



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

MODELO DINÁMICO VECTORIAL APLICADO A LA PREDICCIÓN
DE VELOCIDAD DE VIENTO PARA GENERACIÓN EÓLICA

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELECTRICISTA

FRANCISCO JAVIER CORNEJO CARRASCO

PROFESOR GUÍA:

SR. JORGE SILVA SÁNCHEZ.

PROFESOR CO-GUÍA:

SR. MARCOS ORCHARD CONCHA.

MIEMBRO DE LA COMISIÓN:

SR. HÉCTOR AGUSTO ALEGRÍA.

SANTIAGO DE CHILE

JUNIO 2012

RESUMEN DE LA MEMORIA
PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELECTRICISTA
POR: FRANCISCO CORNEJO CARRASCO
FECHA: 25 DE JUNIO DE 2012
PROF. GUÍA: SR. JORGE SILVA SÁNCHEZ
PROF. CO-GUÍA: SR. MARCOS ORCHARD CONCHA

“MODELO DINÁMICO VECTORIAL APLICADO A LA PREDICCIÓN DE VELOCIDAD DE VIENTO PARA GENERACIÓN EÓLICA”

La necesidad de incrementar el aporte de las Energías Renovables No Convencionales (ERN) dentro del diseño de la matriz energética nacional es cada vez mayor. Este hecho se refleja, particularmente, en el notable aumento del número de centrales que utilizan el recurso eólico. Sin embargo, para la integración de los parques eólicos en los sistemas interconectados se requiere conocer de antemano el nivel de generación que pueden alcanzar. En este contexto, surge la necesidad de contar con herramientas que sean capaces de predecir el comportamiento aleatorio de la velocidad del viento, con un cierto grado de confiabilidad, con el fin de prever cuanta energía podrá entregar un aerogenerador al sistema en un instante determinado.

Para tratar la problemática descrita, este trabajo presenta el diseño e implementación de un método para realizar predicción de velocidad de viento, el cual incluye agrupamiento de datos, análisis estadístico de series de tiempo y la caracterización de éstas a través de modelos paramétricos variantes en el tiempo, para un horizonte de mediano plazo. El modelo se basa en un método auto-regresivo y es desarrollado utilizando el registro histórico de velocidad de viento, durante el año 1990, en la localidad de Punta Lengua de Vaca, región de Coquimbo.

Específicamente, el diseño presentado consiste en un modelo vectorial auto-regresivo con variable externa (V-ARX, *Vector Auto Regressive with eXternal input*), el cual tiene la particularidad de realizar la predicción de un día completo (24 horas) en un solo paso, a diferencia de los algoritmos utilizados normalmente, que realizan predicciones a un paso considerando horizontes de tan sólo una hora; lo que necesariamente implica que si se quiere conseguir horizontes de tiempo más lejanos, las predicciones deben ser realimentadas en el modelo en un proceso iterativo. Por otra parte, la metodología empleada para elaborar el predictor, hace posible identificar componentes periódicas del viento, tanto diarias como estacionales, facilitando la diferenciación de grupos de datos que presenten comportamientos o tendencias similares, permitiendo la creación de un modelo más certero para cada uno de estos intervalos. Con el fin de realizar esta agrupación, se plantea una nueva manera de distinguir patrones estacionales en este tipo de series de tiempo, basándose en la similitud de sus distribuciones de probabilidad. Para ello se propone la divergencia de *Kullback-Leibler* como medida de diferenciación, y el algoritmo *K-means* como método de agrupamiento.

Finalmente, como medio de validación del diseño propuesto, se implementan dos modelos utilizados actualmente en el ámbito de predicción: el modelo de Referencia de Nielsen y un modelo ARMA-GARCH. Al ser éstos comparados con el modelo V-ARX entregan como resultado que, dependiendo de las distintas métricas de error consideradas, dicho modelo vectorial presenta una mejora en la certeza de las predicciones: de un 6% a un 55% en relación al modelo de Referencia de Nielsen, y de un 2% a un 8% respecto al modelo ARMA-GARCH. Los resultados obtenidos permiten concluir que la manera propuesta para realizar predicción, es mejor que algunos de los métodos usados actualmente, pues reduce el error de estimación y en consecuencia, las predicciones son más cercanas a los valores reales.

A mi familia.

Agradecimientos

Quiero agradecer en primer lugar a mis padres Miguel y Nidia, y a mi hermano Miguel, por todo su esfuerzo en brindarme las mejores condiciones para estudiar sin inconvenientes ni preocupaciones externas, y por su constante e incondicional apoyo a mis decisiones durante todos estos años de estudio.

En segundo lugar dar gracias a mis profesores Jorge Silva y Marcos Orchard, quienes siempre mostraron una muy buena disposición para resolver mis inquietudes, y que con sus conocimientos y consejos me guiaron en el desarrollo de esta memoria. También agradecer al profesor Patricio Parada por sus recomendaciones y su constante preocupación por mi trabajo. Todos ellos unos muy buenos docentes y grandes personas.

Además, dar las gracias a mis amigos: Carolina, Cristian, David “Chino”, David “Deibid”, Danilo “Cono”, y tantos otros, por todos los buenos momentos que pasamos y que son dignos de recordar. A mis compañeros de laboratorio: César, Eduardo y Rodrigo, con quienes compartí la mayor parte del tiempo que desarrollé este trabajo y que en más de alguna ocasión me brindaron su ayuda. A Paola por haberme soportado en mis momentos de estrés y por ayudarme a redactar y corregir este documento.

Finalmente, una mención honrosa se la adjudica mi gran amigo y compañero Nicolás Lopez, con quien compartimos tantas victorias, derrotas y sufrimientos durante nuestra enseñanza media, y luego en el paso por la universidad, particularmente en el DIE.

A todos ellos, muchas gracias.

Francisco Cornejo Carrasco

Santiago, Junio 2012

Índice General

1. Introducción	1
1.1. Antecedentes	1
1.2. Objetivo General	3
1.3. Objetivos Específicos	3
1.4. Alcances del trabajo	4
1.5. Estructura del documento	4
2. Predicción de series de tiempo	5
2.1. Serie de tiempo	6
2.2. Modelos estadísticos de predicción	7
2.2.1. Modelos ARMA	7
2.2.1.1. Modelo auto-regresivo (AR)	7
2.2.1.2. Modelo de medias móviles (MA)	8
2.2.1.3. Modelo auto-regresivo de media móvil (ARMA)	9
2.2.2. El modelo GARCH	10
2.2.3. Modelos de Persistencia y de Referencia de Nielsen	12
2.2.4. Filtro de Kalman	13
2.2.4.1. Algoritmo del Filtro de Kalman	15
2.2.4.2. Filtro Extendido de Kalman	16

2.3.	Agrupamiento de datos	18
2.3.1.	Divergencia de Kullback-Leibler	18
2.3.1.1.	Distancia efectiva entre dos distribuciones:	19
2.3.2.	K-means	20
2.3.2.1.	Procedimiento para clusterización con K-means	20
3.	Metodología e implementación del modelo V-ARX	22
3.1.	Análisis de periodicidad diaria	23
3.2.	Análisis de estacionalidad	25
3.3.	Serie residual	27
3.4.	Modelo V-ARX	29
3.4.1.	Proceso de estimación de parámetros auto-regresivos	30
3.5.	Medidas de Desempeño	32
3.5.1.	Error Cuadrático Medio (MSE)	32
3.5.2.	Error Relativo Porcentual Medio (MRPE)	33
3.5.3.	Error Relativo Porcentual Efectivo Medio (MREPE)	33
3.5.4.	Error de Energía Porcentual Medio (MPPE)	34
3.6.	Comparación de modelos	34
4.	Resultados y Discusión	36
4.1.	Determinación de grupos estacionales	36
4.1.1.	Primer grupo: Semanas 1 a 13	38
4.1.2.	Segundo Grupo: Semanas 14 a 30	39
4.1.3.	Tercer Grupo: Semanas 31 a 52	41
4.2.	Resultados de predicción del Modelo V-ARX	43

4.2.1. Desempeño del modelo V-ARX	45
4.3. Resultados de predicción del modelo de Referencia de Nielsen	46
4.3.1. Desempeño del modelo de Referencia de Nielsen	49
4.4. Resultados de predicción del modelo ARMA-GARCH	49
4.4.1. Desempeño del modelo ARMA-GARCH	52
4.5. Comparación de modelos	52
4.5.1. Comparación con modelo de Referencia de Nielsen	52
4.5.2. Comparación con modelo ARMA-GARCH	52
5. Conclusiones y Trabajo Futuro	55
5.1. Conclusiones	55
5.2. Trabajo Futuro	56
Referencias	57

Índice de figuras

2.1.	Primer paso del algoritmo K-means.	20
2.2.	Segundo paso del algoritmo K-means.	21
2.3.	Tercero paso del algoritmo K-means.	21
2.4.	Cuarto paso del algoritmo K-means.	21
3.1.	Función de auto-correlación serie de viento completa.	24
3.2.	Origen de sub-series de 24 datos.	25
3.3.	Representación vectorial de perfiles diarios para magnitud de velocidad de viento.	25
3.4.	Representación vectorial de valores medios y varianza para perfiles de velocidad de viento de una semana.	26
3.5.	Origen de la serie de valores residuales.	28
3.6.	Función de auto-correlación serie de residuos completa.	28
3.7.	Forma de las matrices auto-regresivas.	30
4.1.	Agrupación gráfica obtenida con K-means	37
4.2.	Gráfico de velocidad media y desviación estándar Grupo 1	38
4.3.	Gráfico de función de auto-correlación serie de datos original Grupo 1	39
4.4.	Gráfico de función de auto-correlación serie de datos residuales Grupo 1	39
4.5.	Gráfico de velocidad media y desviación estándar Grupo 2	40

4.6. Gráfico de función de auto-correlación serie de datos original Grupo 2	40
4.7. Gráfico de función de auto-correlación serie de datos residuales Grupo 2 . . .	41
4.8. Gráfico de velocidad media y desviación estándar Grupo 3	42
4.9. Gráfico de función de auto-correlación serie de datos original Grupo 3	42
4.10. Gráfico de función de auto-correlación serie de datos residuales Grupo 3 . . .	43
4.11. Gráfico de predicción a un paso de velocidad de viento. Primera Semana de validación Grupo 1, utilizando el modelo V-ARX	44
4.12. Gráfico de predicción a un paso de velocidad de viento. Primera Semana de validación Grupo 2, utilizando el modelo V-ARX	44
4.13. Gráfico de predicción a un paso de velocidad de viento. Primera Semana de validación Grupo 3, utilizando el modelo V-ARX	45
4.14. Gráfico de predicción a 24 pasos de velocidad de viento. Primera semana de validación Grupo 1, utilizando el modelo de Referencia de Nielsen	47
4.15. Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 2, utilizando el modelo de Referencia de Nielsen	47
4.16. Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 3, utilizando el modelo de Referencia de Nielsen	48
4.17. Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 1, utilizando el modelo ARMA-GARCH	50
4.18. Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 2, utilizando el modelo ARMA-GARCH	50
4.19. Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 3, utilizando el modelo ARMA-GARCH	51

Índice de cuadros

2.1. Algoritmo del Filtro de Kalman	15
2.2. Algoritmo de Filtro Entendido de Kalman	17
4.1. Distribución de grupos estacionales mediante algoritmo K-means	37
4.2. Errores obtenidos con el modelo V-ARX en predicción de 24 horas	46
4.3. Errores obtenidos con el modelo de Referencia de Nielsen en predicción de 24 horas	49
4.4. Errores obtenidos con el modelo de ARMA-GARCH en predicción de 24 horas	52
4.5. Indices de mejora comparando V-ARX con modelo de Referencia de Nielsen	52
4.6. Improvement considerando MAPE como medidor de desempeño	53

Capítulo 1

Introducción

1.1. Antecedentes

Durante los últimos 20 años, la demanda de energía eléctrica en Chile se ha incrementado en forma sostenida, con una tasa promedio anual cercana al 6,7%. Dicha demanda se cubre en la actualidad principalmente a partir de fuentes de energía tradicionales como combustibles fósiles (63%) e hidroelectricidad (34%), correspondiendo sólo el 3% a Energías Renovables No Convencionales (ERNC)[1]. Esta situación genera una gran vulnerabilidad al sistema de suministro eléctrico, ya que se ve afectado por factores climáticos (sequías) o restricciones en el abastecimiento de gas natural [2].

Por otra parte y para fomentar la utilización de energías limpias, el 1° de abril de 2008 entró en vigencia la Ley N° 20.257, que estableció la obligación de las empresas de generación eléctrica de acreditar que la energía equivalente al 10% de sus retiros en cada año haya sido inyectada a partir de ERNC, exigiéndose este cambio de forma paulatina, considerando un 5% entre los años 2010 y 2014, incrementándose posteriormente cada año en un 0,5% hasta llegar al 10% el año 2024 [3].

Es por esto que la utilización de ERNC en la matriz energética nacional es cada vez más necesaria [4], razón por la cual, es imperioso contar con herramientas que permitan predecir su comportamiento con un cierto grado de confiabilidad. Esta necesidad se hace aun más importante en el ámbito de generación eólica, pues la aleatoriedad del viento hace muy difícil

predecir la energía disponible en un instante de tiempo determinado si se considera que el proceso de generación, a partir de éste recurso, es muy irregular. No así en las centrales térmicas en que el combustible necesario para operar es considerado como una variable de control; o en las centrales hidráulicas que aunque dependen de la cantidad de agua disponible, ésta no tiene variaciones muy drásticas en el corto plazo y consecuentemente, es posible determinar la capacidad de generación para horizontes de tiempo cercanos [5].

En este contexto, el uso de herramientas de predicción, al entregar un valor estimado de la generación de las centrales eólicas, facilita la integración de éstas a los sistemas interconectados y permite una eficiente programación tanto técnica como económica del sistema de despacho de carga, pudiendo el operador determinar con menos incertidumbre el nivel de generación y el nivel de reservas necesarias en las demás centrales. Asimismo, si se conocen los periodos de tiempo en los cuales la capacidad de generación es muy baja, es posible coordinar las operaciones de mantenimiento dentro de los parques eólicos, para que sean ejecutadas en esos instantes y no dejen de operar cuando las condiciones de generación son favorables.

Se han desarrollado algunos trabajos que buscan poder predecir el comportamiento de la velocidad de viento, y en particular, calcular la potencia que podría ser generada en un parque eólico a partir de este recurso. En este ámbito, el proyecto ANEMOS (*Development of A NExt generation wind resource forecasting system for large-scale integration of Onshore and offshore wind farms*), desarrollado en la Unión Europea, tiene por finalidad estudiar modelos de predicción utilizados, además de elaborar nuevas metodologías, existiendo una amplia gama de enfoques para enfrentar el problema, entre los que se pueden mencionar: numéricos, físicos, estadísticos y computacionales [6, 7]. Sin embargo la gran mayoría de los trabajos solo hace hincapié en las variables físicas que afectan el problema y los registros históricos de éstas, como la misma velocidad de viento, presión, temperatura, geografía, entre otros; dejando de lado componentes empíricas que si fuesen consideradas podrían aportar información adicional muy relevante a los modelos.

Una de las características que no es considerada de forma directa y que es intrínseca del comportamiento del viento, es la existencia de un patrón dominante en las magnitudes de

su velocidad en determinados periodos del año, no necesariamente marcados por las cuatro estaciones, y que permitiría poder realizar un ajuste más preciso de los modelos implementados.

Por otra parte, existe también una componente periódica diaria, que está ligada a cambios de temperatura o presión, y que se manifiesta en similares magnitudes de velocidad de viento para determinadas horas en días relativamente próximos; lo que es otra característica de interés que puede ser incluida en los modelos de predicción.

1.2. Objetivo General

El objetivo general de este trabajo es desarrollar e implementar un modelo de predicción de velocidad de viento a mediano plazo (un día), basado principalmente en el registro histórico de los valores de esta variable.

1.3. Objetivos Específicos

Los objetivos específicos del trabajo a desarrollar son:

- Diseñar e implementar un modelo de predicción de velocidad de viento basado en un modelo vectorial auto-regresivo con variable externa (V-ARX)
- El modelo debe ser capaz de realizar un pronóstico hacia un horizonte temporal de 24 horas.
- Deber ser capaz de capturar las propiedades periódicas diarias y estacionales de la velocidad de viento.
- Debe tener un desempeño mejor o similar a algunos modelos existentes en la literatura.

1.4. Alcances del trabajo

Esta memoria se enmarca en el fuerte interés mostrado por el Departamento de Ingeniería Eléctrica de la Universidad de Chile hacia el desarrollo de las Energías Renovables No Convencionales, específicamente, este trabajo se realiza en el Grupo de Sistemas de Información y Decisión, y tiene por finalidad apoyar la generación de energía eólica a través de un predictor de velocidad de viento. El cual podría ser utilizado también, en cualquier otro ámbito que requiera estimaciones a futuro de esta variable.

El aporte principal de este trabajo es, por una parte, la proposición de una novedosa manera de distinguir patrones estacionales basados en la información disponible, que permite definir un procedimiento genérico y bien fundamentado, para agrupar datos, evitando basarse en la arbitrariedad o el “gusto” de quien elabore los modelos. Por otra parte, la forma en que la información empírica puede ser incluida para complementar el análisis matemático, de tal forma que los modelos se ajusten de mejor manera a la realidad.

1.5. Estructura del documento

Esta Memoria de Título se organiza de la siguiente manera: En el Capítulo 2 se mencionan y describen algunos modelos que se utilizan actualmente en el ámbito de predicción. Luego, el Capítulo 3 detalla la metodología empleada para resolver el problema presentado, así como describe el modelo propuesto para cumplir con los objetivos trazados. Posteriormente, el Capítulo 4 muestra y compara los resultados obtenidos de los distintos modelos implementados. Finalmente, el Capítulo 5 presenta las conclusiones del trabajo y las posibles líneas de trabajo futuro.

Capítulo 2

Predicción de series de tiempo

Enfrentar el problema de predicción de series de tiempo es una tarea compleja, en especial para datos de velocidad y/o dirección debido a que poseen una gran aleatoriedad. Por esto se han elaborado numerosos trabajos que tratan de estimar este comportamiento utilizando diversas técnicas y enfoques. Entre estos enfoques se encuentran los del tipo fenomenológico, que describen el fenómeno mediante las leyes físicas que dominan los procesos, considerando variables como temperatura, presión, altura. También están los del tipo computacional, cuya principal característica es que son capaces de “aprender” las relaciones existentes entre los datos y de esta manera pueden estimar su comportamiento posterior. Y finalmente se encuentran los del tipo estadístico, que serán los utilizados en este trabajo, y se basan fundamentalmente en la utilización de los valores históricos de la serie de tiempo y en el reconocimiento de patrones característicos que permitan predecir los valores futuros de dicha serie.

En este capítulo se mostrarán algunos modelos del tipo estadístico que se utilizan actualmente para resolver el problema de predicción de velocidad de viento, entregando una descripción de cada uno, señalando sus fundamentos y características principales, además de algunos trabajos realizados en base a ellos.

Los métodos que serán descritos, usualmente trabajan sobre series completas de datos o sobre subconjuntos de éstas, determinados arbitrariamente, lo que no siempre entrega los mejores resultados. Por ello, en este trabajo se propone realizar una agrupación de los datos

disponibles, con fundamentos matemáticos, haciendo uso de herramientas como la divergencia de Kullback-Leibler y del algoritmo K-means para efecto de seleccionar distintos periodos del año en los cuales los perfiles de velocidad de viento horaria son similares. Estos métodos serán descritos y estudiados en detalle al final de este capítulo.

2.1. Serie de tiempo

Previo a describir algún tipo de modelo de predicción que trabaje sobre series de tiempo, es necesario dar una breve definición para dicho concepto.

Una serie de tiempo corresponde a un conjunto de observaciones, que pueden ser valores fijos o variables aleatorias, generadas o medidas secuencialmente en el tiempo a intervalos regulares, e indexadas de acuerdo al momento en que hayan sido registradas [8]. De esta forma, una serie temporal X está expresada como una sucesión de valores ordenados temporalmente.

$$X = (X_1, X_2, X_3, \dots). \quad (2.1)$$

Asimismo, cuando lo que se desea es realizar predicción sobre series de tiempo, se utiliza alguna función o modelo que capture la relación existente entre los mismos valores de la serie (X_i) y entre alguna otra posible variable explicativa asociada (u_i), permitiendo estimar el valor siguiente en la secuencia de valores. Matemáticamente, esta relación se expresa como:

$$X_{t+1} = f(X_1, X_2, X_3, \dots) \quad t \geq 0, \quad (2.2)$$

donde todas las variables hasta el instante t , corresponden a información conocida, y en adelante son estimaciones. Un ejemplo de serie de tiempo se extrae del caso particular de este trabajo, el cual usa una base de datos de mediciones de velocidad de viento, muestreadas a intervalos regulares de una hora.

2.2. Modelos estadísticos de predicción

2.2.1. Modelos ARMA

Un modelo ARMA [8] corresponde a uno de tipo auto-regresivo que, a grandes rasgos, se puede definir como una herramienta matemática que permite estimar $E(X_{t+1})$ de una serie de tiempo $\{X_t : t \geq 0\}$, basándose en el registro histórico de estos valores hasta el instante t . Estos modelos, también conocidos como modelos *Box & Jenkins*, están compuestos por dos partes, una parte auto-regresiva (AR) y una parte de media móvil (MA). Se le conoce más generalmente como ARMA(p,q) donde p y q son los órdenes de la parte auto-regresiva y de media móvil respectivamente.

Para una correcta aplicación de este tipo de modelos, la serie en análisis debe cumplir con la condición de ser estacionaria, esto es, que el valor de su media, varianza y covarianza no varían sistemáticamente en el tiempo.

$$E[X_t] = \mu_0 \quad (2.3)$$

$$Var(X_t) = \sigma_0^2 \quad (2.4)$$

$$Cov(X_t, X_{t+k}) = \gamma_k \quad \forall t, k. \quad (2.5)$$

A continuación, se describe cada componente de este modelo, así como el modelo ARMA en cuestión.

2.2.1.1. Modelo auto-regresivo (AR)

Un modelo auto-regresivo de orden p , cuya notación característica es $AR(p)$, se caracteriza principalmente por representar a la variable X_t , en función de los valores pasados de la serie de tiempo. Matemáticamente, este modelo puede escribirse como:

$$X_t = c + \sum_{i=1}^p \alpha_i X_{t-i} + \epsilon_t, \quad (2.6)$$

donde $\alpha_i, i = 1 \dots p$, son los coeficientes auto-regresivos del modelo, c es una constante que muchas veces se omite por simplicidad y ϵ_t es un término de error asociado al instante t . Otra

manera de representar este tipo de modelos es definiendo un operador de retardo B , el cual opera sobre los valores de la serie de tiempo X_t de la siguiente forma:

$$BX_t = X_{t-1} \quad (2.7)$$

$$B^k X_t = X_{t-k}. \quad (2.8)$$

De esta forma, también es posible representar un modelo auto-regresivo de orden p , en función del operador de retardo B , de la siguiente manera:

$$\left(1 - \sum_{i=1}^p \phi_i B^i\right) X_t = \epsilon_t \quad (2.9)$$

o

$$\phi_p(B)X_t = \epsilon_t. \quad (2.10)$$

en donde $\phi_p(B)$ se conoce como el operador polinomial de retardo de orden p .

2.2.1.2. Modelo de medias móviles (MA)

Adicionalmente al modelo AR , se tiene el modelo de media móvil de orden q , cuya notación es $MA(q)$, se caracteriza fundamentalmente por expresar el valor de la variable X_t en función de los errores que hayan existido en los instantes anteriores a t .

$$X_t = \epsilon_t + \sum_{j=1}^q \beta_j \epsilon_{t-j}, \quad (2.11)$$

donde $\beta_j, j = 1, \dots, q$, son los coeficientes del modelo, y $\epsilon_t, \dots, \epsilon_{t-q}$ corresponden a los términos de error asociados a los instantes $t, \dots, t-q$. Para este tipo de representación, también es posible utilizar el operador de retardo, quedando la ecuación que describe el modelo de la siguiente

manera:

$$X_t = \epsilon_t \left(1 + \sum_{j=1}^q \theta_j B^j \right) \quad (2.12)$$

o

$$X_t = \epsilon_t \theta_q(B). \quad (2.13)$$

2.2.1.3. Modelo auto-regresivo de media móvil (ARMA)

Este modelo, cuya notación es $ARMA(p, q)$, es la combinación de un modelo auto-regresivo con p coeficientes y un modelo de media móvil con q coeficientes. La variable X_t se escribe como función tanto de los valores pasados de la serie de tiempo como del valor de los errores que hayan existido, hecho que se ve reflejado en su ecuación representativa.

$$X_t = c + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \epsilon_{t-j} + \epsilon_t. \quad (2.14)$$

Comúnmente se asume que los términos de error ϵ_t son variables aleatorias independientes idénticamente distribuidas, que siguen una distribución normal de media cero y varianza σ^2 que por lo general se considera unitaria.

$$\epsilon_t \sim N(0, 1). \quad (2.15)$$

Dado que el modelo ARMA es una combinación de los modelos AR y MA , también es posible reescribir la ecuación en términos del operador de retardo, quedando el modelo representado por la ecuación:

$$\left(1 - \sum_{i=1}^p \phi_i B^i \right) X_t = \epsilon_t \left(1 + \sum_{j=1}^q \theta_j B^j \right). \quad (2.16)$$

Para los modelos mostrados, las condiciones que aseguran su estabilidad están relacionadas con el valor de las raíces de los polinomios de retardo $\phi_q(B)$ y $\theta_q(B)$, las cuales deben encontrarse dentro del círculo unitario.

Adicionalmente, existen algunas variaciones, o mejor dicho, generalizaciones de estos modelos que tratan de solucionar problemas como la no-estacionariedad de la una serie de tiempo, lo que en la práctica es bastante común. Así se encuentra al modelo ARIMA (Autoregressive Integrated Moving Average), el cual mediante una diferenciación de la serie busca conseguir la estacionariedad requerida.

Así también, se encuentra el modelo SARIMA, que a diferencia del anterior, éste se utiliza cuando en la serie de tiempo no estacionaria existe además un patrón periódico. Este tipo de modelo busca reconocer esa característica e incluirla en la ecuación que lo representa, sin embargo esta identificación solo permite capturar una de estas características mencionadas.

Las técnicas descritas han sido empleadas en algunos trabajos, por ejemplo: ARMA se utiliza en [9], donde se realiza predicción de velocidad de viento para un horizonte de 10 horas, elaborando un modelo para cada mes del año; Por otra parte, un trabajo que desarrolla un modelo SARIMA para cada semana del año, y que realiza predicción a un horizonte de 5 horas, puede verse en [10].

2.2.2. El modelo GARCH

El modelo GARCH o Generalized AutoRegressive Conditional Heteroskedasticity fue presentado por Bollerslev en 1986 [11], generalizando el modelo propuesto por Engel en 1982 [12], estos modelos tienen como finalidad modelar procesos en que la varianza condicional no es constante, a diferencia de un modelo ARMA en que asume dicho valor como invariante en el tiempo, y se encargan de capturar sus perturbaciones en cada instante. La técnica presentada por Bollerslev consiste en aplicar un proceso ARMA (1,1) a la varianza condicional de los errores, los cuales están modelados de la siguiente forma:

$$\epsilon_t = \tau_t \sqrt{\sigma_t^2}, \quad (2.17)$$

en donde τ_t es ruido blanco, esto es, un proceso independiente idénticamente distribuido (i.i.d.), con distribución de probabilidad normal o gaussiana, de media cero y varianza uni-

taria:

$$\tau_t \sim N(0, 1), \quad (2.18)$$

y σ_t^2 , para un proceso GARCH (1,1) tiene la siguiente forma:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (2.19)$$

Para garantizar la estabilidad del modelo se necesita que los parámetros α_0, α_1 y β cumplan con las siguientes condiciones:

$$\alpha_0 > 0 \quad (2.20)$$

$$0 < \alpha_1 < 1 \quad (2.21)$$

$$-1 < \beta < 1 \quad (2.22)$$

Ahora bien, si en lugar del modelo GARCH(1,1) se pretende utilizar un modelo más general, que permita variar la cantidad de coeficientes auto-regresivos para expresar la varianza condicional, es necesario reescribir las ecuaciones del método para que sea posible realizar este cambio. Las expresiones que dan cuenta de esta modificación y que permiten la obtención del modelo GARCH(p,q) se muestran a continuación:

$$\epsilon_t = \nu \sqrt{\sigma_t^2} \quad (2.23)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \quad (2.24)$$

o bien

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \dots + \alpha_q \epsilon_{t-q}^2 \quad (2.25)$$

$$+ \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2. \quad (2.26)$$

En este caso, las nuevas condiciones sobre los parámetros que garantizan la estabilidad

del sistema son:

$$\alpha_0 > 0 \tag{2.27}$$

$$\alpha_i, \beta_j \geq 0 \tag{2.28}$$

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1. \tag{2.29}$$

Continuando en el contexto de predicción de velocidad de viento, un trabajo que utiliza este método en conjunto con un modelo ARMA, dando origen a modelos ARMA-GARCH, puede verse en [13], en éste se realiza predicción de promedios diarios de velocidad de viento y se comparan 10 diferentes formas de expresar la varianza condicional de los modelos implementados. Por otra parte, un trabajo que utiliza el modelo GARCH con buenos resultados, pero en el ámbito financiero, intentando describir los retornos de un proceso y la volatilidad de éste, puede ser revisado en [14].

2.2.3. Modelos de Persistencia y de Referencia de Nielsen

Dentro de los modelos utilizados en predicción de velocidad de viento, uno de los más sencillos y que comúnmente es usado para compararlo con los nuevos algoritmos predictivos es el Modelo de Persistencia [15].

Este sencillo modelo de predicción determina que el valor del estado en un instante $t+k$, con k natural, es igual al valor que se tenga en el instante t . Si consideramos el estado como la velocidad del viento “V”, se tiene:

$$V(t+k|t) = V(t). \tag{2.30}$$

Pese a su simplicidad, este modelo resulta útil en predicciones a corto plazo dada la lentitud de los cambios atmosféricos. Una generalización de este modelo es considerar, en lugar del valor en el instante t , el promedio de los últimos n valores de la serie de tiempo hasta el instante t .

$$V(t+k|t) = \frac{1}{n} \sum_{j=t-n}^t V(j). \tag{2.31}$$

Este tipo de modelos se les conoce como predictores de media móvil y a medida que n tiende a infinito, se tiende a la media global.

$$V(t+k|t) = \bar{V}(t), \quad (2.32)$$

$$n \rightarrow \infty$$

$\bar{V}(t)$ es el promedio simple de todas las observaciones hasta el instante t . Este último modelo, que no es muy dinámico pues no permite una rápida respuesta ante un cambio brusco en la variable de interés, tiene un desempeño bastante pobre en horizontes de predicción muy cercanos, sin embargo en horizontes más lejanos resulta ser mejor que el modelo de persistencia. Estos dos modelos fueron estudiados por Nielsen [16] y posteriormente se propuso una mezcla de ambos predictores, dando como resultado el modelo de Referencia de Nielsen, cuya ecuación representativa se muestra en (2.33).

$$V(t+k|t) = \rho_k V(t) + (1 - \rho_k) \bar{V}_t \quad (2.33)$$

donde ρ_k corresponde al k -ésimo coeficiente de auto-correlación de la serie de observaciones.

2.2.4. Filtro de Kalman

El filtro de Kalman es una potente herramienta matemática creada en 1960 por Rudolph Emil Kalman [17] con el objeto de filtrar y predecir el comportamiento de sistemas lineales. Básicamente el filtro es un conjunto de ecuaciones, las cuales, mediante un algoritmo recursivo entregan un estimador óptimo del tipo predictor-corrector que minimiza el error cuadrático medio de la aproximación. Este algoritmo procesa todas las medidas disponibles para estimar el valor actual de las variables de interés. Para ello considera:

- El conocimiento previo del sistema y dispositivos de medida.
- La descripción estadística de los ruidos del sistema, errores de medida e incertidumbre en los modelos dinámicos.

- La posible información del estado inicial de las variables de interés.

El filtro de Kalman se enfoca en estimar un estado no observable $x_t \in \mathbb{R}^n$ utilizando las ecuaciones:

$$x_t = Ax_{t-1} + Bu_{t-1} + w_{t-1} \quad (2.34)$$

$$z_t = Hx_t + v_t. \quad (2.35)$$

Donde (2.34) es la ecuación de proceso o transición, que relaciona el estado del sistema en un instante $t-1$ con su estado en un instante posterior t ; y (2.35) es la ecuación de medida, con z_t vector de variables observables en el instante t . Las matrices A,B,H corresponden a:

- A: Matriz que relaciona el estado en el instante $t-1$ con el estado en el instante t en ausencia de señales de control.
- B: Matriz que relaciona las señales de control en el instante $t-1$ con el estado en el instante t .
- H: Matriz que relaciona el estado en el instante t con las observaciones en el mismo instante.

Las variables w, v corresponden a las perturbaciones del sistema y son ruido blanco de proceso y observación respectivamente. Se asumen independientes y con distribución normal.

$$p(w) \sim N(0, R) \quad (2.36)$$

$$p(v) \sim N(0, Q), \quad (2.37)$$

con R y Q matrices de covarianza de ruido de proceso y medida respectivamente. El algoritmo utilizado por este método se describe a continuación:

2.2.4.1. Algoritmo del Filtro de Kalman

Para encontrar la solución al problema de filtrado y predicción, el método de Kalman consta de 2 etapas, como se explica claramente en [18], la etapa de predicción y la etapa de corrección. De la primera se obtiene un estimador a priori del estado y la covarianza del error en el instante t , y en la segunda se “corrigen” estas estimaciones con el uso de la medida registrada en el instante t determinando el estimador a posteriori. Todos los cálculos se fundamentan en encontrar el estimador de estado que minimiza el error cuadrático medio, mediante el cálculo de la esperanza condicional del estimador condicionado a la información disponible. El algoritmo paso a paso, expresado en el caso que las distribuciones de probabilidad de las variables del problema sean Gaussianas, se muestra en el Cuadro 2.1.

<p style="text-align: center;">Algoritmo KF</p> <p style="text-align: center;">Etapas de Predicción</p> <p>1.- Proyección del estado hacia adelante</p> $\hat{x}_t^- = A\hat{x}_{t-1} + Bu_{t-1}$ <p>2.- Proyección de la covarianza del error</p> $P_t^- = AP_{t-1}A^T + Q$ <p style="text-align: center;">Etapas de Actualización</p> <p>3.- Cálculo de la ganancia de Kalman</p> $K_t = P_t^- H^T (HP_t^- H^T + R)^{-1}$ <p>4.- Corrección del estado con la medida z_t</p> $\hat{x}_t = \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-)$ <p>5.- Corrección de la covarianza del error</p> $P = (I - K_t H)P_t^-$
--

Cuadro 2.1: Algoritmo del Filtro de Kalman

El procedimiento descrito se utiliza cuando las ecuaciones involucradas (proceso y medida) son lineales, en caso contrario debe utilizarse una variación llamada Filtro Extendido de

Kalman.

2.2.4.2. Filtro Extendido de Kalman

El Filtro Extendido de Kalman, como se mencionó anteriormente, se utiliza cuando las ecuaciones involucradas son no-lineales, es decir el sistema es de la forma:

$$x_t = f(x_{t-1}, u_{t-1}, w_{t-1}) \quad (2.38)$$

$$z_t = h(x_t, v_t). \quad (2.39)$$

En este caso lo que hace el filtro en primer lugar, es una linealización de las ecuaciones, quedando el sistema expresado como:

$$x_t \approx \tilde{x}_t + A(x_{t-1} - \hat{x}_{t-1}) + Ww_{t-1} \quad (2.40)$$

$$z_t \approx \tilde{z}_t + H(x_t - \tilde{x}_t) + Vv_t, \quad (2.41)$$

donde $\tilde{x} = f(\hat{x}_{t-1}, u_{t-1}, 0)$ y $\tilde{z}_t = h(\tilde{x}_t, 0)$. Las matrices A,W,H,V se derivan de las funciones originales de la siguiente forma:

$$A = \frac{\delta f}{\delta x_t}(\hat{x}_{t-1}, u_{t-1}, 0) \quad (2.42)$$

$$W = \frac{\delta f}{\delta w_t}(\hat{x}_{t-1}, u_{t-1}, 0) \quad (2.43)$$

$$H = \frac{\delta h}{\delta x_t}(\tilde{x}_t, 0) \quad (2.44)$$

$$V = \frac{\delta h}{\delta v_t}(\tilde{x}_t, 0). \quad (2.45)$$

Luego de realizar estos ajustes, se aplica el algoritmo mostrado en el Cuadro 2.2 que es similar al anterior salvo unas pequeñas variaciones.

<p>Algoritmo KF</p> <p>Etapa de Predicción</p> <p>1.- Proyección del estado hacia adelante</p> $\hat{x}_t^- = f(\hat{x}_{t-1}, u_{t-1}, 0)$ <p>2.- Proyección de la covarianza del error</p> $P_t^- = AP_{t-1}A^T + WQW^T$ <p>Etapa de Actualización</p> <p>3.- Cálculo de la ganancia de Kalman</p> $K_t = P_t^- H^T (HP_t^- H^T + VRV^T)^{-1}$ <p>4.- Corrección del estado con la medida z_t</p> $\hat{x}_t = \hat{x}_t^- + K_t(z_t - h(\hat{x}_t^-, 0))$ <p>5.- Corrección de la covarianza del error</p> $P = (I - K_t H)P_t^-$
--

Cuadro 2.2: Algoritmo de Filtro Entendido de Kalman

El Filtro de Kalman ha sido utilizado exitosamente como predictor de velocidad de viento y potencia eólica, algunos trabajos lo implementan individualmente o lo complementan con otros procesos, y en otros es empleado como corrector de predicciones obtenidas de otros modelos. Por ejemplo [19] utiliza este filtro con un modelo ARMA como ecuación de proceso, el cual es perturbado para generar un conjunto de predicciones que son promediadas con el fin de entregar una estimación final; Otro trabajo que utiliza Filtro de Kalman, se encuentra en [20], en este caso se emplea como algoritmo de post-procesamiento, pues recibe estimaciones desde un modelo de predicción físico, y “corrige” los errores asociados; En [21] también se utiliza para corregir el error de las predicciones, pero en este caso se trata de estimación de temperatura, entregando buenos resultados al igual que los dos casos anteriores.

2.3. Agrupamiento de datos

Los modelos presentados anteriormente, al ser utilizados en predicción, usualmente trabajan sobre un conjunto predefinido de valores, ya sea sobre la serie completa de datos o sobre subconjuntos de esta (datos semanales, mensuales, anuales), estos grupos son determinados a priori por quien está realizando el análisis y sin tener un fundamento matemático sólido que respalde la elección de la cantidad de datos a utilizar. Esta situación trae consigo algunas desventajas, pues trabajar, por ejemplo con una serie muy extensa puede generar un ajuste de los parámetros que no sea óptimo, ya que existe una gran probabilidad de que dentro de los datos disponibles existan tendencias que varíen considerablemente a través del tiempo. Para resolver este problema, se hace necesario el uso de herramientas que permitan determinar diferencias en estos comportamientos, así como algoritmos capaces de agrupar los datos de acuerdo a sus similitudes.

En este contexto, un conjunto de herramientas que permiten realizar una agrupación de datos es el formado por la divergencia de *Kullback-Liebler* y por el algoritmo de *K-means*. La primera permite determinar la similitud entre dos variables, específicamente entregando la “diferencia” de sus distribuciones de probabilidad; y el segundo es un algoritmo de agrupamiento o clustering, que sirve para distribuir las variables de acuerdo al “parecido” obtenido a partir del cálculo de la divergencia. Este procedimiento de agrupación de datos se realiza de acuerdo a un criterio de cercanía, el cual se define en términos de una determinada función de distancia, como puede ser la euclidiana. A continuación se describen brevemente ambos métodos.

2.3.1. Divergencia de Kullback-Leibler

La Divergencia de *Kullback-Leibler* (KLD) [22, 23], también conocida como entropía relativa entre dos funciones de densidad de probabilidad $f(x)$ y $g(x)$, es comúnmente utilizada en estadística como medida de “distancia”, siendo aplicada en diversos ámbitos como: reconocimiento de voz e imágenes, clusterización de modelos y optimización. La ecuación que define

a la divergencia se expresa como:

$$D(f||g) = \int f(x) \log \left(\frac{f(x)}{g(x)} \right). \quad (2.46)$$

La KLD cumple con las siguientes propiedades:

- a. Auto-semejanza: $D(f||f) = 0$
- b. Auto-identificación: $D(f||g) = 0 \Leftrightarrow f = g$
- c. Positividad: $D(f||g) \geq 0 \forall f, g$

Particularmente, cuando se tiene que las distribuciones de probabilidad son funciones Gaussianas multivariadas, que llamaremos \hat{f} y \hat{g} , la KLD tiene una forma cerrada y conocida, definida por:

$$\hat{f} \sim N(\mu_{\hat{f}}, \Sigma_{\hat{f}}) \quad (2.47)$$

$$\hat{g} \sim N(\mu_{\hat{g}}, \Sigma_{\hat{g}}) \quad (2.48)$$

$$D(f||g) = \frac{1}{2} \left[\log \frac{\Sigma_{\hat{g}}}{\Sigma_{\hat{f}}} + Tr \left[\Sigma_{\hat{g}}^{-1} \Sigma_{\hat{f}} \right] - d + (\mu_{\hat{f}} - \mu_{\hat{g}})^T \Sigma_{\hat{g}}^{-1} (\mu_{\hat{f}} - \mu_{\hat{g}}) \right] \quad (2.49)$$

En donde $\mu_{\hat{f}}$ y $\mu_{\hat{g}}$ corresponden a los vectores de medias de las respectivas distribuciones, $\Sigma_{\hat{f}}$ y $\Sigma_{\hat{g}}$ corresponden a las matrices de covarianza y d corresponde a la dimensión de los vectores.

2.3.1.1. Distancia efectiva entre dos distribuciones:

Dado que la KLD no es simétrica, se define una “distancia efectiva” dada por:

$$d(f, g) = D(f||g) + D(g||f) - D(f||f) - D(g||g) \quad (2.50)$$

La cual satisface las condiciones de simetría y de distancia nula entre distribuciones iguales.

$$d(f, g) = d(g, f) \quad (2.51)$$

$$d(f, f) = 0. \quad (2.52)$$

2.3.2. K-means

Este algoritmo, propuesto por Mac Queen en 1967 [24], es uno de los más simples y conocidos algoritmos de agrupamiento, pues se basa en una forma muy fácil para dividir una base de datos dada, en una cantidad de grupos predeterminada (k). La idea principal de este método es definir k *centroides* (uno para cada grupo), en los cuales se distribuyen los datos, siendo éste un proceso iterativo que finaliza cuando se obtiene alguna condición pre-establecida. El procedimiento, paso a paso, para agrupar los elementos se detalla a continuación:

2.3.2.1. Procedimiento para clusterización con K-means

Para ejemplificar el algoritmo se utilizan tres grupos($k=3$), cuyos centroides son representados por círculos.

- a. Inicialmente se seleccionan K objetos (datos) del conjunto de entrada. Estos k objetos serán los centroides iniciales de los k -grupos.

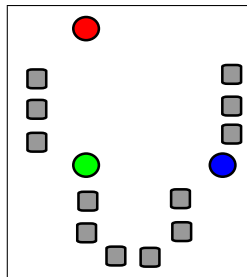


Figura 2.1: Primer paso del algoritmo K-means.

- b. Se calculan las distancias de todos los objetos a cada uno de los centroides, y se agrupan con aquel que presente la menor distancia.

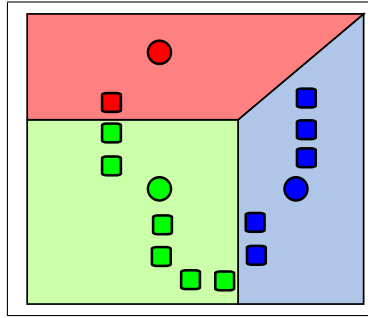


Figura 2.2: Segundo paso del algoritmo K-means.

- c.* Se actualizan los centroides como el valor medio de todos los datos asignados a ese grupo.

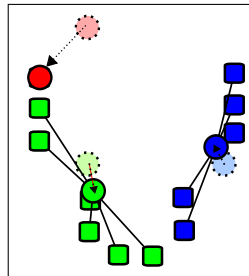


Figura 2.3: Tercero paso del algoritmo K-means.

- d.* Se repite el proceso hasta que se satisface algún criterio de convergencia, por ejemplo, algún porcentaje de variación entre el cálculo de un centroide con el valor obtenido de su iteración anterior.

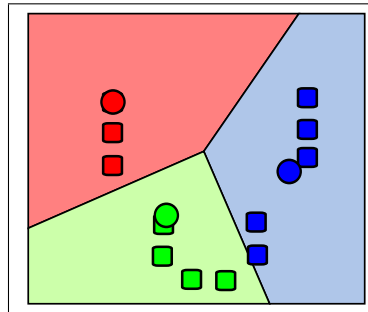


Figura 2.4: Cuarto paso del algoritmo K-means.

Capítulo 3

Metodología e implementación del modelo V-ARX

En este capítulo se detalla la metodología e implementación de un modelo vectorial de predicción de velocidad de viento para un horizonte temporal equivalente a un día, que considere las características de periodicidad y estacionalidad que posee el viento y que fueron mencionadas en el Capítulo 1.

Los datos utilizados para llevar a cabo el trabajo fueron obtenidos del proyecto EOLO [25], los cuales corresponden a una serie de mediciones de velocidad de viento durante el año 1990 y 1991, en la localidad de Punta Lengua de Vaca, región de Coquimbo. Estos datos, conforman una serie de tiempo de 8760 muestras consecutivas, con una tasa de muestreo de 1 hora, y están tomados a una altura de 10 m.s.n.m. (metros sobre el nivel del mar).

El trabajo desarrollado se basa principalmente en la creación e implementación de un modelo auto-regresivo vectorial con variable externa, que llamaremos V-ARX (*Vector Auto-Regressive with eXternal input*). El cual tiene la característica de no utilizar directamente la serie de valores observados, si no que usa una transformación de éstos que llamamos serie residual, lo que permite eliminar las características periódicas y obtener propiedades de estacionariedad que se requieren en los modelos auto-regresivos.

Con el fin de cuantificar el impacto que esta propuesta tiene desde el punto de vista del error de predicción, se implementan dos modelos de comparación: el modelo de Referencia

de Nielsen [16] y un modelo ARMA-GARCH [13], ambos utilizados con buenos resultados en el ámbito de predicción de series de tiempo. Finalmente se utilizan medidores de desempeño para los modelos implementados y un indicador de comparación entre ellos, para determinar que tan exactas son las predicciones realizadas y si mejoran con respecto a los algoritmos ya conocidos.

A continuación se explica cada etapa del proceso que da origen al modelo V-ARX, comenzando por el análisis de componentes periódicas, que son el fundamento del modelo vectorial. Se continúa con el análisis de comportamientos estacionales presentes en la serie de datos. Luego se muestra la manera en que se obtiene la serie de valores residuales a partir de los datos reales. Se sigue con la estructura del modelo V-ARX y la descripción de cada uno de sus componentes, para finalmente mostrar los indicadores de desempeño y comparación.

3.1. Análisis de periodicidad diaria

Con la finalidad de validar matemáticamente la existencia de una componente de periodicidad en los perfiles diarios de velocidad de viento, que será uno de los fundamentos del modelo presentado, se calcula la función de auto-correlación de la serie de datos. De esta manera es posible comprobar si efectivamente existe alguna relación entre las horas del día y la magnitud de la velocidad de viento en dicha hora. El gráfico de esta función se muestra a continuación en la Figura 3.1.

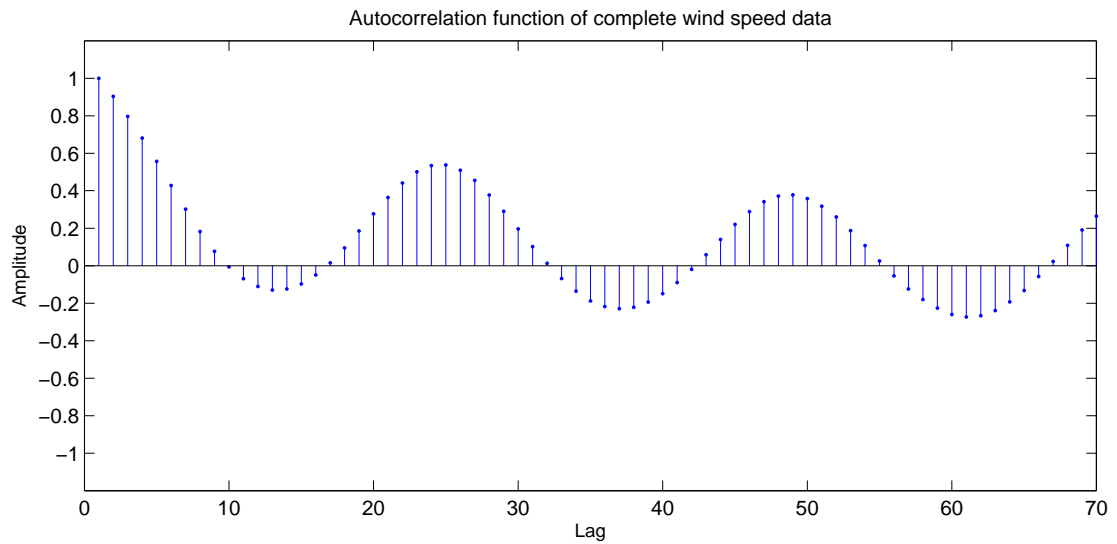


Figura 3.1: Función de auto-correlación serie de viento completa.

En el gráfico mostrado se puede ver claramente que existe una fuerte correlación entre los valores de la serie de datos, la cual aumenta cada 24 muestras aproximadamente, lo que da cuenta de la presencia de una componente de periodicidad con esa duración de tiempo y que ratifica que la velocidad de viento tiene un comportamiento similar cada 24 horas.

Teniendo esto claro, se realiza una transformación de la serie de datos original, de modo que en lugar de ser una serie correlativa de valores escalares, será una serie de vectores diarios, cada uno con 24 componentes, correspondientes a cada hora del día, comenzando desde las 1 AM hasta las 24 PM. Esta transformación da origen a 365 vectores de 24 componentes, uno para cada día del año. La siguiente imagen ejemplifica la división de la serie original en sub-series de 24 valores cada una, creando una para cada día del año.

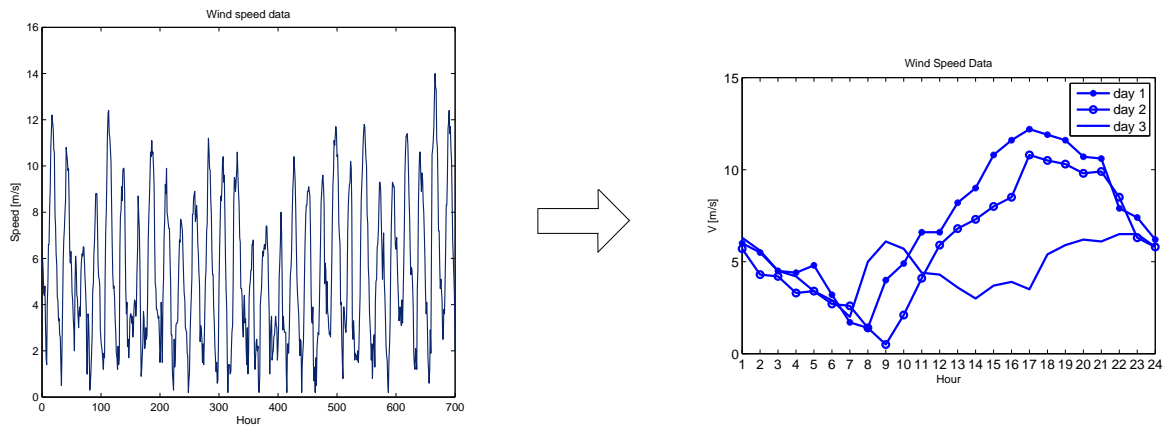


Figura 3.2: Origen de sub-series de 24 datos.

Finalmente, cada una de estas sub-series se convierte en un vector de velocidades diarias, en que cada componente corresponde al valor de velocidad de viento en una hora específica.

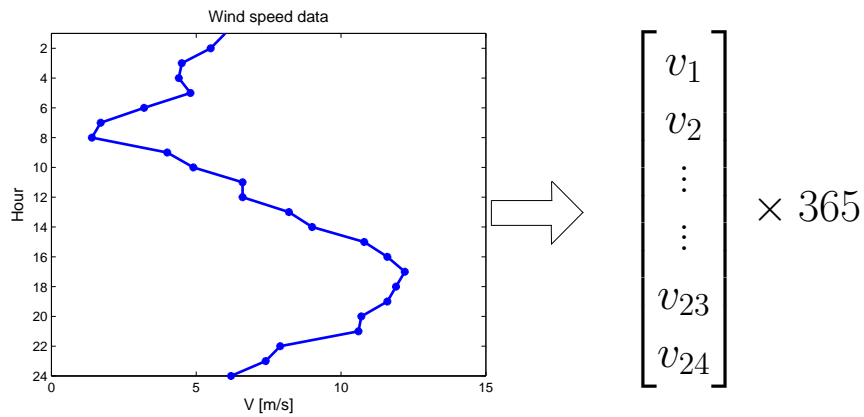


Figura 3.3: Representación vectorial de perfiles diarios para magnitud de velocidad de viento.

3.2. Análisis de estacionalidad

Para verificar las propiedades estacionales de la magnitud de velocidad de viento se busca la manera de agrupar los vectores en determinados intervalos, cada uno con un comportamiento similar considerando la media y la varianza de los valores en los vectores. Para esto,

inicial y arbitrariamente se agrupan los vectores diarios en grupos semanales (7 vectores) y se les calcula su vector de media y su vector de varianza. Esto se realiza para todas las semanas del año. La Figura 3.4 muestra esta acción, en la cual se ven representadas las series de datos correspondientes a la velocidad de viento de los 7 días de una semana, y la determinación de la media y desviación estándar de los valores horarios.

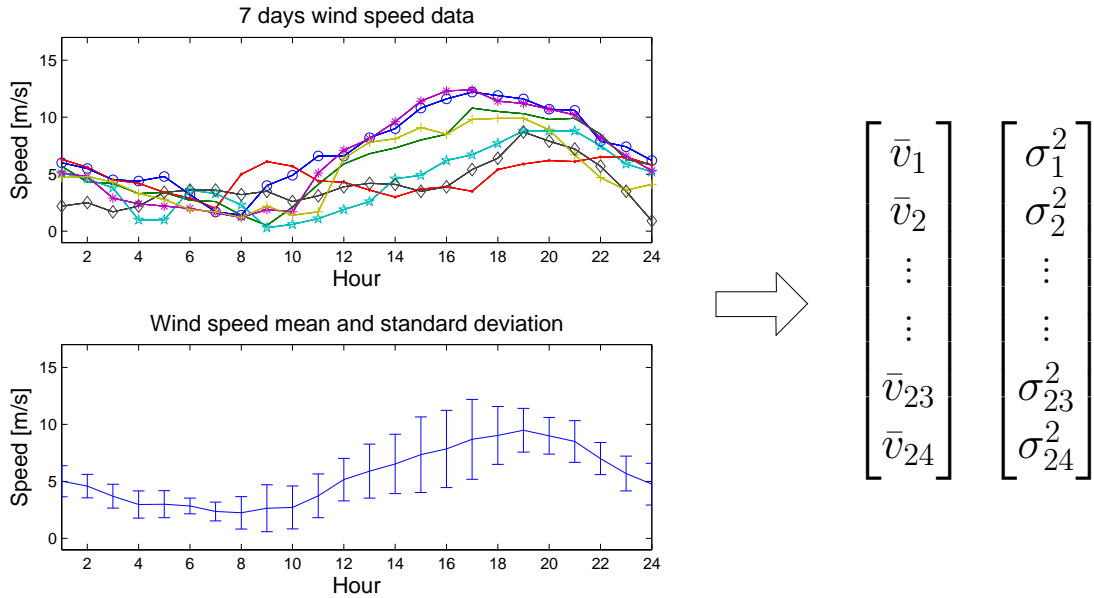


Figura 3.4: Representación vectorial de valores medios y varianza para perfiles de velocidad de viento de una semana.

Posteriormente, mediante la utilización de la divergencia de *Kullback-Liebler*[22, 23] y considerando ambos vectores de media y varianza semanal, es posible determinar la similitud entre las distintas semanas del año de acuerdo a sus distribuciones estadísticas, generando una matriz de distancias “D”, de dimensiones 52×52 , en la cual cada componente d_{ij} indica la “diferencia” entre una semana y otra. Es importante mencionar que para este cálculo se asume que los vectores obtenidos están compuestos por variables aleatorias independientes idénticamente distribuidas con distribución de probabilidad Gaussiana.

Finalmente, y para llevar a cabo la agrupación, se ocupa el algoritmo K-means [24] sobre la matriz de “distancias” obtenida con la divergencia de *Kullback-Liebler*, con el que se obtiene

el agrupamiento o *clustering* de las distintas semanas de acuerdo a cuan “parecidas” sean sus distribuciones, obteniendo los grupos de datos con los que se desarrollará el modelo.

Una observación importante para esta etapa del análisis es que, aparte de utilizar la divergencia de *Kullback-Liebler* y *K-means* como base para la agrupación, existe una componente empírica al momento de tomar la decisión de formar los grupos. Basándose principalmente en la continuidad temporal que se necesita para poder originar los modelos, de esta manera es que se excluyen o incluyen algunas semanas a los distintos grupos originados a partir del algoritmo *K-means*. Cabe mencionar que para este trabajo se obtiene una segmentación de los datos en tres grupos, y en cada uno de ellos se utilizará un tercio de la información para el entrenamiento o cálculo de los parámetros, y los dos tercios restantes serán utilizados para validación de los resultados.

3.3. Serie residual

Para originar esta serie de datos, es necesario obtener en primer lugar los *vectores residuales*, que corresponden a los valores obtenidos a partir de la transformación de cada uno de los vectores que contienen las mediciones de magnitud de velocidad de viento. Éstos nuevos vectores se calculan restando a cada vector de valores diarios el vector de medias horarias correspondiente al grupo al que pertenece. Luego la concatenación de estos *vectores residuales* da origen a la *serie residual* de velocidad de viento.

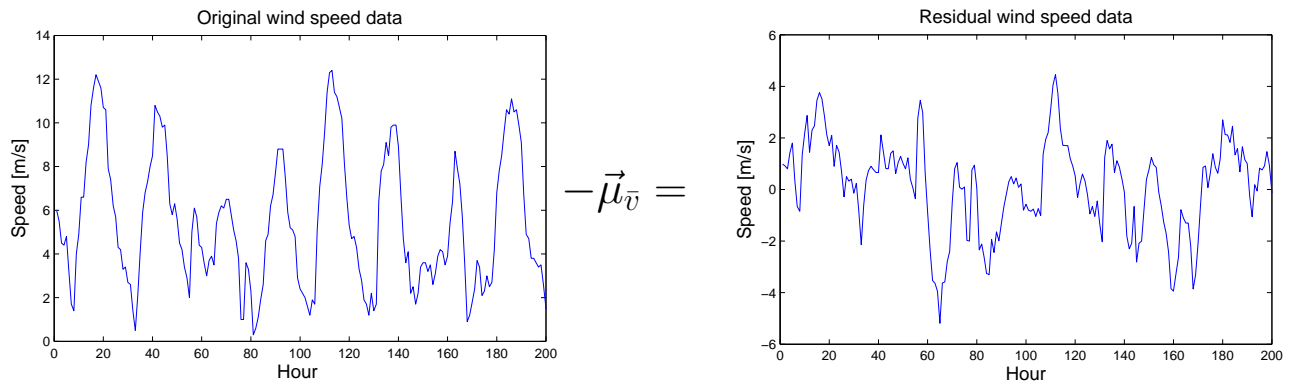


Figura 3.5: Origen de la serie de valores residuales.

Esta nueva serie tiene la característica de poseer media cero y además, reducir o incluso eliminar la componente periódica que se encuentra en la serie original de datos, tal como se puede ver en su función de auto-correlación.

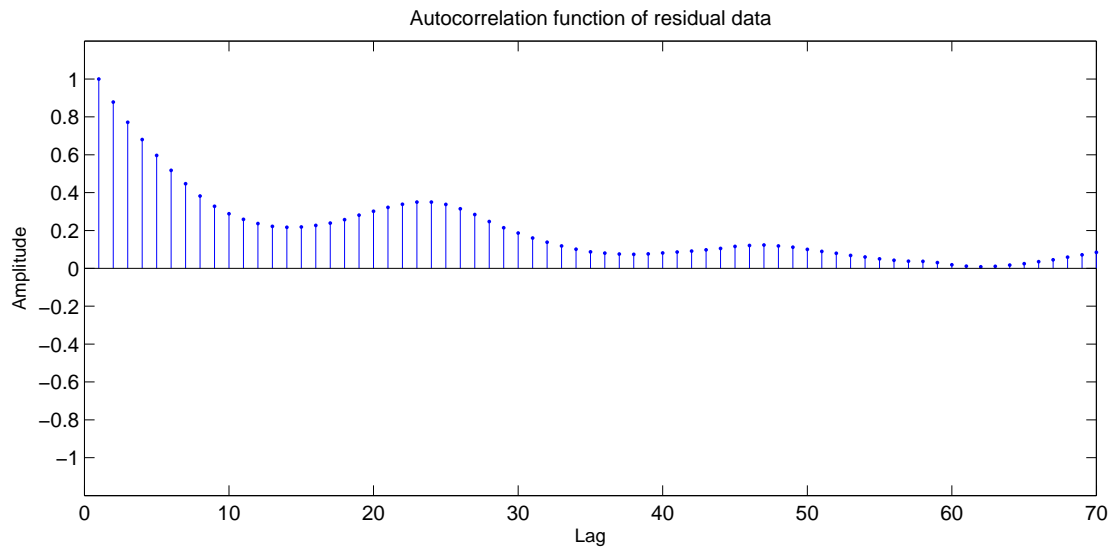


Figura 3.6: Función de auto-correlación serie de residuos completa.

3.4. Modelo V-ARX

Como ya se mencionó, el modelo de predicción que se propone es un modelo vectorial auto-regresivo con variable externa. Este utilizará la información conocida hasta el día n , y dará origen a las predicciones para el día $n+1$. La ecuación que define este predictor se muestra a continuación:

$$\vec{V}(n+1) = A_1\vec{V}_r(n) + A_2\vec{V}_r(n-1) + A_3\vec{V}_r(n-2) + \vec{V}_{ext}, \quad (3.1)$$

donde sus componentes se definen como:

- **Vector de predicción \vec{V}**

El vector \vec{V} , es aquel que entregará el valor de las predicciones para el día $n+1$, es un vector columna de 24 componentes en donde cada fila corresponde a una hora del día, comenzando desde las 1 AM hasta las 12 PM.

- **Vector de residuo \vec{V}_r**

El vector \vec{V}_r , es el vector de valores residuales de los días n , $n-1$, $n-2$, y al igual que el vector de predicción es un vector columna de 24 componentes en donde cada fila corresponde a una hora del día, comenzando desde las 1 AM hasta las 12 PM.

- **Variable exógena \vec{V}_{ext}**

La variable \vec{V}_{ext} , que corresponde a la variable externa del modelo, se obtiene a partir de los grupos determinados en la etapa de clusterización, y corresponde al vector de medias horarias o perfil promedio del grupo en que se esté trabajando. Dada la forma en que se obtiene, éste es un vector columna de 24 componentes.

- **Matrices auto-regresivas A_i**

Estas 3 matrices, A_1 , A_2 , A_3 , son las matrices auto-regresivas del modelo y operan directamente sobre los vectores residuales. Se construyen como matrices triangulares superiores, esto se debe a que a priori se sabe que la primera fila de esta matriz dará origen a la primera predicción, y de la función de auto-correlación de la serie de residuos,

sabemos que el primer valor estimado es el que podrá capturar la mayor cantidad de información conocida, pues está más correlacionado con los valores cercanos, en cambio el más lejano temporalmente dispondrá de la menor cantidad de información. Dicho esto, es que la primera fila de la matriz tiene 24 coeficientes, para utilizar la mayor cantidad de información contenida en el vector de residuos, la segunda fila de la matriz tiene 23 coeficientes, y así sucesivamente hasta la última fila que posee sólo un coeficiente. Dada esta característica, las dimensiones de esta matriz son 24×24 .

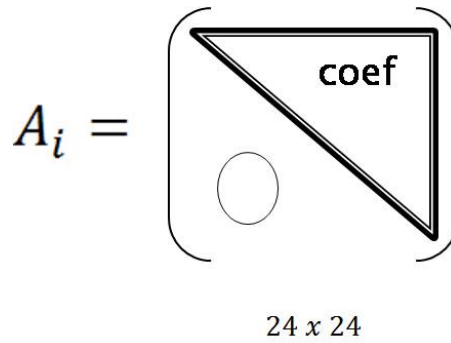


Figura 3.7: Forma de las matrices auto-regresivas.

Los coeficientes de estas matrices, deben ser calculados para cada fila por separado y son estimados mediante las ecuaciones de Yule-Walker [23], las cuales trabajan con los valores de la auto-covarianza de la serie que se quiere estimar, que en este caso es la serie residual. A modo de ejemplo, se detalla el procedimiento para calcular los coeficientes de la primera fila de la matriz, lo que dará origen a la primera predicción.

3.4.1. Proceso de estimación de parámetros auto-regresivos

- a. En primer lugar se debe escribir el valor a estimar en función de los coeficientes del modelo, aún no determinados, y de las observaciones ya registradas que correspondan.

$$\begin{aligned}
 V_{r,1}^{n+1} &= \alpha_{24}^1 V_{r,1}^n + \alpha_{23}^1 V_{r,2}^n + \alpha_{22}^1 V_{r,3}^n + \dots + \alpha_1^1 V_{r,24}^n \\
 &+ \alpha_{24}^2 V_{r,1}^{n-1} + \alpha_{23}^2 V_{r,2}^{n-1} + \alpha_{22}^2 V_{r,3}^{n-1} + \dots + \alpha_1^2 V_{r,24}^{n-1} \\
 &+ \alpha_{24}^3 V_{r,1}^{n-2} + \alpha_{23}^3 V_{r,2}^{n-2} + \alpha_{22}^3 V_{r,3}^{n-2} + \dots + \alpha_1^3 V_{r,24}^{n-2}.
 \end{aligned} \tag{3.2}$$

En esta notación, la variable $V_{r,t}^{n+1}$ corresponde al valor de la predicción de velocidad de viento para la serie de residuos en la hora t del día $n+1$; el coeficiente α_j^k corresponde al j -ésimo coeficiente de la matriz auto-regresiva A_k , considerado de derecha a izquierda; la variable $V_{r,t}^{n-k}$; $k = 0, 1, 2$ corresponde al valor de la serie residual en el día n , $n-1$, $n-2$ para la hora t . Para cualquier otra hora que se quiera estimar, la expresión que permite obtener el valor está dada por:

$$V_{r,t}^{n+1} = \sum_{i=t}^{24} \alpha_{25-i}^1 V_{r,t}^n + \sum_{i=t}^{24} \alpha_{25-i}^2 V_{r,t}^{n-1} + \sum_{i=t}^{24} \alpha_{25-i}^3 V_{r,t}^{n-2}. \quad (3.3)$$

- b. Luego, se multiplica a ambos lados de la ecuación por cada uno de los valores reales que se utilizarán, dando origen a un sistema de ecuaciones, con tantas ecuaciones como parámetros deban ser estimados.

$$\begin{aligned} V_{r,1}^{n+1} V_{r,1}^n &= \alpha_{24}^1 V_{r,1}^n V_{r,1}^n + \alpha_{23}^1 V_{r,2}^n V_{r,1}^n + \dots + \alpha_1^1 V_{r,24}^n V_{r,1}^n \\ V_{r,1}^{n+1} V_{r,2}^n &= \alpha_{24}^2 V_{r,1}^{n-1} V_{r,2}^n + \alpha_{23}^2 V_{r,2}^{n-1} V_{r,2}^n + \dots + \alpha_1^2 V_{r,24}^{n-1} V_{r,2}^n \\ &\vdots \\ V_{r,1}^{n+1} V_{r,24}^{n-2} &= \alpha_{24}^3 V_{r,1}^{n-2} V_{r,24}^{n-2} + \alpha_{23}^3 V_{r,2}^{n-2} V_{r,24}^{n-2} + \dots + \alpha_1^3 V_{r,24}^{n-2} V_{r,24}^{n-2}. \end{aligned} \quad (3.4)$$

- c. Una vez determinado el sistema, se calcula la esperanza de cada término de éste, con lo cual aparecen los valores de covarianza de la serie utilizada.

$$\begin{aligned} \gamma_1 &= \alpha_{24}^1 \gamma_0 + \alpha_{23}^1 \gamma_1 + \alpha_{22}^1 \gamma_2 + \dots + \alpha_1^3 \gamma_{71} \\ \gamma_2 &= \alpha_{24}^1 \gamma_1 + \alpha_{23}^1 \gamma_0 + \alpha_{22}^1 \gamma_1 + \dots + \alpha_1^3 \gamma_{70} \\ &\vdots \\ \gamma_{72} &= \alpha_{24}^1 \gamma_{71} + \alpha_{23}^1 \gamma_{70} + \alpha_{22}^1 \gamma_{69} + \dots + \alpha_1^3 \gamma_0. \end{aligned} \quad (3.5)$$

d. Finalmente se despeja el vector de coeficientes buscado.

$$\begin{bmatrix} \alpha_{24}^1 \\ \alpha_{23}^1 \\ \alpha_{22}^1 \\ \vdots \\ \alpha_1^3 \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{71} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{70} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{69} \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ \gamma_{71} & \gamma_{70} & \gamma_{69} & \cdots & \gamma_0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \vdots \\ \gamma_{72} \end{bmatrix} \quad (3.6)$$

Con este procedimiento se calculan todas las filas de la matriz, teniendo en consideración que cada vez que la estimación se aleje en una hora, se utilizará un coeficiente menos de auto-regresión.

3.5. Medidas de Desempeño

Para una serie de tiempo el error de predicción se define como la diferencia entre el valor real de dicha serie en un instante $t+k$, y el valor estimado por el modelo para ese mismo instante $t+k$. Por lo tanto el error de predicción para un determinado instante futuro se define como:

$$e(t+k|t) = V(t+k|t) - \hat{V}(t+k|t). \quad (3.7)$$

A diferencia del error de predicción que se utiliza para un instante en particular, existen medidas de desempeño que permiten un análisis cuantitativo del modelo de predicción para todas las estimaciones obtenidas [21]. En este trabajo se utilizan los siguientes:

3.5.1. Error Cuadrático Medio (MSE)

Para un horizonte de predicción de k pasos, el error cuadrático medio (Mean Square Error) se define como:

$$\begin{aligned} MSE(k) &= \frac{1}{k} \sum_{i=1}^k e(t+k|t)^2 \\ &= \frac{1}{k} \sum_{i=1}^k V(t+k|t) - \hat{V}(t+k|t). \end{aligned} \quad (3.8)$$

Una observación sobre este tipo de indicador, es que al utilizar el cuadrado de los errores, intrínsecamente se le está dando más “peso.” a los errores más grandes al momento de promediarlos.

3.5.2. Error Relativo Porcentual Medio (MRPE)

Para un horizonte de predicción de k pasos, el error relativo porcentual (Mean Relative Percentage Error), se define como:

$$\begin{aligned} MRPE(k) &= \frac{1}{k} \sum_{i=1}^k \frac{|e(t+k|t)|}{|V(t+k|t)|} * 100 \\ &= \frac{1}{k} \sum_{i=1}^k \frac{|V(t+k|t) - \hat{V}(t+k|t)|}{|V(t+k|t)|} * 100. \end{aligned} \quad (3.9)$$

Esta medida se consigue comparando el error de predicción y el valor de la señal medida para todos los instantes entre $t+1$ y $t+k$, para finalmente calcular su promedio. Como observación de este indicador se tiene que los errores existentes en los instantes de tiempo donde el valor real es pequeño, serán promediados con un “peso” mayor a los errores donde el dato real tiene una magnitud mayor.

3.5.3. Error Relativo Porcentual Efectivo Medio (MREPE)

Para un horizonte de predicción de k pasos, el error relativo porcentual efectivo (Mean Relative Effective Percentage Error) se define como:

$$\begin{aligned} MREPE(k) &= \frac{1}{k} \sum_{i=1}^k \frac{|e(t+k|t)|}{|\bar{V}_{1,k}|} * 100 \\ &= \frac{1}{k} \sum_{i=1}^k \frac{|V(t+k|t) - \hat{V}(t+k|t)|}{\left| \frac{1}{k} \sum_{i=1}^k V(t+k|t) \right|} * 100. \end{aligned} \quad (3.10)$$

Este error se consigue comparando el error de predicción y el valor del promedio de la señal medida entre t y $t+k$, para todos los instantes entre $t+1$ y $t+k$, para finalmente calcular su

promedio. En este caso, la utilización del valor promedio de los datos reales para normalizar los porcentajes calculados, intenta que éstos sean ponderados de forma equivalente y así, reducir el efecto de los “pesos” presente en el error absoluto porcentual.

3.5.4. Error de Energía Porcentual Medio (MP EE)

Para un horizonte de predicción de k pasos, el error de energía porcentual (Mean Percentage Energy Error) se define como:

$$\begin{aligned} MP EE(k) &= \frac{1}{k} \sum_{i=1}^k \frac{e(t+k|t)^2}{V(t+k|t)^2} * 100 \\ &= \frac{1}{k} \sum_{i=1}^k \frac{(V(t+k|t) - \hat{V}(t+k|t))^2}{V(t+k|t)^2} * 100. \end{aligned} \quad (3.11)$$

Este error se consigue comparando la energía de la señal de error de predicción y la energía de la señal medida para todos los instantes entre $t+1$ y $t+k$, para finalmente calcular su promedio.

3.6. Comparación de modelos

Finalmente, para comparar los resultados obtenidos con cada modelo y determinar si existe algún grado de mejora en las predicciones realizadas se implementa el siguiente indicador de comparación, que llamaremos IMP (del ingles *improvement*, “mejora”) y que nos indica porcentualmente cuanto se mejora la predicción obtenida de un determinado modelo en relación a un modelo de referencia:

$$IMP_{e_{ref},e}(k) = \frac{e_{ref}(k) - e(k)}{e_{ref}} * 100. \quad (3.12)$$

En donde $e_{ref}(k)$ corresponde al error obtenido del modelo de referencia para un horizonte de predicción de k pasos y $e(k)$ corresponde al error obtenido del modelo que se quiere estudiar. En palabras simples, si el valor del IMP resulta un valor positivo, significa que el

modelo evaluado tiene un mejor desempeño que el modelo de referencia y por el contrario, si resultase negativo, indica que el modelo de referencia es más preciso que el modelo estudiado. Cabe mencionar que el error considerado puede ser cualquiera de los mostrados anteriormente.

Capítulo 4

Resultados y Discusión

En este capítulo se dan a conocer los resultados obtenidos con el modelo V-ARX propuesto, así como los obtenidos mediante la implementación del modelo de Referencia de Nielsen y de un modelo ARMA-GARCH, como modelos de comparación. Además, se indican específicamente los distintos grupos estacionales que son determinados mediante clusterización, para los cuales se decide crear un modelo de predicción representativo de cada caso en particular. Se recalculan también las funciones de auto-correlación de los datos originales y de los datos residuales, para verificar que las condiciones de periodicidad de las series, que son el fundamento del modelo, siguen presentes al momento de hacer la agrupación y posterior implementación de los algoritmos.

4.1. Determinación de grupos estacionales

Como se mencionó en el capítulo anterior, esta etapa del trabajo se realiza para agrupar las distintas semanas del año en grupos que presenten un comportamiento similar, o mejor dicho, un patrón de velocidades diarias similar. Para esto es necesario calcular previamente la matriz de distancias mediante la divergencia de *Kullback-Leibler*, para luego aplicar sobre ésta el algoritmo *K-means*, el cual agrupa las semanas independientemente de su ubicación temporal durante el año, es por esto que se utiliza un criterio empírico para que estos grupos de datos sean formados por datos que se encuentren correlativos temporalmente. La agrupación obtenida una vez realizada la clusterización con *K-means* entrega el siguiente resultado:

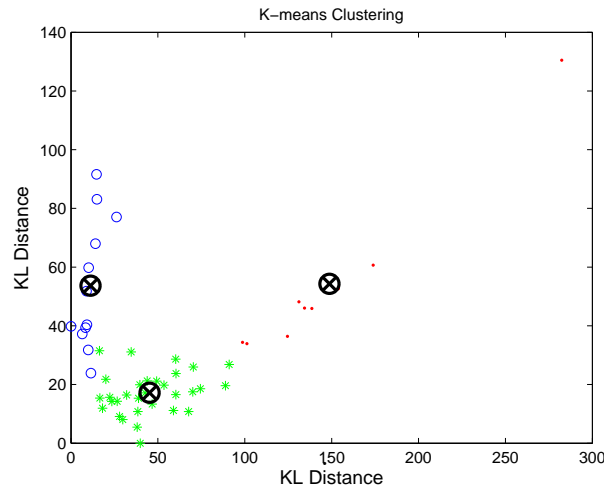


Figura 4.1: Agrupación gráfica obtenida con K-means

En la figura anterior donde cada punto representa una semana del año, se aprecian los tres grupos determinados para dividir los datos, pues la elección de una cantidad mayor da origen a segmentos con muy pocas semanas, lo que se contradice con la finalidad principal de realizar la agrupación, la cual es minimizar la cantidad de modelos necesarios.

Para entender de mejor manera la distribución obtenida, se muestra en detalle la agrupación realizada, indicando cada semana y el grupo al que está asociada. El Grupo A corresponde a los asteriscos, el Grupo B corresponde a los círculos y el Grupo C a los puntos.

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13
Grupo	A	A	A	A	A	A	B	A	B	B	C	B	B
Semana	14	15	16	17	18	19	20	21	22	23	24	25	26
Grupo	C	B	B	C	B	C	C	C	B	B	C	B	C
Semana	27	28	29	30	31	32	33	34	35	36	37	38	39
Grupo	C	B	B	C	B	B	B	B	C	B	B	B	B
Semana	40	41	42	43	44	45	46	47	48	49	50	51	52
Grupo	B	B	A	A	B	A	B	A	B	B	A	B	B

Cuadro 4.1: Distribución de grupos estacionales mediante algoritmo K-means

Adicionalmente, debido a que se requiere que estos grupos estén formados por semanas correlativas, se utiliza el resultado obtenido con *K-means* como estructura base de la agru-

pación y con ayuda de un análisis empírico de los datos se redefinen los grupos, los cuales se detallan a continuación:

4.1.1. Primer grupo: Semanas 1 a 13

El primer grupo está formado por el intervalo de semanas desde la 1 a la 13 (1 enero - 1 Abril). El patrón característico de media y desviación estándar, de la velocidad del viento en este periodo, puede apreciarse en la Figura 4.2.

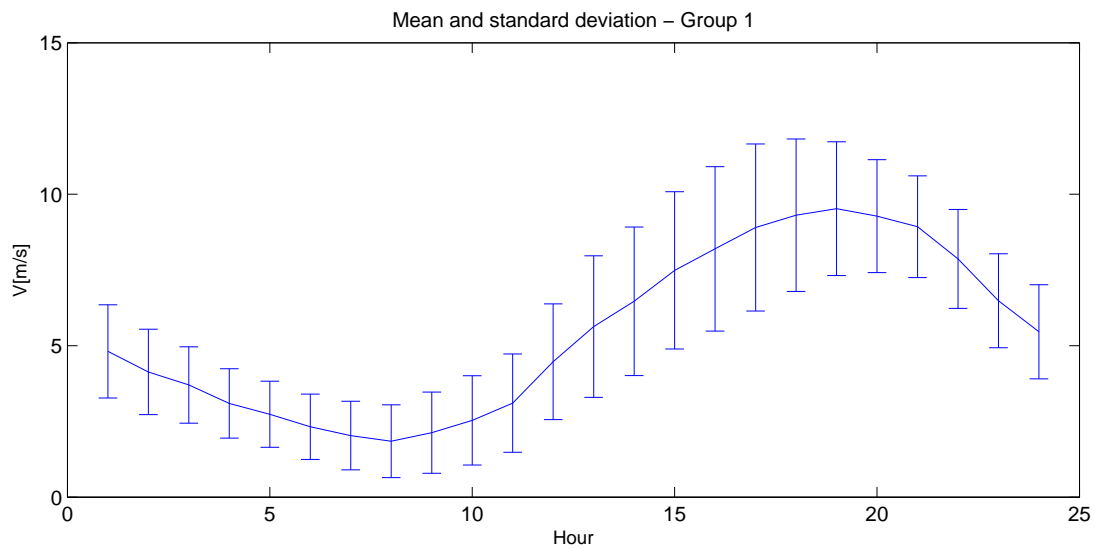


Figura 4.2: Gráfico de velocidad media y desviación estándar Grupo 1

Asimismo, tanto en la Figura 4.3 como en la Figura 4.4, se muestran las funciones de auto-correlación de la serie original y de la serie residual respectivamente, demostrando que los fundamentos del modelo siguen presentes. Esto es, una muy marcada periodicidad en la serie original, y en la serie residual una componente periódica muy leve y con valores de correlación muy pequeños.

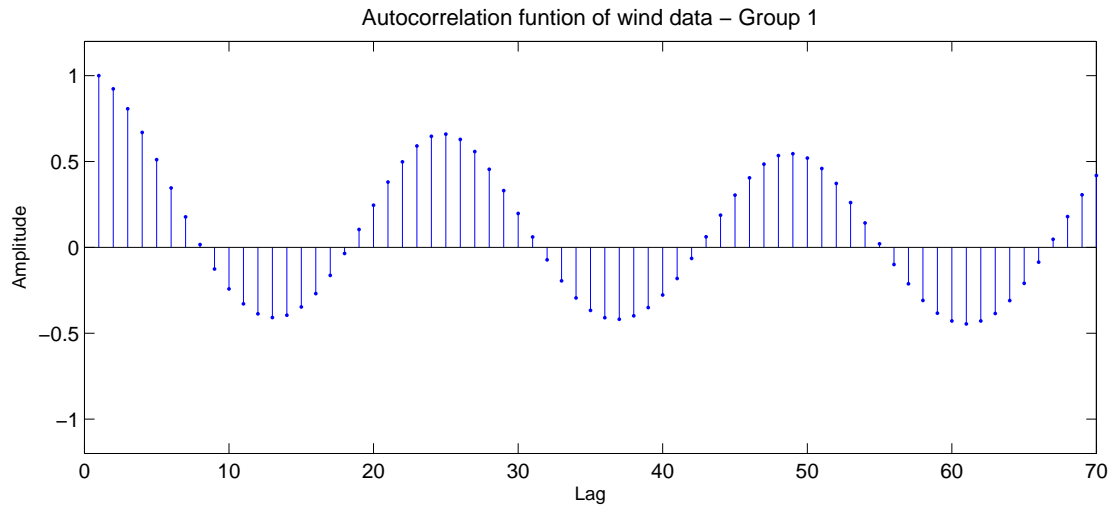


Figura 4.3: Gráfico de función de auto-correlación serie de datos original Grupo 1

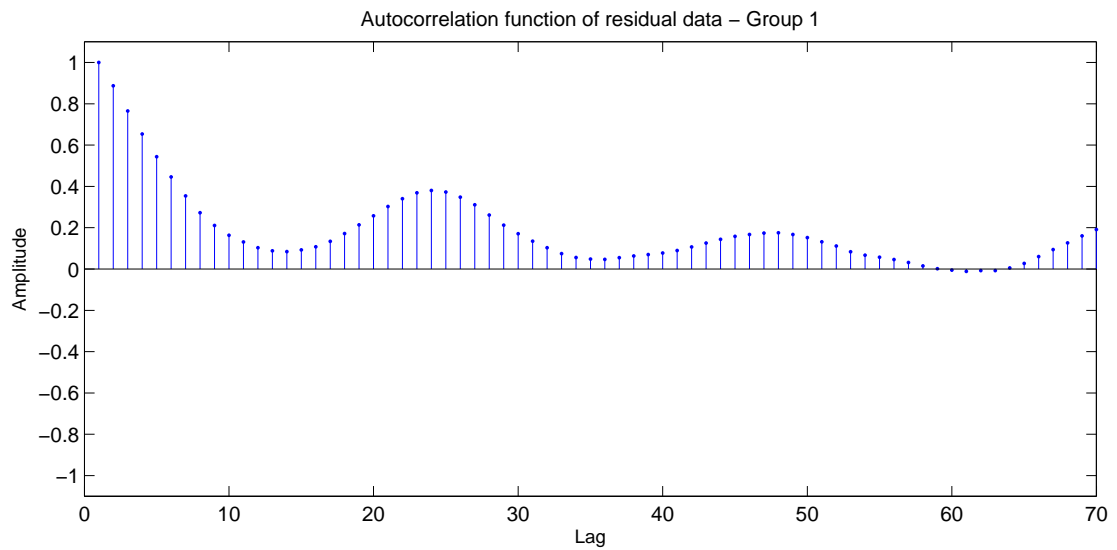


Figura 4.4: Gráfico de función de auto-correlación serie de datos residuales Grupo 1

4.1.2. Segundo Grupo: Semanas 14 a 30

El segundo grupo determinado, comprende al conjunto de semanas desde la 14 a la 30 (2 Abril - 29 Julio). A continuación, en la Figura 4.5, se muestra el patrón característico de la velocidad de viento, considerando su media y desviación estándar.

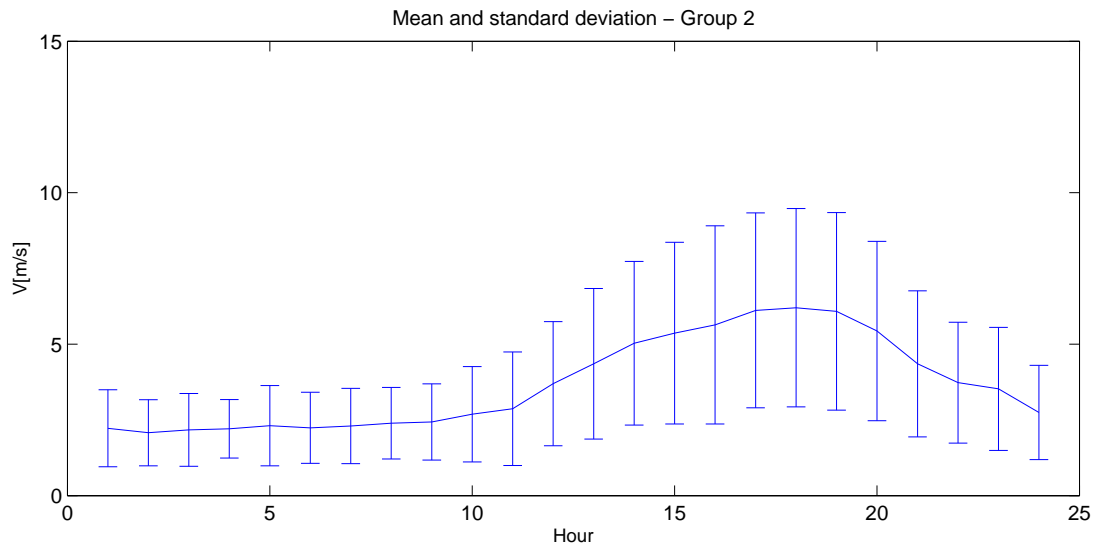


Figura 4.5: Gráfico de velocidad media y desviación estándar Grupo 2

Asimismo, la función de auto-correlación de los datos originales mostrada en la Figura 4.6 y de los datos residuales mostrada en la Figura 4.7, al igual que en el grupo anterior, muestran que las bases del modelo se mantienen al darse las mismas condiciones sobre la periodicidad de las series.

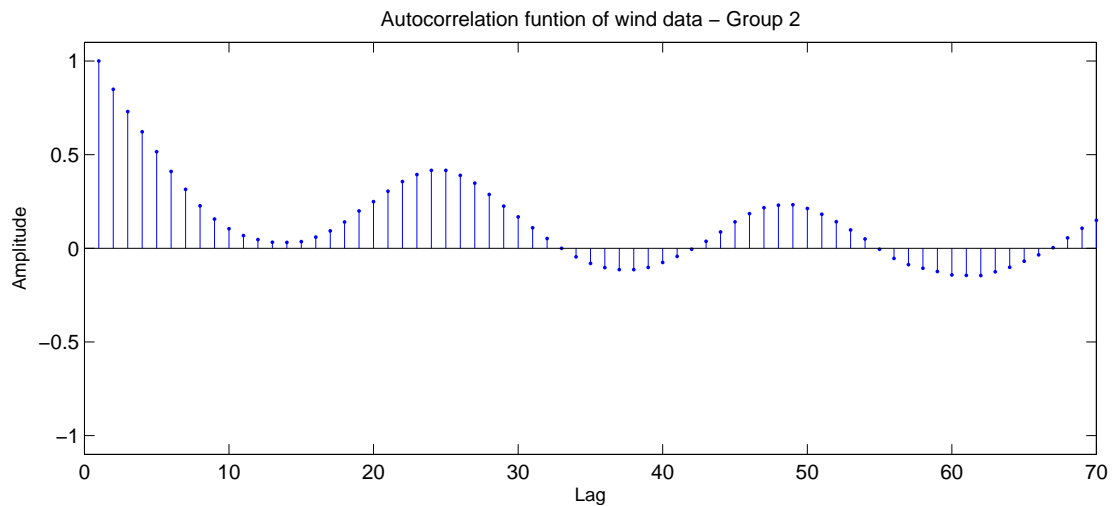


Figura 4.6: Gráfico de función de auto-correlación serie de datos original Grupo 2

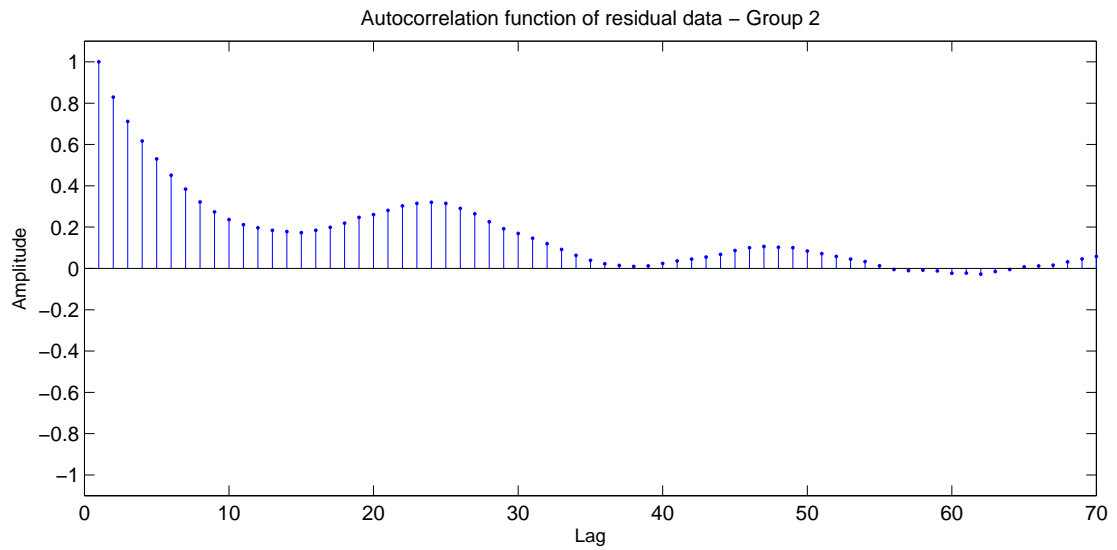


Figura 4.7: Gráfico de función de auto-correlación serie de datos residuales Grupo 2

4.1.3. Tercer Grupo: Semanas 31 a 52

El último grupo establecido, incluye el resto de las semanas que no fueron consideradas en los grupos anteriores (30 Julio - 31 Diciembre), las cuales corresponden al intervalo desde la Semana 31 a la Semana 52. Para este grupo también se muestra el patrón característico de media y desviación estándar, para la velocidad de viento. Estas características se pueden apreciar en la Figura 4.8.

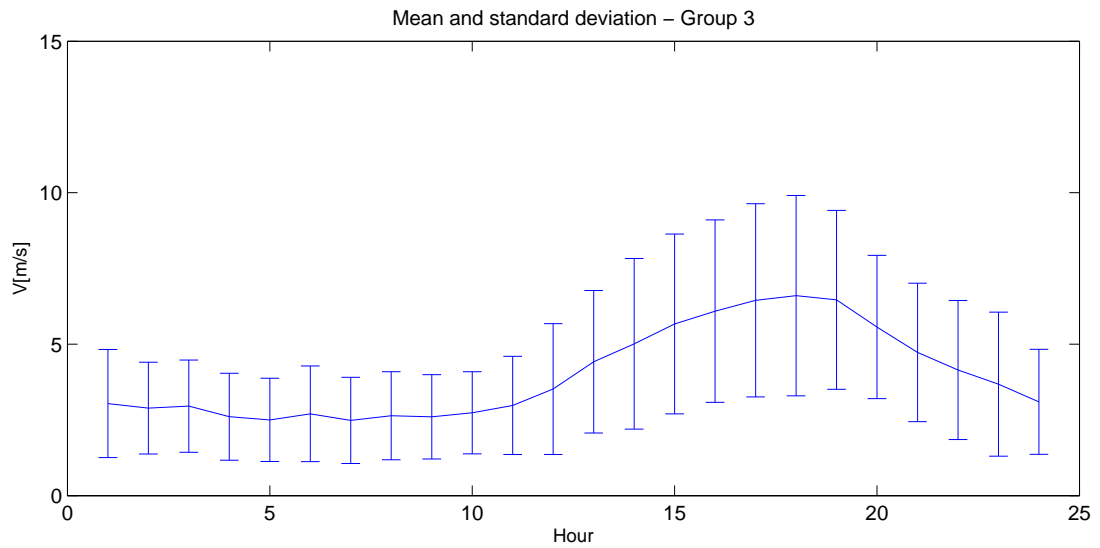


Figura 4.8: Gráfico de velocidad media y desviación estándar Grupo 3

Para este grupo, al igual que en los anteriores, se muestran las funciones de auto-correlación de la serie original y de la serie residual, en la Figura 4.9 y en la Figura 4.10 respectivamente, mostrando nuevamente que los fundamentos del modelo siguen presentes. En este caso, se observa también una fuerte componente periódica en la serie de datos originales, y una periodicidad casi nula en la serie de datos residuales.

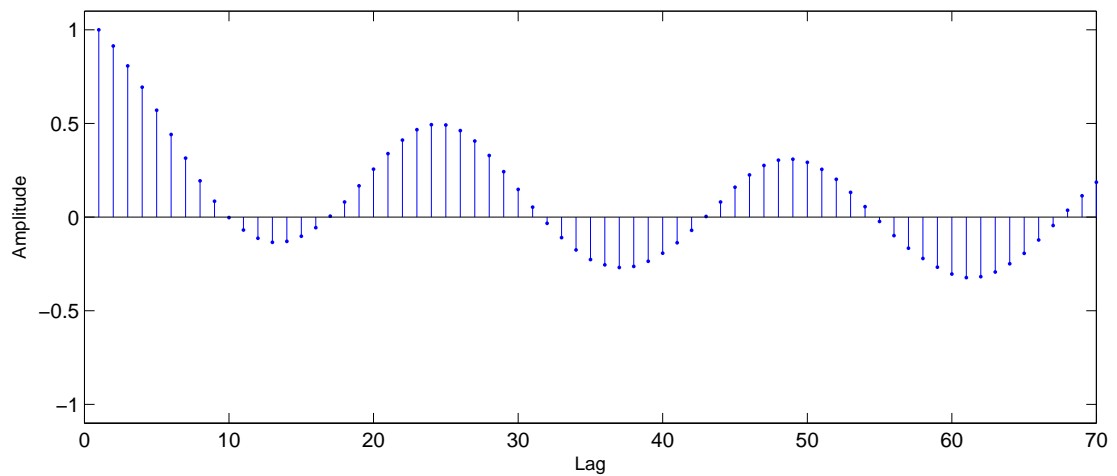


Figura 4.9: Gráfico de función de auto-correlación serie de datos original Grupo 3

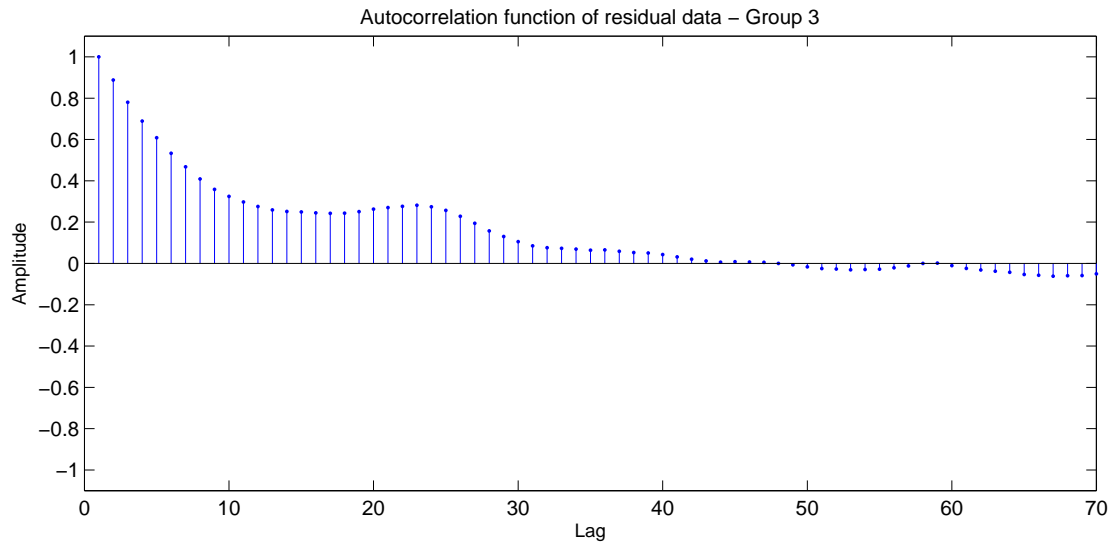


Figura 4.10: Gráfico de función de auto-correlación serie de datos residuales Grupo 3

Al observar los gráficos de los 3 grupos descritos, se aprecia que al realizar la agrupación de los datos, los patrones periódicos de las series originales se mantienen presentes y de forma muy marcada. Por otra parte, las series residuales si bien no presentan la ausencia completa de periodicidad, ésta es mucho menos notoria que en la serie de valores observados, y en el último grupo es casi inexistente. Lo que permite continuar con el siguiente paso del problema.

4.2. Resultados de predicción del Modelo V-ARX

Una vez que los grupos estacionales son determinados, se implementa el modelo V-ARX descrito en el capítulo anterior, utilizando como ya se mencionó, un tercio de los datos reales para el cálculo de los parámetros y los dos tercios restantes como datos de validación. Para ejemplificar los resultados obtenidos, se grafica la predicción realizada para los primeros 7 días, es decir las primeras 168 horas, del periodo de validación de cada uno de los 3 grupos. Es importante recordar que la predicción se realiza a un horizonte de 24 horas, por lo tanto, al transcurrir dicho periodo de tiempo, los valores ingresados al modelo para realizar la predicción siguiente son los valores reales. Las Figuras siguientes dan cuenta de lo

mencionado, y contraponen los valores reales de velocidad de viento con los valores estimados por el modelo V-ARX, para cada uno de los grupos.

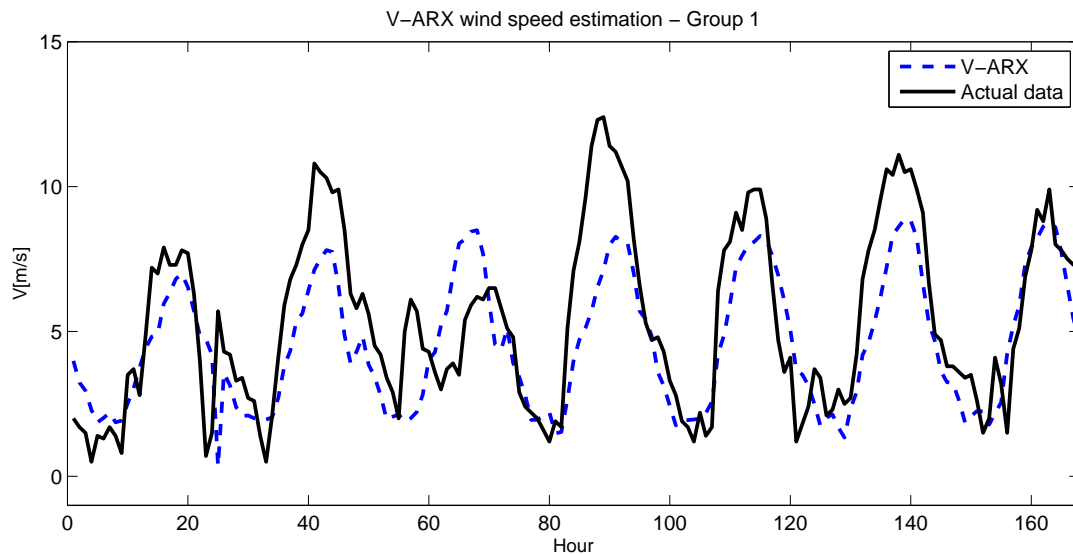


Figura 4.11: Gráfico de predicción a un paso de velocidad de viento. Primera Semana de validación Grupo 1, utilizando el modelo V-ARX

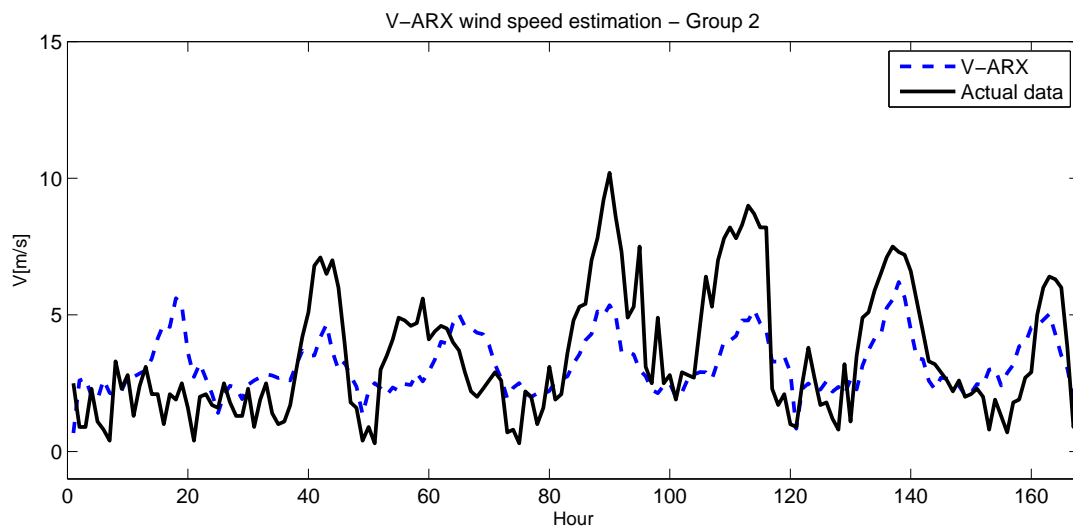


Figura 4.12: Gráfico de predicción a un paso de velocidad de viento. Primera Semana de validación Grupo 2, utilizando el modelo V-ARX

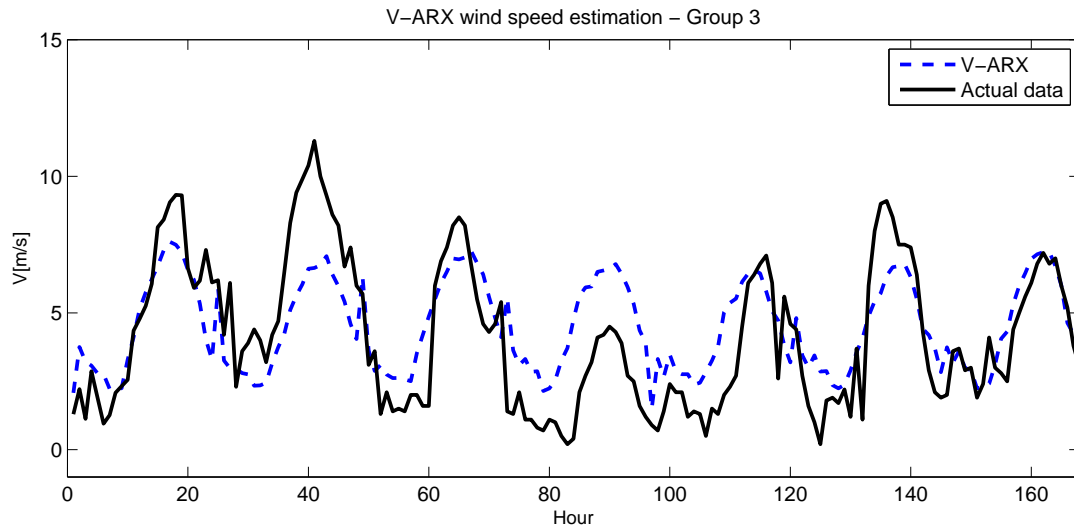


Figura 4.13: Gráfico de predicción a un paso de velocidad de viento. Primera Semana de validación Grupo 3, utilizando el modelo V-ARX

Los gráficos mostrados dan cuenta claramente que el modelo V-ARX sigue muy bien la tendencia de la serie original, esto se debe principalmente a la utilización del patrón de velocidades medias, representado en la variable externa de la ecuación que define el modelo. Sin embargo, debido a la fuerte influencia de este parámetro en los valores estimados, se tiene que en el caso de que la velocidad real se “escape” de los valores promedios, se va a incurrir en un error de predicción mayor, ya que la parte auto-regresiva no será capaz de compensar o mejor dicho, de estimar de forma correcta esa gran variación de magnitud. Este hecho se ve ejemplificado en la Figura 4.12 entre la hora 1 y la hora 40, y en la Figura 4.13 entre las horas 70 y 110 aproximadamente, en que los valores reales son menores a los valores medios de la serie de datos. Finalmente y para cada uno de los grupos estacionales, una vez realizadas las predicciones sobre los datos de validación, se calculan los indicadores de desempeño detallados en el capítulo anterior.

4.2.1. Desempeño del modelo V-ARX

A continuación se entrega el resultado obtenido para los distintos indicadores de desempeño, al realizar predicción de velocidad de viento a un horizonte de 24 horas con el modelo

V-ARX.

Grupo	MSE	MRPE	MREPE	MPEE
1	3.86	68.10	42.90	13.25
2	4.79	89.83	61.19	25.53
3	5.23	83.94	50.49	18.15

Cuadro 4.2: Errores obtenidos con el modelo V-ARX en predicción de 24 horas

4.3. Resultados de predicción del modelo de Referencia de Nielsen

Tal como se mencionó anteriormente, se implementan modelos conocidos para poder comparar el desempeño del predictor propuesto en este trabajo. Inicialmente se utiliza el modelo de Referencia de Nielsen, con un horizonte de predicción de 24 horas para poder realizar una comparación de resultados bajo las mismas circunstancias que el modelo anterior. Cabe mencionar que tanto los datos de entrenamiento como de validación siguen siendo los mismos que para el modelo V-ARX. Los gráficos obtenidos para los tres grupos utilizando este modelo, y que, al igual que el caso anterior, representan sólo los primeros 7 días de los datos de validación se muestran a continuación:

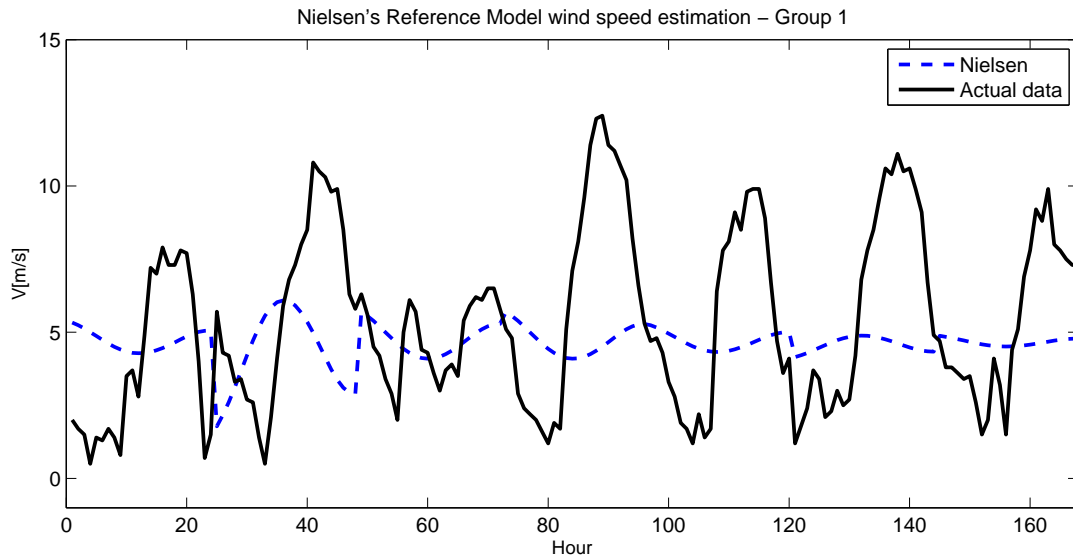


Figura 4.14: Gráfico de predicción a 24 pasos de velocidad de viento. Primera semana de validación Grupo 1, utilizando el modelo de Referencia de Nielsen

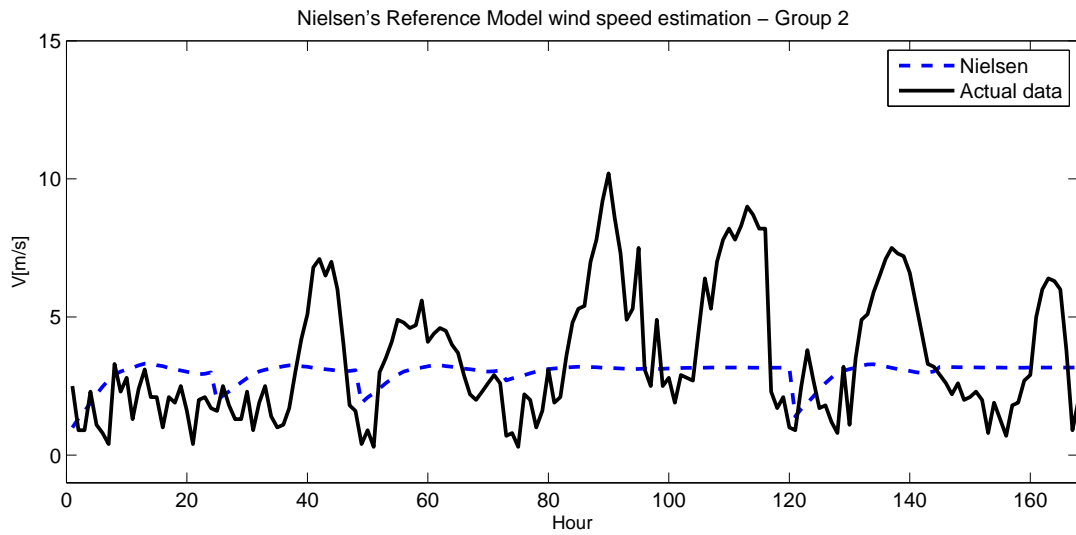


Figura 4.15: Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 2, utilizando el modelo de Referencia de Nielsen

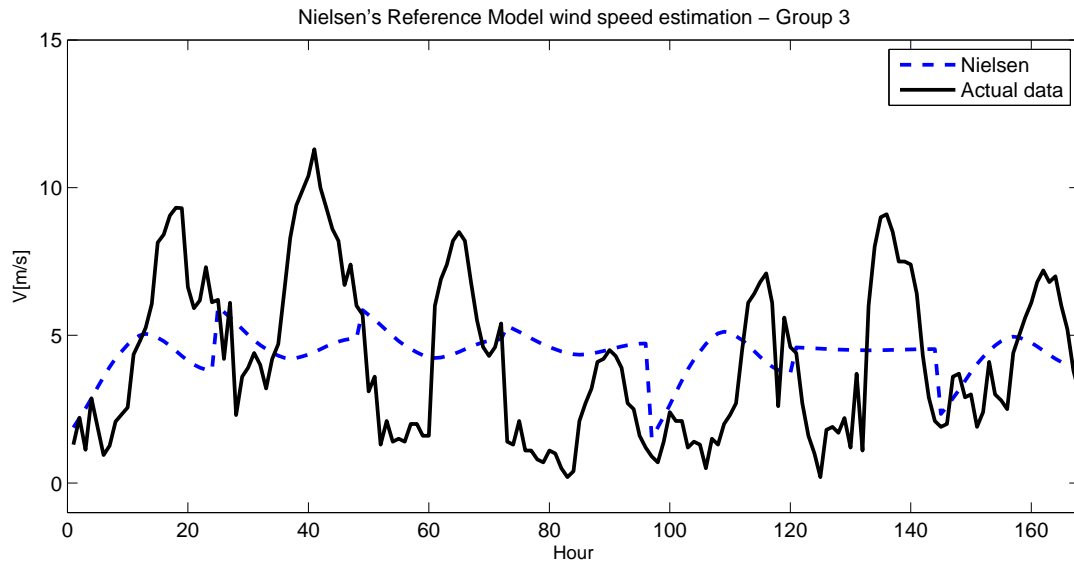


Figura 4.16: Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 3, utilizando el modelo de Referencia de Nielsen

De los gráficos obtenidos, se observa que el Modelo de Referencia de Nielsen no presenta un buen desempeño al realizar predicción a 24 pasos. Esto se debe principalmente a la ecuación que lo conforma, la cual tiene una componente de media que en este caso corresponde a la media de los valores históricos de la serie y no a una media horaria como el caso del modelo V-ARX, por lo tanto las predicciones realizadas a un horizonte relativamente lejano tenderán a parecerse a dicho valor, pues la parte encargada de capturar los cambios utiliza los valores de auto-correlación de la serie y estos son muy pequeños a medida que se aumenta el desfase.

Por el contrario, si se observan con detalle las predicciones realizadas a instantes cercanos de tiempo, es decir, las primeras horas de un día, que gráficamente corresponde a las horas 1,2,3 y sus desplazamientos en 24 muestras, se tiene que las estimaciones se acercan bastante a los valores reales, esto se debe a la fuerte correlación que existe entre horas cercanas y a la utilización del valor real en el instante anterior al que se comienza a realizar la estimación. Para este caso también, una vez realizadas las predicciones sobre los datos de validación, se calculan los indicadores de desempeño detallados en el capítulo anterior.

4.3.1. Desempeño del modelo de Referencia de Nielsen

La siguiente tabla muestra los valores obtenidos para los distintos indicadores de desempeño al realizar predicción de velocidad de viento con el modelo de Referencia de Nielsen en un horizonte temporal de 24 horas:

Grupo	MSE	MRPE	MREPE	MPEE
1	8.74	118.31	64.55	30.00
2	6.08	96.31	68.95	32.42
3	8.10	102.31	62.82	28.10

Cuadro 4.3: Errores obtenidos con el modelo de Referencia de Nielsen en predicción de 24 horas

4.4. Resultados de predicción del modelo ARMA-GARCH

Adicionalmente al modelo de Referencia de Nielsen, se implementa el modelo ARMA(p,q)-GARCH(1,1) también con un horizonte de predicción de 24 horas para ser comparable con los modelos ya usados. Cabe mencionar que tanto los datos de entrenamiento como de validación son los mismos que han sido utilizados previamente.

Es importante destacar también, que el modelo ARMA-GARCH en este caso es aplicado sobre la serie residual, por lo que también se dice que tiene una variable externa, correspondiente al valor medio de la hora a la cual se le está realizando la predicción. La cantidad de parámetros p y q utilizada para cada grupo en la parte ARMA es distinta y se determina calculando el mínimo error cuadrático medio obtenido de variar p y q entre 1 y 6. Los gráficos obtenidos para este modelo y que al igual que en los dos casos anteriores, representan sólo los primeros 7 días de los datos de validación se muestran a continuación:

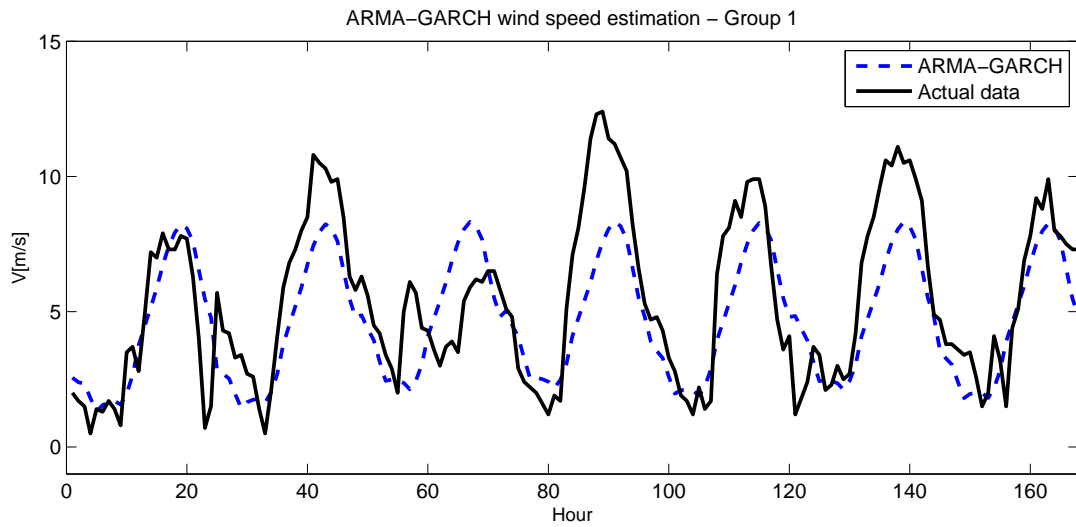


Figura 4.17: Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 1, utilizando el modelo ARMA-GARCH

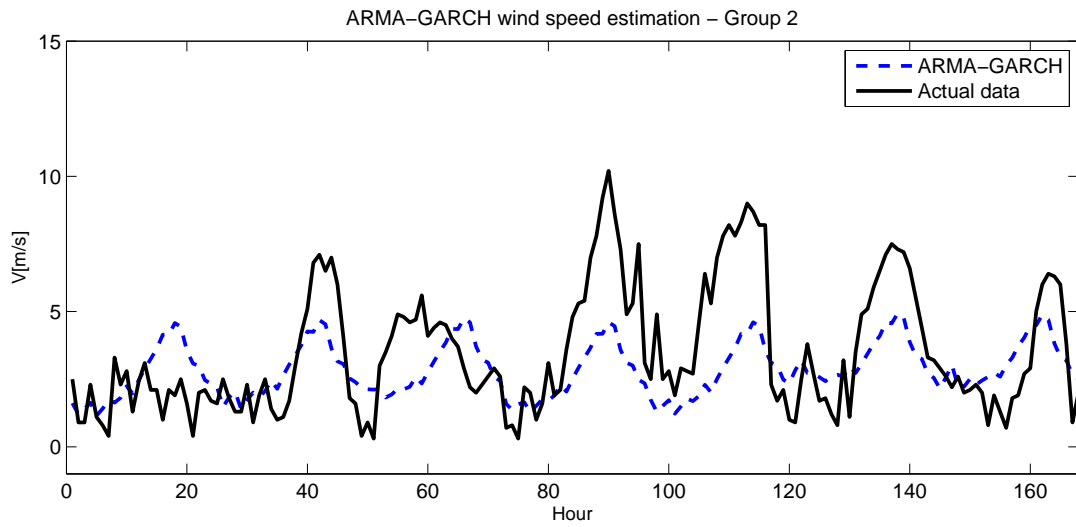


Figura 4.18: Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 2, utilizando el modelo ARMA-GARCH

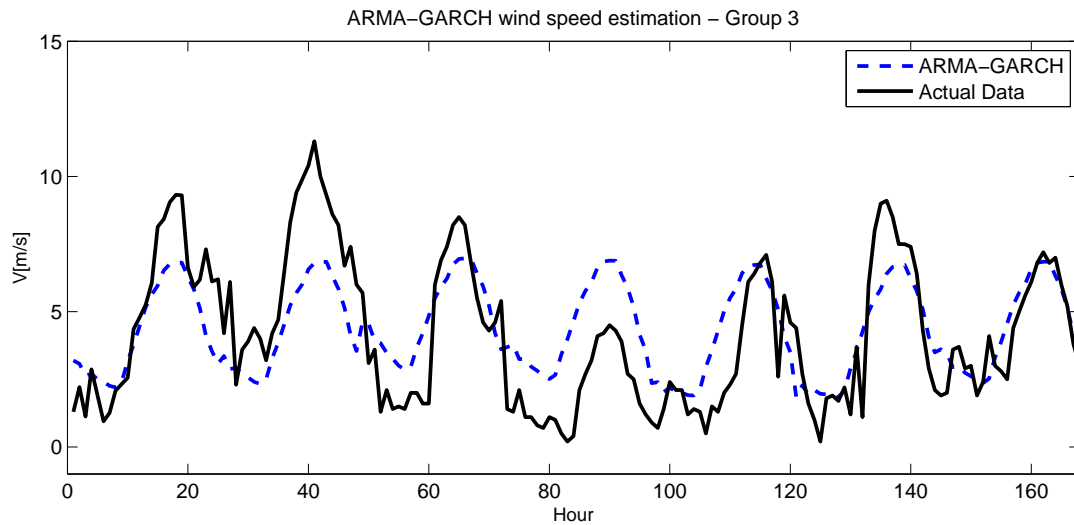


Figura 4.19: Gráfico de predicción a 24 pasos de velocidad de viento. Primera Semana de validación Grupo 3, utilizando el modelo ARMA-GARCH

En los gráficos mostrados se observa que los valores estimados por el modelo ARMA-GARCH, trabajando sobre la serie residual y usando los valores de media horaria, predicen de buena forma la tendencia de la serie de valores observados, nuevamente la componente de velocidades medias juega un papel fundamental en la realización de las predicciones, las que a simple vista son muy similares a las obtenidas con el modelo V-ARX, sin embargo una comparación de los errores podrá establecer cual modelo de predicción tiene mejor precisión.

Además, si se utiliza el modelo ARMA-GARCH directamente sobre la serie de valores observados y no sobre la serie residual, el desempeño del predictor es muy similar al obtenido con el modelo de Referencia de Nielsen. Particularmente en el sentido de que las predicciones a un horizonte cercano (máximo 3 a 4 horas) son muy precisas, pero con horizontes más lejanos la certeza de la estimación disminuye notoriamente. En este caso también, una vez realizadas las predicciones sobre los datos de validación, se calculan los indicadores de error detallados en el capítulo anterior.

4.4.1. Desempeño del modelo ARMA-GARCH

A continuación se entrega el resultado obtenido para los distintos indicadores de desempeño, al realizar predicción de velocidad de viento a un horizonte de 24 horas con el modelo ARMA-GARCH.

Grupo	MSE	MRPE	MREPE	MPEE
1	4.22	73.91	44.85	14.48
2	5.24	91.61	63.95	27.89
3	5.60	87.82	52.24	19.43

Cuadro 4.4: Errores obtenidos con el modelo de ARMA-GARCH en predicción de 24 horas

4.5. Comparación de modelos

Para poder determinar si un modelo tiene mejor desempeño que otro, se hace uso del indicador de mejora o *Improvement* (IMP), el cual se detalló en el capítulo anterior, este indicador se calcula para cada uno de los distintos medidores de desempeño calculados.

4.5.1. Comparación con modelo de Referencia de Nielsen

Los índices de mejora obtenidos al comparar el modelo propuesto con el modelo de Referencia de Nielsen se muestran en la siguiente tabla:

Grupo	IMP - Base Nielsen [%]			
	MSE	MRPE	MAEPE	MPEE
1	55.84	42.44	33.55	55.84
2	21.25	6.72	11.26	21.25
3	35.42	17.99	19.67	35.48

Cuadro 4.5: Índices de mejora comparando V-ARX con modelo de Referencia de Nielsen

4.5.2. Comparación con modelo ARMA-GARCH

Los índices de mejora obtenidos al comparar el modelo propuesto con el modelo ARMA-GARCH se muestran en la siguiente tabla:

Grupo	IMP - Base ARMA-GARCH [%]			
	MSE	MRPE	MAEPE	MPEE
1	8.57	7.87	4.36	8.53
2	8.46	1.94	4.36	8.46
3	6.60	4.42	3.35	6.60

Cuadro 4.6: Improvement considerando MAPE como medidor de desempeño

Del análisis de valores obtenidos es claro que el modelo V-ARX propuesto presenta un mejor desempeño tanto al compararlo con el Modelo de Referencia de Nielsen como con el modelo ARMA-GARCH, sin importar el tipo de error que se considere.

Una primera observación que se puede extraer de estos valores es la efectividad del modelo propuesto ante la variabilidad de los valores horarios medios de velocidad de viento, se distingue una mayor precisión de las predicciones en el Grupo 1, el cual tiene un patrón de media mucho más “variable”, en el cual la diferencia entre el valor máximo y el mínimo es mayor, en cambio la precisión en el Grupo 3, donde esta diferencia no se observa tan acentuada es menor, y disminuye aun más en el Grupo 2, en donde el patrón dominante no presenta una variación muy significativa.

Por otra parte, si comparamos el modelo V-ARX sólo con el Modelo de Referencia de Nielsen, estamos en presencia de una mejora significativa en la precisión de las predicciones, provocada fundamentalmente por la utilización del patrón horario que está presente en la serie de velocidad de viento y que permite utilizar información relevante para predicciones a un horizonte lejano en donde la incertidumbre es mucho mayor, hecho que no puede ser capturado de forma efectiva por un modelo lineal como el de Nielsen, provocando que al realizar estimaciones con este método a un horizonte temporal de 24 pasos como es este caso, no sea posible capturar la información necesaria para realizar una buena aproximación. Finalmente, si se analiza la mejora con respecto al modelo ARMA-GARCH, que a pesar de no ser tan significativa, de igual forma indica que el modelo V-ARX es ligeramente más preciso, ya no se puede atribuir este hecho a la utilización del patrón de velocidades medias, pues ambos modelos hacen uso de él, por lo tanto la diferencia en este caso se da en la manera

en que se resuelve la parte auto-regresiva, resultando que la manera matricial utilizada por el modelo propuesto presenta un mejor desempeño que un método lineal. Un motivo por el cual se produce esta mejora es que en el primer caso se está utilizando toda la información disponible de una sola vez para obtener el vector de predicciones sin incurrir en errores asociados al tener que reincorporar al modelo valores estimados, en cambio con el método lineal, al utilizar una realimentación de los valores predichos para realizar las siguientes, en cada paso se está aumentando el nivel de incertidumbre existente y por lo tanto del error asociado a una estimación.

Capítulo 5

Conclusiones y Trabajo Futuro

5.1. Conclusiones

El trabajo aquí presentado consiste en el desarrollo de un predictor vectorial de velocidad de viento para un horizonte de 24 horas, éste fue diseñado de tal forma que es capaz de capturar las características periódicas del viento, tanto diarias como estacionales, demostrando que tomar en cuenta esta información en la elaboración de un modelo entrega mejoras sustanciales en relación a los modelos comúnmente utilizados y que no las incorporan como fuente de información, o de hacerlo toman en cuenta solo una.

Otro punto a destacar, se obtiene de realizar la predicción de las 24 horas del día en una sola iteración utilizando un modelo vectorial, pues hacerlo de este modo permite disminuir el error obtenido al utilizar toda la información relevante de una sola vez. Por el contrario, si se utilizara un modelo lineal para poder realizar la predicción con ese horizonte temporal, el cálculo debe ser realizado 24 veces, uno por cada hora adicional, y para esto, las estimaciones obtenidas deben ser realimentadas en la ecuación del modelo, incluyendo además los errores asociados, los que se ven incrementados en cada iteración. En este contexto, también es importante reconocer la gran utilidad de trabajar sobre la serie de residuos y de utilizar el patrón de velocidades medias en el modelo, pues esto permite que, incluso al realizar predicciones en horizontes de tiempo lejanos, se disponga de información relevante que favorece a que la estimación sea más cercana a los valores reales.

Por otra parte, los valores obtenidos del indicador de mejora (IMP), dan cuenta que el modelo vectorial presentado entrega mejores resultados que ambos modelos de comparación utilizados. Esto permite concluir que el diseño propuesto tiene un gran potencial, pues actualmente solo utiliza un proceso AR para resolver la parte auto-regresiva, y con esto ya supera a algunos algoritmos conocidos, sin embargo, existen algoritmos más precisos que si fuesen adaptados a la forma vectorial podrían mejorar aun más el desempeño del predictor y entregar estimaciones mucho más acertadas.

También es importante destacar que la manera en que se definieron los grupos de datos, la cual no había sido presentada en trabajos anteriores, entregó buenos resultados. En primer lugar, el hecho de haber utilizado la divergencia de *Kullback-Leibler* como medida de discriminación y el posterior agrupamiento con *K-means*, permitió realizar una buena identificación, fundamentada matemáticamente, de los distintos patrones que gobiernan el fenómeno del viento durante el año, pudiendo crear un modelo más preciso para cada uno de estos grupos de datos. En segundo lugar, esta forma de realizar la diferenciación permitió encontrar una cantidad adecuada de grupos para crear los modelos, pues haber tenido solo un predictor para todo el año podría aumentar significativamente los errores asociados, o por el contrario, haber tenido modelos para todas las semanas podría no ser tan representativo pues no siempre se repite el mismo patrón en un periodo tan corto de tiempo de un año para otro.

Finalmente, el desarrollo de esta memoria permitió conocer la existencia de un gran número de métodos de estimación utilizados para abordar el problema aquí tratado, así como la cantidad de trabajos al respecto, lo que da cuenta de la gran importancia que ha ido adquiriendo el desarrollo de las ERNC en la actualidad.

5.2. Trabajo Futuro

Como posibles mejoras al modelo propuesto, o líneas de investigación que puedan surgir de este trabajo de título, es posible considerar las siguientes alternativas:

- Utilizar una mayor cantidad de datos para entrenar y validar el modelo, como el registro histórico de 3 o más años, para incluir mayor cantidad de información y minimizar eventos “extraños” que puedan ocurrir durante algún periodo particular y que no sean representativos de los valores históricos.
- Utilizar un modelo matricial auto-regresivo más completo, por ejemplo un ARMA en lugar de sólo un AR, o modelos no lineales que sean capaces de procesar de mejor forma la información existente en horizontes de tiempos más lejanos.
- Realizar la estimación de los intervalos de confianza de las predicciones.
- Adaptar un filtro de Kalman a este problema matricial, como corrector de predicciones, el cual ha demostrado buenos resultados al ser utilizado sobre modelos lineales.

Referencias

- [1] Ministerio de Energía, “Estrategia nacional de energía 2012 - 2030,” Febrero 2012.
- [2] Programa de Estudios e Investigaciones en Energía del Instituto de Asuntos Públicos de la Universidad de Chile, Núcleo Milenio de Electrónica Industrial y Mecatrónica del Centro de Innovación en Energía de la Universidad Técnica Federico Santa María, “Aporte potencial de energías renovables no convencionales y eficiencia energética a la matriz eléctrica, 2008 - 2025,” Junio 2008.
- [3] Ministerio de Economía, Fomento y Reconstrucción, “Ley 20257, introduce modificaciones a la ley general de servicios eléctricos respecto de la generación de energía eléctrica con fuentes de energía renovables no convencionales,” Abril 2008.
- [4] J. Quintanilla, “Energías Renovables No Convencionales y la puesta en marcha del modelo de cuotas en el sistema eléctrico chileno,” Abril 2010. [Online]. Available: <http://www.core.cl/wp/wp-content/uploads/2010/04/enrc-y-la-puesta-en-marcha-del-modelo-de-cuotas-en-el-sistema-electrico-chileno.pdf>
- [5] A. Gutiérrez, J. Cataldo, and P. Santoro, “Avances sobre la predicción de corta duración de energía la eléctrica generada en parques eólicos.” Julio 2008.
- [6] G. Kariniotakis, P. Pinson, and N. Siebert, “The state of the art in short-term prediction of wind power-from an offshore perspective,” p. 20–21, 2004. [Online]. Available: http://www.see.asso.fr/clubs_techniques/se/xmedia/Club_Tech_SE-2001-2008/2004/Eolien_ELEC_2004/paper_prediction_GKariniotk-1-sur-2.pdf

- [7] Instituto para la Diversificación y Ahorro de la Energía (IDAE), “Project anemos, estudio sobre predicción eólica en la unión europea,” 2007.
- [8] G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*, 3rd ed. Prentice Hall PTR, 1994.
- [9] J. Torres, A. García, M. D. Blas, and A. D. Francisco, “Forecast of hourly average wind speed with arma models in navarre (spain),” *Solar Energy*, vol. 79, no. 1, pp. 65 – 77, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0038092X04002877>
- [10] S.Norambuena, “Predicción de corto plazo de potencia generada en un aerogenerador usando modelo sarima: Memoria para optar al título de ingeniero civil electricista de la universidad de chile,” 2011.
- [11] T. Bollerslev, “Generalized autoregressive conditional heteroskedasticity,” *Journal of Econometrics*, vol. 31, no. 3, pp. 307 – 327, 1986. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0304407686900631>
- [12] R. F. Engel, “Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation,” *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982. [Online]. Available: <http://www.jstor.org/stable/1912773>
- [13] H. Liu, E. Erdem, and J. Shi, “Comprehensive evaluation of arma–garch(-m) approaches for modeling the mean and volatility of wind speed,” *Applied Energy*, vol. 88, no. 3, pp. 724 – 732, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261910003934>
- [14] F.Tobar, “Inferencia de la volatilidad de retornos financieros usando filtro de partículas”. tesis para optar al grado de magíster en ciencias de la ingeniería, mención eléctrica; memoria para optar al título de ingeniero civil electricista de la universidad de chile,” 2010.

- [15] H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. S. Nielsen, “Standardizing the performance evaluation of shortterm wind power prediction models,” *Wind Engineering*, vol. 29, no. 6, pp. 475–489, 2005. [Online]. Available: <http://multi-science.metapress.com/openurl.asp?genre=article&id=doi:10.1260/030952405776234599>
- [16] T. S. Nielsen, A. Joensen, H. Madsen, L. Landberg, and G. Giebel, “A new reference for wind power forecasting,” *Wind Energy*, vol. 1, no. 1, pp. 29–34, 1998. [Online]. Available: [http://dx.doi.org/10.1002/\(SICI\)1099-1824\(199809\)1:1<29::AID-WE10>3.0.CO;2-B](http://dx.doi.org/10.1002/(SICI)1099-1824(199809)1:1<29::AID-WE10>3.0.CO;2-B)
- [17] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Transactions of the ASME—Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.
- [18] G. Welch and G. Bishop, “An introduction to the kalman filter,” Chapel Hill, NC, USA, Tech. Rep., 1995.
- [19] Z. Wei and W. Weimin, “Wind speed forecasting via ensemble kalman filter,” in *Advanced Computer Control (ICACC), 2010 2nd International Conference on*, vol. 2, march 2010, pp. 73 –77.
- [20] 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.
- [21] R. Libonati, I. Trigo, and C. C. DaCamara, “Correction of 2 m-temperature forecasts using kalman filtering technique,” *Atmospheric Research*, vol. 87, no. 2, pp. 183 – 197, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S016980950700155X>
- [22] S. Kullback, *Information Theory and Statistics*. Dover Publications Inc., 1968.

- [23] J. Hershey and P. Olsen, “Approximating the kullback leibler divergence between gaussian mixture models,” in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 4, april 2007, pp. IV–317 –IV–320.
- [24] J. B. MacQueen, *Some methods for classification and analysis of multivariate observations*. University of California Press, 1967, vol. 1, no. 281-297, pp. 281–297. [Online]. Available: http://www.projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.bsmsp/1200512992
- [25] P. Aceituno, “Proyecto eolo,” 2003. [Online]. Available: <http://www.dgf.uchile.cl/eolo/>
- [26] A. Dorvlo, “Estimating wind speed distribution,” *Energy Conversion and Management*, vol. 43, no. 17, pp. 2311 – 2318, 2002.
- [27] I. Erbetta, “Predicción de la potencia para la operación de parques eólicos: Memoria para optar al título de ingeniero civil electricista de la universidad de chile,” 2010.
- [28] S. P. Kani and M. Ardehali, “Very short-term wind speed prediction: A new artificial neural network–markov chain model,” *Energy Conversion and Management*, vol. 52, no. 1, pp. 738 – 745, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0196890410003638>
- [29] G. Li and J. Shi, “On comparing three artificial neural networks for wind speed forecasting,” *Applied Energy*, vol. 87, no. 7, pp. 2313 – 2320, 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261909005364>
- [30] J. Catalão, H. Pousinho, and V. Mendes, “Short-term wind power forecasting in portugal by neural networks and wavelet transform,” *Renewable Energy*, vol. 36, no. 4, pp. 1245 – 1251, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0960148110004477>
- [31] X. Basogain, *Curso Redes Neuronales Artificiales y sus Aplicaciones*. Depto. de Sistemas y Automática, Escuela Superior de Ingeniería de Bilbao, 2010.

- [32] J. L. Deng, “Introduction to grey system theory,” *J. Grey Syst.*, vol. 1, no. 1, pp. 1–24, Nov. 1989. [Online]. Available: <http://dl.acm.org/citation.cfm?id=90757.90758>
- [33] T. El-Fouly, E. El-Saadany, and M. Salama, “Improved grey predictor rolling models for wind power prediction,” *IET Generation, Transmission and Distribution*, vol. 1, no. 6, pp. 928–937, 2007. [Online]. Available: <http://link.aip.org/link/?GTD/1/928/1>
- [34] 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. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0957417409007258>
- [35] C. Valenzuela, “Predicción de largo plazo para la generación eólica mediante modelos grises: Memoria para optar al título de ingeniero civil electricista de la universidad de chile,” 2011.
- [36] G. Galanis, P. Louka, P. Katsafados, I. Pytharoulis, and G. Kallos, “Applications of kalman filters based on non-linear functions to numerical weather predictions,” *Annales Geophysicae*, vol. 24, no. 10, pp. 2451–2460, 2006. [Online]. Available: <http://www.ann-geophys.net/24/2451/2006/>
- [37] I. Sánchez, “Short-term prediction of wind energy production,” *International Journal of Forecasting*, vol. 22, no. 1, pp. 43 – 56, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207005000622>
- [38] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, “A review on the forecasting of wind speed and generated power,” *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915 – 920, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032108000282>
- [39] R. M. Gray and L. D. Davisson, *An Introduction to Statistical Signal Processing*. Liège: Cambridge Univ. Press, 2004.

- [40] C. Monteiro, H. Keko, R. Bessa, V. Miranda, A. Botterud, J. Wang, , and G. Conzelmann, “A quick guide to wind power forecasting: State-of-the-art, 2009,” 2009. [Online]. Available: <http://www.osti.gov/bridge>
- [41] E.Pavéz, “Evaluación de medidas de discriminación entre hidden markov models para reconocimiento fonético,” 2010, Trabajo final del curso EL4101 - Procesos Aleatorios en Comunicaciones y Procesamiento de Señales.
- [42] L. M. Molinero, “Análisis de series temporales,” 2004. [Online]. Available: <http://www.seh-lelha.org/tseries.htm>
- [43] Z. Huang and Z. Chalabi, “Use of time-series analysis to model and forecast wind speed,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 56, no. 2–3, pp. 311 – 322, 1995. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/016761059400093S>
- [44] R. Billinton, H. Chen, and R. Ghajar, “Time-series models for reliability evaluation of power systems including wind energy,” *Microelectronics Reliability*, vol. 36, no. 9, pp. 1253 – 1261, 1996. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0026271495001549>