\right]$$

Al despejar la exponencial y aplicar dos veces logaritmo natural, se obtiene:

$$\ln(-\ln(1 - F_W(v))) = k \ln(v) - k \ln(c)$$

Lo cual se puede tomar como una relación lineal de la forma $y = mx + n$ en que:

$$y = \ln(-\ln(1 - F_W(v))) \quad x = \ln(v) \quad m = k \quad n = -k \ln(c)$$

Los parámetros m y n se determinan por mínimos cuadrados y a partir de éstos se calculan los parámetros k y c de la distribución.

Método basado en resúmenes climatológicos

Este método consiste en la obtención de los parámetros de Weibull a partir de la información que puede ser extraída de resúmenes climatológicos. En el caso chileno se puede utilizar la información incluida en los anuarios de la Dirección Meteorológica de Chile. Estos contienen para cada estación, el porcentaje de calmas, la velocidad media, y la frecuencia de observación de velocidades superiores a un cierto valor (20 ó 30 nudos). El porcentaje de calmas es igual a la probabilidad de ocurrencia de velocidades inferiores a la mínima que el instrumento puede medir V_0 .

$$\% \text{ Calmas} = \int_0^{V_0} p(V) dV = 1 - \exp\left(-\left(\frac{V_0}{c}\right)^k\right)$$

donde $p(\cdot)$ es la f.d.p. de Weibull y V_0 es la velocidad umbral de detección del instrumento (típicamente $0.5 \frac{m}{s}$). De igual manera el porcentaje de excedencia sobre una cierta velocidad V_0 es:

$$\% \text{ Excedencia} = \int_{V_0}^{\infty} p(V) dV = \exp\left(-\left(\frac{V_0}{c}\right)^k\right)$$

y la velocidad media:

$$\bar{V} = c\Gamma\left(1 + \frac{1}{k}\right)$$

Usando dos de las tres ecuaciones anteriores, se determinan los parámetros de la f.d.p. de Weibull.

3.2.3. Distribución de Rayleigh

Esta técnica de estimación para la velocidad del viento, al igual que el caso de la distribución de Weibull, corresponde a un refinamiento de la técnica de estimación basada en

histogramas, en donde esta vez se tiene como hipótesis que la velocidad del viento sigue una distribución de Rayleigh.

Una variable aleatoria v sigue una distribución de Rayleigh de parámetro v_m si su función de densidad de probabilidad es:

$$f_R(v) = \left(\frac{\pi}{2}\right)\left(\frac{v}{v_m^2}\right)\exp\left[-\left(\frac{\pi}{4}\right)\left(\frac{v}{v_m}\right)^2\right] \quad (3.6)$$

para $v \in [0, +\infty)$. Con valor esperado:

$$E[v^n] = \int_0^{+\infty} v^n f_W(v) dv = c^n \Gamma\left(1 + \frac{n}{k}\right) \quad (3.7)$$

en donde $\Gamma(\cdot)$ corresponde a la función Gamma.

Esta distribución corresponde a un caso particular de la distribución de Weibull y en general también ha presentado un buen ajuste a la distribución que tiene la velocidad del viento para un intervalo de tiempo de un mes aproximadamente. Para ajustar los valores del parámetro de la distribución, los métodos aplicados a la distribución de Weibull siguen siendo válidos para este caso, tomando en cuenta que se tiene $k=2$ como condición fija. Mayor información al respecto se puede encontrar en los textos actuales tales como [10–12].

3.2.4. Modelo de Persistencia

Un modelo propuesto en el ámbito de la meteorología corresponde al modelo de persistencia, el cual consiste en hacer una estimación de la velocidad del viento h -pasos hacia adelante suponiendo que la última medición de velocidad de viento persiste en el tiempo. Matemáticamente, este modelo está dado por:

$$\hat{u}_{t+h} = u_t. \quad (3.8)$$

3.3. Modelos Lineales Estacionarios

Los modelos lineales estacionarios son una importante herramienta para pronósticos, y son la base de muchas ideas fundamentales del análisis de series de tiempo [13]. Este tipo de técnicas han sido utilizadas con éxito por si solas en el ámbito de predicción de series de

tiempo de potencia generada en parques eólicos y velocidad de viento en [14,15], en predicción de energía consumida en [16]. Estos modelos también han sido empleados en problemas de predicción de tiempos de viaje en [17], demanda de combustible en [18] y en conjunto con otras técnicas en [19,20]. A continuación se detallan los modelos lineales más básicos.

3.3.1. Proceso Auto Regresivo (AR)

Una serie de tiempo Z_t se dice que es un proceso auto regresivo de orden p (abreviado $AR(p)$) si el proceso cumple con:

$$Z_t = \sum_{k=1}^p \phi_k Z_{t-k} + e_t \quad (3.9)$$

Es decir, el proceso en el tiempo t es una combinación lineal de los p valores anteriores más un ruido de media cero y varianza σ_e^2 .

3.3.2. Proceso de Media Móvil (MA)

Una serie de tiempo Z_t se dice que es un proceso de media móvil de orden q (abreviado $MA(q)$ por *Moving Average*) si cumple con:

$$Z_t = \sum_{k=1}^q \theta_k e_{t-k} + e_t. \quad (3.10)$$

Es decir, el proceso en el tiempo t es una combinación lineal de los q valores de ruido anteriores, en donde el ruido es de media cero y varianza constante σ_e^2 .

Proceso Auto Regresivo de Media Móvil

Una clase de modelo que combina los modelos anteriores, es el modelo auto regresivo de media móvil. Si el modelo cuenta con p términos auto regresivos y q términos para la media móvil, entonces el modelo autorregresivo de media móvil es abreviado por $ARMA(p, q)$ (por sus siglas en inglés *Autoregressive Moving Average*) y se encuentra dado por:

$$Z_t = \sum_{k=1}^q \theta_k e_{t-k} + e_t + \sum_{k=1}^p \phi_k Z_{t-k} \quad (3.11)$$

Proceso Auto Regresivo Integral de Media Móvil

Una clase de modelo más general corresponde al modelo auto regresivo integral de media móvil ARIMA (por sus siglas en inglés *Autoregressive Integrated Moving Average*). En la práctica, la mayoría de las series de tiempo son no-estacionarias por lo que no se pueden aplicar directamente los modelos AR, MA ó ARMA ya que estos corresponden a modelos para procesos estacionarios [13]. Por esta razón, los modelos ARIMA buscan obtener un proceso estacionario por medio de la diferenciación de la serie de tiempo, es decir, se plantea el modelo:

$$(1 - \sum_{k=1}^p \phi_k B^k)(1 - B)^d Z_t = \sum_{k=1}^q \theta_k e_{t-k} + e_t \quad (3.12)$$

en donde B corresponde al operador de atraso, es decir:

$$BZ_t = Z_{t-1}$$

$$B^d Z_t = Z_{t-d}.$$

De esta manera, además de tener el grado de libertad del polinomio autoregresivo y el polinomio de media móvil, también se tiene el nivel de diferenciación bajo el cual se aplica el modelo.

Proceso Auto Regresivo Integral de Media Móvil Estacional

Si las series de tiempo son periódicas, con un periodo de tiempo s , entonces ellas pueden seguir un tipo de modelo llamado SARIMA (por sus siglas en inglés *Seasonal ARIMA*), los cuales son capaces de incorporar dicha periodicidad en el modelo matemáticamente. De esta manera, un modelo SARIMA se encuentra dado por:

$$(1 - \sum_{k=1}^p \phi_k B^k)(1 - B)^d(1 - B^s)^D Z_t = \sum_{k=1}^q \theta_k e_{t-k} + e_t \quad (3.13)$$

De manera más general, el modelo SARIMA multiplicativo es:

$$\phi(B)\Phi(B^s)(1 - B)^d(1 - B^s)^D Z_t = \Theta(B^s)\theta(B)e_t \quad (3.14)$$

En donde:

$$\begin{aligned}\phi(B) &= 1 - \sum_{k=1}^p \phi_k B^k \\ \Phi(B^s) &= 1 - \sum_{k=1}^p \Phi_k B^{ks} \\ \theta(B) &= 1 + \sum_{k=1}^p \Theta_k B^k \\ \Theta(B^s) &= 1 + \sum_{k=1}^p \Theta_k B^{ks}\end{aligned}$$

Luego, para datos que presentan cierto nivel de periodicidad resulta más conveniente usar este tipo de modelos en vez de los presentados en primera instancia [21].

Identificación de sistemas

En las estructuras presentadas anteriormente existen un conjunto de parámetros a determinar dependiendo del modelo que se desee implementar. De esta manera, se hace necesario tener un criterio con el cual obtener dichos parámetros, que típicamente es la minimización del error cuadrático medio. En consecuencia, para cada modelo se obtiene una expresión particular para los parámetros en función de la serie de tiempo, en especial de los estadísticos de ésta [21]. Una vez que se obtienen los parámetros, se debe verificar si el modelo propuesto se ajusta a la serie de tiempo y a su vez, si cumple con las hipótesis bajo las cuales estos se rigen. Ejemplos de esto último se pueden encontrar en [21], [14], [22]. En [21] se presenta el siguiente procedimiento para la identificación de sistemas:

1. Detectar estacionalidad por medio de la función de autocorrelación.
2. Detectar estacionariedad por medio de la función de la función de autocorrelación y el espectro de frecuencias.
3. Diferenciación estacionaria.

4. Diferenciación estacional.
5. Determinación de los órdenes del modelo p y q .
6. Análisis de la forma de la función de autocorrelación.

3.3.3. Filtro de Kalman

El filtro de Kalman fue propuesto en 1960 por R. E. Kalman [23], quien describió un método recursivo para la resolución del problema de estimación del estado de un sistema lineal con ruido aditivo Gaussiano cuando este no es observable. Las mediciones en un sistema en términos genéricos, se corrompen por alguna perturbación externa del sistema, o por errores de medición en alguna(s) etapa(s) de éste. El filtro de Kalman es una herramienta que permite obtener la mejor estimación de una o más variables de un sistema lineal de forma recursiva, lo que permite que se pueda implementar en sistemas de tiempo real. En la actualidad es de interés de muchos investigadores como método de predicción y corrección de datos y se usa especialmente en el área de control automático y en procesamiento estadístico de señales [24].

El filtro de Kalman permite estimar, de manera recursiva, el vector de estado $x \in \mathbb{R}^n$ de un proceso que es modelado por la siguiente ecuación diferencial discreta y estocástica, llamada ecuación de estado:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}, \quad (3.15)$$

en donde:

- x es el vector que contiene toda la información del estado actual del sistema, sin embargo no se puede medir directamente.
- A corresponde a una matriz de $n \times n$ constante, que relaciona la variable de estado x en el paso $k - 1$ con la del paso k .
- B corresponde a una matriz constante de $n \times l$, que relaciona la entrada de control, $u \in \mathbb{R}^l$, con el estado x . Este término es opcional.

- w es el proceso aleatorio que modela la incertidumbre del proceso.

Luego, si existe un vector de observaciones $z \in \mathbb{R}^m$ que depende linealmente del vector de estado, entonces éste se puede utilizar para estimar el valor de x . El vector de observaciones z y el vector de estado x se relacionan por medio de la llamada ecuación de observación:

$$z_k = Hx_k + v_k, \quad (3.16)$$

en donde:

- H es una matriz de $m \times n$ constante, que relaciona el vector x con la medida z .
- v es la variable aleatoria que representa la incertidumbre del proceso de observación.

Para poder aplicar el filtro de Kalman en un modelo lineal, primero se deben especificar las covarianzas de los ruidos de proceso w y de observación v . Sin embargo, no debe existir correlación entre ellos. Las matrices de covarianza Q y R de los ruidos de proceso y de observación, respectivamente. Si la operación $E[\cdot]$ corresponde a la esperanza de la variable correspondiente, entonces las matrices de covarianza son:

$$Q = E[w_{k-1}w_{k-1}^T] \quad (3.17)$$

$$R = E[v_kv_k^T] \quad (3.18)$$

La deducción de las ecuaciones que rigen el filtro Kalman-Bucy se derivan de los siguientes resultados fundamentales.

Teorema 1. - Teorema Fundamental de la Teoría de Estimación [25]. *Dados los parámetros aleatorios θ y z , tal que*

$$z(k) = H(k)\theta + v(k)$$

Y dadas las observaciones $z(1), z(2), \dots, z(k)$, el estimador del parámetro θ que minimiza el error cuadrático medio es:

$$\hat{\theta}_{MS} = E[\theta|Z(k)] \quad (3.19)$$

Demostración. Sea $\hat{\theta}_{MS}$ un estimador para θ . Luego, dado $Z = Z(k)$ tenemos que el error cuadrático medio (*MSE - Mean Square Error*) es:

$$\begin{aligned} MSE &= E[(\theta - \hat{\theta}_{MS})^T(\theta - \hat{\theta}_{MS})|Z] \\ &= E[\theta^T\theta - \theta^T\hat{\theta}_{MS} - \hat{\theta}_{MS}^T\theta + \hat{\theta}_{MS}^T\hat{\theta}_{MS}|Z] \\ &= E[\theta^T\theta|Z] - E[\theta^T|Z]\hat{\theta}_{MS} - \hat{\theta}_{MS}^TE[\theta|Z] + \hat{\theta}_{MS}^T\hat{\theta}_{MS} \\ &= E[\theta^T\theta|Z] + (\hat{\theta}_{MS} - E[\theta|Z])^T(\hat{\theta}_{MS} - E[\theta|Z]) + E[\theta^T|Z]E[\theta|Z] \end{aligned}$$

Luego, el error cuadrático es mínimo si:

$$\hat{\theta}_{MS} = E[\theta|Z(k)]$$

y el error cuadrático medio mínimo (*MMSE - Minimum Mean Square Error*) es:

$$MMSE = E[\theta^T\theta|Z] + E[\theta^T|Z]E[\theta|Z].$$

□

Teorema 2. Principio de Ortogonalidad [25]. *Para cualquier función medible $f[Z(k)]$ de los datos $Z(k)$, el error en el estimador de mínimo error cuadrático medio es ortogonal a $f[Z(k)]$ en el sentido que*

$$E[(\theta - \hat{\theta}_{MS})^T f[Z(k)]] = 0.$$

Demostración. De la teoría de probabilidades se tiene que para dos variables aleatorias α y β se cumple:

$$E[\alpha g(\beta)] = E[E[\alpha|\beta]g(\beta)]$$

En donde el operador de esperanza exterior de la mano derecha es respecto a β . Luego, en base a este resultado se tiene.

$$\begin{aligned} E[(\theta - \hat{\theta}_{MS})^T f[Z(k)]] &= E[E[(\theta - \hat{\theta}_{MS})^T|Z]f[Z(k)]] \\ &= E[(\hat{\theta}_{MS} - \hat{\theta}_{MS})^T f[Z(k)]] \\ &= 0. \end{aligned}$$

Con lo cual queda demostrado el principio de ortogonalidad. □

Luego, para el proceso lineal descrito y usando los resultados anteriores y propiedades asociadas a la densidad Gaussiana multivariable, tenemos que:

$$\hat{x}(k+1|k+1) = E[x(k+1)|Z(k+1)] = E[x(k+1)|Z(k), z(k+1)] = E[x(k+1)|Z(k), \tilde{z}]$$

Con $\tilde{z} = z(k+1) - E[z(k+1)|Z(k)]$

Desarrollando la expresión, obtenemos:

$$\hat{x}(k+1|k+1) = E[x(k+1)|Z(k)] + E[x(k+1)|\tilde{z}(k+1|k)] - m_x(k+1)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + E[x(k+1)|\tilde{z}(k+1|k)] - m_x(k+1) \quad (3.20)$$

Como resultado preliminar, se busca desarrollar la expresión anterior para obtener una relación de la forma:

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)\tilde{z}(k+1|k) \quad (3.21)$$

De acuerdo a la relación 3.21, se obtiene que la matriz de covarianza del error de estimación es:

$$\begin{aligned} E[(x(k+1) - \hat{x}(k+1|k+1))(x(k+1) - \hat{x}(k+1|k+1))^T] &= P(k+1|k+1). \\ P(k+1|k+1) &= [I - K(k+1)H(k+1)]P(k+1|k)[I - K(k+1)H(k+1)]^T + K(k+1)R(k+1)K^T(k+1). \end{aligned} \quad (3.22)$$

Por lo cual, desarrollando el segundo término de la derecha de 3.20 se obtiene:

$$E[x(k+1)|\tilde{z}(k+1|k)] = m_x(k+1) + P_{x\tilde{z}}(k+1, k+1|k)P_{\tilde{z}\tilde{z}}^{-1}(k+1|k)\tilde{z}(k+1|k).$$

Reemplazando en 3.20, se obtiene:

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + P_{x\tilde{z}}(k+1, k+1|k)P_{\tilde{z}\tilde{z}}^{-1}(k+1|k)\tilde{z}(k+1|k). \quad (3.23)$$

Por lo que se define

$$K(k+1) = P_{x\tilde{z}}(k+1, k+1|k)P_{\tilde{z}\tilde{z}}^{-1}(k+1|k). \quad (3.24)$$

Luego, se calcula cada una de estas matrices de covarianza.

$$\begin{aligned} P_{x\tilde{z}}(k+1, k+1|k) &= E[(x(k+1) - m_x(k+1))\tilde{z}^T(k+1|k)] \\ &= E[x(k+1)\tilde{z}^T(k+1|k)]. \end{aligned}$$

Porque \tilde{z} es de media cero. Sustituyendo $\tilde{z}(k+1|k)$ se obtiene:

$$\begin{aligned} P_{x\tilde{z}}(k+1, k+1|k) &= E[x(k+1)\tilde{x}^T(k+1|k)]H^T(k+1) \\ &= E[(\tilde{x}(k+1|k) + \hat{x}(k+1|k))\tilde{x}^T(k+1|k)]H^T(k+1) \\ &= P(k+1|k)H^T(k+1). \end{aligned}$$

Además, se tiene que:

$$\begin{aligned}
P_{\tilde{z}\tilde{z}}(k+1|k) &= E[\tilde{z}(k+1)\tilde{z}^T(k+1)|k] \\
&= E[(H(k+1)\tilde{x}(k+1|k) + v(k))(H(k+1)\tilde{x}(k+1|k) + v(k))^T|k] \\
&= H(k+1)P(k+1|k)H^T(k+1) + R(k+1).
\end{aligned}$$

Con lo cual sustituyendo estos resultados en 3.24 se obtiene la matriz de ganancia de Kalman:

$$K(k+1) = P(k+1|k)H^T(k+1)[H(k+1)P(k+1|k)H^T(k+1) + R(k+1)]^{-1} \quad (3.25)$$

Al sustituir esta matriz en 3.22, se obtiene

$$\begin{aligned}
P(k+1|k+1) &= (\mathbf{I} - \mathbf{KH})\mathbf{P}(\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T \\
&= (\mathbf{I} - \mathbf{KH})\mathbf{P} - \mathbf{PH}^T\mathbf{K}^T + \mathbf{KHPH}^T\mathbf{K}^T + \mathbf{KRK}^T \\
&= (\mathbf{I} - \mathbf{KH})\mathbf{P} - \mathbf{PH}^T\mathbf{K}^T + \mathbf{K}(\mathbf{HPH} + \mathbf{R})\mathbf{K}^T \\
&= (\mathbf{I} - \mathbf{KH})\mathbf{P} - \mathbf{PH}^T\mathbf{K}^T + \mathbf{PH}^T\mathbf{K}^T \\
&= (\mathbf{I} - \mathbf{KH})\mathbf{P}.
\end{aligned}$$

El filtro de Kalman estima el estado del proceso en cada paso de la iteración, a través de retroalimentación, por medio de un algoritmo Predictor-Corrector. Las ecuaciones de predicción permiten proyectar el estado anterior y el error de covarianza del proceso, para obtener una estimación a priori x^- del vector de estado x del paso actual. A continuación se indican las ecuaciones de predicción.

$$x_k^- = Ax_{k-1} + Bu_{k-1} \quad (3.26)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (3.27)$$

donde P es una matriz que indica la estimación del error de covarianza en las componentes del vector de estado. Las ecuaciones de corrección actualizan las estimaciones a priori, para generar estimaciones a posteriori. Para corregir se crea la ganancia de Kalman K , que tiene que ver con la ponderación de la nueva medición z .

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (3.28)$$

$$x_k = x_k^- + K_k(z_k - Hx_k^-) \quad (3.29)$$

$$P_k = (I - K_kH)P_k^- \quad (3.30)$$

El primer paso de la actualización es calcular la ganancia de Kalman en la ecuación 3.28. La actualización del vector de estado x en la ecuación 3.29, se realiza con el valor real de la medición z del proceso, la que contiene la señal alterada. El último paso es obtener el valor actualizado del error de la covarianza P en la ecuación 3.30. Una vez terminadas la predicción y la corrección del vector de estado, se continúa con la siguiente iteración, hasta que el algoritmo converge al resultado.

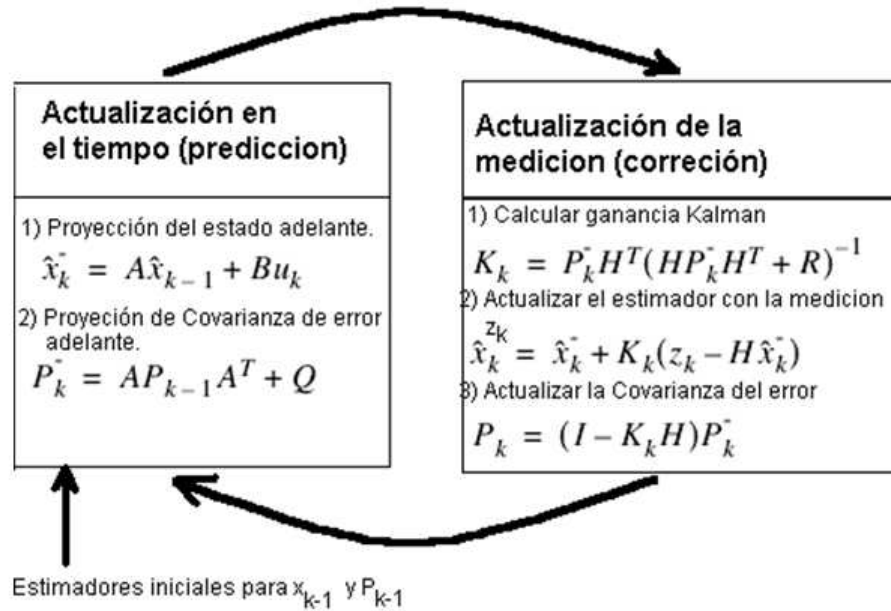


Figura 3.2: Algoritmo de Filtrado de Kalman

Dado que en general los sistemas no cumplen con las hipótesis de linealidad del sistema, existe una variación de este filtro denominado Filtro Extendido de Kalman (EKF) el cual consiste en la aplicación del filtro de Kalman al proceso no lineal mediante una expansión de Taylor entorno al punto de operación actual del sistema conservando la restricción de gaussianidad de las variables de incertidumbre [25]. En este caso, se tiene que las ecuaciones de estado y observación son:

$$x_k = f(x_{k-1}, u_{k-1}) + w_{k-1} \quad (3.31)$$

$$z_k = h(x_k) + v_k. \quad (3.32)$$

Condiciones con las cuales el filtro extendido de Kalman no garantiza optimalidad en el sentido de minimización del error cuadrático medio. Para este caso se tiene que las ecuaciones de predicción y corrección son:

Predicción

$$x_k^- = f(x_{k-1}, u_{k-1}) \quad (3.33)$$

$$P_k^- = F_{k-1} P_{k-1} F_{k-1}^T + Q \quad (3.34)$$

Corrección

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R)^{-1} \quad (3.35)$$

$$x_k = x_k^- + K_k (z_k - h(x_k^-)) \quad (3.36)$$

$$P_k = (I - K_k H_k) P_k^- \quad (3.37)$$

En donde

$$F_k = \frac{\partial f}{\partial x} \Big|_{x_{k-1}, u_{k-1}} \quad (3.38)$$

$$H_k = \frac{\partial h}{\partial x} \Big|_{x_k}. \quad (3.39)$$

Son las matrices Jacobianas de las funciones de los procesos de predicción y observación, evaluadas en los puntos que corresponde.

De esta manera, teniendo un modelo para la evolución de la velocidad del viento (ecuación de estado) y un modelo observacional (ecuación de observación) tomando como variable de observación la potencia generada y las velocidades de viento anteriores es posible aplicar esta técnica de filtrado estadístico y su posterior uso en predicción, siempre y cuando se conserven las restricciones impuestas sobre las incertidumbres de los modelos. Cabe mencionar que esta técnica de estimación posee la ventaja de entregar el valor esperado para la potencia generada en el paso $n+k$ y también entrega la matriz de covarianza asociada a la predicción.

El filtro de Kalman y algunas de sus variantes han sido exitosamente utilizadas en el ámbito de predicción de potencia generada en parques eólicos en [26, 27], así como también en predicción de series caóticas en [28, 29]. Además, otro tipo de filtrado también usado en predicción de potencia generada en parques eólicos es el filtrado adaptativo utilizado en [30].

3.4. Métodos no Paramétricos

3.4.1. Redes Neuronales Artificiales

El concepto de red neuronal artificial es derivado originalmente de las células biológicas nerviosas, llamadas neuronas, las cuales están interconectadas en un alto grado y son capaces de generar una respuesta específica dependiendo del nivel de actividad de sus vecinas. Las redes neuronales artificiales son una amplia clase de modelos matemáticos los cuales simulan parcialmente el comportamiento de estos sistemas biológicos.

El nivel de actividad de una neurona está determinado por la integración de las señales de entrada provenientes de otras neuronas, o bien, por ser las entradas al sistema. Por lo tanto, en términos matemáticos su actividad es descrita por una sumatoria con pesos relativos de las señales de entrada de la neurona.

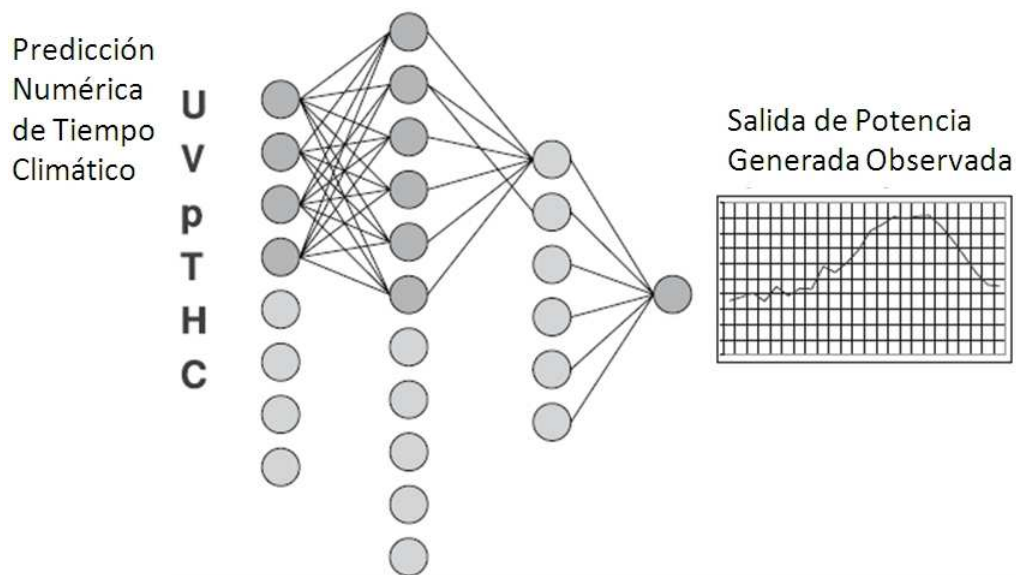


Figura 3.3: Estructura de red neuronal artificial.

Esta técnica ha sido ampliamente utilizada en predicción de potencia generada en parques eólicos, tanto sola en [31–34] como en conjunto con otras técnicas tales como lógica difusa en [35–37], descomposición wavelet [20], algoritmos genéticos en [38] y teoría de información en [39, 40]. Mayor información al respecto se puede encontrar en [6].

3.4.2. Modelo Gris y otros métodos

La teoría de sistemas grises es un área científica interdisciplinaria introducida por primera vez por el profesor J. Deng en 1982. Desde entonces, la teoría se ha vuelto popular por su habilidad de trabajar en sistemas que tienen parámetros parcialmente conocidos. Poseen como ventaja sobre los modelos estadísticos, el necesitar una limitada cantidad de datos para estimar el comportamiento de sistemas desconocidos [41]. Este tipo de técnica ha sido utilizada exitosamente en predicción de potencia generada en parques eólicos en [42] [43] y series de tiempo en [44].

En el mismo ámbito, es posible encontrar otro tipo de técnicas tales como *Support Vector Machine* en [45], Combinación de pronósticos en [46], mínimos cuadrados en [47], filtros de respuesta finita e infinita al impulso en [48] y modelo autorregresivo condicional heteroscedasticidad generalizado en [49].

3.5. Modelos Numéricos de Predicción de Tiempo Climático

Los modelos numéricos de predicción de tiempo climático (*NWP Numerical Weather Prediction*) simulan la dinámica atmosférica por medio de integración numérica de las ecuaciones no lineales que rigen el movimiento del viento, partiendo del estado atmosférico actual. Para este propósito, se discretiza el espacio en una malla tridimensional de puntos.

Debido a las limitaciones de capacidad computacional, la resolución de la malla numérica es finita, por lo que los sistemas NWP no pueden simular directamente procesos en una escala más pequeñas (bajo los 10[km]). Esto se mantiene en particular para la influencia de estructuras orográficas del terreno y de procesos térmicos tales como la brisa marina.

Los procesos asociados a esta sub-malla tienen influencia en la evolución de gran escala, como por ejemplo la turbulencia o la formación de nubes, por lo que este fenómeno tiene que ser parametrizado, lo cual significa que su efecto macroscópico es modelado sin considerar los detalles microscópicos. Por lo tanto, los modelos NWP sirven para describir el compor-

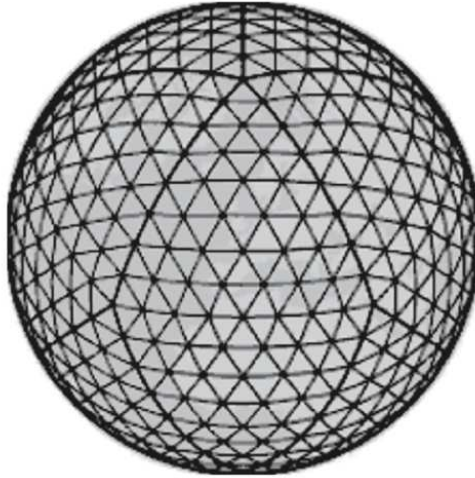


Figura 3.4: Modelo NWP del servicio metereológico alemán con resolución espacial de $60 \times 60 [km^2]$

tamiento atmosférico de gran escala tales como mapas sinópticos de extensiones cuyo orden van desde 10 [km] hasta 1000 [km]. Adicionalmente, la resolución finita conllevan a que las variables calculadas en cada puntos representan promedios sobre la correspondiente celda de la grilla. Por lo tanto, el valor predicho puede resultar que no sea la mejor estimación para todos los lugares al interior de la celda.

Debido a la complejidad del fenómeno y a la gran cantidad de datos que se requieren para simular estos modelos, estos son operados típicamente por servicios metereológicos nacionales. En general, ellos usan un modelo global con una resolución desde $100 \times 100 [km^2]$ hasta $50 \times 50 [km^2]$ para capturar el desarrollo global de los sistemas climáticos. Estos modelos globales son usados para imponer condiciones de borde sobre modelos locales de predicción, los cuales conllevan una mayor resolución espacial, del orden de $50 \times 50 [km^2]$ hasta $7 \times 7 [km^2]$, pero que cubren una superficie menor.

Fijar el estado inicial para realizar una predicción de tiempo climatológico requiere de una gran cantidad de datos. El estado atmosférico es medido con una base regular de una gran cantidad de bases sinópticas, radio sondas, satélites y aviones distribuidos por el mundo. De esta manera, el formato de los datos y el ciclo de medida son estandarizados por la Organización Meterológica Mundial con lo que la información metereológica se encuentra

disponible para que los servicios meteorológicos nacionales fijen las condiciones iniciales a sus modelos NWP.

Este tipo de modelos se ha sido utilizado exitosamente en predicción de mediano plazo de potencia generada en parques eólicos en [26] [50]. Mayor información al respecto se puede encontrar en [6].

Capítulo 4

Metodología de Trabajo

En el presente capítulo se detalla la metodología con la cual se aborda la problemática de predicción de corto plazo de potencia generada en parques eólicos, tomando como base la técnica de modelos SARIMA.

Las metodologías a implementar corresponden a un modelo autorregresivo integral de media móvil y el modelo de persistencia, que será utilizado como patrón de comparación para evaluar el desempeño del modelo SARIMA

En primer lugar se trabaja con series de tiempo de viento obtenidas desde el proyecto EOLO [11], las cuales corresponden a mediciones de velocidad de viento que se realizaron durante los años 1990 y 1991. Dicho proyecto contó con mediciones de velocidad de viento en distintas localidades en donde se tuvieron distintos periodos de muestreo, por lo que se seleccionaron 4 localidades cuya tasa de muestreo era cada 1[h].

A partir de estos datos de velocidad de viento, se obtienen datos de potencia generada en un aerogenerador usando la curva de potencia del aerogenerador NM82 fabricado por NEG Micon que es utilizado en el parque eólico Canela II.

Las mediciones de velocidad de viento en general se encontraban incompletas en el sentido que el registro de datos no era correlativo en la mayoría de las localidades estudiadas en el informe. Por ello, se trabajó con las secuencias de datos consecutivas máximas posibles en cada localidad. Las coordenadas de estas localidades, el número de datos así como también los metros sobre el nivel del mar al cual se encuentran se muestran a continuación:

Localidad	Ubicación	Número de datos	Metros sobre el nivel del mar
Punta Lengua de Vaca	30°14'26"S,71°37'46" O	8760	10
Summit	29°27'00"S,71°14'42" O	5088	80
Los Vilos	31°50'48"S,71°30'48" O	1464	60
Paranal	24°37'24"S,70°24'05" O	2524	2665
Cruz Grande	29°27'S,71°19Ó	5808	10

Cuadro 4.1: Características de las localidades seleccionadas.

Las técnicas se implementan usando solamente los datos de la localidad de Punta Lengua de Vaca, ya que se cuenta con mediciones de velocidad de viento durante un año completo sin interrupciones. Los datos de las otras localidades son usados para el análisis del desempeño de cada técnica de predicción.

A partir de los datos originales, se construyen subconjuntos de datos correspondientes a las mediciones de una semana completa y por medio de un pre-procesamiento, se obtienen las siguientes series de tiempo:

- Velocidad de viento sin pre-procesamiento.
- Diferencias consecutivas de velocidad de viento.
- Suma acumulada de primer orden de velocidad de viento.
- Resta entre velocidad de viento y su media respectiva.
- Resta entre diferencias consecutivas de velocidad de viento y su media respectiva.
- Resta entre suma acumulada de primer orden de velocidad de viento y su media respectiva.

De este modo, se aplican los modelos a estas nuevas series de tiempo para luego analizar el desempeño de las técnicas de predicción. Además, cada serie de tiempo se divide de manera que los datos de los primeros cinco días son usados para entrenamiento y los datos dos días restantes son utilizados para validación de cada modelo.

4.1. Análisis estadístico y de frecuencia

El primer análisis que se realiza a cada una de las series de tiempo corresponde a un análisis estadístico en donde se calcula la función de autocorrelación empírica de cada serie de tiempo, para determinar si existe dependencia relevante entre algunos de los términos que componen dicha serie.

Luego se le calcula la transformada a tiempo discreto de Fourier a cada una de las series de tiempo que se tienen, para determinar si existen componentes periódicas presentes en el comportamiento de las series de tiempo, de manera de sacar provecho de este comportamiento para resolver el problema de predicción, dejándolo claramente expresado en el modelo de predicción obtenido.

4.2. Desarrollo de un modelo SARIMA para predicción

Utilizando el software MATLAB se implementan funciones para realizar la identificación del sistema, usando como criterio de selección de parámetros, la minimización del error cuadrático medio en la base de datos de entrenamiento. Luego, dado que en los modelos ARMA se debe definir el grado de auto regresión, el óptimo se halla ajustando modelos para $p = 1, \dots, 27$ dado que en general la memoria de un proceso no llega más allá de eso [21] y hallando cual fue la combinación de parámetros p, q que obtuvo el menor error. De todos los modelos obtenidos, se selecciona el que obtiene un mejor compromiso entre el error cuadrático medio en la base de datos de entrenamiento y el error cuadrático medio en la base de validación, ya que de esta manera se evita el sobreajuste del modelo a los datos de entrenamiento.

Un diagrama de bloques del funcionamiento de las funciones programadas se muestra en la figura 4.1.

En un comienzo a la serie de tiempo se le resta su media y se inicializan los ordenes del modelo SARIMA $(p, q, d) \times (0, 0, D)_s$. A continuación se realizan las diferenciaciones

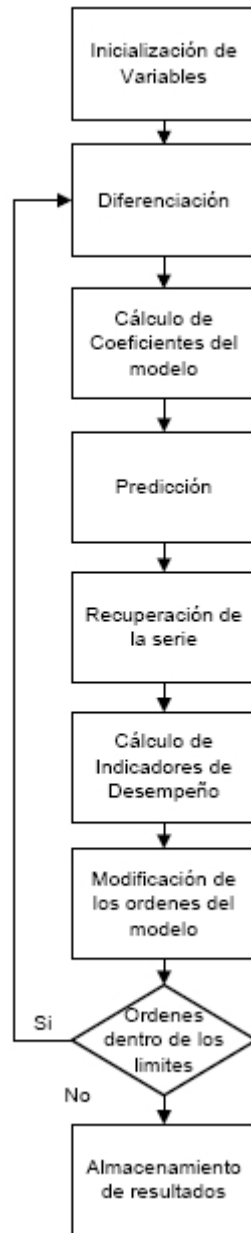


Figura 4.1: Diagrama de Bloques de las funciones programadas

respectivas:

$$\begin{aligned}
 (1 - B^s)^D(u_t - \bar{u}) &= v_t \\
 (1 - B)^d v_t &= x_t
 \end{aligned}
 \tag{4.1}$$

Realizada la diferenciación, se ha eliminado la parte periódica y la parte no-estacional por lo que solo queda estimar los parámetros del modelo ARMA sobre x_t . Los p -parámetros de la parte autorregresiva se estiman de acuerdo a las ecuaciones de Yule-Walker en base a las autocovarianzas [21]:

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & \dots \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \dots \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \vdots \end{pmatrix} \quad (4.2)$$

Una vez estimados los parámetros de la parte autorregresiva, se estiman las primeras $q+1$ autocovarianzas $c'_j (j = 0, \dots, q)$ de la nueva serie:

$$x'_t = x_t - \sum_{i=1}^p \varphi_i x_{t-i} \quad (4.3)$$

Con estas covarianzas, de acuerdo al modelo planteado se deduce que [21]:

$$\begin{aligned} \sigma_a^2 &= \frac{c'_0}{1 + \theta_1^2 + \dots + \theta_q^2} \\ \theta_j &= -\left(\frac{c'_j}{\sigma_a^2} - \theta_1\theta_{j+1} - \theta_2\theta_{j+2} - \dots - \theta_{q-j}\theta_q\right) \end{aligned} \quad (4.4)$$

Este último sistemas de ecuaciones no lineales se resuelve numéricamente utilizando el método simplex de Nelder-Mead.

Finalmente, una vez que se obtienen los parámetros del modelo ARMA, se procede a realizar la estimación de la serie de tiempo. Para esto último se necesitan los primeros p -valores de la serie de tiempo como condición inicial. Luego, el estimador que minimiza el error cuadrático medio es:

$$\begin{aligned} \hat{x}_t &= E[x_t | (x_{t-1}, \dots, x_1)] \\ &= \sum_{i=1}^p \varphi_i x_{t-i} + \sum_{j=1}^q \theta_j e_{t-j}, \end{aligned} \quad (4.5)$$

en donde e_{t-j} es el error de estimación cometido j instantes anteriores.

Después de realizar la estimación sobre la serie de tiempo de validación y de entrenamiento, los valores obtenidos corresponden a una estimación de las diferencias de la serie original, por lo que se debe volver a reconstruir la serie original por medio de las siguientes relaciones:

$$\begin{aligned}
(1 - B)^d v_t &= x_t \\
\left(\sum_{k=0}^d \binom{d}{k} 1^{n-k} (-B)^k \right) v_t &= x_t \\
\left(1 + \sum_{k=1}^d (-B)^k \right) v_t &= x_t \\
v_t + \sum_{k=1}^d (-1)^k v_{t-k} &= x_t \\
v_t &= x_t - \sum_{k=1}^d (-1)^k v_{t-k} \tag{4.6}
\end{aligned}$$

$$\begin{aligned}
(1 - B^s)^D (u_t - \bar{u}) &= v_t \\
\left(\sum_{k=0}^D \binom{D}{k} 1^{n-k} (-B^s)^k \right) (u_t - \bar{u}) &= v_t \\
\left(1 + \sum_{k=1}^D \binom{D}{k} (-B^s)^k \right) (u_t - \bar{u}) &= v_t \\
u_t - \bar{u} + \sum_{k=1}^D \binom{D}{k} (-1)^k u_{t-sk} &= v_t \\
u_t &= v_t - \sum_{k=1}^D \binom{D}{k} (-1)^k u_{t-sk} + \bar{u} \tag{4.7}
\end{aligned}$$

Con esto último, se tiene finalmente una estimación de la serie original de datos con lo cual se hace posible realizar el cálculo de indicadores de desempeño. Estos resultados se almacenan en el disco duro del PC utilizado para la realización de experimentos.

4.3. Indicadores de desempeño

Para analizar cuantitativamente el desempeño de los modelos propuestos en el contexto de la predicción de velocidad de viento, se define el siguiente conjunto de indicadores.

4.3.1. Error Cuadrático Medio

Para una predicción a N -pasos, el error cuadrático medio de dicha predicción se encuentra dado por:

$$e_{MSE} = \frac{1}{N} \sum_{i=1}^N (u_i - \hat{u}_i)^2. \quad (4.8)$$

4.3.2. Error Cuadrático Medio con pesos

Para una predicción a N -pasos, el error cuadrático medio con pesos de dicha predicción se encuentra dado por:

$$e_{WMSE} = \frac{1}{N} \sum_{i=1}^N \omega(u_i - \hat{u}_i) \cdot [u_i - \hat{u}_i]^2. \quad (4.9)$$

en donde

$$\omega(x) = \begin{cases} 0.8 & x \geq 0 \\ 1 & x < 0. \end{cases} \quad (4.10)$$

4.3.3. Error Absoluto Medio

Para una predicción a N -pasos, el error absoluto medio de dicha predicción se encuentra dado por:

$$e_{MAE} = \frac{1}{N} \sum_{i=1}^N |u_i - \hat{u}_i|. \quad (4.11)$$

4.3.4. Error Absoluto Porcentual Medio

Para una predicción a N -pasos, el error absoluto porcentual medio de dicha predicción se encuentra dado por:

$$e_{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|u_i - \hat{u}_i|}{u_i}. \quad (4.12)$$

4.3.5. Error Absoluto Porcentual Efectivo

Para una predicción a N -pasos, el error absoluto porcentual efectivo de dicha predicción se encuentra dado por:

$$e_{RMSEP} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (u_i - \hat{u}_i)^2}}{\frac{1}{N} \sum_{i=1}^N u_i}. \quad (4.13)$$

4.3.6. Criterio de Información Bayesiano

El criterio de información bayesiano (*BIC* por *Bayesian Information Criterion*) para una predicción a N -pasos, se encuentra dado por:

$$\text{BIC} = -2\ln(L) + k\ln(n). \quad (4.14)$$

en donde:

1. L = Verosimilitud
2. k = Número de parámetros libres
3. n = Número de datos de la muestra

En especial, para el caso del modelo ARMA cuyos residuos poseen una varianza estimada σ_e^2 , se obtiene:

$$\text{BIC} = N\ln(\hat{\sigma}_e^2) + (p + q)\ln(N). \quad (4.15)$$

4.3.7. Criterio de Información de Akaike

El criterio de información de Akaike (*AIC* por *Akaike's Information Criterion*) para una predicción a N -pasos, se encuentra dado por:

$$AIC = 2k - 2\ln(n). \quad (4.16)$$

en donde:

1. L = Verosimilitud
2. k = Número de parámetros libres
3. n = Número de datos de la muestra

En especial, para el caso del modelo ARMA cuyos residuos poseen una varianza estimada σ_e^2 , se obtiene:

$$AIC = 2k + N\ln(\hat{\sigma}_e^2). \quad (4.17)$$

Capítulo 5

Resultados y Discusión

En este capítulo se presentan, comparan, y evalúan, los resultados de las técnicas propuestas para predecir la potencia generada en un aerogenerador. Las técnicas de predicción implementadas corresponden a un modelo SARIMA y al modelo de persistencia. En base a los modelos anteriores se realiza predicción a un paso y a cinco pasos utilizando datos de series de velocidad de viento reales, con el propósito de analizar el desempeño de cada modelo bajo distintos horizontes de tiempo. La evaluación de los resultados obtenidos es realizada por medio de un conjunto de indicadores propuestos.

La plataforma para todas las simulaciones y estimaciones fue MATLAB 7.10 (R2010a) corriendo sobre Mac OS X 10.6.5 en un Intel Core 2 Duo de 2.16GHz (32bits). El cálculo de los parámetros de los modelos se realizó mediante algoritmos programados en MATLAB que pueden ser encontrados en el Apéndice.

El capítulo está organizado como sigue. En la sección 5.1 se grafican series de tiempo de velocidad de viento con el objetivo de inspeccionar el comportamiento temporal de la serie de tiempo. Luego, en la sección 5.2 se grafica la función de autocorrelación de las series de tiempo de modo de identificar el grado de dependencia lineal existente entre los términos de la serie de tiempo. A continuación, en la sección 5.3, se grafica el espectro de frecuencias de los datos para identificar alguna componente en frecuencia que este influenciando el comportamiento temporal de la serie de tiempo. Una vez realizados los análisis anteriores, en la sección 5.4 se muestran los resultados de predicción obtenidos mediante un modelo SARIMA y en la sección

5.5 se muestran los resultados obtenidos mediante el modelo de persistencia. Finalmente en la sección 5.6 se realiza una comparación de ambos métodos.

5.1. Series de tiempo

En la figura 5.1 se grafican series de tiempo de Punta Lengua de Vaca, cada una correspondiente a los datos de velocidad de viento promedio por hora durante 7 días.

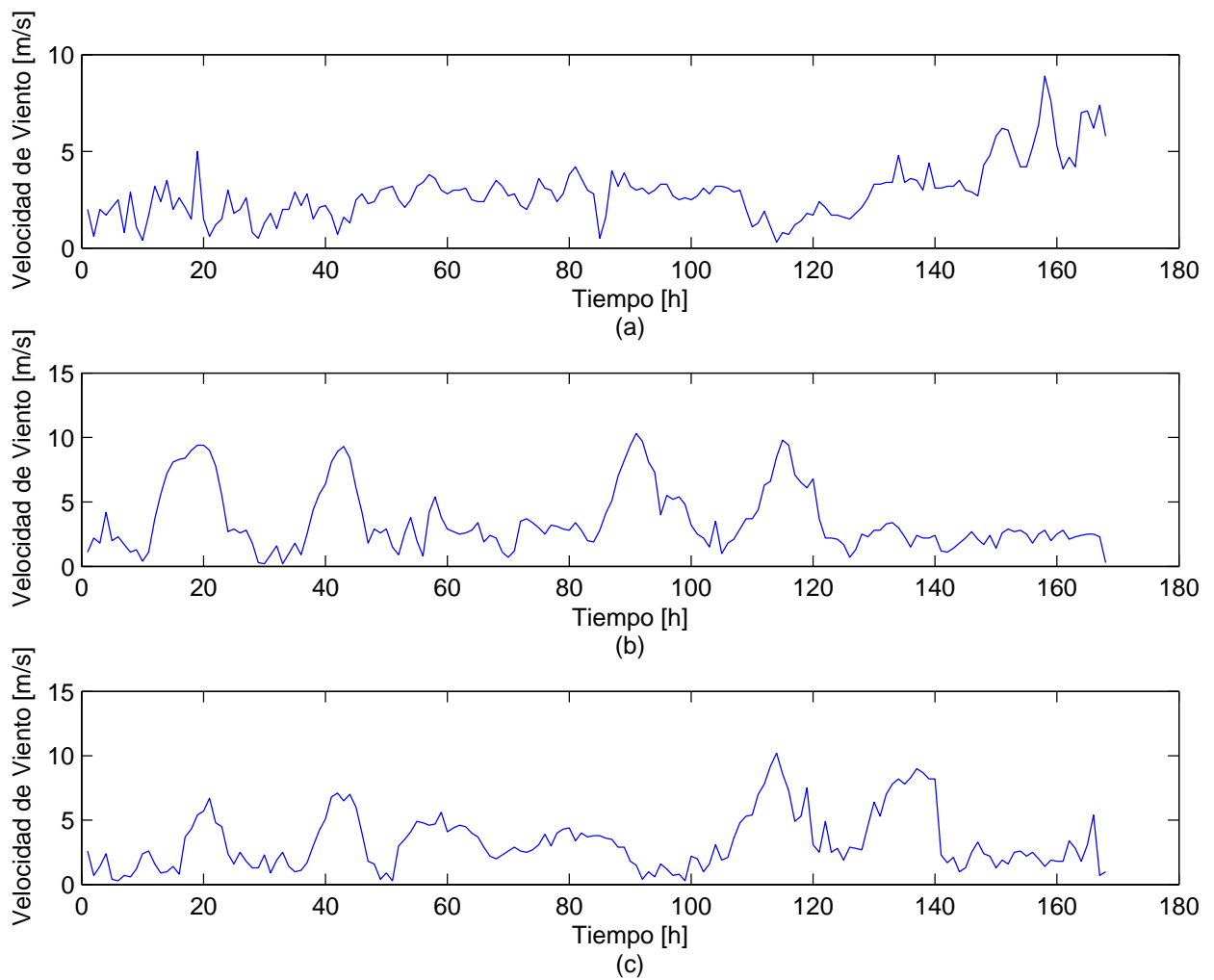
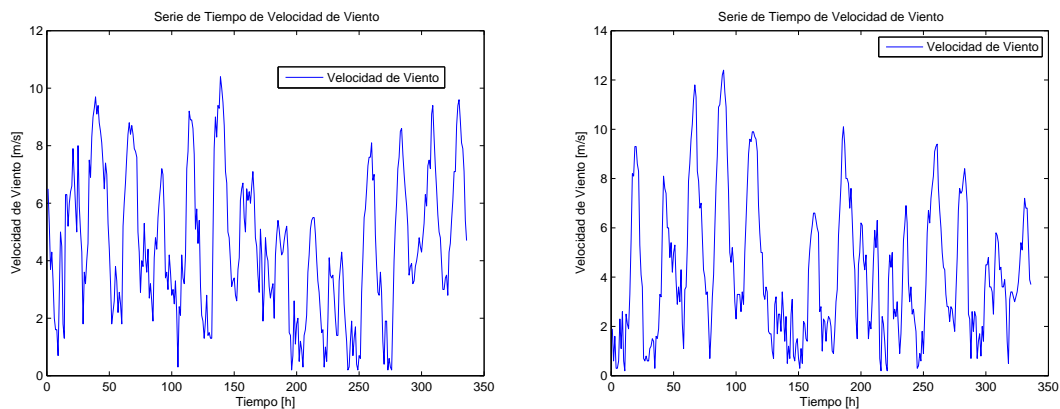


Figura 5.1: Series de tiempo de Punta Lengua de Vaca: (a) Mediciones tomadas en la semana 24 de 1990. (b) Mediciones tomadas en la semana 12 de 1990. (c) Mediciones tomadas en la semana 14 de 1990.

En la figura 5.1 se puede observar claramente que existe un comportamiento aproxima-

damente periódico, dado que cada $24[h]$ la velocidad del viento toma valores similares. Esto último se explica por la naturaleza del fenómeno que se estudia, ya que como la velocidad del viento se encuentra influenciada por los niveles de temperatura y consecuentemente por el nivel de radiación solar, variables que poseen un comportamiento periódico y que causan el comportamiento periódico de la velocidad del viento.

Además, se decide trabajar con series de tiempo cuyo largo corresponde a los datos de siete días puesto que, en general, para series de tiempo de velocidad de viento de largos mayores, se tiene que el comportamiento del viento varía considerablemente, por lo que se deja de tener un comportamiento estacionario. Ejemplos de esto último se muestran en las figuras 5.2a y 5.2b.



(a) Serie de tiempo de Punta Lengua de Vaca con datos tomados en la semana 25 y 26 de 1990 (b) Serie de tiempo de Punta Lengua de Vaca con datos tomados en la semana 28 y 29 de 1990

Figura 5.2: Series de tiempo en Punta Lengua de Vaca con mayor número de muestras

5.2. Función de Autocorrelación Empírica

Luego de analizar el comportamiento temporal de la serie de tiempo y de elegir un largo adecuado para la serie, se realiza un análisis estadístico de la serie de tiempo, graficando la función de autocorrelación empírica de la serie para así identificar dependencia lineal entre los términos de la serie.

La función de autocorrelación empírica para tres series de tiempo se muestran en la figura 5.3:

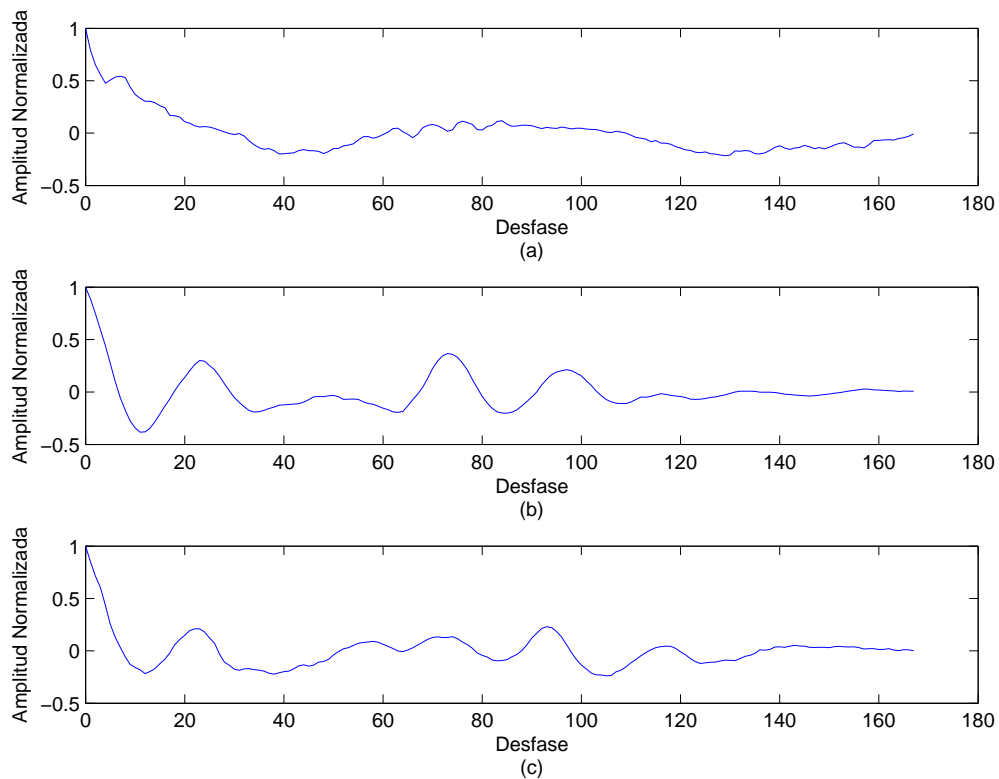


Figura 5.3: Función de Autocorrelación empírica para series de tiempo de Punta Lengua de Vaca para datos tomados en: (a) Semana 24 de 1990. (b) Semana 12 de 1990. (c) Semana 14 de 1990.

Observando la figura 5.3, se puede detectar un alto grado de dependencia lineal entre los términos consecutivos de la serie, dependencia que va decayendo a medida que aumenta el desfase entre los términos pero que vuelve a incrementarse cada 24 términos. Esto último sugiere una periodicidad en el comportamiento de la serie de tiempo, por lo que se justifica la utilización de un modelo SARIMA cuyo período es $d = 24$.

5.3. Espectro de Frecuencias

Una vez realizado el análisis estadístico, se procede a graficar el espectro de frecuencias de los datos, para así identificar componentes periódicas que existan en la serie. Los resultados obtenidos se muestran en la figura 5.4:

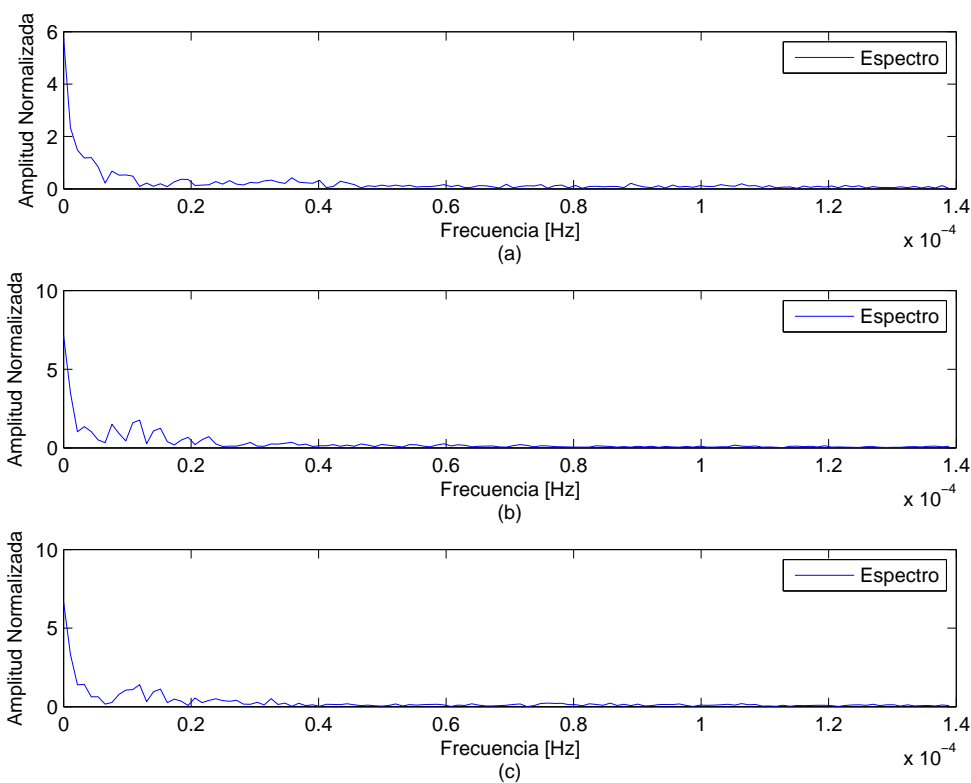


Figura 5.4: Espectro de frecuencias para series de tiempo de Punta Lengua de Vaca para datos tomados en: (a) Semana 24 de 1990. (b) Semana 12 de 1990. (c) Semana 14 de 1990.

En la figura 5.4, se puede notar claramente que existe una componente de frecuencia que sobresale por sobre las demás. Haciendo un análisis visual, se obtiene que dicha frecuencia se encuentra asociada a un período de 24[h], por lo cual se tiene otra evidencia de que el comportamiento de la serie de tiempo es periódico y que el período asociado a dicho comportamiento es igual a 24[h].

5.4. Resultados de modelo SARIMA

Después de confirmar la existencia de una componente periódica en el comportamiento de la serie de tiempo, se implementa el modelo SARIMA de acuerdo a lo explicado en el capítulo anterior y se analizan los resultados.

En general, el mínimo error cuadrático medio se encontró para:

$$q = 1 \quad d = 0 \quad D = 1.$$

Luego, el orden de autorregresión se determina observando las curvas de error cuadrático medio sobre la base de datos de entrenamiento y de validación en función del orden del autorregresión, curvas que se muestra en la figura 5.5.

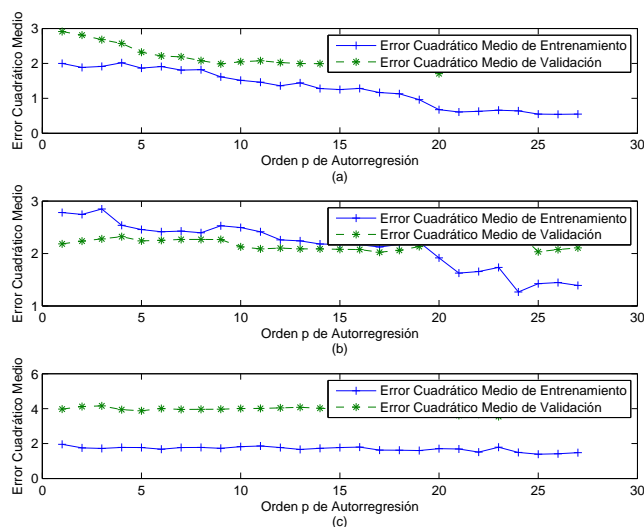


Figura 5.5: Error cuadrático medio de modelo SARIMA de ordenes $q = 1$ $d = 0$ $D = 1$ sobre series de tiempo en Punta Lengua de Vaca con datos tomados en: (a) Semana 24 de 1990. (b) Semana 12 de 1990. (c) Semana 14 de 1990.

Observando las curvas mostradas en la figura 5.5, para cada serie de tiempo se decide utilizar:

$$p_1 = 20 \quad p_2 = 25 \quad p_3 = 23.$$

En las figuras 5.6 a 5.8 se muestran los resultados de predicción a 1 paso de viento y potencia sobre la base de datos de entrenamiento y validación:

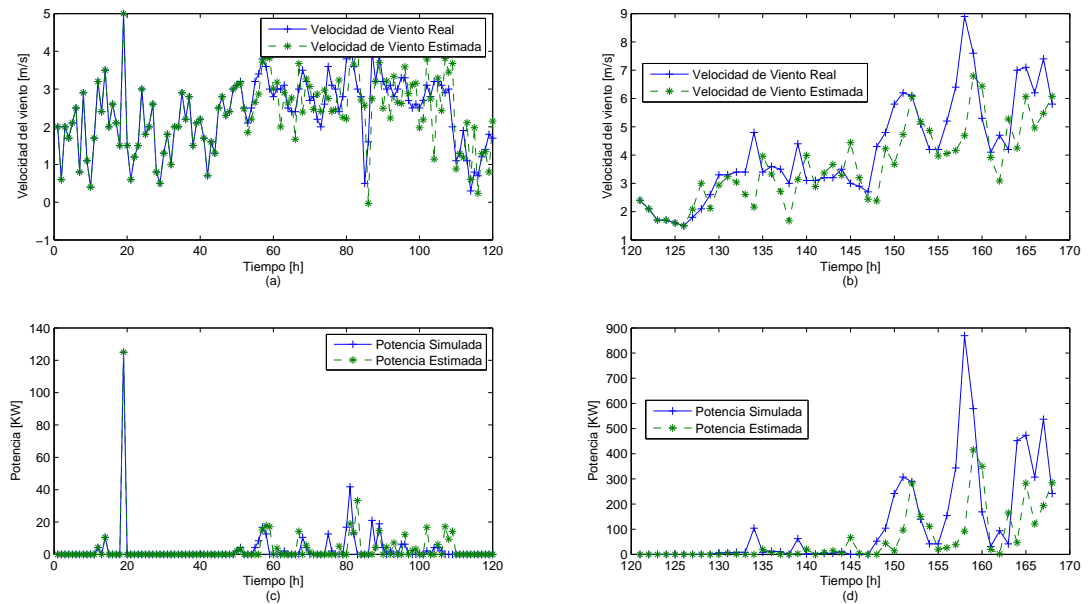


Figura 5.6: Resultados de predicción a un paso con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

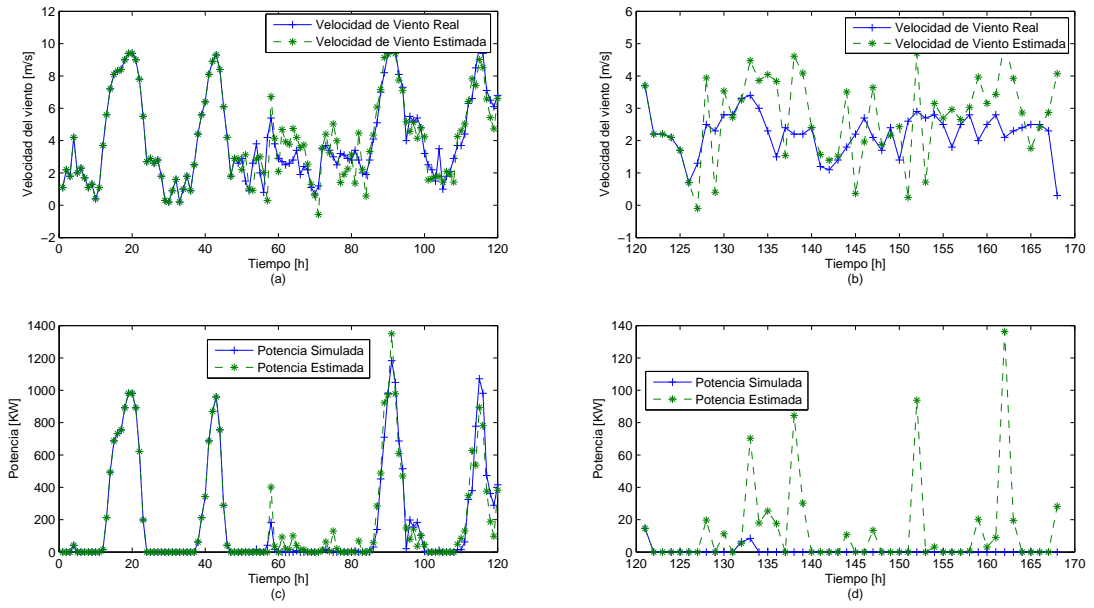


Figura 5.7: Resultados de predicción a un paso con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

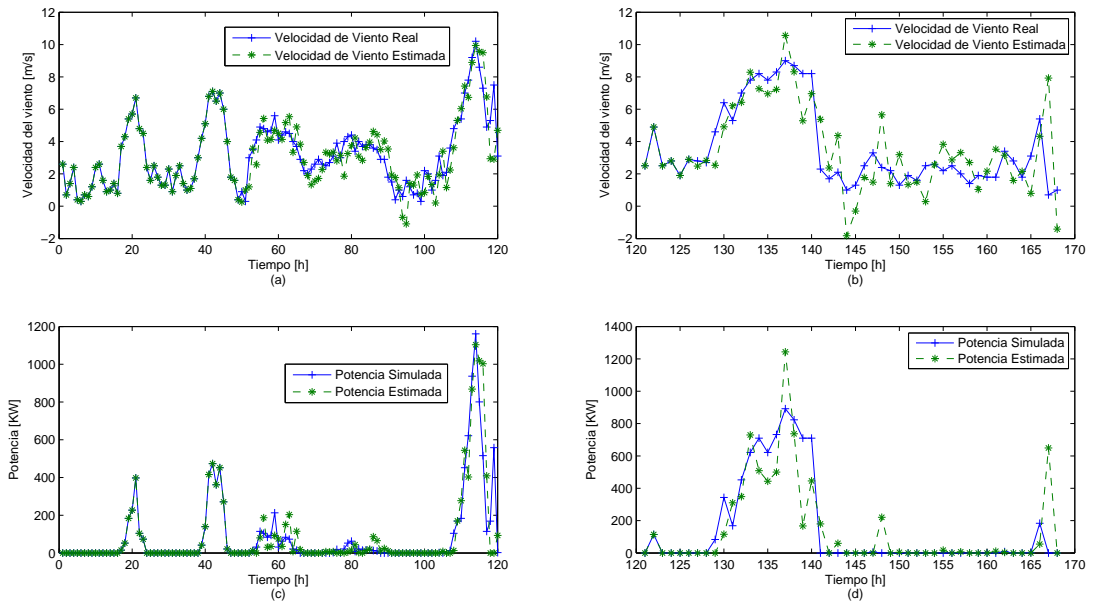


Figura 5.8: Resultados de predicción a un paso con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

Analizando las figuras 5.6 a 5.8, se puede notar que el modelo SARIMA sigue la tendencia de la serie de tiempo, pero con cierto error asociado al valor predicho en cada paso. Además se aprecia que transcurridas 24[h], el modelo obtenido presenta mayor errores que en un principio, lo cual sugiere que el comportamiento estadístico de la velocidad del viento presenta leves variaciones transcurridas las 24[h]. Para el caso de la predicción de potencia, se observa claramente un aumento en la diferencia entre la potencia simulada y la potencia predicha por el modelo, lo cual es consecuencia de la transformación de la velocidad del viento y en especial de la no-linealidad de esta.

Además en las figuras 5.6 a 5.8, se puede notar que el modelo SARIMA sigue la tendencia de los datos de entrenamiento, lo cual era de esperarse dado que éste fue diseñado para seguir la estadística de los datos de entrenamiento. De hecho, se observa que el error asociado a la predicción en cada paso, es menor en el caso de los datos de entrenamiento que sobre los datos de validación.

Nuevamente, se observa que transcurridas las 24[h] el error en la predicción aumenta levemente, lo cual se infiere que es debido a que la estadística de la velocidad del viento para este tipo de datos cambia levemente. Para el caso de la predicción de potencia, se observa claramente un aumento en la diferencia entre la potencia simulada y la potencia predicha por el modelo, lo cual es consecuencia de la transformación de la velocidad del viento y en especial de la no-linealidad de esta.

Con el fin de analizar el comportamiento del modelo SARIMA en el ámbito de predicción con un horizonte de tiempo distinto al anterior, en las figuras 5.9 a 5.11 se muestran los resultados de predicción y potencia a 5 pasos sobre los datos de entrenamiento y validación:

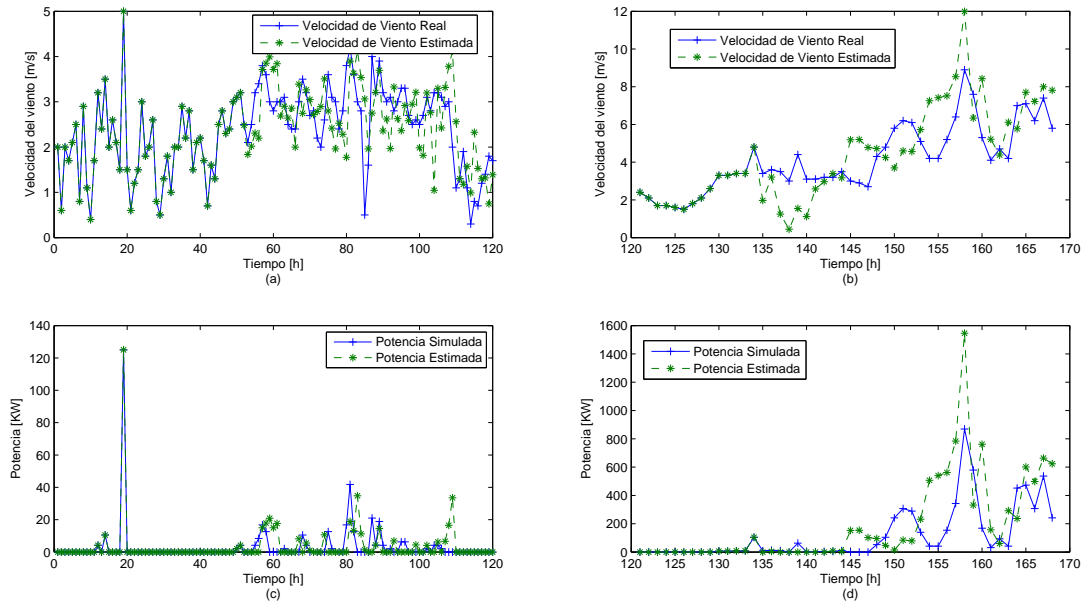


Figura 5.9: Resultados de predicción a cinco pasos con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

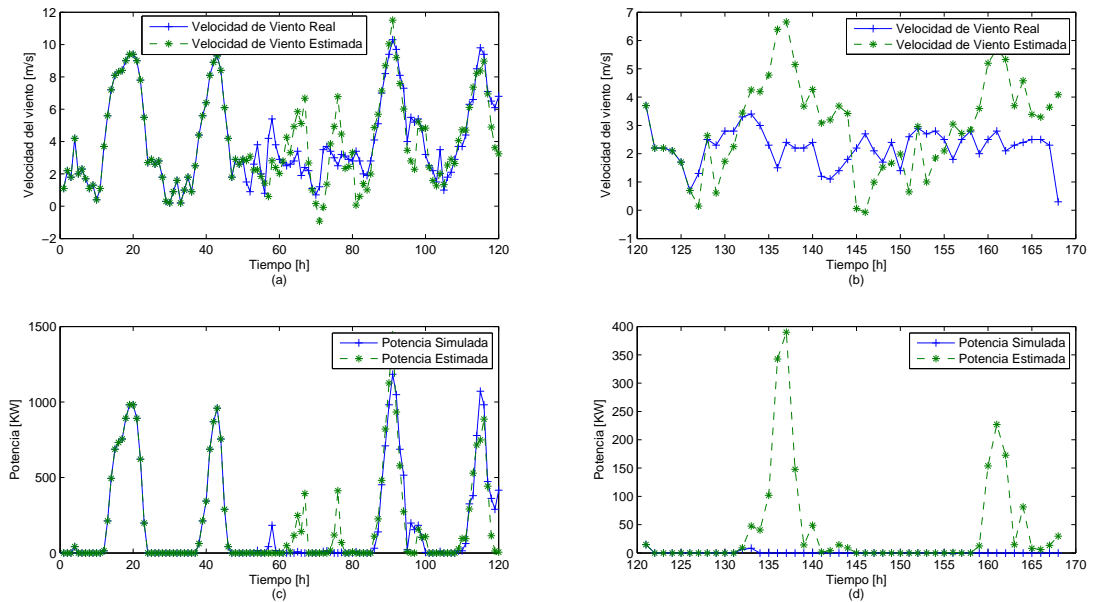


Figura 5.10: Resultados de predicción a cinco pasos con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

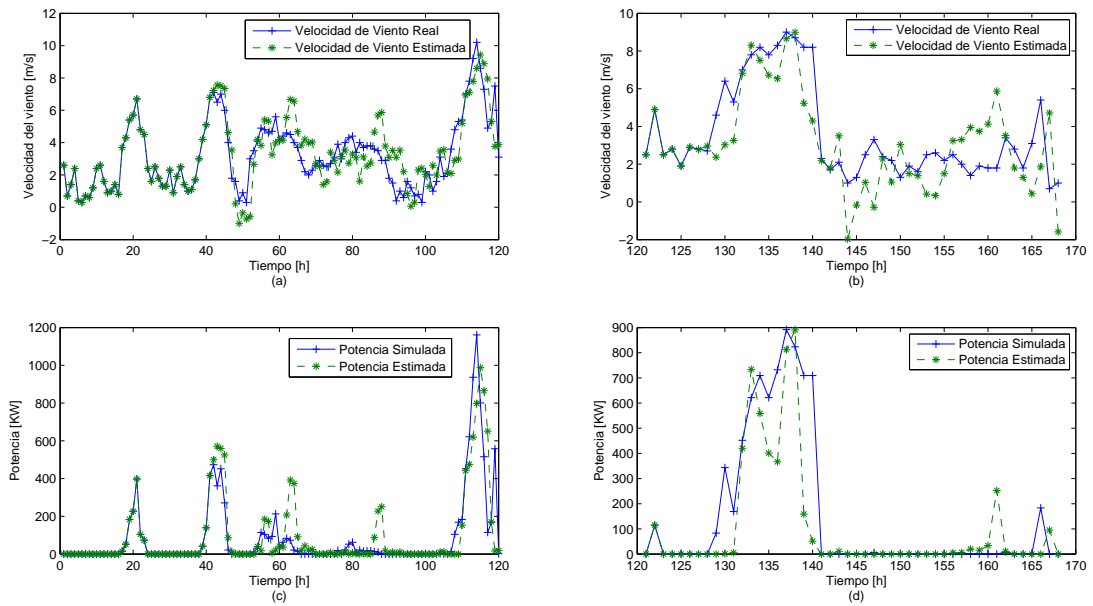


Figura 5.11: Resultados de predicción a cinco pasos con modelo SARIMA sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

A través del análisis de las figuras 5.9 a 5.11, se puede notar que nuevamente el modelo SARIMA es capaz de predecir las tendencias de los datos de validación, pero con cierto error asociado al valor predicho. Esto último ocurre debido a que los parámetros del modelo SARIMA se ajustan de acuerdo a propiedades de la estadística de los datos de entrenamiento, la cual varía levemente con los datos de validación, razón por la que se tienen errores de predicción.

Para el caso de la potencia, se observa que los errores aumentan en comparación con el caso de la predicción a 1 paso, lo cual es consecuencia del aumento del error en la predicción de viento así como también de la amplificación de este error debido a la transformación de la curva de potencia. Además se aprecia como el modelo disminuye su error al cabo de 5 pasos, resultado que se justifica en que al hacer predicción a 5 pasos, se tiene que cada cinco pasos se utilizan los valores reales de velocidad de viento y no los predichos, con la consecuente disminución del error.

Además, por medio del análisis de las figuras 5.9 a 5.11, se puede notar que nuevamente

el modelo SARIMA es capaz de predecir las tendencias de los datos de entrenamiento, pero con cierto error asociado al valor pronosticado. El error asociado a los valores pronosticados aumenta en comparación al caso de la predicción a un paso, lo cual es de esperarse puesto a que en general la incertidumbre o varianza asociada al proceso aumenta a medida que aumenta el horizonte de predicción.

Para el caso de la potencia, se observa que los errores aumentan en comparación con el caso de la predicción a 1 paso, lo cual es consecuencia del aumento del error en la predicción de viento así como también de la amplificación de este error debido a la transformación de la curva de potencia.

5.5. Resultados de modelo de persistencia

Se implementa el modelo de persistencia en predicción a uno y cinco pasos, con el fin de tener un *benchmark* con el cual comparar el modelo SARIMA propuesto para predicción de potencia generada en un aerogenerador. Los resultados de predicción a un paso de velocidad de viento y potencia generada sobre la base de datos de entrenamiento y validación del modelo SARIMA se muestran en las figuras 5.12 a 5.14.

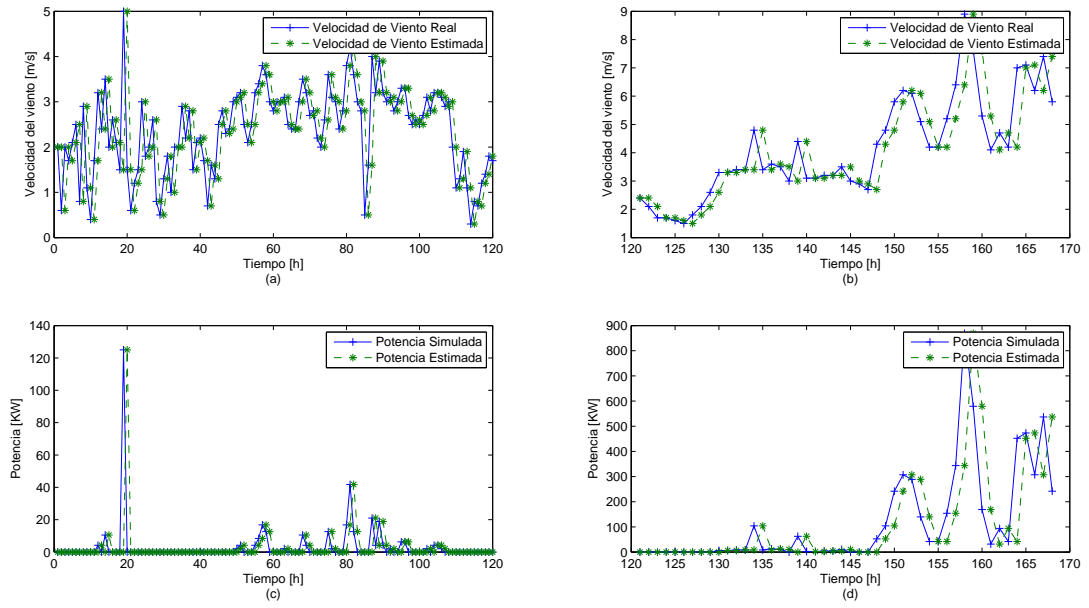


Figura 5.12: Resultados de predicción a un paso con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

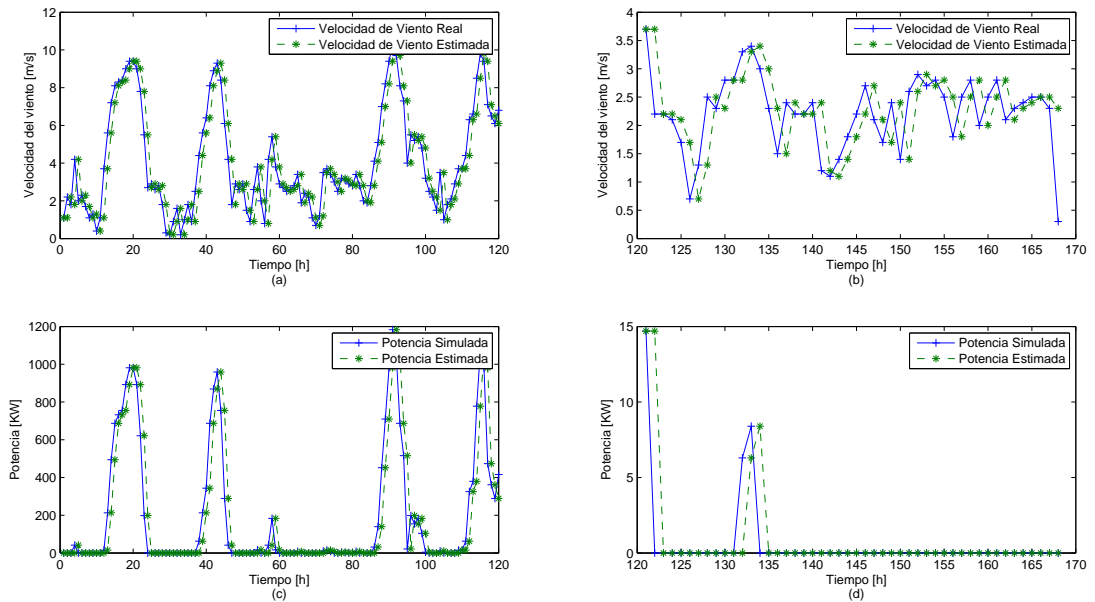


Figura 5.13: Resultados de predicción a un paso con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

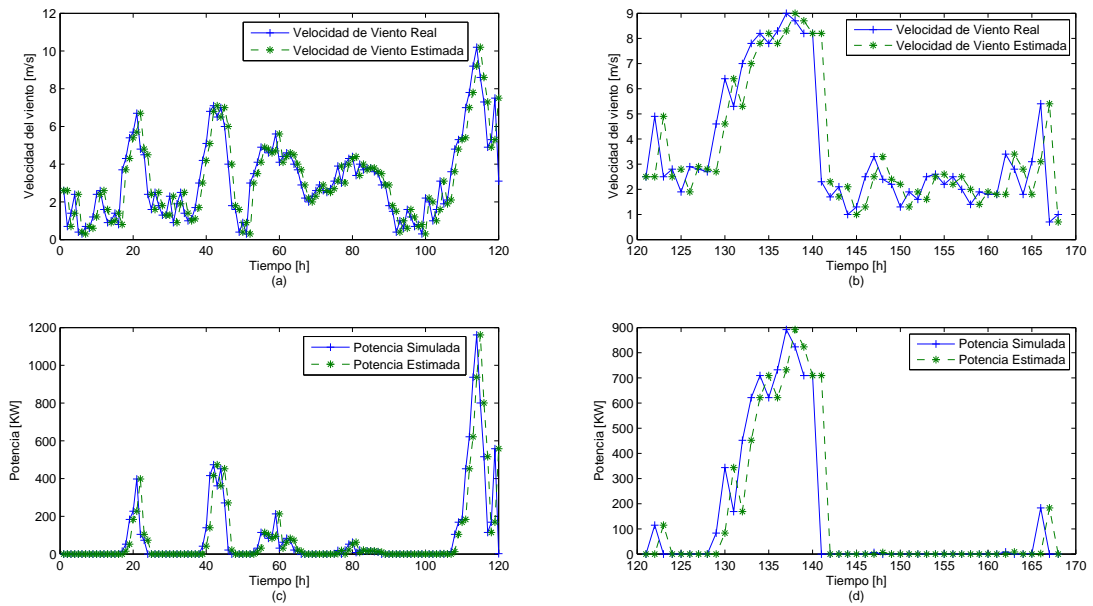


Figura 5.14: Resultados de predicción a un paso con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

Tal como se puede ver en las figuras 5.12 a 5.14, el modelo de persistencia sigue la tendencia de los datos de validación sin mayores problemas, lo cual se debe a que en el caso de predicción a un paso de velocidad de viento y potencia generada en un aerogenerador, los valores consecutivos de la serie de tiempo no tienen variaciones que sean acentuadas, sino más bien son graduales, por lo que este tipo de modelo no tiene problemas en seguir la tendencia de los datos. A pesar de esto último, existe un error asociado debido a que los valores no cumplen con la hipótesis de persistencia que requiere el modelo. El error en la predicción de velocidad del viento se incrementa una vez que se simula la potencia que genera el aerogenerador, debido a la transformación no-lineal que representa la curva de potencia del aerogenerador.

Además, se muestra que el modelo de persistencia tiene un desempeño similar tanto en los datos de entrenamiento como en los de validación del modelo SARIMA. Esto se justifica en el hecho de que el modelo de persistencia no requiere datos de entrenamiento, por lo que no se ajusta para una serie de datos dada.

Con el fin de analizar el comportamiento del modelo persistencia en el ámbito de predicción con un horizonte de tiempo distinto al anterior, a continuación se muestran los resultados de predicción y potencia a 5 pasos sobre los datos de entrenamiento y validación del modelo SARIMA:

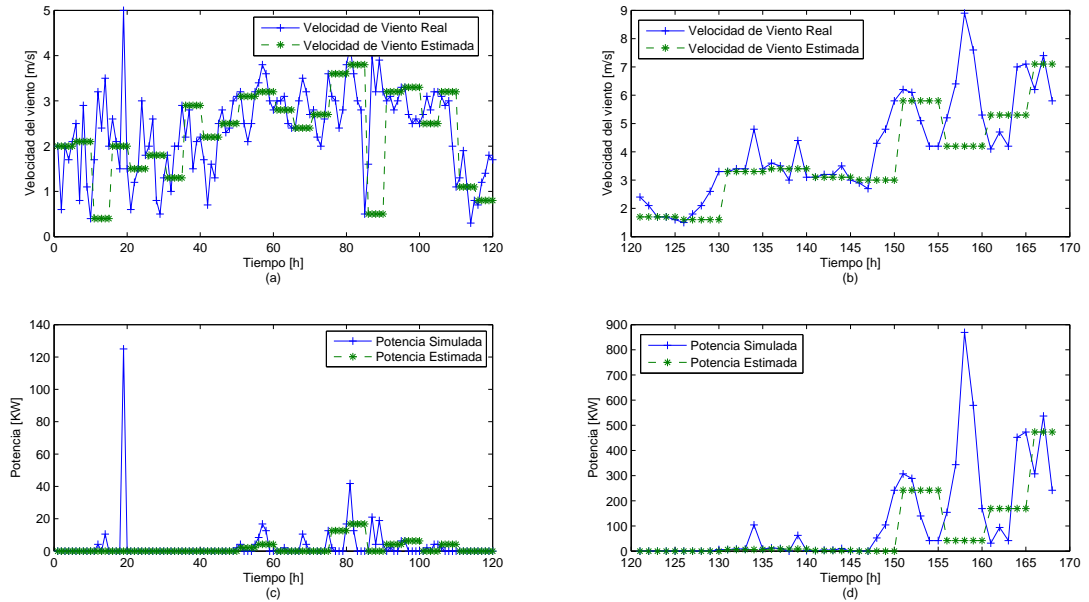


Figura 5.15: Resultados de predicción a cinco pasos con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 24 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

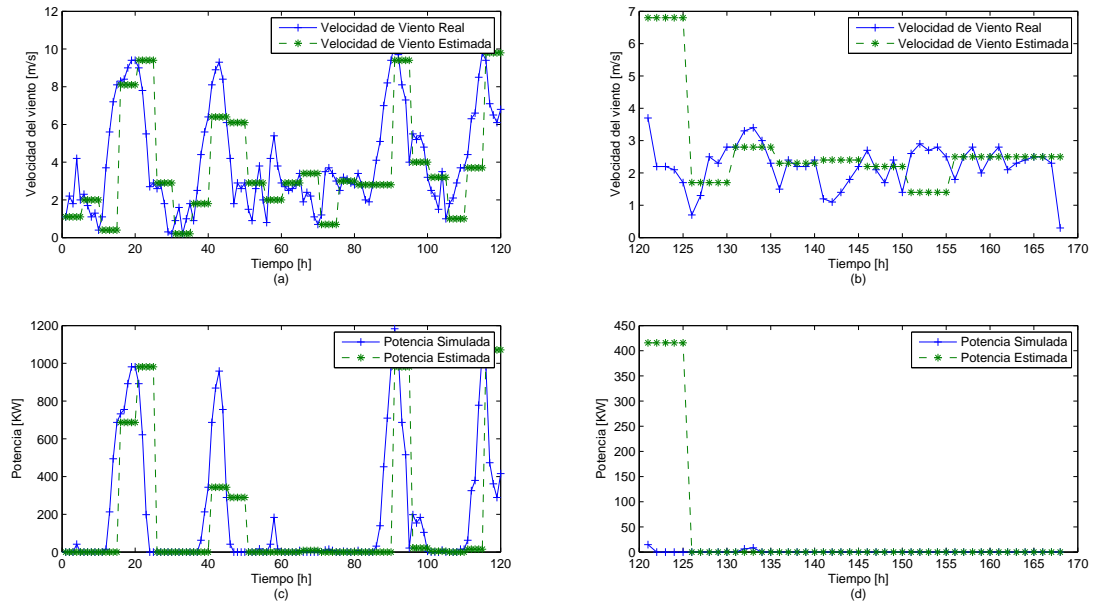


Figura 5.16: Resultados de predicción a cinco pasos con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 12 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

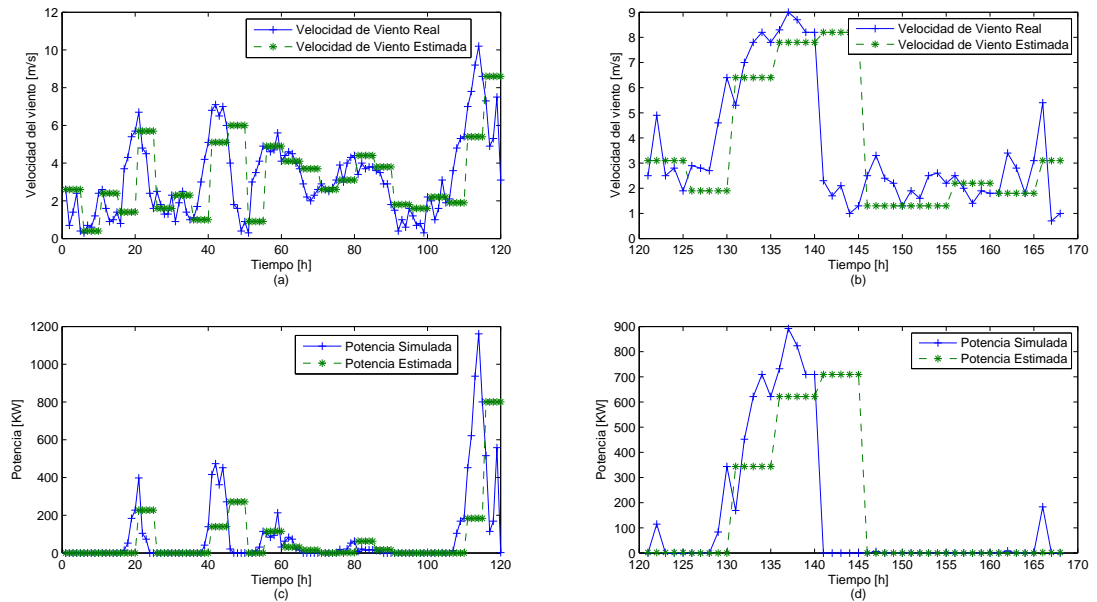


Figura 5.17: Resultados de predicción a cinco pasos con modelo de persistencia sobre serie de tiempo de viento muestreada en Punta Lengua de Vaca en la semana 14 de 1990. (a) Datos de entrenamiento. (b) Datos de validación. (c) Potencia simulada con datos de entrenamiento. (d) Potencia simulada con datos de validación.

A través de las figuras 5.15 a 5.17 se puede notar que en el caso de la predicción a cinco pasos, el modelo de persistencia no es capaz de describir la dinámica de la velocidad del viento por lo que en las predicciones realizadas, claramente no se sigue la tendencia de los datos. Para los datos de entrenamiento y validación los resultados se explican por la tendencia de los datos, debido a la dinámica del viento, la cual no logra ser descrita por medio del modelo de persistencia. Además, se nota que para el caso de la potencia, el error se incrementa más aún, por la no-linealidad que agrega la transformación de la curva de potencia.

5.6. Comparación entre los modelos implementados

Una vez obtenidos los resultados de predicción a un paso y cinco pasos usando modelo SARIMA y modelo de persistencia, se realiza una comparación entre ambos desde el punto de vista del error cuadrático medio. Para esto último se calcula el siguiente índice de desviación estándar porcentual (*DEP*) con respecto a la media de los datos que se está prediciendo:

$$DEP = \frac{\sqrt{MSE} \cdot 100\%}{\text{Media de los datos}} \quad (5.1)$$

Las desviaciones estándar porcentuales en predicción a un paso de ambos modelos, se muestran en el siguiente cuadro:

Serie de Tiempo	Horizonte de Predicción	
	SARIMA - 1 Paso	Persistencia - 1 Paso
Entrenamiento 1	31.1	37.1
Entrenamiento 2	27.5	30.1
Entrenamiento 3	35.2	33.3
Validación 1	31.5	23.8
Validación 2	52.1	29.4
Validación 3	40.2	35.2

Cuadro 5.1: Desviaciones estándar porcentuales obtenidas en predicción a un paso para modelo SARIMA y modelo de persistencia.

En el cuadro 5.1 se puede notar que en general el modelo SARIMA obtuvo un mejor desempeño en el sentido del error cuadrático medio que el modelo de persistencia en predicción a un paso sobre los datos de entrenamiento del modelo SARIMA, pero no así en la predicción a un paso sobre los datos de validación, en donde el modelo de persistencia presentó un mejor desempeño. Esto se justifica en el hecho que las expresiones bajo las cuales se ajustan los parámetros del modelo SARIMA derivadas de la minimización del error cuadrático medio, son equivalentes al ajuste de la función de autocorrelación, por lo que en el fondo el modelo SARIMA se sustenta en propiedades estadísticas de la serie de tiempo y al realizar predicciones a un paso el modelo busca conservar dicha estadística, lo cual en la realidad no es estrictamente cierto entre los datos de entrenamiento y validación. Además, dado que en predicción a un paso la dinámica del viento puede ser seguida por el modelo de persistencia de manera aproximada, se obtuvo que el desempeño de este modelo fue relativamente similar para la serie de datos de validación y entrenamiento del modelo SARIMA.

Las desviaciones estándar porcentuales en predicción a cinco pasos de ambos modelos, se muestran en el siguiente cuadro:

Horizonte de Predicción	SARIMA - 5 Pasos	Persistencia - 5 Pasos
Serie de Tiempo		
Entrenamiento 1	35.4	45.5
Entrenamiento 2	40.7	62.4
Entrenamiento 3	42.1	61.8
Validación 1	44.2	46.4
Validación 2	52.2	46.1
Validación 3	52.5	69.4

Cuadro 5.2: Desviaciones estándar porcentuales obtenidas en predicción a un paso para modelo SARIMA y modelo de persistencia.

En el cuadro 5.2 se puede notar que en general el modelo SARIMA obtuvo un mejor desempeño en el sentido del error cuadrático medio que el modelo de persistencia tanto en predicción a cinco pasos sobre los datos de entrenamiento del modelo SARIMA, como en predicción a cinco pasos sobre los datos de validación, a excepción de la segunda serie de tiempo. Esto se justifica debido a que el horizonte de predicción es tal que la dinámica del viento no puede ser correctamente descrita por el modelo de persistencia pero si lo es descrita de mejor manera por el modelo SARIMA. Esto era de esperarse dado que es claro que para un horizonte de predicción de cinco pasos, la hipótesis del modelo de persistencia no es válida.

Además se nota que para ambos modelos, el error cuadrático medio obtenido en predicción a cinco pasos es mayor que el error cuadrático medio obtenido en predicción a un paso, lo cual también es de esperarse debido a que mientras mayor es el horizonte de predicción, se posee mayor incertidumbre en el proceso, con el consecuente aumento en la varianza de cada estimador.

Capítulo 6

Conclusiones

El correcto desarrollo de un modelo SARIMA para predicción de potencia generada en un aerogenerador indica que el objetivo principal de este trabajo de título ha sido alcanzado satisfactoriamente. El cumplimiento de esta meta significa a su vez, el cumplimiento de todos los objetivos específicos planteados al comienzo de este trabajo de título.

La investigación teórica previa permitió conocer, en un primer acercamiento, métodos de predicción existentes, las áreas donde estos han sido aplicados exitosamente, sus limitaciones y el desempeño que han mostrado. Asimismo, también permitió conocer la dinámica del viento en la atmósfera y de que manera se relaciona la potencia generada por un aerogenerador con la velocidad y dirección del viento del lugar en el cual se encuentra instalado.

Mediante la introducción de índices de desempeño, fue posible cuantificar el desempeño de los modelos propuestos para predicción a uno y cinco pasos. Basado en dichos índices, es posible afirmar que los resultados al problema de predicción dependen del horizonte de tiempo que se está tomando. Para el caso de predicción a un paso, el modelo de persistencia obtuvo mejores resultados sobre la base de datos de validación que el modelo SARIMA y para el caso de predicción a cinco pasos, el modelo SARIMA resulta superior al modelo de persistencia.

El desempeño de los modelos SARIMA y de persistencia se encuentran condicionados al horizonte de predicción utilizados y depende de si dichos modelos son capaces de seguir la dinámica del fenómeno que se está estudiando. En especial, se obtuvo que el modelo SARIMA

presenta mejor desempeño cuando la estadística de los datos no varía mayormente a lo largo de la serie de tiempo y es bajo esta condición en donde el modelo sirve. Aún así, se concluye que el modelo SARIMA es factible de implementar para realizar predicción de corto plazo de potencia generada en un aerogenerador y que la fiabilidad de sus predicciones depende del horizonte de tiempo al cual se encuentra trabajando.

La curva de potencia de un aerogenerador resultó ser una herramienta de gran utilidad para predecir la potencia generada por un aerogenerador debido a que al relacionar la velocidad del viento con la potencia generada, hizo posible que se trabajara con datos de velocidad de viento los cuales presentaron propiedades estadísticas que fueron aprovechadas para ajustar un modelo SARIMA de predicción.

Finalmente, la implementación como tal permitió desarrollar habilidades en el ámbito de toma de decisiones por medio de la utilización de herramientas estadísticas que caracterizan el problema enfrentado y que dan un respaldo desde el punto de vista teórico a la toma de decisiones.

Los trabajos futuros que podrían generar extensiones del sistema aquí presentado corresponden a:

- Utilizar filtro de Kalman para predicción utilizando el modelo SARIMA obtenido como la ecuación de estado, caracterizando el desempeño de este sistema conjunto para distintos horizontes de tiempo.
- Implementar un modelo SARIMA más general que no realice mayores simplificaciones.
- Utilizar en conjunto datos reales de velocidad de viento y potencia generada en un aerogenerador, para así analizar que tan exacta es la relación que da la curva de potencia del aerogenerador.
- Analizar una posible transformación de la serie de datos con la cual se trabaja, con el fin de tomar en cuenta la altura y la rugosidad del terreno en donde fueron tomados los datos.

Referencias

- [1] J. Zhong, Y. Hou, and F. F. Wu, “Wind power forecasting and integration to power grids,” in *International Conference on Green Circuits and Systems (ICGCS)*, Shanghai, 2010, pp. 555–560.
- [2] B. Ernst, B. Oakleaf, M. L. Ahlstrom, M. Lange, C. Moehrlen, B. Lange, U. Focken, and K. Rohrig, “Predicting the wind,” *IEEE Power Energy Mag.*, vol. 5, no. 6, pp. 78–89, November-December 2007.
- [3] Y. Zhang and K. W. Chang, “The impact of wind forecasting in power system reliability,” in *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, Nanjuing, 2008, pp. 2781–2785.
- [4] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, “A review of wind power and wind speed forecasting methods with different time horizons,” in *North American Power Symposium*, Arlington, TX, 2010, pp. 1–8.
- [5] C. de Despacho Económico del Sistema Interconectado Central, www.cdec-sic.cl, 2010.
- [6] M. Lange and U. Focken, *Physical Approach to Short-Term Wind Power Prediction*, 1st ed. Berlin, Germany: Springer, 2005.
- [7] K. W. Arnaud, “Memoria para optar al título de Ingeniero Civil Eléctrico de la Universidad de Chile: Simulación dinámica para distintos escenarios de alta penetración eólica en el SIC,” 2008.

- [8] P. B. Rosenmann, “Memoria para optar al título de Ingeniero Civil Eléctrico de la Universidad de Chile:diseño de aerogeneradores con imanes permanentes para aplicaciones en electrificación rural,” 2007.
- [9] A. M. Foley, P. G. Leahy, and E. J. McKeogh, “Wind power forecasting & prediction methods,” in *9th International Conference on Environment and Electrical Engineering*, Prague, Czech Republic, 2010, pp. 61–64.
- [10] A. S. Dorvlo, “Estimating wind speed distribution,” *Energy Conversion and Management*, vol. 43, no. 17, pp. 2311–2318, November 2002.
- [11] P. Aceituno, “Proyecto EOLO,” <http://www.dgf.uchile.cl/eolo/>.
- [12] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. Wiley-Interscience, 1994, vol. 1.
- [13] C. Chatfield, *Time-Series Forecasting*. Boca Raton, Florida: Chapman & Hall/CRC, 2000.
- [14] B. G. Brown, R. W. Katz, and A. H. Murphy, “Time series models to simulate and forecast wind speed and wind power,” *Journal of Climate and Applied Meteorology*, vol. 23, pp. 1184–1195, August 1984.
- [15] S.-J. Huang and K.-R. Shih, “Short-term load forecasting via ARMA model identification including non-Gaussian process considerations,” *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 673–679, 2005.
- [16] P. Damrongkulkamjorn and P. Churueang, “Monthly energy forecasting using decomposition method with application of seasonal ARIMA,” in *The 7th International Power Engineering Conference, 2005. IPEC 2005.*, Singapore, 2005, pp. 1–6.
- [17] A. Guin, “Travel time prediction using a seasonal autoregressive integrated moving average time series model,” in *Intelligent Transportation Systems Conference*, Toronto, Ont., 2006, pp. 493–498.

- [18] V. Ediger and S. Akar, “ARIMA forecasting of primary energy demand by fuel in Turkey,” *Energy Policy*, vol. 35, no. 3, pp. 1701–1708, 2006.
- [19] C. Lei and L. Ran, “Short-term wind speed forecasting model for wind farm based on wavelet decomposition,” in *Electric Utility Deregulation and Restructuring and Power Technologies International Conference*, April 2008, pp. 2525 – 2529.
- [20] L. Wang, L. Dong, Y. Hao, and X. Liao, “Wind power prediction using wavelet transform and chaotic characteristics,” in *World Non-Grid-Connected Wind Power and Energy Conference*, September 2009, pp. 1–5.
- [21] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*. San Francisco, California: Holden-Day, 1976.
- [22] L. Kamal and Y. Z. Jafri, “Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan,” *Solar Energy*, vol. 61, no. 1, pp. 23–32, March 1997.
- [23] R. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [24] G. Welch and G. Bishop, “An introduction to the Kalman filter,” University of North Carolina at Chapel Hill, Tech. Rep. TR 95-041, 2006.
- [25] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications and Control*. Upper Saddle River, New Jersey: Prentice Hall, 1995.
- [26] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, and G. Kallos, “Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 12, pp. 2348–2362, 2008.
- [27] E. Bossanyi, “Short-term wind prediction using Kalman filters,” *Wind Engineering*, vol. 9, no. 1, pp. 1–8, 1985.

- [28] X. Wu and Z. Song, "Online chaotic time-series prediction with the derivative-free extended Kalman filter," in *7th World Congress on Intelligent Control and Automation, 2008. WCICA 2008.*, Chongqing, June 2008, pp. 2360 – 2364.
- [29] J. Ma and J.-F. Teng, "Predict chaotic time-series using unscented Kalman filter," in *Third International Conference on Machine Learning and Cybernetics*, Shanghai, 2004, pp. 687–690.
- [30] M. Khalid and A. V. Savkin, "Adaptative filtering based short-term wind power prediction with multiple observations points," in *IEEE International Conference on Control and Automation, 2009. ICCA 2009*, Christchurch, 2009, pp. 1547–1552.
- [31] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural networks models," *IEEE Transactions on Energy Conversion*, vol. 11, no. 4, pp. 762–767, 1996.
- [32] S. Li, "Wind power prediction using recurrent multilayer perceptron neural networks," in *Power Engineering Society General Meeting, IEEE*, Toronto, Ontario, 2003, pp. 2325–2330.
- [33] A. K. Mishra and L. Ramesh, "Application of neural networks in wind power (generation) prediction," in *International Conference on Sustainable Power Generation and Supply. SUPERGEN '09*, Nanjing, 2009, pp. 1–5.
- [34] L. Xin, L. Bin, X. Jianyuan, and T. Yun, "A novel power predicting model of wind farm based on double anns," in *Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, Chengdu, 2010, pp. 1–4.
- [35] G. Sideratos and N. D. Hatziargyriou, "An advanced statical method for wind power forecasting," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 258–265, 2007.

- [36] M. Negnevitsky and C. W. Potter, “Innovative short-term wind generation prediction techniques,” in *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, Atlanta, GA, 2006, pp. 60–65.
- [37] N. M., P. Mandal, and S. A.K., “Machine learning applications for load, price and wind power prediction in power systems,” in *15th International Conference on Intelligent System Applications to Power Systems, 2009. ISAP '09.*, Curitiba, 2009, pp. 1–6.
- [38] X. Weidong, L. Yibing, and L. Xingpei, “Short-term forecasting of wind turbine power generation based on genetic neural network,” in *8th World Congress on Intelligent Control and Automation (WCICA)*, Jinan, 2010, pp. 5943–5946.
- [39] R. Bessa, V. Miranda, and J. Gama, “Wind power forecasting with entropy-based criteria algorithms,” in *Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems. PMAPS '08*, Rincon, 2008, pp. 1–7.
- [40] —, “Improvement in Wind Power Forecasting Based on Information Entropy-Related Concepts,” in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, IEEE*, Pittsburgh, PA, 2008, pp. 1–6.
- [41] J. Deng, “Introduction to grey system theory,” *The Journal of Grey System*, pp. 1–24, 1989.
- [42] T. H. M. El-Fouly, E. F. El-Saadany, and M. M. A. Salama, “Grey predictor for wind energy conversion systems output power prediction,” *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1450–1452, August 2006.
- [43] T. H. M. El-Fouly, E. F. El-Saadany, and M. M. A. Salama, “Improved grey predictor rolling models for wind power prediction,” *IET Generation Transmission and Distribution*, vol. 1, no. 6, pp. 928–937, 2007.
- [44] E. Kayacan, B. Ulutas, and O. Kaynak, “Grey system theory-based models in time series prediction,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 1784–1789, 2010.

- [45] J. Shi, Y. Yang, P. Wang, Y. Liu, and S. Han, “Genetic algorithm-piecewise support vector machine model for short term wind power prediction,” in *8th World Congress on Intelligent Control and Automation*, Jinan, 2010, pp. 2254–2258.
- [46] C. Ye, G. Li, and M. Zhou, “A combined prediction method of wind farm power,” in *5th International Conference on Critical Infrastructure (CRIS), 2010*, Beijing, 2010, pp. 1–5.
- [47] S. Dutta and T. J. Overbye, “Prediction of short term power output of wind farms based on least squares method,” in *Power and Energy Society General Meeting*, Minneapolis, MN, 2010, pp. 1–6.
- [48] N. Abdel-Karim, M. Small, and M. Ili?, “Short-term wind speed prediction by finite and infinite impulse response filters: A state space model representation using discrete markov process,” in *PowerTech Conference*, Bucharest, 2009, pp. 1–8.
- [49] J. W. Taylor, P. E. McSharry, and R. Buizza, “Wind power density forecasting using ensemble predictions and time series models,” *IEEE Transactions on Energy Conversion*, vol. 24, no. 3, pp. 775–782, 2009.
- [50] L. Landberg, “Short-term prediction of the power production from wind farms,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 80, no. 1-2, pp. 207–220, 1999.
- [51] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time signal processing*. Upper Saddle River, New Jersey: Prentice-Hall, 1999.

Apéndices

A . Función de Autocorrelación

Una serie de tiempo es un conjunto de observaciones generadas secuencialmente en el tiempo [21]. Un proceso estocástico z_t discreto y estacionario es aquel en donde el muestreo de la señal se hace cada cierto intervalo de tiempo y además la distribución de probabilidad que describe al proceso en un instante de tiempo fijo es la misma para todos los instantes de tiempo. En consecuencia, parámetros tales como la media y la varianza, si existen, no varían a lo largo del tiempo o la posición. La media para un proceso de esta naturaleza se encuentra definida por:

$$\mu = E[z_t] = \frac{1}{N} \sum_{i=1}^N z_i. \quad (6.1)$$

en donde n corresponde al número de datos de la serie de tiempo. De igual modo, la varianza se define por:

$$\sigma_z^2 = E[(z_t - \mu)^2] = \frac{1}{N} \sum_{i=1}^N (z_i - \mu)^2. \quad (6.2)$$

De este modo se puede interpretar la varianza de una muestra de datos como una medida de la desviación que tienen las observaciones con respecto a su media [21]. También se define la autocovarianza como:

$$\gamma_k = E[(z_t - \mu)(z_{t+k} - \mu)] \quad (6.3)$$

Finalmente se define la autocorrelación por:

$$\rho_k = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E[(z_t - \mu)^2]E[(z_{t+k} - \mu)^2]}} = \frac{\gamma_k}{\gamma_0} \quad (6.4)$$

La autocorrelación entrega una medida de que tan relacionadas se encuentra la medida de la variable z realizada en el instante t con aquella medida realizada en el instante $t + k$. Luego, si no estuviesen relacionadas, se tiene que $\rho_k = 0$ y por el contrario, si ambas medidas estuviesen perfectamente relacionadas (existe una relación lineal entre ambas variables) se tiene que $\rho_k = 1$ [21].

Lo anterior puede ser utilizado para el análisis de series de tiempo de viento y/o potencia, ya que al graficar la función de autocorrelación para distintos espacios de tiempo, se puede ver si existe dependencia entre las mediciones realizadas en diferentes instantes de tiempo. Luego, en caso de existir relación, el problema se reduce a encontrar dicha relación de manera tal que sea utilizada para realizar predicciones de potencia.

B . Transformada a Tiempo Discreto de Fourier

La transformada discreta a tiempo discreto de Fourier es una herramienta matemática para obtener el espectro en frecuencia de una señal, es decir, de las componentes de frecuencia que componen a dicha señal. Para el caso de las series de tiempo, la aplicación de la transformada de Fourier a los datos ayuda a identificar cierta periodicidad en el comportamiento de la variable estudiada. La transformada discreta de Fourier para tiempo discreto está definida por [51]:

$$Z(\omega) = \sum_{n=-\infty}^{\infty} z[n] e^{-i\omega n} \quad (6.5)$$

C . Códigos

En este anexo se incluyen los códigos de los programas utilizados para las simulaciones.

C .1. `simular`

```
function y = simular()
```

```

load historial_datos.mat

n_cruz = floor(length(cruz)/168);
n_vaca = floor(length(leng_vaca)/168);
n_los = floor(length(los)/168);
n_paranal = floor(length(paranal)/168);
n_sumit = floor(length(sumit)/168);

for i =1:n_vaca
    nombre_archivo = strcat(strcat('resultados_vaca_',num2str(i)),'.mat');
    busqueda_ordenes_sarima2(leng_vaca(((i-1)*168+1):(i*168)),nombre_archivo);
end

for i =1:n_cruz
    nombre_archivo = strcat(strcat('resultados_cruz_',num2str(i)),'.mat');
    busqueda_ordenes_sarima2(cruz(((i-1)*168+1):(i*168)),nombre_archivo);
end

for i =1:n_los
    nombre_archivo = strcat(strcat('resultados_vilos_',num2str(i)),'.mat');
    busqueda_ordenes_sarima2(los(((i-1)*168+1):(i*168)),nombre_archivo);
end

for i =1:n_paranal
    nombre_archivo = strcat(strcat('resultados_paranal_',num2str(i)),'.mat');
    busqueda_ordenes_sarima2(paranal(((i-1)*168+1):(i*168)),nombre_archivo);
end

for i =1:n_sumit
    nombre_archivo = strcat(strcat('resultados_sumit_',num2str(i)),'.mat');
    busqueda_ordenes_sarima2(sumit(((i-1)*168+1):(i*168)),nombre_archivo);
end

y = 1e10;

```

C .2. busqueda_ordenes_sarima2

```

function w = busqueda_ordenes_sarima2(datos,filename)

datos = generar_serie();

media = mean(datos);

datos_con_media = datos;
datos = datos-media;

s=24;

mse_min_train = 1e20;
wmse_min_train = 1e20;
bic_min_train = 1e20;
aic_min_train = 1e20;
mape_min_train = 1e20;

mse_min_val = 1e20;
wmse_min_val = 1e20;
bic_min_val = 1e20;
aic_min_val = 1e20;
mape_min_val = 1e20;

```

```

p_chico_max = 13;
q_max = 13;
d_max = 2;
D_max = 2;
iter_max = 1;

matriz_mse_train = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_wmse_train = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_mape_train = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_bic_train = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_aic_train = zeros(p_chico_max,q_max,d_max+1,D_max);

matriz_mse_val = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_wmse_val = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_mape_val = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_bic_val = zeros(p_chico_max,q_max,d_max+1,D_max);
matriz_aic_val = zeros(p_chico_max,q_max,d_max+1,D_max);

p_nan = [];
q_nan = [];
d_nan = [];
D_nan = [];

for p_chico=1:p_chico_max
    for q=1:q_max
        for d = 0:d_max
            if(d~=0)
                nuevos_datos = diff(datos,d);
            else
                nuevos_datos = datos;
            end

            for D=1:D_max

                promedio_mse_train =0;
                promedio_mse_val =0;

                promedio_mape_train =0;
                promedio_mape_val =0;

                promedio_bic_train = 0;
                promedio_bic_val = 0;

                promedio_aic_train = 0;
                promedio_aic_val = 0;

                promedio_wmse_train =0;
                promedio_wmse_val =0;

            for iteraciones = 1:iter_max

                data_D = zeros(length(nuevos_datos)-D*s,1);
                for i=(D*s+1):length(nuevos_datos)
                    data_D(i-D*s) = nuevos_datos(i);
                    for j=1:D
                        data_D(i-D*s) = data_D(i-D*s)//
                            nchoosek(D,j)*(-1)^(j)*nuevos_datos(i-j*s);
                    end
                end

                [estimacion_arma estimacion_arma_train , ~, sigma_arma ~] = boxjenkins3(data_D,p_chico,q);
            end
        end
    end
end

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Se recupera la serie diferenciada

estimacion_diferencial = zeros(length(nuevos_datos),1);
estimacion_diferencial(1:D*s) = nuevos_datos(1:D*s);

for i=(D*s+1):length(nuevos_datos)
    estimacion_diferencial(i) = estimacion_arma(i-D*s);
    for j =1:D
        estimacion_diferencial(i) = estimacion_diferencial(i)-//
            nchoosek(D,j)*(-1)^(j)*estimacion_diferencial(i-j*s);
    end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Se recupera la serie original

if(d>0)

    estimacion_datos_sin_media = zeros(length(datos),1);
    estimacion_datos_sin_media(1:d) = datos(1:d);

    for i = (d+1):length(datos)
        estimacion_datos_sin_media(i) = estimacion_diferencial(i-d);
        for j=1:d
            estimacion_datos_sin_media(i) = estimacion_datos_sin_media(i)+//
                nchoosek(d,j)*(-1)^(j+1)*estimacion_datos_sin_media(i-j);
        end
    end

else

    estimacion_datos_sin_media = estimacion_diferencial;

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Se recupera la serie diferenciada para
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% entrenamiento

largo_train = floor(length(nuevos_datos)*120/168);

estimacion_diferencial_train = zeros(largo_train,1);
estimacion_diferencial_train(1:D*s) = nuevos_datos(1:D*s);

for i=(D*s+1):largo_train
    estimacion_diferencial_train(i) = estimacion_arma_train(i-D*s);
    for j =1:D
        estimacion_diferencial_train(i) = estimacion_diferencial_train(i)-//
            nchoosek(D,j)*(-1)^(j)*estimacion_diferencial_train(i-j*s);
    end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Se recupera la serie original

```

```

if(d>0)

    estimacion_datos_sin_media_train = zeros(largo_train,1);
    estimacion_datos_sin_media_train(1:d) = datos(1:d);

    for i = (d+1):largo_train
        estimacion_datos_sin_media_train(i) = estimacion_diferencial_train(i-d);
        for j=1:d
            estimacion_datos_sin_media_train(i) = estimacion_datos_sin_media_train(i)+//
                nchoosek(d,j)*(-1)^(j+1)*estimacion_datos_sin_media_train(i-j);
        end
    end

else

    estimacion_datos_sin_media_train = estimacion_diferencial_train;

end

%%%%%%%%%%%%

estimacion_datos_con_media_val = estimacion_datos_sin_media + media;
estimacion_datos_con_media_train = estimacion_datos_sin_media_train+media;

largo_datos = length(datos);

i = 1;
while(abs(estimacion_datos_con_media_train(i) - datos_con_media(i))<0.1)
    i = i+1;
end

n_datos_train = length(estimacion_datos_con_media_train)+1-i;

i = 121;
while(abs(estimacion_datos_con_media_val(i) - datos_con_media(i))<0.01)
    i = i+1;
end

n_datos_val = length(estimacion_datos_con_media_val)+1-i;

mse_train = sum((datos_con_media(1:largo_train)-estimacion_datos_con_media_train).^2) //
n_datos_train;
mse_val = sum((datos_con_media(121:largo_datos)-//
estimacion_datos_con_media_val(121:largo_datos)).^2) //
n_datos_val;

mape_train =sum(abs(datos_con_media(1:largo_train)-estimacion_datos_con_media_train)./ //
datos_con_media(1:largo_train))/n_datos_train;
mape_val =sum(abs(datos_con_media(121:length(datos))- //
estimacion_datos_con_media_val(121:largo_datos)).//
datos_con_media(121:length(datos)))/n_datos_val;

bic_train =largo_train*log(sigma_arma)+(p_chico+q+1)*log(largo_train);
bic_val =48*log(sigma_arma)+(p_chico+q+1)*log(48);

aic_train =2*(p_chico+q+1)+largo_train*log(sigma_arma);
aic_val = 2*(p_chico+q+1)+48*log(sigma_arma);

wmse_train = 0;

```

```

wmse_val = 0;

for i=1:largo_train
    penalizacion = estimacion_datos_con_media_train(i)-datos_con_media(i);
    if(penalizacion>=0)
        wmse_train = wmse_train + penalizacion^2;
    else
        wmse_train = wmse_train + 0.8*penalizacion^2;
    end
end

for i=121:length(estimacion_datos_con_media_val)
    penalizacion = estimacion_datos_con_media_val(i)-datos_con_media(i);
    if(penalizacion>=0)
        wmse_val = wmse_val + penalizacion^2;
    else
        wmse_val = wmse_val + 0.8*penalizacion^2;
    end
end

wmse_train = wmse_train/n_datos_train;
wmse_val = wmse_val/n_datos_val;

promedio_mse_train = ((iteraciones-1)*promedio_mse_train+mse_train)/iteraciones;
promedio_mse_val = ((iteraciones-1)*promedio_mse_val +mse_val)/iteraciones;

promedio_mape_train = ((iteraciones-1)*promedio_mape_train+mape_train)/iteraciones;
promedio_mape_val = ((iteraciones-1)*promedio_mape_val+mape_val)/iteraciones;

promedio_bic_train = ((iteraciones-1)*promedio_bic_train+bic_train)/iteraciones;
promedio_bic_val = ((iteraciones-1)*promedio_bic_val+bic_val)/iteraciones;

promedio_aic_train = ((iteraciones-1)*promedio_aic_train+aic_train)/iteraciones;
promedio_aic_val = ((iteraciones-1)*promedio_aic_val+aic_val)/iteraciones;

promedio_wmse_train = ((iteraciones-1)*promedio_wmse_train+wmse_train)/iteraciones;
promedio_wmse_val = ((iteraciones-1)*promedio_wmse_val +wmse_val)/iteraciones;

end

matriz_mse_train(p_chico,q,d+1,D) = promedio_mse_train;
matriz_wmse_train(p_chico,q,d+1,D) = promedio_wmse_train;
matriz_mape_train(p_chico,q,d+1,D) = promedio_mape_train;
matriz_bic_train(p_chico,q,d+1,D) = promedio_bic_train;
matriz_aic_train(p_chico,q,d+1,D) = promedio_aic_train;

matriz_mse_val(p_chico,q,d+1,D) = promedio_mse_val;
matriz_wmse_val(p_chico,q,d+1,D) = promedio_wmse_val;
matriz_mape_val(p_chico,q,d+1,D) = promedio_mape_val;
matriz_bic_val(p_chico,q,d+1,D) = promedio_bic_val;
matriz_aic_val(p_chico,q,d+1,D) = promedio_aic_val;

if(promedio_mse_train<mse_min_train)
    mse_min_train = promedio_mse_train;
    mejor_estimacion_mse_train = estimacion_datos_con_media_train;
end

if(promedio_wmse_train<wmse_min_train)
    wmse_min_train = promedio_wmse_train;
    mejor_estimacion_wmse_train = estimacion_datos_con_media_train;
end

if(promedio_mape_train<mape_min_train)

```



```

        mape_min_train = promedio_mape_train;
        mejor_estimacion_mape_train = estimacion_datos_con_media_train;
    end

    if(promedio_bic_train<bic_min_train)
        bic_min_train = promedio_bic_train;
        mejor_estimacion_bic_train = estimacion_datos_con_media_train;
    end

    if(promedio_aic_train<aic_min_train)
        aic_min_train = promedio_aic_train;
        mejor_estimacion_aic_train = estimacion_datos_con_media_train;
    end

    if(promedio_mse_val<mse_min_val)
        mse_min_val = promedio_mse_val;
        mejor_estimacion_mse_val = estimacion_datos_con_media_val;
    end

    if(promedio_wmse_val<wmse_min_val)
        wmse_min_val = promedio_wmse_val;
        mejor_estimacion_wmse_val = estimacion_datos_con_media_val;
    end

    if(promedio_mape_val<mape_min_val)
        mape_min_val = promedio_mape_val;
        mejor_estimacion_mape_val = estimacion_datos_con_media_val;
    end

    if(promedio_bic_val<bic_min_val)
        bic_min_val = promedio_bic_val;
        mejor_estimacion_bic_val = estimacion_datos_con_media_val;
    end

    if(promedio_aic_val<aic_min_val)
        aic_min_val = promedio_aic_val;
        mejor_estimacion_aic_val = estimacion_datos_con_media_val;
    end
end
end
end

w = 0;
save(filename)

```

C .3. boxjenkins3

```

function [serie_estimada_val serie_estimada_train theta sigma phi] = boxjenkins3(datos_total,p,q)

% El 120/168 tambien se debe cambiar en la funcion busqueda ordenes y funcion
% error phi
n = floor(length(datos_total)*120/168);
datos_train = datos_total(1:n);
datos_val =datos_total((n+1):length(datos_total));

q=0;

y = iddata(datos_train,[],1);

ordenes = [p q];

```

```

m = armax(y,ordenes);

theta_aux = (m.c(2:(q+1)));

phi_aux = -1*(m.a(2:(p+1)));

theta = theta_aux;
%theta = 0;
phi = phi_aux;

%n1 = length(theta);

%for i=1:n1
%   theta(i) = -1*theta_aux(n1+1-i);
%end

%n1 = length(phi);

%for i=1:n1
%   phi(i) = phi_aux(n1+1-i);
%end

sigma = m.NoiseVariance;

serie_estimada = zeros(length(datos_val),1);
maximo = p;

ruido = zeros(length(datos_val)+q);

for i=1:length(serie_estimada)
    for j=1:q
        serie_estimada(i) = serie_estimada(i)+theta(j)*ruido(q+i-j);
    end
    for j=1:p
        serie_estimada(i) = serie_estimada(i)+phi(j)*datos_total((n+i)-j);
    end

    ruido(q+i) = datos_total(n+i)-serie_estimada(i);
end

serie_estimada_val = [datos_train;serie_estimada];
serie_estimada_train = datos_train;

ruido = zeros(length(datos_train)+q,1);

for i=((maximo+1):length(datos_train))
    serie_estimada_train(i) = 0;
    for j=1:q
        serie_estimada_train(i) = serie_estimada_train(i)+theta(j)*ruido(q+i-j);
    end
    for j=1:p
        serie_estimada_train(i) = serie_estimada_train(i)+phi(j)*datos_train(i-j);
    end

    ruido(q+i) = datos_train(i)-serie_estimada_train(i);
end

```