



UNIVERSIDAD DE VALLADOLID

Departamento de Ingeniería de Sistemas y Automática

**NUEVOS MODELOS DE PREDICCIÓN EÓLICA
BASADOS EN SERIES TEMPORALES**

Tesis Doctoral

Marta Poncela Blanco

Ingeniero Industrial

2012

DEPARTAMENTO DE INGENIERÍA DE SISTEMAS Y AUTOMÁTICA

Escuela de Ingenierías Industriales

**NUEVOS MODELOS DE PREDICCIÓN EÓLICA
BASADOS EN SERIES TEMPORALES**

Autor:

Marta Poncela Blanco

Ingeniero Industrial

Directores:

Dr. D. José Ramón Perán González

Catedrático de Universidad

Dra. Dña. María del Pilar Poncela Blanco

Profesora Titular de Universidad

Valladolid, 2012

Tribunal

Presidente: _____

Vocal: _____

Vocal: _____

Vocal: _____

Secretario: _____

Calificación de la Tesis: _____

Dedicatoria

Quiero dedicar esta tesis a las personas que han marcado mi vida de algún modo u otro. Son muchas, pero aquí va un esbozo de las más importantes.

Empiezo por mi abuela María y mi tía María Jesús porque el cariño que se da a un niño permanece en su corazón para siempre.

No puedo olvidarme de mis padres y hermanos. Especialmente quiero dedicar la tesis a mi hermana María porque no podría entender muchos aspectos de mi carácter y mi vida sin ella. Otra dedicatoria muy especial es para mi hermana Pilar, porque además de directora de tesis lleva vigilando mis pasos de cerca desde que tengo uso de razón. Por esa vigilancia, gracias.

Quienes me conocen saben que no pueden faltar Luis, Marta y Rodrigo. Luis además de ser un gran compañero quiero agradecerle la paciencia que ha demostrado conmigo no sólo durante la realización de la tesis sino a lo largo de nuestra vida en común. Marta y Rodrigo son la luz que ilumina cada mañana.

También quiero dedicar esta tesis a Macus y a todas aquellas personas que cada día se enfrentan a las adversidades con una capacidad asombrosa.

Y por último, para mi primo Javier con el que he pasado grandísimos momentos y que siempre está ahí.

Agradecimientos

Esta tesis representa parte de mi trabajo realizado en CARTIF, por lo que estos agradecimientos empiezan a esta institución en abstracto y en particular a D. José Ramón Perán y D. Luis Javier de Miguel por darme la oportunidad de trabajar en dicha institución y por transmitirme “el gusanillo” de la investigación.

Además quiero agradecer a mis directores de tesis D. José Ramón Perán González y Dña Pilar Poncela Blanco por su apoyo técnico y académico en esta tesis y en mi carrera profesional dentro y fuera de CARTIF.

No puedo olvidarme del Departamento de Ingeniería de Sistemas y Automática de la Universidad de Valladolid, por las facilidades que siempre nos ha brindado a todos los que intentamos realizar un trabajo de Tesis Doctoral compaginándolo con nuestra actividad laboral fuera de la Universidad.

Esta tesis no podría haberse materializado sin la colaboración desinteresada de Me-teoGalicia, suministrándonos los datos meteorológicos y resolviendo cuantas dudas se nos planteaban en su uso.

Por último, agradezco al profesor Ismael Sánchez sus sugerencias en la orientación de las líneas de trabajo.

Resumen

El desarrollo de la energía eólica en los últimos años ha sido exponencial, tanto en lo que se refiere en potencia instalada, como en desarrollo tecnológico. Para que esta tendencia no se vea frenada es necesario dotar a estos sistemas de herramientas que permitan su integración óptima en la red eléctrica sin poner en peligro su estabilidad ni garantía de suministro. Un instrumento fundamental hoy en día es la predicción de la potencia generada con suficiente antelación, de manera que pueda dimensionarse y coordinar adecuadamente el funcionamiento de los generadores convencionales. Por otro lado, las herramientas de predicción son necesarias para limitar el uso de reserva rodante y con ello contribuir aún más a los beneficios medioambientales de este tipo de energía. Desde el punto de vista del gestor del parque eólico, la predicción de la producción es necesario para, entre otras cosas, acudir al mercado eléctrico.

El trabajo desarrollado en esta tesis se centra en la mejora de los modelos estadísticos presentes en herramientas de predicción eólica actualmente existentes. Para ello, se partirá de un caso particular como es Sipreólico. Las principales aportaciones de este trabajo son, en primer lugar, partiendo de los modelos basados en el filtro de Kalman hasta ahora desarrollados, la estimación de las matrices del sistema por máxima verosimilitud en vez de imponerlas de manera ad hoc. De esta manera se incorporan, en cierto modo, las características del parque eólico a la predicción y se adapta la herramienta de predicción a cada parque.

En segundo lugar, se estudia la aplicación de técnicas multivariantes que permitan captar, a través de las correlaciones entre parámetros, relaciones no recogidas explícitamente en los modelos.

Como el problema objeto de estudio es no lineal y variante con el tiempo, es usual en predicción eólica recurrir a un conjunto de modelos operando en paralelo, ya que cada uno de ellos funcionará adecuadamente en determinados instantes de tiempo, dado que cada modelo capta mejor unas dinámicas que otras. Por ello se plantean, de manera novedosa, distintas alternativas multivariantes de reducción de datos como estrategia

de combinación de predicciones para la obtención de la predicción final.

Los resultados obtenidos en la aplicación práctica de las estrategias planteadas a lo largo de esta tesis se muestran para el parque eólico de Sotavento.

Por último, todo el trabajo se ha enmarcado en el desarrollo de una herramienta de predicción operando en tiempo real y con la intervención mínima del experto.

Abstract

Wind power has experienced an enormous growth in the last years, not only in its installed capacity, but also in its great technological development. Today, wind power forecasts are necessary for the large-scale integration of wind turbines into the electrical grid due to this intermittency nature. Forecasting tools are also necessary to limit the use of spinning reserves, increasing even more the environmental benefits of this kind of energy. Wind power forecasts are necessary to participate in the electricity markets too.

The work developed in this thesis is devoted to improve the performance of statistical models used in actual wind energy forecasting tools. To do so, we have begun from a particular case - Sipreolico. The main contributions of this thesis are the following: first, we estimate the parameters of the models by maximum likelihood through the Kalman filter and smoother instead of imposing them ad-hoc. This parameter estimation allows us to capture the operating conditions or characteristics of each wind farm to some degree and obtain significantly lower root mean square error (RMSE). Second, we propose new specifications for the state equation, including multivariate models, to capture, through the correlation of the elements of the state vector, the effect caused by omitted inputs.

The relationship between wind and generated power is highly nonlinear and time-varying, so it is very common to use alternative models based on different assumptions about the variables involved. We proposed a novel forecast combination procedure through multivariate dimension reduction techniques.

We check the strategies shown throughout this thesis in the Sotavento wind farm.

At last, all the work developed is included as part of a real time forecasting tool with minimal expert intervention.

Índice general

Resumen	I
Abstract	III
Índice de contenidos	v
Índice de figuras	IX
Índice de tablas	xv
1. Introducción	1
1.1. Introducción	1
1.2. Descripción del problema	2
1.3. Objetivos de esta tesis	4
1.4. Estado Tecnológico Actual	6
1.5. Estructura y contenidos de la tesis	10
2. Presentación de los datos y modelos de partida	11
2.1. Introducción	11
2.1.1. Presentación de los datos de trabajo	11
2.2. Evaluación de los modelos	14
2.2.1. Introducción	14
2.2.2. Modelos de Referencia	14
2.2.3. Medidas de Error	15

2.2.4.	Resultados de los modelos de referencia	17
2.3.	Modelos de partida	24
2.3.1.	Introducción	24
2.3.2.	Análisis de resultados con los modelos de partida	26
2.4.	Anexo: Gráficas de los resultados de los modelos de predicción	29
2.4.1.	Modelos de partida	29
3.	Optimización de los modelos y propuesta de nuevos modelos	43
3.1.	Introducción	43
3.2.	Algoritmo EM	44
3.2.1.	Introducción	44
3.2.2.	Resultados obtenidos	49
3.3.	Nuevos modelos multivariantes para el estado	55
3.3.1.	Introducción	55
3.3.2.	Resultados obtenidos	57
3.4.	Estrategia de funcionamiento en tiempo real.	60
3.5.	Bootstrap de los parámetros	64
3.6.	Apéndice: Resultados empíricos	71
3.6.1.	Resultados de la optimización mediante el algoritmo EM	71
3.6.2.	Resultados numéricos del remuestreo mediante bootstrap.	86
4.	Combinación de predicciones	87
4.1.	Introducción	87
4.2.	Técnicas multivariantes de reducción de datos	90
4.2.1.	Componentes principales	90
4.2.2.	Mínimos cuadrados parciales	91
4.3.	Resultados obtenidos	92
4.3.1.	Modelo de referencia	93
4.3.2.	Componentes principales	95

4.3.3. Mínimos cuadrados parciales	101
4.4. Conclusiones	103
5. Conclusiones	105
5.1. Conclusiones	105
5.2. Desarrollos futuros	106
5.3. Difusión de resultados	109
Bibliografía	111

Índice de figuras

2.1.	Potencia eólica de salida.	12
2.2.	Diagrama de cajas de los errores de predicción del modelo persistente para cada horizonte de predicción.	18
2.3.	Histograma de los errores de predicción. Modelo persistente.	18
2.4.	Correlograma del error de predicción a horizonte 1. Modelo persistente.	19
2.5.	Diagrama de cajas de los errores de predicción para el modelo media en función del horizonte de predicción.	20
2.6.	Histograma de los errores de predicción. Modelo media.	20
2.7.	Correlograma del error de predicción a horizonte 1. Modelo media.	21
2.8.	Diagrama de cajas de los errores de predicción del modelo AR(1) de parámetros constantes para cada horizonte de predicción.	21
2.9.	Histograma de los errores de predicción. Modelo AR(1) de parámetros constantes.	22
2.10.	Correlograma de los errores de predicción a horizonte 1. Modelo AR(1) de parámetros constantes.	22
2.11.	Estadísticos de los errores de predicción de los modelos de referencia.	23
2.12.	Diagrama de cajas de los errores de predicción del modelo M1.	29
2.13.	Histograma de los errores de predicción. Modelo M1.	29
2.14.	Correlograma de los errores de predicción. Modelo M1.	30
2.15.	Diagrama de cajas de los errores de predicción del modelo M2.	30
2.16.	Histograma de los errores de predicción. Modelo M2.	31
2.17.	Correlograma de los errores de predicción. Modelo M2.	31

2.18.	Diagrama de cajas de los errores de predicción del modelo M3.	32
2.19.	Histograma de los errores de predicción. Modelo M3.	32
2.20.	Correlograma de los errores de predicción. Modelo M3.	33
2.21.	Diagrama de cajas de los errores de predicción del modelo M4.	33
2.22.	Histograma de los errores de predicción. Modelo M4.	34
2.23.	Diagrama de cajas de los errores de predicción del modelo M5.	34
2.24.	Histograma de los errores de predicción. Modelo M5.	35
2.25.	Correlograma de los errores de predicción. Modelo M5.	35
2.26.	Diagrama de cajas de los errores de predicción del modelo M6.	36
2.27.	Histograma de los errores de predicción. Modelo M6.	36
2.28.	Correlograma de los errores de predicción. Modelo M6.	37
2.29.	Diagrama de cajas de los errores de predicción del modelo M7.	37
2.30.	Histograma de los errores de predicción. Modelo M7.	38
2.31.	Correlograma de los errores de predicción. Modelo M7.	38
2.32.	Sesgo de los modelos de referencia y los modelos de partida.	39
2.33.	EMAN de los modelos de referencia y los modelos de partida.	39
2.34.	RECMN de los modelos de referencia y los modelos de partida.	40
2.35.	Sesgo de los modelos M.	40
2.36.	EMAN de los modelos M.	41
2.37.	RECMN de los modelos M.	41
2.38.	Coefficiente de mejora respecto al EMAN de los modelos de partida. . .	42
2.39.	Coefficiente de mejora respecto a la RECMN de los modelos de partida.	42
3.1.	Sesgo de los errores de predicción para el mejor de los modelos de referencia, los modelos de partida (etiquetados como iniciales), y los modelos con las matrices estimadas por máxima verosimilitud. El modelo cuyos parámetros siguen un paseo aleatorio se ha etiquetado como RQ-diagonal.	50

3.2.	RECMN de los errores de predicción para el mejor de los modelos de referencia, los modelos de partida (etiquetados como iniciales), y los modelos univariantes con matrices estimadas por máxima verosimilitud. El modelo cuyos parámetros siguen un paseo aleatorio se ha etiquetado como RQdiagonal.	51
3.3.	Sesgo de los errores de predicción para los modelos de referencia persistente y nueva referencia (etiquetado como nr), los modelos de partida (etiquetados como iniciales), y los modelos univariantes con matrices estimadas por máxima verosimilitud. El modelo con la restricción $\phi(1, 1) = 1$ se ha etiquetado como RQ ϕ diagonal $\phi(1, 1) = 1$	53
3.4.	RECMN de los errores de predicción para el mejor de los modelos de referencia, los modelos de partida (etiquetados como iniciales), y los modelos univariantes con matrices estimadas por máxima verosimilitud. El modelo con la restricción $\phi(1, 1) = 1$ se ha etiquetado como RQ ϕ diagonal $\phi(1, 1) = 1$	54
3.5.	Sesgo de los nuevos modelos multivariantes propuestos. El modelo SUC se ha etiquetado como RQ y el modelo VAR(1) como RQ ϕ . También se han incluido los modelos de referencia e iniciales.	58
3.6.	RECMN de los nuevos modelos multivariantes propuestos. El modelo SUC se ha etiquetado como RQ y el modelo VAR(1) como RQ ϕ . También se han incluido los modelos de referencia e iniciales.	59
3.7.	Comparación de resultados. Sesgo del error de predicción de todos los modelos propuestos.	61
3.8.	Comparación de resultados de los modelos propuestos. Coeficiente de mejora de la RECMN respecto a los modelos de partida.	62
3.9.	Mejora de la RECMN de los modelos optimizados para cada horizonte de predicción respecto al modelo con matrices estimadas unicamente para $h=1$	63
3.10.	RECMN para los modelos recursivos.	64

3.11.	Histograma de la distribución bootstrap de los hiperparámetros del modelo SUC M1. Con línea roja continua se representa el valor óptimo del hiperparámetro obtenido mediante el algoritmo EM. Con línea roja discontinua se ha representado el intervalo de confianza obtenido mediante la aproximación normal y con línea verde discontinua el intervalo obtenido mediante el método de los percentiles.	69
3.12.	QQplot de la distribución bootstrap de los hiperparámetros del modelo SUC M1.	70
4.1.	Sesgo de los modelos de partida y distintas alternativas de referencia de combinación de predicciones. OLS1 y OLS2 denotan la combinación de predicciones realizada mediante mínimos cuadrados ordinarios con y sin término independiente respectivamente.	94
4.2.	RECMN de los modelos de partida y distintas alternativas de referencia de combinación de predicciones. OLS1 y OLS2 denotan la combinación de predicciones realizada mediante mínimos cuadrados ordinarios con y sin término independiente respectivamente.	94
4.3.	Sesgo de la combinación de predicciones mediante componentes principales en función del número de factores seleccionados; f1, f2, f3, y f4 denotan uno, dos, tres y cuatro componentes principales respectivamente.	96
4.4.	RECMN de la combinación de predicciones mediante componentes principales en función del número de factores seleccionados; f1, f2, f3, y f4 denotan uno, dos, tres y cuatro componentes principales respectivamente.	96
4.5.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la primera componente principal a horizonte 1.	97
4.6.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la primera componente principal a horizonte 36.	97
4.7.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la primera componente principal a horizonte 36 para los modelos M1 a M7.	98
4.8.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la segunda componente principal a horizonte 1.	99
4.9.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la segunda componente principal a horizonte 36.	99

4.10.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la segunda componente principal a horizonte 36 para los modelos M1 a M7.	100
4.11.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la tercera componente principal a horizonte 1.	100
4.12.	Evolución temporal del coeficiente de ponderación de cada predictor en la formación del tercer componente principal a horizonte 36.	101
4.13.	Sesgo de la combinación de predicciones mediante mínimos cuadrados parciales en función del número de factores seleccionados.	102
4.14.	RECMN de la combinación de predicciones mediante mínimos cuadrados parciales en función del número de factores seleccionados.	102

Índice de cuadros

1.1. Comparativa entre modelos estadísticos y físicos.	7
3.1. Optimización de R y Q diagonales.	72
3.2. Optimización de R, Q y Φ diagonales.	73
3.3. Optimización de R, Q y Φ diagonales con $\Phi(1, 1) = 1, 0$	75
3.4. Optimización de R y Q.	77
3.5. Optimización de R, Q Y Φ	80
3.6. Resultados bootstrap modelo M1 modelo SUC.	86

Capítulo 1

Introducción

En este capítulo se muestra el panorama actual, en cuanto a generación eólica se refiere, en España. Se plantea la necesidad de herramientas de predicción a corto plazo y su utilidad tanto para la gestión y operación eficiente del sistema eléctrico como para la presentación de ofertas en el mercado de energía eléctrica. Se describe el estado tecnológico actual de dichas herramientas de predicción y, por último, se fijan los objetivos y se resume la estructura y contenido de la tesis.

1.1. Introducción

La energía eólica ha sido, en la última década, la tecnología energética de más rápido crecimiento. En países como Alemania, Dinamarca y España ha alcanzado una alta penetración en el sistema eléctrico. En el caso español, es la segunda tecnología que más potencia instalada posee, sólo por detrás del ciclo combinado, cubriendo el 16,4% de la demanda eléctrica, según datos de Red Eléctrica de España (REE). Así mismo, se ha consolidado como la tercera tecnología que más aporta al sistema tras los ciclos combinados y la nuclear. En el año 2010 se instalaron 1.516 MW eólicos, con una potencia que eleva el total acumulado en el territorio nacional a 20.676 MW, según los datos recopilados por la Asociación Empresarial Eólica que confirman a España como cuarta potencia eólica mundial, sólo por detrás de China, Estado Unidos y Alemania. Castilla y León es la Comunidad Autónoma con mayor potencia instalada. Así, en 2010 sumó 917 MW, por lo que cuenta con algo más de 4.800 MW acumulados. Las otras dos comunidades con mayor potencia acumulada son Castilla-La Mancha con 3.709 MW (aunque en 2010 apenas instaló 6 MW) y Galicia que sumó 54,80 MW alcanzado un total de 3.290 MW. Sin embargo, tal cantidad de capacidad eólica instalada puede

tener efectos negativos sobre el sistema eléctrico, que serán más acusados a medida que aumente el grado de penetración de estos sistemas en la red. La principal dificultad es que la energía eólica no puede ser programada y, por lo tanto, siempre hay incertidumbre sobre su producción, lo que puede disminuir la eficacia en la operación del sistema eléctrico.

El sector eólico español ha cumplido con los objetivos que se fijó el Plan de Fomento de las Energías Renovables (PER 2005-2010) y ahora se enfrenta al nuevo reto de alcanzar los 35.750 MW para 2020 que establece el PER 2011-2020. De todos es reconocido que sin el desarrollo de las herramientas de predicción eólica, ni se habrían alcanzado los objetivos del PER 2005-2010, ni se podrán alcanzar los objetivos previstos para el año 2020 ([53]).

El presente capítulo se ha organizado de la siguiente manera: en la sección 1.2 se abordan los aspectos más relevantes de la problemática de la generación y gestión de la energía eléctrica mediante energía eólica. A continuación, en la sección 1.3 se detallan las metas marcadas en este proyecto de tesis. Posteriormente, en la sección 1.4 se describe el estado tecnológico actual junto con una recopilación bibliográfica de trabajos relevantes en este área. Para finalizar, en la sección 1.5 se detalla la estructura de esta tesis.

1.2. Descripción del problema

Durante la operación normal de la red eléctrica, los operadores controlan los centros de generación y reparten las cargas con el objetivo, por un lado, de minimizar el coste de la generación eléctrica y, por otro, maximizar la fiabilidad del sistema. Cada generador de energía se carga en función del coste del combustible que maneja, su rapidez de arranque, su disponibilidad, las pérdidas de transmisión asociadas al mismo, y la velocidad de respuesta. Los parques eólicos utilizan el viento como fuente de energía para generar electricidad. La ventaja de estos sistemas es que no tienen costes de energía primaria y los costes de operación y mantenimiento son bajos, lo que conlleva un muy bajo coste de generación eléctrica. Esto ha motivado, junto con aspectos medioambientales, que estos sistemas sean preferidos para la generación eléctrica y, obligatoriamente, deben ser los primeros en cargarse y operar al 100 % de su capacidad mientras el recurso eólico lo permita. Sin embargo, a diferencia de los generadores convencionales, el operador de la red no tiene control sobre la disponibilidad y producción de estos sistemas, siendo las condiciones climáticas las que dictan las variaciones en la producción de estas fuentes intermitentes de energía. Los sistemas de control convencionales están constantemente regulando la potencia de salida de los generadores eléctricos para seguir las variacio-

nes en la demanda. La inclusión de generadores intermitentes en el sistema implica que los generadores convencionales no sólo deben seguir los cambios en la demanda, sino también las variaciones de la producción de los parques eólicos, lo que provoca un importante impacto en la operación de la red, siendo el resultado un aumento de la variabilidad de la curva de carga a seguir.

Es por ello que, en los últimos años, a medida que aumenta el uso de la energía eólica, va cobrando mayor relevancia conocer de antemano y con cierta precisión, la potencia eólica que va a ser generada. Dada la variabilidad del viento, la generación eólica posee un alto grado de incertidumbre del que carecen otras fuentes de energía cuya generación puede ser más fácilmente controlada y planificada. Esta posibilidad de control y planificación es indispensable para llevar una correcta gestión del sistema eléctrico.

Para intentar subsanar estos inconvenientes, que podrían llegar a frenar el crecimiento de esta energía en nuestro país, el 27 de marzo de 2004 se publicó el Real Decreto 436/2004 de 12 de marzo por el que se estableció la metodología para la actualización y sistematización del régimen jurídico y económico de la actividad de producción de energía eléctrica en régimen especial. En este Real Decreto se establece por primera vez la obligatoriedad, a partir de 1 de enero de 2006, de realizar previsiones de la producción para las instalaciones de generación acogidas al régimen especial cuya potencia instalada sea mayor de 10 MW, tanto en el caso de vender la energía mediante tarifa regulada¹ como en el caso de vender la energía en el mercado eléctrico. Además, este Real Decreto fija penalizaciones económicas de las desviaciones de la previsión respecto a la producción real mayores del 20%. Es este el punto de partida y es en este contexto donde se ha reconocido el desarrollo de herramientas de predicción eólica con un horizonte de hasta dos días, como un factor clave para la integración a gran escala de los sistemas eólicos.

Por su parte, el Real Decreto 661/2007 introduce en lo que se refiere a la predicción, el pago de todas las desviaciones desapareciendo el umbral del 20% sin penalización.

¹Artículo 22. Mecanismos de retribución de la energía eléctrica producida en régimen especial. Para vender su producción o excedentes de energía eléctrica, los titulares de instalaciones a los que resulte de aplicación este real decreto deberán elegir entre una de las dos opciones siguientes:

1. Ceder la electricidad a la empresa distribuidora de energía eléctrica. En este caso, el precio de venta de la electricidad vendrá expresado en forma de tarifa regulada, única para todos los períodos de programación, expresada en céntimos de euro por kilovatio-hora.
2. Vender la electricidad libremente en el mercado, a través del sistema de ofertas gestionado por el operador de mercado, del sistema de contratación bilateral o a plazo o de una combinación de todos ellos. En este caso, el precio de venta de la electricidad será el precio que resulte en el mercado organizado o el precio libremente negociado por el titular o el representante de la instalación, complementado por un incentivo y, en su caso, por una prima, ambos expresados en céntimos de euro por kilovatio-hora.

En resumen, las ventajas principales del desarrollo de herramientas de predicción eólica son las siguientes:

1. Integrar de manera eficaz más potencia eólica en el sistema eléctrico, sin poner a este en situaciones de inestabilidad.
2. Reducir los costes de programación y operación, tanto del gestor de la red de transporte, como de las compañías distribuidoras.
3. Obtener mayores ingresos debido a una mejor negociación en el mercado eléctrico.
4. A medida que aumenten la fiabilidad y exactitud en las predicciones, las compañías explotadoras de la red permitirán un mayor grado de penetración de esta energía en el sistema eléctrico debido a que la gestión de la red se verá facilitada.
5. Optimizar las tareas de operación y mantenimiento no sólo de los parques eólicos sino también de las redes de transporte a las que están conectados.

1.3. Objetivos de esta tesis

Hasta hace relativamente poco tiempo las series de viento y potencia eólica no han sido objeto de estudio. Entre las razones podemos citar:

- Lo tedioso que resultaba la manipulación de estas series y sus modelos.
- La carencia de herramientas informáticas software adecuadas.
- La falta de incentivos económicos en el sistema que justificase la inversión en su estudio.
- Los grandes volúmenes de información que generan.
- Son series que han aparecido recientemente, y todavía más reciente es la disponibilidad de las mismas y la necesidad de estimar la producción energética.

Sobre la base de todo lo anterior se justifica un esfuerzo tendente a un mejor conocimiento de las características esenciales de estas series, y el desarrollo de técnicas de modelado que permita un tratamiento simple, general y lo más automático posible.

El objetivo de esta tesis doctoral, en línea con lo anterior, es encontrar una familia de modelos que permita que entorno a ella se pueda desarrollar una metodología automática de modelado de series horarias de generación eléctrica en parques eólicos. Se pretende

conseguir modelos de altas prestaciones en la predicción a corto plazo (1 a 2 días). Todo ello dentro de un marco general de desarrollo y análisis.

La consecución del objetivo anterior obliga a un proceso de investigación que se puede estructurar en una serie de objetivos parciales que, sintetizando, aparecen recogidos en la siguiente enumeración:

- Explorar las principales características particulares que presentan las series objeto de estudio.
- Analizar la bondad de los resultados de la aplicación a estas series de las metodologías y técnicas actualmente existentes.
- Estudiar nuevos planteamientos y enfoques de modelado buscando el desarrollo de una metodología completa aplicada a herramientas que operan en tiempo real y con intervención mínima por parte del operario.
- Dentro de esa metodología general, profundizar en el estudio de aquellas características particulares de especial impacto. Se tendrán que diagnosticar aquellos aspectos principales a analizar de forma diferenciada, es decir, el tratamiento dentro del modelo del
 1. efecto de la predicción meteorológica,
 2. la combinación de predicciones,
 3. el ensayo de nuevas familias de modelos multivariantes
- Proceder al desarrollo y programación de todo el marco anterior de análisis, con la mirada puesta en la mayor automatización posible de todo el proceso, de forma que permita de forma simple un tratamiento masivo de series.
- Mostrar la bondad de los resultados en un parque eólico como el de Sotavento situado en terreno complejo y con la dificultad añadida de estar formado por aerogeneradores de distinta tecnología.

El estudio se iniciará tomando como base modelos de series temporales actualmente en operación para la realización de predicciones de parques eólicos, a partir de los cuales identificar las posibles fuentes de mejora y profundizar en su modelización.

1.4. Estado Tecnológico Actual

Una primera clasificación de los modelos de predicción es en función de la aplicación a la que se destinen, o lo que es lo mismo, según el horizonte de predicción y el periodo de muestreo de las series (véase [18]).

- Modelos de muy corto plazo. El objetivo al que se destinan es el control de los aerogeneradores. El periodo de muestreo o de monitorización es de segundos o minutos y los modelos estadísticos más simples (persistentes y autorregresivos) son los más comúnmente empleados.
- Modelos a corto plazo. El objetivo es la integración eficiente en la red eléctrica para la gestión óptima de ésta o bien la participación en el mercado eléctrico. El horizonte de predicción es de uno o dos días y el periodo de muestreo más comúnmente utilizado es el periodo horario aunque hay algún modelo con periodos de muestreo de media hora o 15 minutos.
- Modelos a largo plazo. El objetivo principal de estos modelos es la estimación anual de la producción energética de los parques o el apoyo a las tareas de planificación del mantenimiento. El periodo de muestreo suele ser horario, aunque los primeros desarrollos comenzaron con series de datos diarios.

Esta tesis se va a centrar en el estudio de las herramientas de predicción en el corto plazo (uno o dos días) por lo que serán estos modelos los que analicemos a continuación. Aunque existe una variedad de metodologías de predicción empleadas, todas comparten características similares. Su denominador común es que emplean datos proporcionados por los pronósticos de tiempo de algún Instituto de Meteorología y datos operativos de los parques eólicos. Los dos conjuntos de datos se combinan para proporcionar una predicción de la producción eólica. Los métodos de predicción eólica tienen como objetivo transformar las predicciones meteorológicas de velocidad de viento en predicciones de potencia eólica. Es de destacar que las predicciones originales de los modelos numéricos de predicción meteorológica (MNPM) presentan errores considerables y sistemáticos en las capas próximas al suelo ya que estos modelos son típicamente de escala regional, nacional, continental o incluso global y, por consiguiente, la resolución no es la óptima para satisfacer las necesidades de la industria eólica. En la actualidad, existen varios modelos meteorológicos desarrollados. Un modelo numérico de predicción meteorológica permite, partiendo de un estado inicial con unas características atmosféricas dadas, simular la evolución atmosférica mediante la traducción, a través de métodos numéricos, de las leyes de la mecánica de fluidos y de la termodinámica.

Existen además metodologías de distinta naturaleza para alcanzar la transformación entre las predicciones de viento del MNPM y las predicciones de potencia necesarias para el parque eólico. La clasificación clásica de los distintos métodos los agrupa en tres grandes categorías ([17])

- Modelos físicos
- Modelos estadísticos y
- Una combinación de ambos.

En la actualidad existen varias revisiones muy completas y extensas del estado del arte. Destacamos [5], [16], [26] y [33], por lo que únicamente presentaremos un breve resumen de este aspecto. Los modelos físicos tienen como objetivo principal mejorar la resolución del MNPM original, es decir, utilizan las predicciones meteorológicas para inicializar un modelo físico a escala local que origine predicciones de velocidad de viento y de producción eólica específicas para el emplazamiento en cuestión, introduciendo información sobre la topografía (orografía, rugosidad, obstáculos, etc) y los efectos físicos de microescala. Los modelos estadísticos se basan en minimizar alguna función de las diferencias sistemáticas entre los datos reales y los predichos o ajustados dentro del modelo. En la actualidad lo más empleado es una combinación de ambos. El resultado del modelo físico se utiliza como variable de entrada junto con los datos históricos en un modelo estadístico. Según el método utilizado, estadístico o físico, se pueden encontrar una serie de ventajas e inconvenientes, que se resumen en el cuadro 1.1.

CRITERIO	MODELO ESTADISTICO	MODELO FISICO
Requisitos informáticos	No requiere mucha potencia de cálculo	Es esencial una gran potencia de cálculo
Implementación	Adaptable y transferible entre distintos emplazamientos	Es necesario un estudio particularizado para cada emplazamiento
Tiempos de cálculo	Segundos/minutos. Optimo para mercados intradiarios.	Horas. Más conveniente para mercados diarios
Precisión	Muy buena precisión a muy corto plazo (hasta 6-12 horas)	Precisión constante a lo largo de todo el periodo de programación
Autoaprendizaje	Implementado automáticamente en los algoritmos	Necesita intervención del analista meteorólogo para realizar los cambios

Cuadro 1.1: Comparativa entre modelos estadísticos y físicos.

Los sistemas de predicción más conocidos basados en modelos físicos son el denominado Previento ([3], [13]) y el denominado Prediktor ([23] y [24]). Recientemente,

propiciado por el marco legislativo en España, se han desarrollado varios modelos físicos entre los que podemos citar el GH Forecaster desarrollado por la empresa Garrad Hassan ([34]), Casandra ([14] y [12]) desarrollado por la empresa Gamesa, LocalPred desarrollado por CENER ([27]), y RAMS desarrollado por Iberinco ([29]). En nuestro caso, nos decantaremos por los modelos estadísticos por varias razones. En primer lugar, porque los modelos físicos necesitan un desarrollo particularizado a cada emplazamiento que impide la realización de su ajuste o calibración de una manera automática; en segundo lugar, porque la relación entre la velocidad de viento y la potencia generada es altamente no lineal y variante con el tiempo (esta relación depende de otras muchas variables como son la dirección del viento, la densidad del aire, las variaciones locales de temperatura, efectos locales de nubes, lluvia, etc.). Muchas de estas variables son muy difíciles de medir cuando no imposible, por lo que es muy complicado incluirlas de manera implícita en el modelo. En [28] se analiza la mejora de las predicciones de velocidad de viento y potencia empleando el filtro de Kalman a la predicción de velocidad de viento obtenida a través de dos modelos meteorológicos distintos y además a distintas resoluciones espaciales. La conclusión es que la aplicación del filtro de Kalman mejora los resultados de los modelos (sesgo, valor absoluto del sesgo y desviación típica del sesgo) para todos los horizontes de predicción en todos los modelos. Así mismo, se obtiene que la aplicación del filtro de Kalman a la predicción de velocidad de viento obtenida con el modelo físico de resolución espacial de 12 km es mejor que la predicción obtenida por el modelo físico de resolución espacial de 0.5 km, con la consecuente disminución del tiempo de ejecución de los modelos. Una de sus conclusiones es, por tanto, que la ejecución de algoritmos meteorológicos para mallas de resolución menor de 6 km, que exigen gran recurso y tiempo de cálculo, pueden evitarse empleando métodos estadísticos adaptativos como el filtro de Kalman. La utilización de modelos estadísticos adaptativos posee además otra ventaja ya que es capaz de detectar y 'seguir' cambios en las relaciones entre las diversas variables, tales como variaciones en la rugosidad en el terreno, o en el modelo numérico de predicción meteorológica. A modo de ejemplo, cabe citar el cambio que experimentó Hirlam en Septiembre de 1997, que tuvo efectos muy significativos sobre la precisión de las predicciones [20].

A continuación se describen brevemente los modelos estadísticos que son en los que se basa esta tesis.

En cuanto a los modelos basados en técnicas estadísticas, el primer desarrollo se denominó Wind Power Prediction Tool (WPPT), aunque en la actualidad se denomina Zephyr([25], [38], [39] y [15]). En España, se ha desarrollado Sipleólico ([46]) que opera en el Centro de Control de Red Eléctrica de España (REE). También existen

varios intentos de modelado con redes neuronales. Los trabajos más destacables son los realizados por G. Kariniotakis aunque enfocados a sistemas eléctricos aislados, lo cual simplifica el objetivo respecto a los desarrollos anteriores por cuanto el horizonte de predicción se reduce a 8-10 horas ([22]).

Zephyr utiliza un modelo no paramétrico, basado en regresiones polinómicas locales (véase [21] y [37]), que emplea la información de la potencia observada y las predicciones de la velocidad y dirección de viento. En Sipleólico se emplea un sistema más sofisticado. En vez de utilizar un único modelo se utilizan nueve modelos alternativos que se estiman recursivamente mediante Mínimos Cuadrados Recursivos y Filtro de Kalman, es decir, para cada parque se estiman 18 modelos. La predicción final se realiza mediante una combinación lineal de los predictores que se ponderan en función de su actual error de predicción (véase [46], [47]). Todos estos modelos necesitan de predicciones meteorológicas numéricas para la estimación de la potencia eólica. Sipleólico, Zephyr y Prediktor emplean las predicciones Hirlam que proporcionan los Institutos Meteorológicos Europeos. En todos los casos, los errores de predicción aumentan al aumentar el horizonte de predicción. Todos los autores justifican estos errores en base a las limitaciones que tienen estas predicciones meteorológicas. Para intentar subsanar estas deficiencias se ha desarrollado un módulo denominado LocalPred ([31] y [32]) que intenta mejorar la predicción Hirlam, sobre todo en terreno complejo, a partir de la información extraída de otras variables Hirlam como diferencias de presiones o temperaturas mediante componentes principales. En Zephyr se utiliza un predictor no paramétrico al igual que en los modelos denominados M8 y M9 en Sipleólico. Estos modelos tienen como desventajas para el caso que nos ocupa, primero, que su ajuste necesita mucho tiempo de cálculo, por lo que en ambos casos lo que se realiza es una evaluación de las funciones únicamente en determinados puntos y para el resto de valores de las entradas se interpolan las funciones; en segundo lugar, es necesaria una calibración para cada emplazamiento utilizando los datos históricos, por lo que se descartan inicialmente en nuestro objetivo de desarrollar una herramienta de modelado lo más automática posible. Todas las herramientas recurren a una estimación adaptativa recursiva. Las dos razones principales son, por un lado, el comportamiento no lineal que provoca que los modelos sean únicamente aproximaciones válidas en un cierto entorno de los datos, y por otro lado, las curvas de potencia cambian con el tiempo en respuesta a cambios meteorológicos y por el desgaste propio de las máquinas, por lo que los parámetros no pueden considerarse constantes, y deben adaptarse a la última información disponible del proceso.

1.5. Estructura y contenidos de la tesis

La tesis consta de 5 capítulos. En este primer capítulo se introduce y motiva el problema a abordar y se fijan los objetivos de la tesis. Los principales resultados de la tesis se encuentran en los capítulos 2 y, especialmente, 3 y 4. En el capítulo 2 se presentan los datos y modelos de partida así como los estadísticos y modelos de referencia empleados para la evaluación de nuevas herramientas de predicción y para la comparación entre varias de ellas. En el capítulo 3 se estiman los parámetros de los modelos de partida mediante máxima verosimilitud así como se presentan nuevos modelos de naturaleza multivariante. En ambos casos se demuestra su mejor comportamiento frente a los modelos de partida. El capítulo 4 presenta distintas técnicas multivariantes de reducción de datos como estrategia de combinación de predicciones para la obtención de una única predicción final. El objetivo es conseguir que la predicción final, obtenida mediante combinación, tenga mejor comportamiento (menor error cuadrático medio y sesgo) que las predicciones de partida. Por último, en el capítulo 5, se resumen las conclusiones así como las principales contribuciones de esta tesis para finalmente identificar y argumentar las posibles líneas de investigación futura.

Capítulo 2

Presentación de los datos y modelos de partida

Este capítulo se dedica a analizar el funcionamiento de los modelos de partida. Para ello, primero se hace una presentación de los datos de trabajo y se describen los modelos de referencia empleados en predicción eólica como modelos extremadamente sencillos a batir. Estos modelos, aunque muy simples, son difíciles de superar, especialmente en los primeros horizontes de predicción, como se verá a lo largo de este capítulo. Se describen las medidas del error comúnmente empleadas para medir el grado de precisión de cada uno de los modelos y se presentan los resultados de los modelos de referencia respecto a dichas medidas del error. Por último, se describen los modelos basados en filtro de Kalman presentes en Sipleólico y los resultados de estos modelos y su comparación respecto a los modelos de referencia.

2.1. Introducción

2.1.1. Presentación de los datos de trabajo

Para el desarrollo de esta tesis se han analizado los datos procedentes del parque eólico de Sotavento que se encuentra ubicado en Galicia, concretamente en A Serra da Loba, al norte de las provincias de Lugo y A Coruña, con vientos medios entre 6.5 y 7.5 m/s. Este parque comienza a funcionar como planta de producción de energía eólica en el año 2000 con 24 aerogeneradores de diferentes tecnologías (5) y fabricantes (9), con una potencia instalada total de más de 17.56 MW. Como parque experimental, el papel de Sotavento trasciende la producción energética para servir como entorno

para la educación e investigación sobre las energías renovables en general y la energía eólica en particular. Tanto la descripción del parque como de los datos meteorológicos se muestran en [41]. Es de destacar que este parque eólico, dada la orografía y la presencia de aerogeneradores de diferentes tecnologías, presenta una curva de parque (relación entre la potencia eólica producida y la velocidad de viento incidente) más dispersa que la de otros parques, por lo que puede considerarse a efectos de modelado y /o predicción como complejo. Se dispone de dos conjuntos de datos. El primer grupo de datos procede de medidas in situ en el propio parque y es la potencia de salida en kW, (ver figura 2.1).

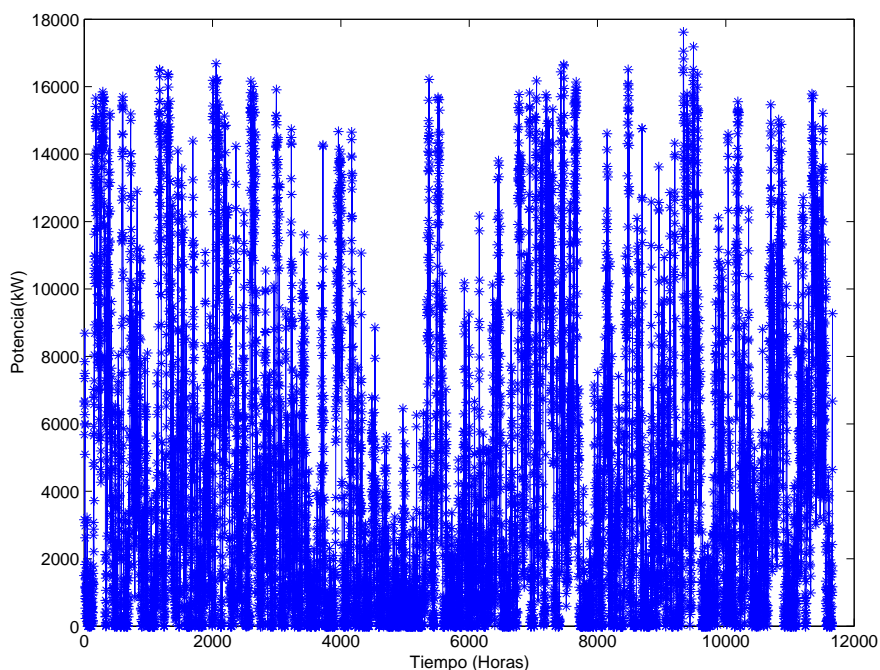


Figura 2.1: Potencia eólica de salida.

El segundo grupo de datos son las predicciones de velocidad (m/s) y dirección (grados) del viento obtenidas mediante un sistema de predicción numérica de alta resolución (10 km) con el modelo ARPS (Advanced Regional Prediction System) en dos pasadas diarias (00Z y 12Z) incluidas en un ciclo de asimilación de datos 6-horario y proporcionadas por MeteoGalicia. Los datos se presentan en valores medios horarios y en horario GMT (Greenwich Mean Time) que se indica por el sufijo Z en el valor numérico de las horas. Se suministran estas variables en los 4 puntos de la malla que rodean al parque eólico. El modelo ARPS es un modelo atmosférico desarrollado en el Center for Analysis and Prediction of Storms (CAPS) de la Universidad de Oklahoma. Este modelo no hidrostático incluye las parametrizaciones físicas suficientes para la descripción de los movimientos atmosféricos en una topografía tan compleja como la gallega. Siguiendo las indicaciones de MeteoGalicia se utilizan los datos del punto 4 de la malla que es el

punto meteorológicamente más próximo al parque eólico (influye no sólo la distancia geométrica entre el parque y el nodo de la malla, sino también la similitud orográfica, principalmente, la altitud).

Para el tratamiento de estos datos hay que tener en cuenta que los datos de potencia son valores medios horarios y se suministran en horario oficial y que los datos procedentes del modelo meteorológico son valores en horario GMT.

Es decir, el conjunto de datos para un instante t determinado es el siguiente:

- Potencia del parque eólico p_t en kW.
- Matriz de predicciones meteorológicas

$$\begin{pmatrix} v_{t+1,t} & v_{t+2,t} & v_{t+3,t} & \dots & v_{t+36,t} \\ d_{t+1,t} & d_{t+2,t} & d_{t+3,t} & \dots & d_{t+36,t} \end{pmatrix}$$

donde $v_{t+i,t}$ y $d_{t+i,t}$ con $i = 1, 2, \dots, 36$ son las predicciones de velocidad y dirección de viento proporcionadas por el modelo de predicción meteorológico, disponibles en t para el instante $t + i$. Como aclaración para entender mejor la operativa de cálculo del modelo meteorológico dentro de la herramienta de predicción eólica, hay que tener en cuenta que este modelo numérico de predicción meteorológica, se ejecuta dos veces al día (a las 00Z y 12Z horas) y tarda en realizar los cálculos entre 8 y 11 horas, es decir, los primeros valores de predicción no son válidos en el momento en que están disponibles pues corresponden a datos ya pasados. En cada periodo de muestreo t , seleccionamos las predicciones meteorológicas del último fichero meteorológico recibido para los instantes $t+1$ hasta $t+36$ que denotamos por $v(t+1, t)$, $v(t+2, t)$, $v(t+3, t)$, \dots , $v(t+36, t)$, aunque respecto al modelo meteorológico y refiriéndonos al horizonte de predicción respecto al cálculo realizado por dicho modelo, este horizonte es mayor (e igual a la diferencia entre el tiempo actual + h y las 00Z o 12Z horas). Por tanto, el menor horizonte de predicción de la estimación de velocidad y dirección de viento es, en realidad, 1 hora más el tiempo de ejecución del modelo meteorológico (entre 8 y 11 horas).

Después de depurar los ficheros de datos históricos se trabaja con una muestra de datos horarios comprendida entre el 17 de febrero de 2004 a las 00 horas hasta el 30 de abril de 2005 a las 11 horas. Para el ajuste de los modelos se ha dividido el conjunto en un primer tramo de 8760 datos (un año) que se utilizarán en la estimación de los modelos y el resto se ha empleado para la etapa de predicción.

2.2. Evaluación de los modelos

2.2.1. Introducción

Dado el coste de los sistemas de predicción y el impacto económico que tiene su precisión, se ha detectado la necesidad, por parte de los usuarios finales, de establecer una metodología estándar para la evaluación de dichos sistemas. Dicha metodología se ha desarrollado dentro del proyecto europeo ANEMOS, puede consultarse en [30] y es la que se va a seguir en esta tesis. Tal y como se recomienda en dicha publicación, para poder comparar los resultados obtenidos por los distintos modelos en diferentes parques eólicos, se divide la potencia entre la potencia nominal del parque y la velocidad de viento entre la velocidad máxima, así todos los valores están comprendidos entre cero y uno y, por tanto, son independientes del tamaño del parque. En esta sección, se describen los modelos de referencia que se emplean como punto de partida y los estadísticos y medidas del error comúnmente empleados en predicción eólica que conforman la base del sistema de evaluación de los modelos de predicción.

2.2.2. Modelos de Referencia

Es útil desarrollar una herramienta de predicción si el modelo obtenido es capaz de mejorar los resultados que se obtienen con los modelos de referencia, los cuales son el resultado de consideraciones simples. El modelo de referencia más comúnmente empleado en el marco de las herramientas de predicción eólica es el modelo persistente (que viene a ser el paseo aleatorio sin deriva). Este modelo tan sencillo establece que la predicción de potencia es igual al valor de la última potencia medida, es decir,

$$\hat{p}_{t+h,t} = p_t. \quad (2.1)$$

Pese a la extrema sencillez de este modelo, es difícil mejorarlo en los primeros horizontes de predicción (hasta 6 horas) debido a que la escala temporal de cambios en la atmósfera es muy lenta, del orden de varios días. Una generalización del modelo persistente son los predictores de medias móviles, en el cual se reemplaza la última potencia medida por la media de las n últimas medidas disponibles

$$\hat{p}_{t+h,t} = \frac{1}{n} \sum_{i=0}^{n-1} p_{t-i}. \quad (2.2)$$

Asintóticamente, este modelo tiende a la media global del proceso, es decir, $\hat{p}_{t+h,t} \rightarrow$

\bar{p} siendo \bar{p} la media del proceso. En los primeros horizontes de predicción suele tener una precisión muy pobre, pero mejora conforme aumenta el horizonte de predicción. La exactitud de estos modelos ha sido analizada por [36], y a raíz de este estudio los autores propusieron el siguiente modelo como referencia:

$$\hat{p}_{t+h,t} = a_h p_t + (1 - a_h) \bar{p}_t \quad (2.3)$$

siendo a_h el coeficiente de correlación entre p_t y p_{t+h} y \bar{p}_t la media del proceso hasta el instante t . Los valores \bar{p}_t y a_h se determinan con la serie de datos de entrenamiento. A lo largo de este trabajo nos referiremos a este modelo con las siglas nr (nueva referencia).

2.2.3. Medidas de Error

El criterio fundamental para evaluar el grado de precisión de un modelo de predicción se basa en examinar las diferencias entre las producciones predichas y las reales. Para cada horizonte de predicción h se define el error de predicción como

$$e_{t+h,t} = p_{t+h} - \hat{p}_{t+h,t}. \quad (2.4)$$

Para independizar esta medida del tamaño del parque eólico, se normaliza como porcentaje de la potencia nominal del mismo si los datos no han sido previamente normalizados.

$$\epsilon_{t+h,t} = \frac{p_{t+h} - \hat{p}_{t+h,t}}{P_{instalada}}. \quad (2.5)$$

Una vez definidos los errores de predicción, existen diversos estadísticos que se pueden utilizar para caracterizar el error para cada horizonte de predicción h . Los que se van a utilizar en esta tesis y que son los comúnmente empleados en predicción eólica son:

- Sesgo: Es una medida del error sistemático y se define como

$$sesgo_{t+h,t} = \bar{e}_{t+h} = \bar{e}(h) = \frac{1}{N_h} \sum_{t=1}^{N_h} e_{t+h,t} \quad (2.6)$$

siendo N_h el número de datos de error.

- Sesgo normalizado

$$sesgo_normalizado_{t+h,t} = \bar{\epsilon}_{t+h} = \bar{\epsilon}(h) = \frac{1}{N_h} \sum_{t=1}^{N_h} \epsilon_{t+h,t}. \quad (2.7)$$

- Error Medio Absoluto (EMA)

$$EMAN_{t+h,t} = \frac{\sum_{t=1}^{N_h} |e_{t+h,t}|}{N_h}. \quad (2.8)$$

- Error Medio Absoluto Normalizado (EMAN) o Error de Estado

$$EMAN_{t+h,t} = \frac{\sum_{t=1}^{N_h} |\epsilon_{t+h,t}|}{N_h}. \quad (2.9)$$

- Error Cuadrático Medio (ECM)

$$ECMN_{t+h,t} = \frac{\sum_{t=1}^{N_h} (e_{t+h,t})^2}{N_h} \quad (2.10)$$

- Error Cuadrático Medio Normalizado (ECMN)

$$ECMN_{t+h,t} = \frac{\sum_{t=1}^{N_h} (\epsilon_{t+h,t})^2}{N_h}. \quad (2.11)$$

- Raíz del Error Cuadrático Medio (RECM)

$$RECM_{t+h,t} = \sqrt{\frac{\sum_{t=1}^{N_h} (e_{t+h,t})^2}{N_h}}. \quad (2.12)$$

- Raíz del Error Cuadrático Medio Normalizado (RECMN)

$$RECMN_{t+h,t} = \sqrt{\frac{\sum_{t=1}^{N_h} (\epsilon_{t+h,t})^2}{N_h}}. \quad (2.13)$$

Estadísticamente, el sesgo está asociado con el momento de orden uno del error de predicción, por lo que es una medida relacionada directamente con la energía producida. El error cuadrático medio está asociado con el momento de orden dos, y por lo tanto con la varianza del error de predicción. En este último estadístico, los mayores errores de predicción tienen mayor influencia.

Es interesante poder comparar la mejora que representa uno de los modelos desarrollados frente a los modelos que se emplean como referencia. La ganancia, medida como

mejora respecto al modelo de referencia, se denota como

$$mejora_{ref,CE}(h) = \frac{CE_{ref}(h) - CE(h)}{CE_{ref}(h)} \quad (2.14)$$

donde CE denota Criterio de Evaluación, que puede ser EMAN, ECMN, o bien, RECMN.

2.2.4. Resultados de los modelos de referencia

Los resultados de estos modelos se estudian en los siguientes apartados. En primer lugar, se analiza cada uno de los modelos individualmente y, posteriormente, se comparan diversos estadísticos del error. Para el análisis individual se evalúan los siguientes aspectos:

1. Diagrama de cajas de los errores.
2. Histograma de los errores y se compara éste con la distribución normal.
3. Correlograma de los errores.

Por razones de espacio y brevedad se ilustran únicamente las gráficas correspondientes a los horizontes de predicción 1, 2, 24 y 36.

Resultados del modelo persistente

El diagrama de cajas de los errores se muestra en la figura 2.2. En dicha figura puede comprobarse que los errores parecen centrados en torno a cero aumentando al rango intercuartílico y el número de valores extremos conforme aumenta el horizonte de predicción.

En los histogramas de los errores de predicción (figura 2.3) puede verse cómo la distribución de los mismos es simétrica con colas más pesadas que las de la distribución normal, lo que indica cierto grado de curtosis. Además, la distribución de los mismos cambia a medida que aumenta el horizonte de predicción. Para horizontes bajos los errores aparecen más concentrados en torno a cero, siendo las distribuciones más aplanadas (aumenta la dispersión) a horizontes de predicción mayores.

En el correlograma de los errores de predicción a horizonte 1 (véase la figura 2.4) se observa cierta autocorrelación, consecuencia lógica de este tipo de modelos donde no se explota la estructura autocorrelación existente en los datos.

2.2. Evaluación de los modelos

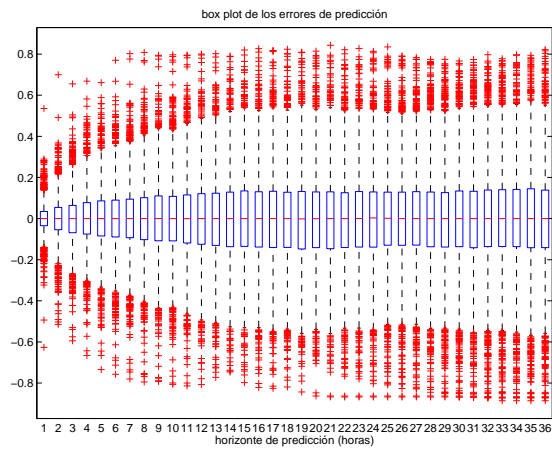


Figura 2.2: Diagrama de cajas de los errores de predicción del modelo persistente para cada horizonte de predicción.

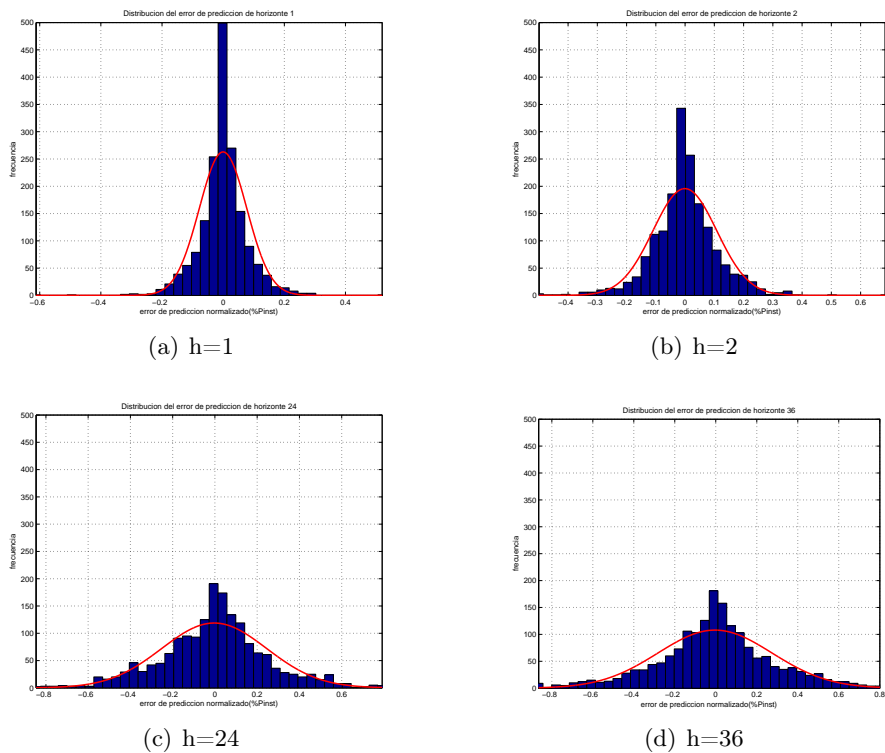


Figura 2.3: Histograma de los errores de predicción. Modelo persistente.

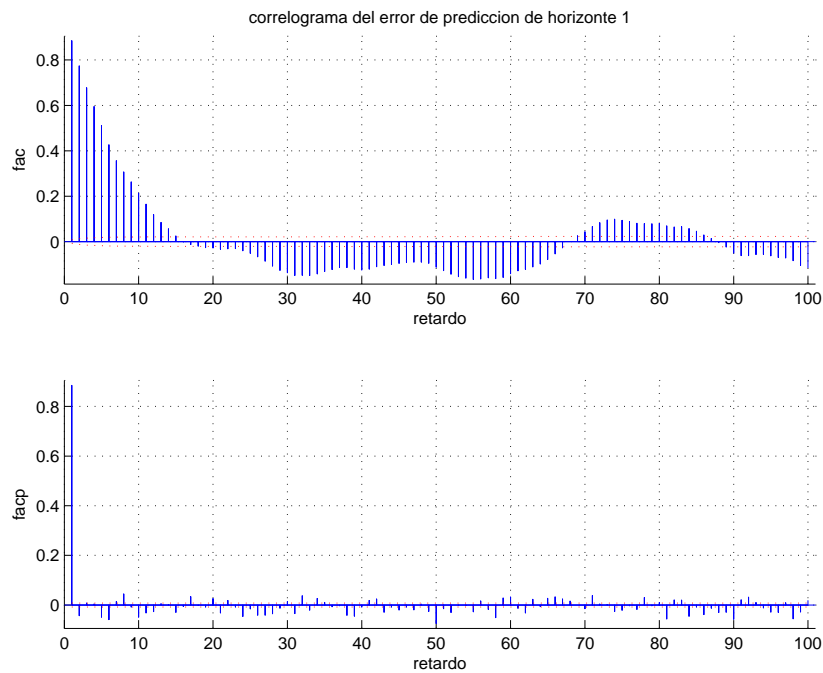


Figura 2.4: Correlograma del error de predicción a horizonte 1. Modelo persistente.

Resultados del modelo media

El diagrama de cajas de los errores se muestra en la figura 2.5. En dicha figura puede comprobarse cómo la distribución de los errores es independiente del horizonte de predicción (consecuencia del hecho de que el modelo es constante e idéntico para todos los horizontes). Lo mismo se observa en los histogramas (figura 2.6) y en el correlograma de los errores a horizonte 1 (figura 2.7). Además, los errores no están centrados en torno a cero, es decir, presentan cierto sesgo positivo, y su distribución no es simétrica. Como en el caso anterior, el modelo no explota la naturaleza autorregresiva de los datos originales, por lo que los errores presentan autocorrelación como se puede observar en la figura 2.7.

Resultados del modelo nr

El diagrama de cajas de los errores de predicción (figura 2.8) muestra cómo aumenta el rango intercuartílico con el horizonte de predicción. En las gráficas de los histogramas (véase la figura 2.9) se puede observar el aumento del sesgo y de la varianza con el horizonte de predicción. Asimismo, se observa que la distribución de los errores es no simétrica siendo la cola de la derecha de valores más elevados (en valor absoluto), mientras que la cola de la izquierda presenta menores valores absolutos pero mayor

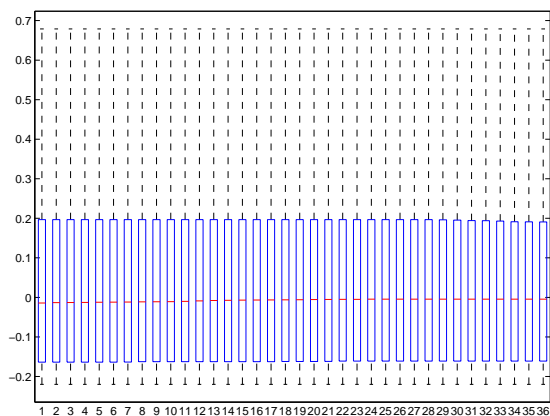


Figura 2.5: Diagrama de cajas de los errores de predicción para el modelo media en función del horizonte de predicción.

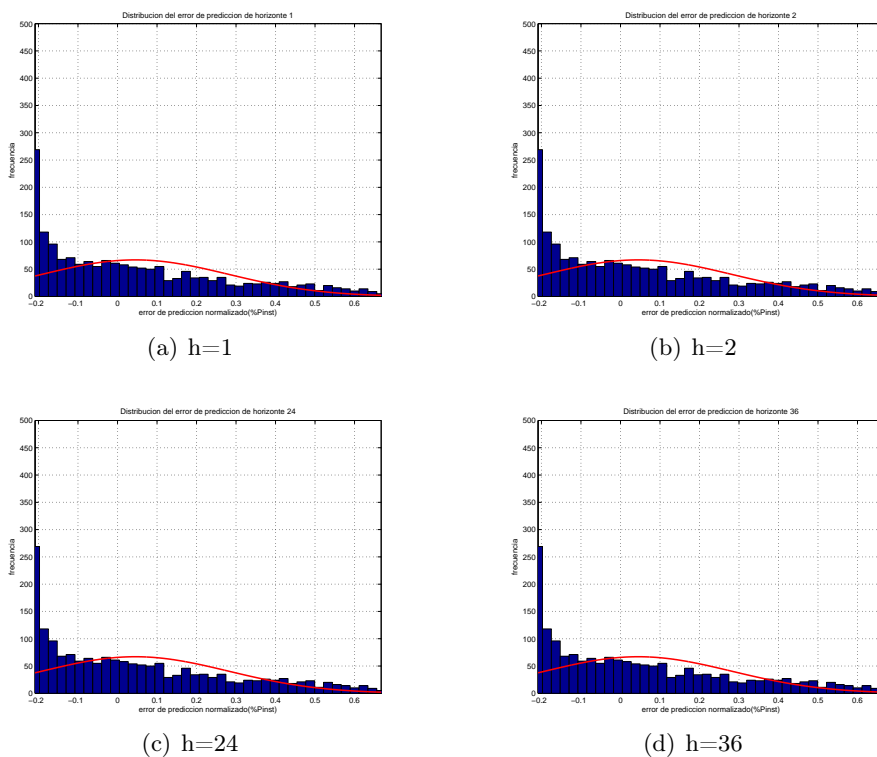


Figura 2.6: Histograma de los errores de predicción. Modelo media.

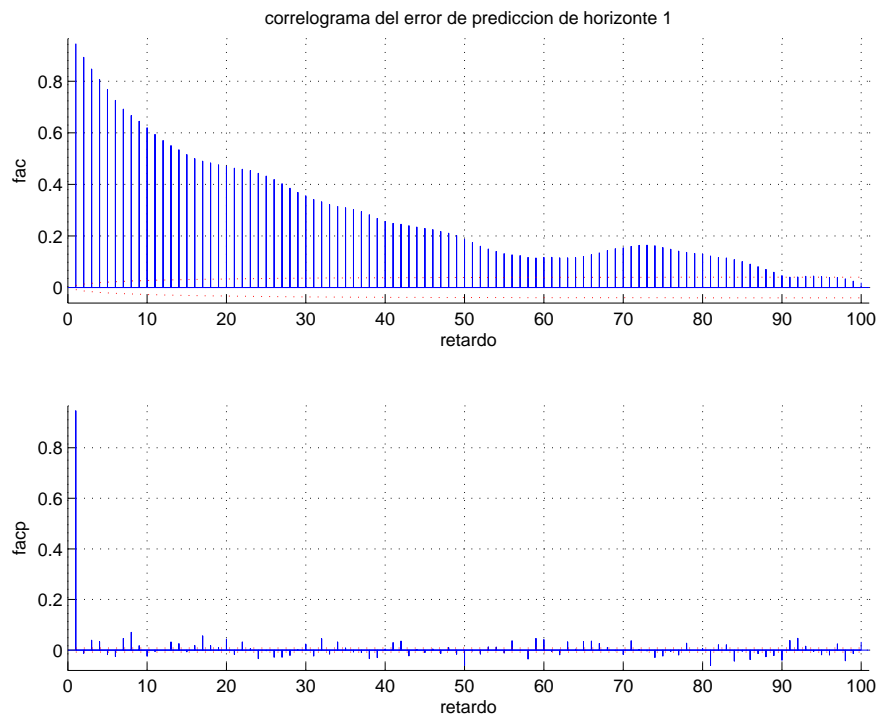


Figura 2.7: Correlograma del error de predicción a horizonte 1. Modelo media.

frecuencia, es decir, los modelos sobrestiman el valor de la potencia más veces aunque con un error de menor magnitud. En el correlograma de los errores de predicción a horizonte 1 (figura 2.10) se puede observar, al igual que en los dos casos anteriores, que existe autocorrelación en los mismos.

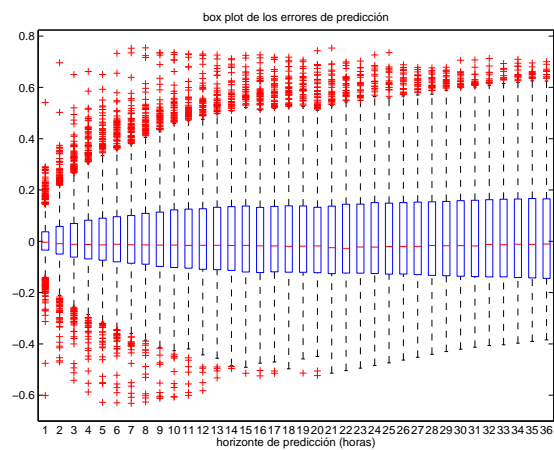


Figura 2.8: Diagrama de cajas de los errores de predicción del modelo AR(1) de parámetros constantes para cada horizonte de predicción.

2.2. Evaluación de los modelos

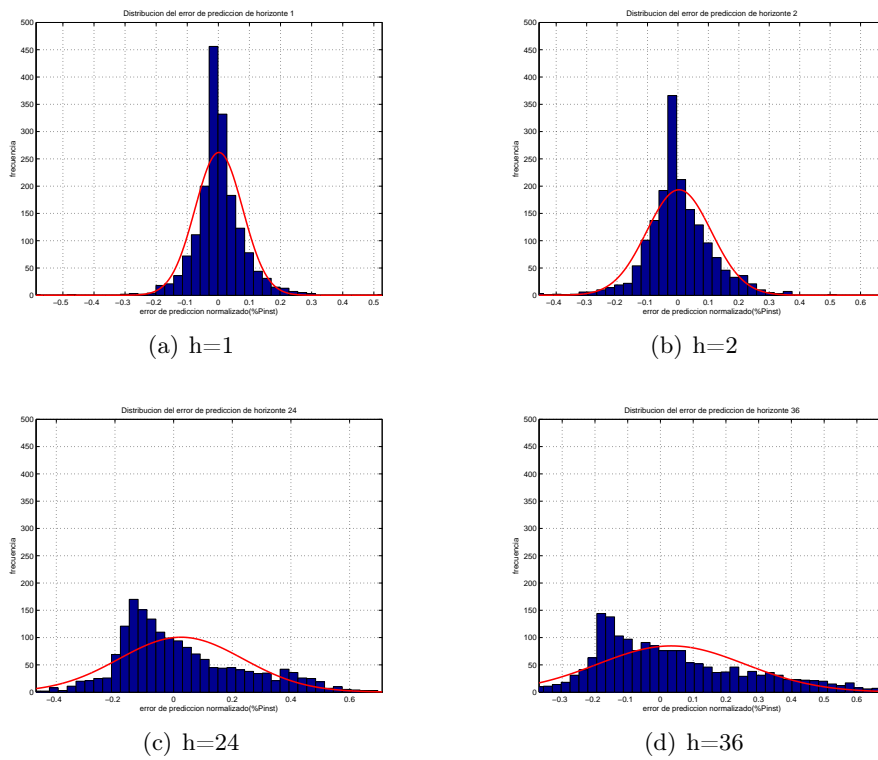


Figura 2.9: Histograma de los errores de predicción. Modelo AR(1) de parámetros constantes.

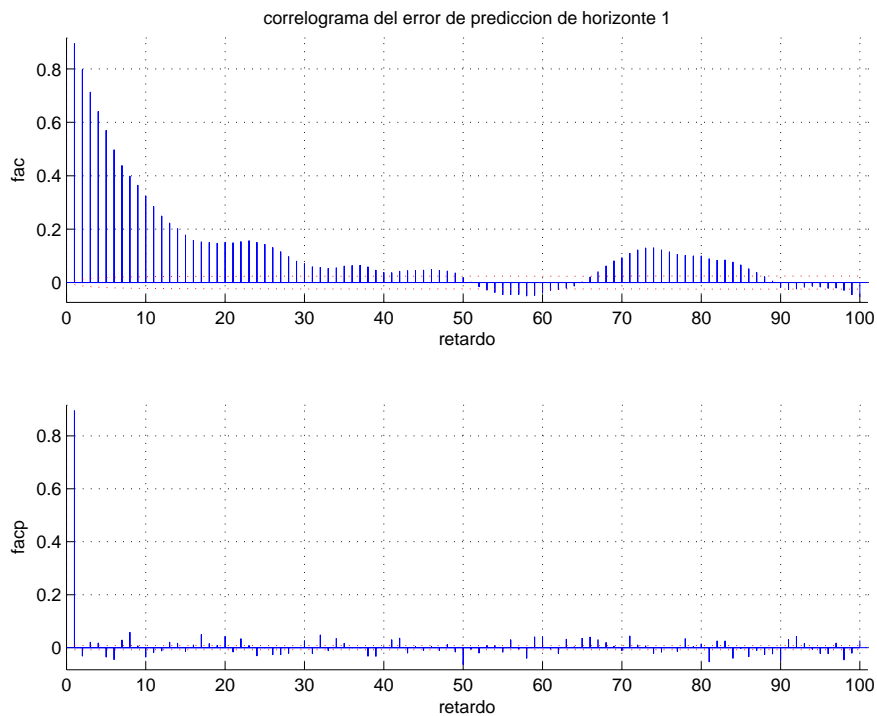


Figura 2.10: Correlograma de los errores de predicción a horizonte 1. Modelo AR(1) de parámetros constantes.

Comparación de resultados de los modelos de referencia.

A continuación se van a comparar de manera gráfica los resultados que se obtienen con los modelos de referencia. En primer lugar, se muestra la capacidad predictiva de los mismos analizando tres estadísticos: sesgo, EMAN (Error Medio Absoluto Normalizado) y RECMN (Raíz del Error Cuadrático Medio Normalizado) representados en la figura 2.11 junto con el rango intercuartílico de los errores de predicción. En estos gráficos se puede observar que el modelo persistente presenta sesgo prácticamente nulo (los errores están centrados en torno a cero, característica ya observada en el diagrama de cajas) aumentando ligeramente con el horizonte de predicción. Para el modelo media, el sesgo es independiente del horizonte de predicción, y para el modelo nr, el sesgo aumenta con el horizonte de predicción. Las gráficas del EMAN y de la RECMN son cualitativamente iguales. En ambos casos se ve que en el modelo media el valor del estadístico es independiente del horizonte de predicción. En los otros dos modelos de referencia su valor aumenta con el horizonte de predicción, siendo mejor el modelo nr que el modelo persistente, consecuencia lógica ya que el primero de ellos estima parcialmente la autocorrelación de la potencia en vez de imponer el valor unitario como ocurre en el modelo persistente. El rango intercuartílico aumenta con el horizonte de predicción presentando características similares para el modelo persistente y el modelo nr.

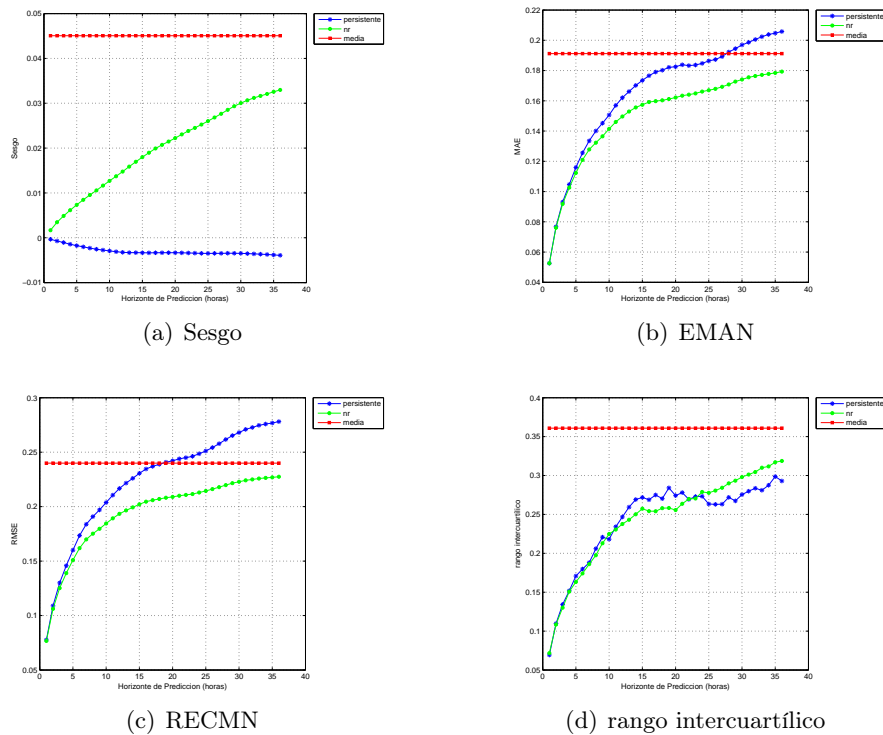


Figura 2.11: Estadísticos de los errores de predicción de los modelos de referencia.

2.3. Modelos de partida

2.3.1. Introducción

El empleo de modelos estadísticos adaptativos para la realización de predicciones de energía eólica se aplica por primera vez en el sistema Zephyr (véase [16], [25] y [40]). El empleo de modelos adaptativos está justificado porque la relación entre la velocidad del viento y la potencia generada es altamente no lineal y además los sistemas de predicción que se desarrollen tienen el riesgo de ser únicamente fiables en un cierto rango de los datos. Esta relación es variante con el tiempo porque depende además de otras variables como la dirección del viento, la densidad del aire, las variaciones de temperatura, los efectos locales de las nubes y la lluvia, etc. Debido a que algunas de estas variables son muy difíciles de prever o incluso de medir, no pueden ser incluidas de una manera apropiada en el modelo. Por todo ello, un modelo de parámetros constante no es satisfactorio. Algunos de los modelos de Sipleólico son similares a los formulados en Zephyr, aunque en este último sistema se basan en regresiones polinómicas locales. En esta tesis se parte de los modelos paramétricos basados en filtro de Kalman y formulados en el espacio de los estados desarrollados en Sipleólico para cada horizonte de predicción (véase [46]) y que se describen a continuación:

- Modelo M1. Es un modelo autorregresivo de parámetros cambiantes, univariante, de la forma:

1. Ecuación de salida:

$$p_{t+h} = \beta_{0t} + P_t(k, c) + u_{t+h|t}^{(M1)} \quad (2.15)$$

con

$$P_t(k, c) = \sum_{i=1}^k \beta_{it} p_{t+1-i} + \beta_{k+1} p_{t+h-c} \quad (2.16)$$

donde p_{t+h} es la potencia en el instante $t + h$ siendo h el horizonte de predicción, $u_{t+h|t}^{(M1)}$ es el error de predicción de horizonte h del modelo M1 y β_{it} , $i = 0, 1, \dots, k+1$ son los parámetros del modelo, variantes con el tiempo. Se supone que $u_{t+h|t}^{(M1)} \rightsquigarrow N(0, R)$. El término p_{t+h-c} en (2.16) trata de captar el ciclo diario. Si $h < 24 - k$ entonces $c = 24$, lo que implica que p_{t+h-24} es la potencia generada 24 horas antes que el instante que se quiere predecir. Si $h \geq 24 - k$, entonces $c = h + k$ y el modelo es simplemente un $AR(k+1)$. En Sipleólico inicialmente se empleaba el valor $k = 4$ y posteriormente se pasó a $k = 3$.

2. Ecuación de estado: Se supone que los parámetros del modelo siguen un paseo aleatorio

$$\beta_{it} = \beta_{it-1} + \eta_{it} \quad (2.17)$$

con $\eta_{it} \rightsquigarrow N(0, \sigma_i^2)$, $cov(\eta_{it}, \eta_{j\tau}) = 0, \forall i \neq j$, y $t \neq \tau$.

Este modelo es de especial interés cuando no se tengan disponibles las predicciones meteorológicas y sí los datos en tiempo real de la producción de potencia del parque.

- Modelos M2, M3, y M4: Se emplea la misma parametrización que en M1 incorporando la información de la predicción del módulo de la velocidad de viento en la ecuación de salida (la ecuación de estado es idéntica a la planteada en 2.17, extendiéndola también a los parámetros que acompañan a la velocidad de viento)

$$p_{t+h} = \beta_{0t} + P_t(k, c) + W_{t+h|t}(d) + u_{t+h|t}^{(Mm)}; m = 2, 3, 4; \quad (2.18)$$

con

$$W_{t+h|t}(d) = \sum_{i=1}^d \beta_{k+1+i,t} \hat{v}_{t+h|t}^i \quad (2.19)$$

donde $\hat{v}_{t+h|t}$ es la predicción meteorológica del módulo de la velocidad del viento disponible en el instante t para el periodo $t+h$. En M_2 , $d = 1$; en M_3 , $d = 2$ y en M_4 , $d = 3$ que modelan relación lineal, cuadrática y cúbica, respectivamente, entre potencia y velocidad.

- Modelos M5, M6 y M7. Se introduce la información de la predicción meteorológica de la dirección de viento. La dirección de viento es importante por varias razones. Una de ellas es que los parques eólicos suelen presentar unas direcciones de viento predominantes en las que la velocidad suele moverse en un rango estrecho, por lo que se puede tomar como un predictor de la velocidad. Además, el funcionamiento de cada aerogenerador depende de la dirección, debido a los efectos del terreno y al efecto sombra de las turbinas colindantes. Para la velocidad, M5, M6 y M7 emplean la misma parametrización que M2, M3 y M4 respectivamente. Estos modelos son

$$\hat{p}_{t+h} = \beta_{0t} + P_t(k, c) + W_{t+h|t}(d) + D_{t+h|t} + u_{t+h|t}^{(Mm)}; m = 5, 6, 7; \quad (2.20)$$

con

$$D_{t+h|t} = \beta_{k+d+2,t} \sin\left(\frac{2\pi \hat{d}_{t+h|t}}{360}\right) + \gamma_{k+d+3,t} \cos\left(\frac{2\pi \hat{d}_{t+h|t}}{360}\right) \quad (2.21)$$

donde $\hat{d}_{t+h|t}$ es la predicción meteorológica de la dirección de viento. Como en (2.19), el término $W_{t+h|t}(d)$ en el modelo *M5* emplea $d = 1$, *M6* utiliza $d = 2$ y *M7* toma el valor $d = 3$. Estos modelos obtienen muy buenos resultados cuando la predicción de la velocidad del viento es de baja calidad ya que se ve mejorada significativamente por la predicción de la dirección.

2.3.2. Análisis de resultados con los modelos de partida

En este apartado se van a analizar los resultados que se obtienen con los modelos descritos anteriormente. Para ello debe tenerse en cuenta que para poder comparar los resultados de unos parques con otros se trabaja con los valores de potencia normalizados, es decir, divididos por la potencia nominal, por lo que dichos valores siempre están comprendidos entre 0 y 1. En el caso que nos ocupa la potencia nominal es de 17,56 MW. Para el análisis de los distintos modelos estudiaremos los siguientes aspectos:

- En los errores de predicción se analiza su sesgo, EMAN, RECMN, histograma, correlograma y diagrama de cajas.
- Respecto a la predicción se calculará el parámetro denominado mejora definido en la sección 2.2.3.

En la figura 2.12 se muestra el diagrama de cajas de los errores de predicción en función del horizonte de predicción para el modelo *M1*. Es de destacar que existe un número apreciable de datos que se toman como atípicos o valores extremos por estar fuera de los bigotes del diagrama de cajas. Estas características se detectan en las gráficas posteriores donde se muestran los histogramas de los errores de predicción (figura 2.13). Puede verificarse que la distribución de los errores varía con el horizonte de predicción: la distribución a horizonte 1 es más estrecha y a medida que aumenta el horizonte, la curtosis o apuntalamiento de la distribución disminuye y, por otro lado, aumenta el rango de los datos del error. Además, se puede comprobar cómo la distribución de los errores presenta unas colas más pesadas que la distribución normal, es decir, los errores presentan una varianza mayor.

En la figura 2.14 se muestran los correlogramas de los errores de predicción para el modelo *M1*. En estos gráficos se puede observar que los errores de predicción a horizonte uno no pueden considerarse ruido blanco. Tampoco puede determinarse claramente un modelo para los mismos. A partir de horizonte 2, puede verse cómo los errores siguen un patrón $MA(k-1)$ siendo k el horizonte de predicción hasta un cierto valor a partir del cual el orden del modelo de media móvil no se incrementa.

El modelo M2 es el que resulta de incorporar al modelo M1 la predicción del módulo de la velocidad de viento estimada para ese horizonte de predicción por el modelo meteorológico. Los resultados se muestra gráficamente desde la figura 2.15 hasta la figura 2.17. Al igual que en el modelo M1 podemos observar cómo aumenta el coeficiente de asimetría con el horizonte de predicción, así como una disminución de la curtosis. En la figura 2.17 se muestran los correlogramas de los residuos. En estos gráficos se puede observar que los errores de predicción a horizonte 1 no pueden considerarse ruido blanco aunque no se encuentra un patrón claro en los mismos. Para horizontes mayores, los errores presentan estructura MA(k-1) siendo k el horizonte de predicción, aunque el orden del modelo MA tiende a estabilizarse alcanzado cierto horizonte de predicción. El mismo patrón de comportamiento se observa en los modelos M3, M4, M5, M6 y M7 lo que se muestra gráficamente en las figuras 2.18 hasta la 2.31

Resultados comparativos de los modelos

En las figuras 2.32, 2.33 y 2.34 se muestra gráficamente el comportamiento de los modelos de partida respecto a los modelos de referencia en términos de sesgo, EMAN y RECMN de los errores. En dichas gráficas puede observarse:

- (i) Respecto al sesgo, los modelos de partida no logran superar al modelo persistente para ningún horizonte de predicción. Respecto al modelo nr, el modelo M1 no consigue superarlo para ningún horizonte de predicción, los modelos M2, M3, M4 y M5 son ligeramente mejores que éste, y los modelos M6 y M7 no lo superan hasta horizonte 20. En todos los casos, se ve el aumento del sesgo con el horizonte de predicción para todos los modelos de partida.
- (ii) Respecto al EMAN y la RECMN, los modelos de partida superan claramente a los modelos de referencia persistente y media. Respecto al modelo nr, el modelo M1 presenta un comportamiento similar a éste, aunque con valores ligeramente superiores. El resto de modelos de partida no logran superar a este modelo hasta horizonte 20.

Por cuestiones de claridad, en las figuras 2.35, 2.36 y 2.37, se vuelven a representar gráficamente estas medidas de capacidad predictiva únicamente de los modelos de partida para analizarlos entre sí. En estas gráficas se distingue el comportamiento del modelo M1, autorregresivo puro, de los modelos M2, M3 y M4 que son modelos donde además de la parte autorregresiva se ha introducido la variable predicción meteorológica del módulo de la velocidad de viento, de los modelos M5, M6 y M7 que son los modelos en

los que además del módulo se ha introducido la predicción meteorológica de la dirección de la velocidad del viento como variable explicativa. Respecto al EMAN y la RECMN, la inclusión de la velocidad presenta una mejora en la predicción, aunque no puede apreciarse una mejora significativa entre los mismos, es decir, apenas existe diferencia entre el modelo M2 que tiene como variable de entrada la predicción meteorológica de la velocidad, M3 que tiene como entrada la predicción de la velocidad y de la velocidad al cuadrado y M4 que introduce la predicción de velocidad de viento, el cuadrado y el cubo de dicha variable. La dirección de viento también mejora significativamente los modelos y puede apreciarse como esta mejora es mayor conforme aumenta el horizonte de predicción.

En las figuras 2.38 y 2.39 se muestra el coeficiente de mejora de los modelos de partida respecto al modelo de referencia planteado en [36] y se observa lo siguiente:

- Respecto al EMAN, los modelos de partida no mejoran el modelo de referencia hasta horizontes mayores que 23. El modelo M1 no lo mejora en ningún caso.
- Respecto a la RCECM los modelos de partida superan al modelo de referencia a partir de horizontes mayores que 20. El modelo M1, al igual que en el EMAN, no supera al modelo de referencia en ningún caso.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

2.4.1. Modelos de partida

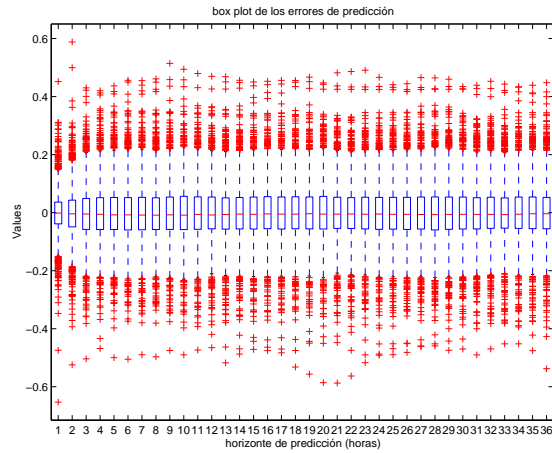


Figura 2.12: Diagrama de cajas de los errores de predicción del modelo M1.

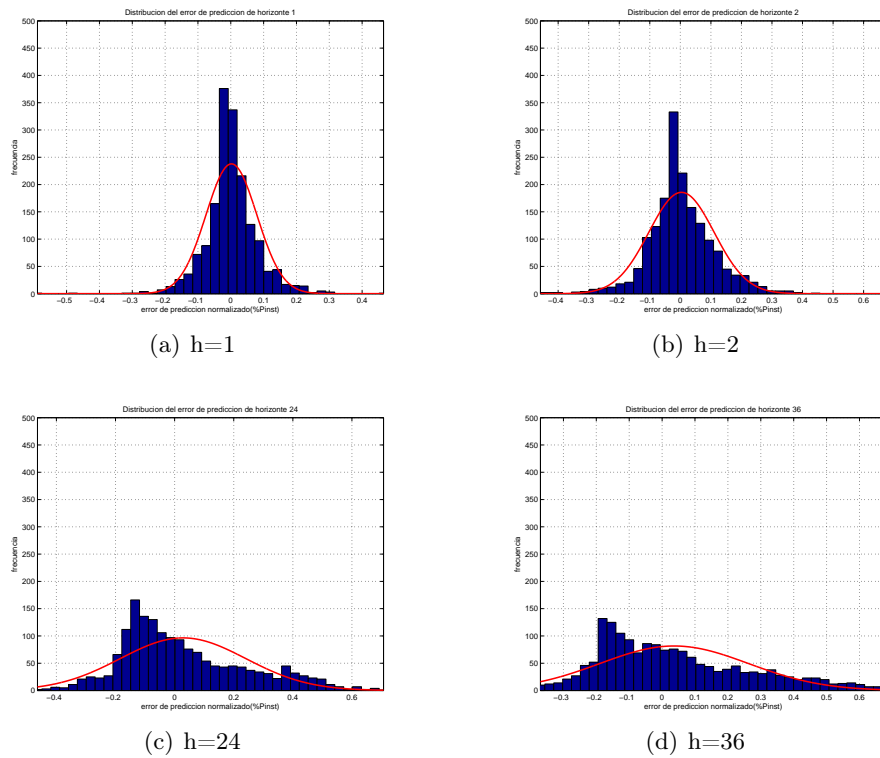


Figura 2.13: Histograma de los errores de predicción. Modelo M1.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

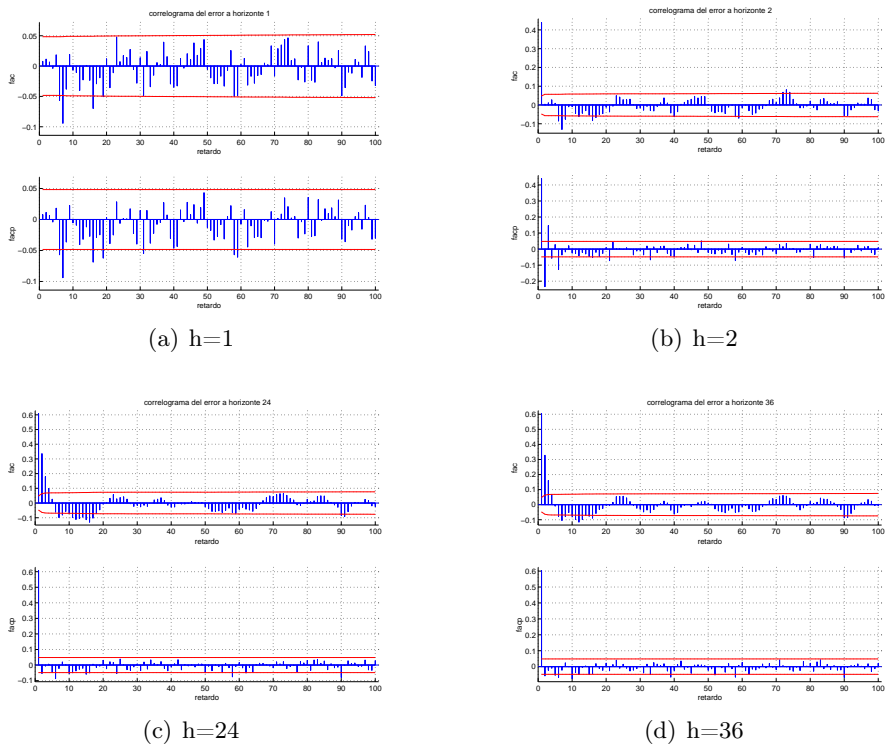


Figura 2.14: Correlograma de los errores de predicción. Modelo M1.

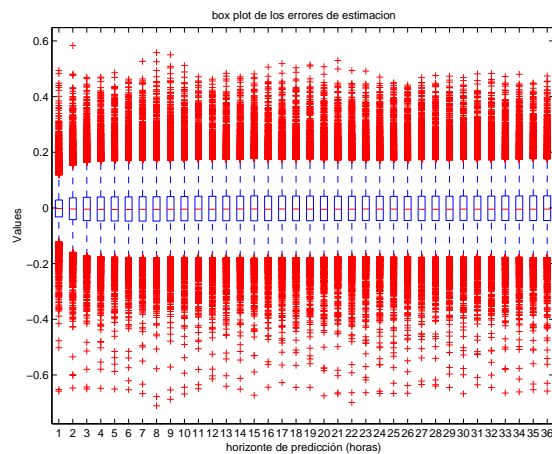


Figura 2.15: Diagrama de cajas de los errores de predicción del modelo M2.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

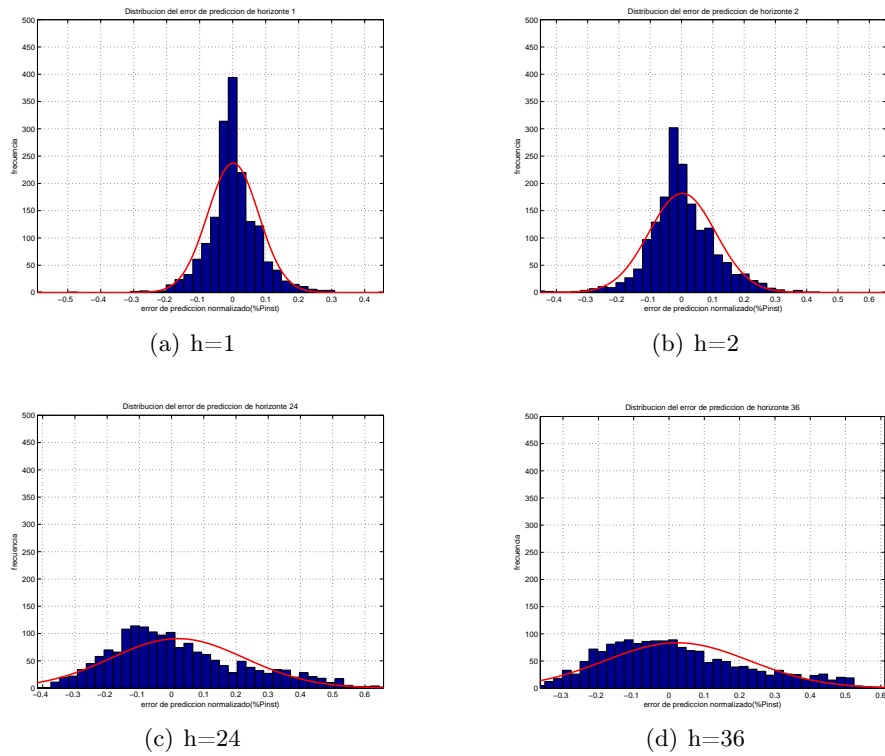


Figura 2.16: Histograma de los errores de predicción. Modelo M2.

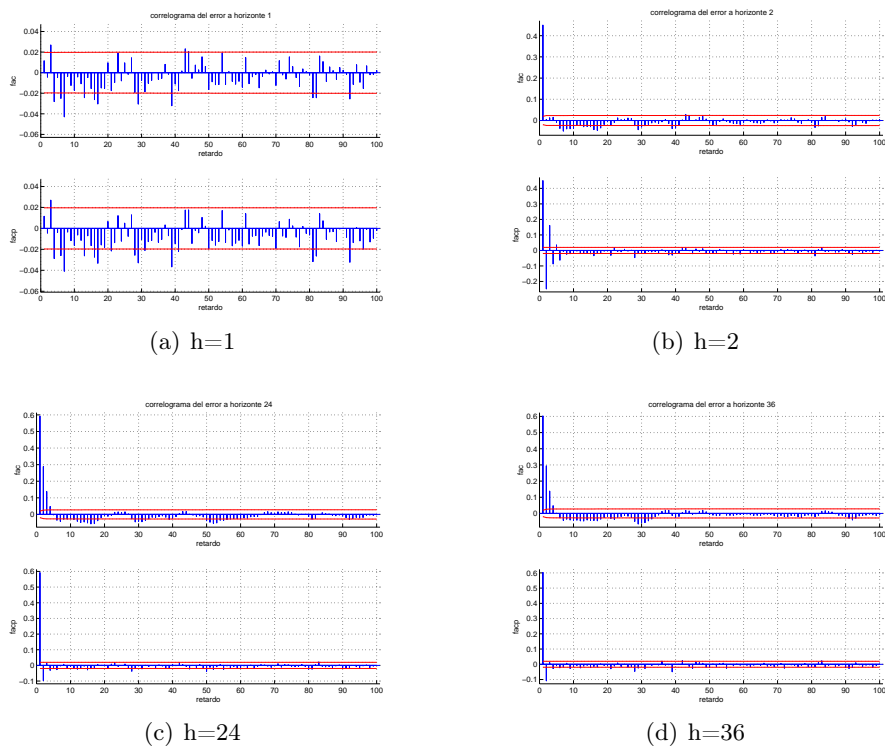


Figura 2.17: Correlograma de los errores de predicción. Modelo M2.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

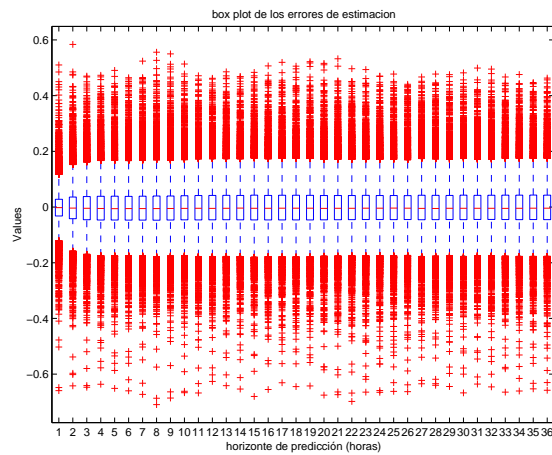


Figura 2.18: Diagrama de cajas de los errores de predicción del modelo M3.

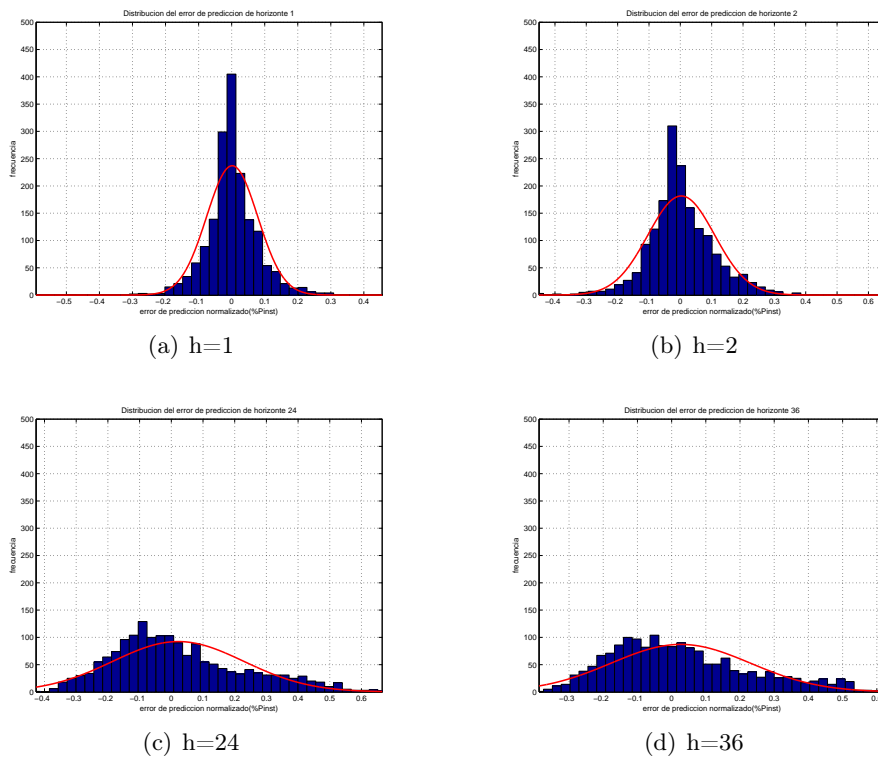


Figura 2.19: Histograma de los errores de predicción. Modelo M3.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

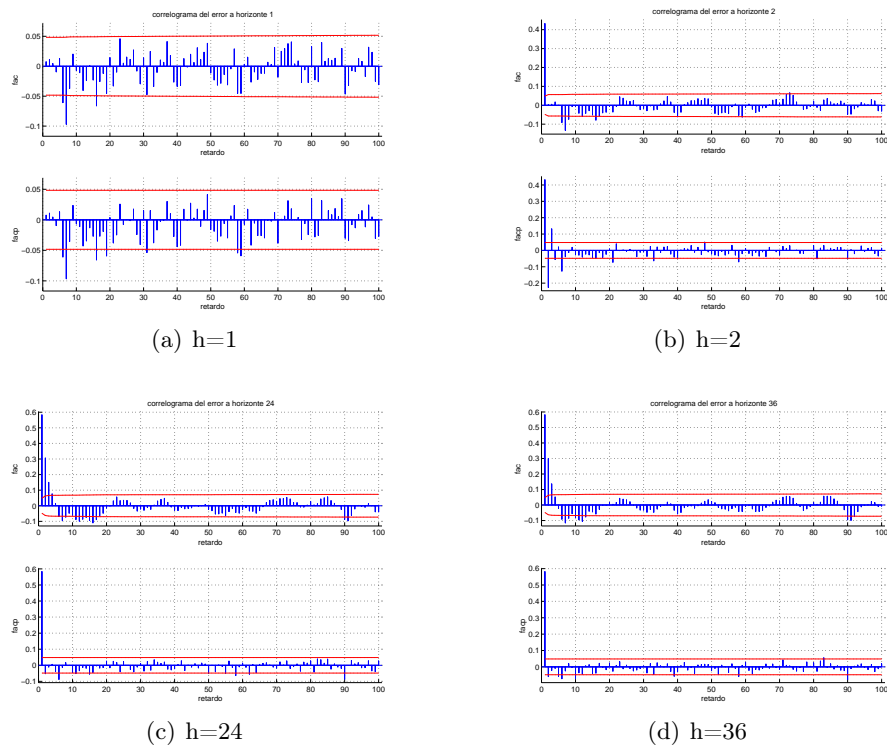


Figura 2.20: Correlograma de los errores de predicción. Modelo M3.

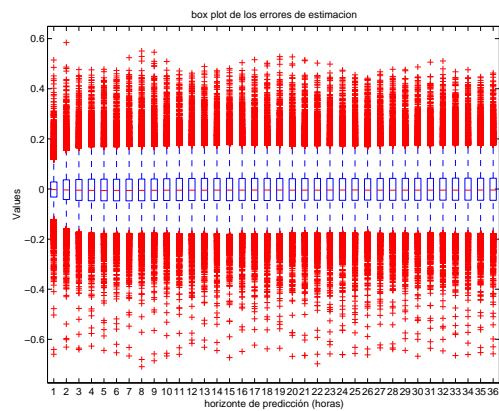


Figura 2.21: Diagrama de cajas de los errores de predicción del modelo M4.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

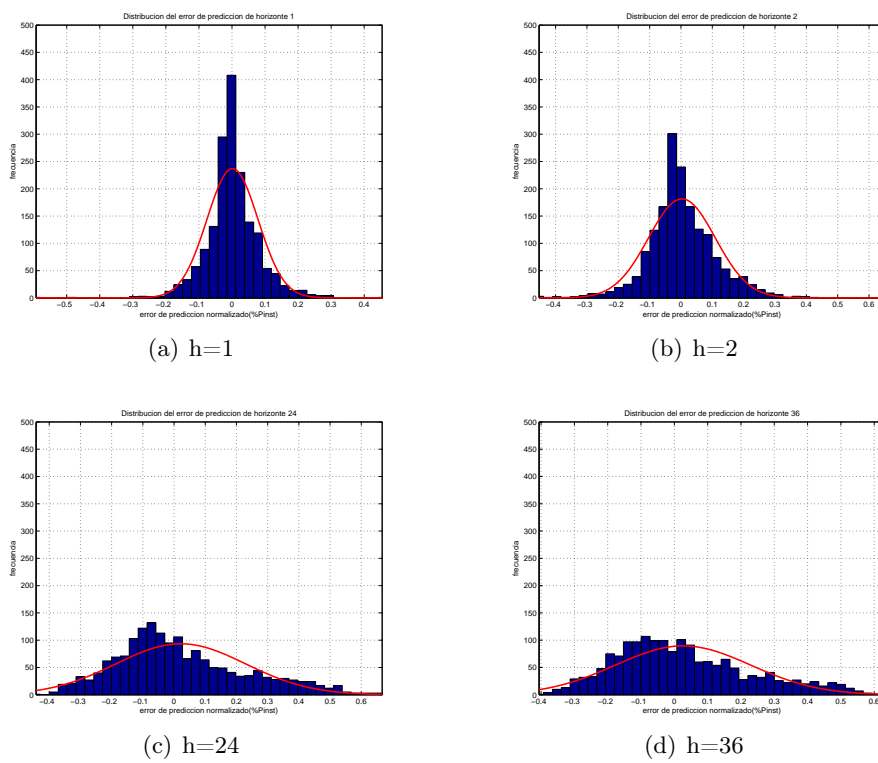


Figura 2.22: Histograma de los errores de predicción. Modelo M4.

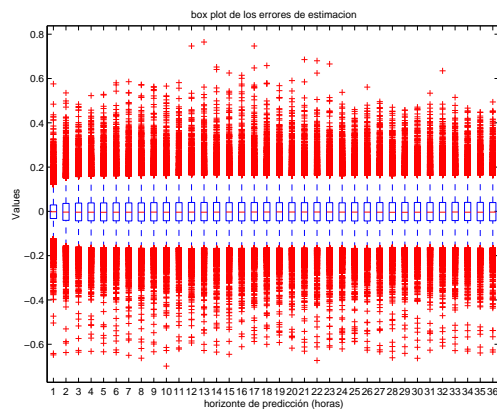


Figura 2.23: Diagrama de cajas de los errores de predicción del modelo M5.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

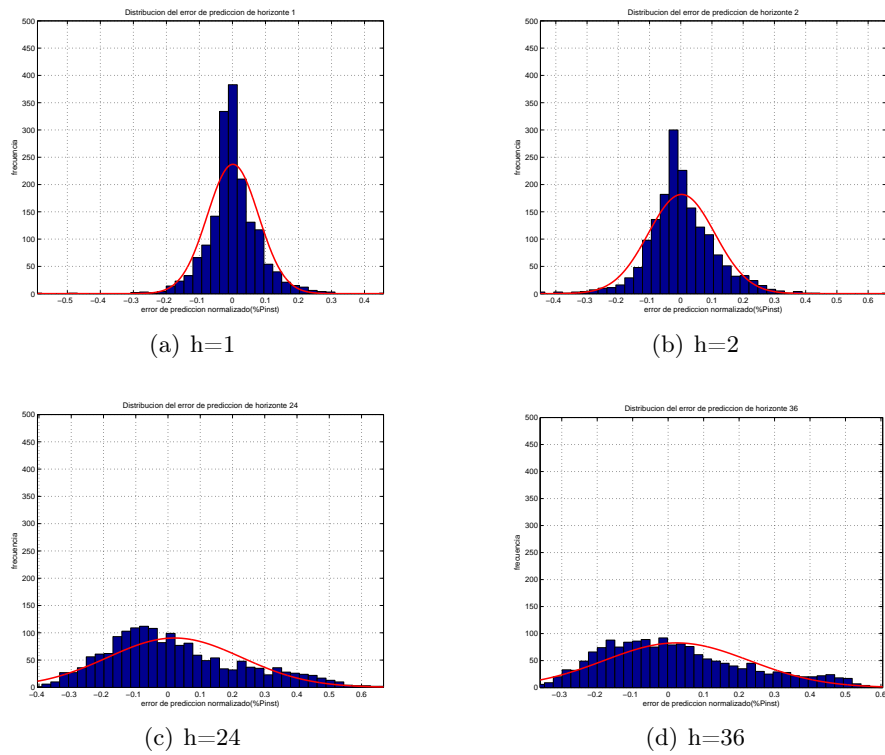


Figura 2.24: Histograma de los errores de predicción. Modelo M5.

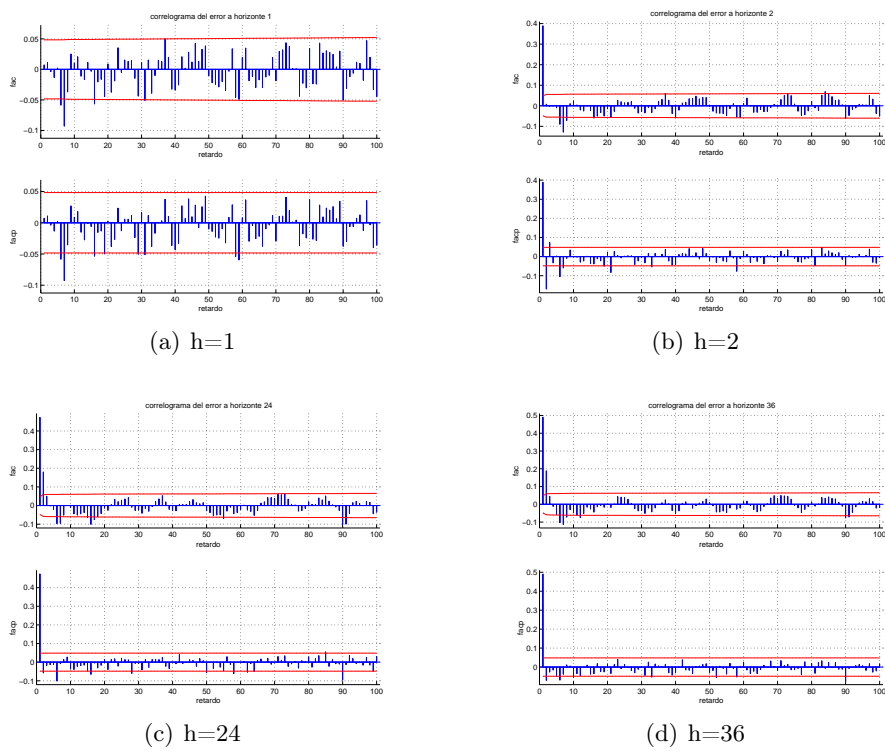


Figura 2.25: Correlograma de los errores de predicción. Modelo M5.

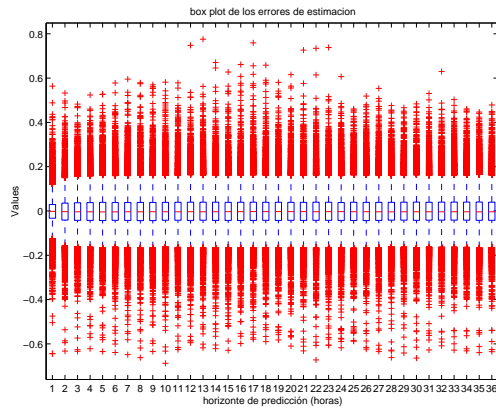


Figura 2.26: Diagrama de cajas de los errores de predicción del modelo M6.

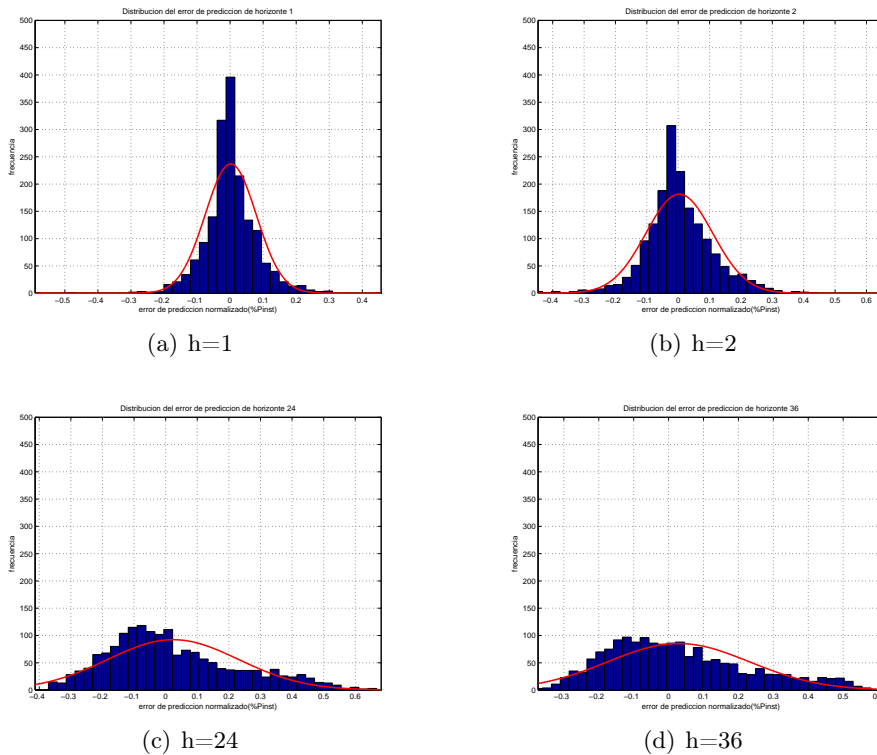


Figura 2.27: Histograma de los errores de predicción. Modelo M6.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

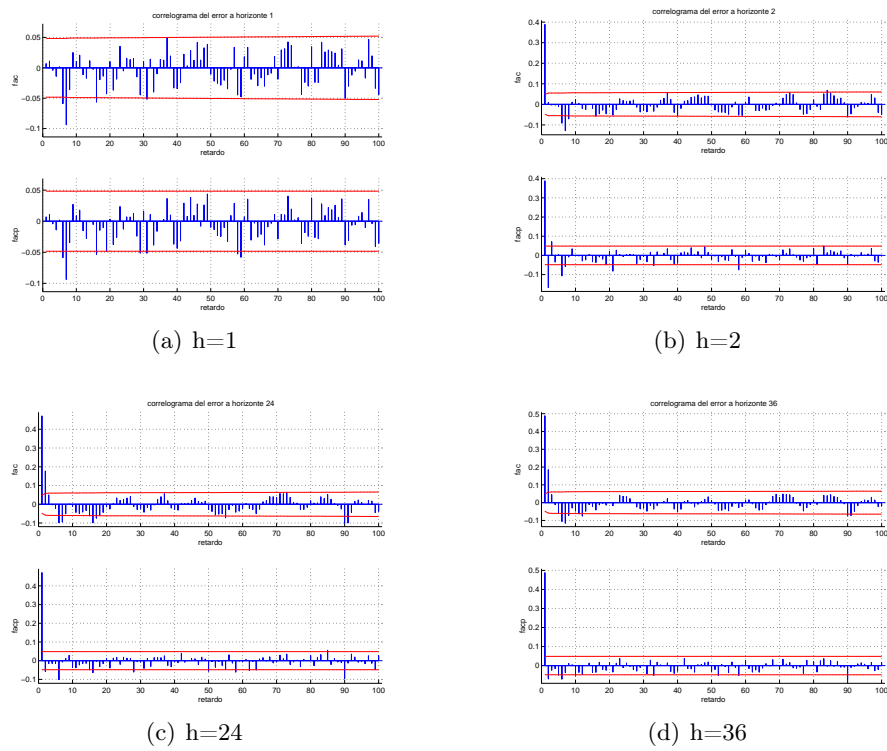


Figura 2.28: Correlograma de los errores de predicción. Modelo M6.

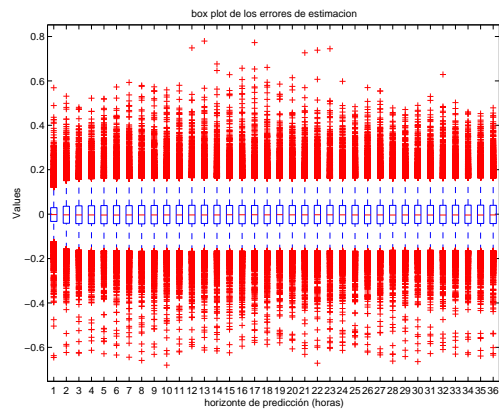


Figura 2.29: Diagrama de cajas de los errores de predicción del modelo M7.

2.4. Anexo: Gráficas de los resultados de los modelos de predicción

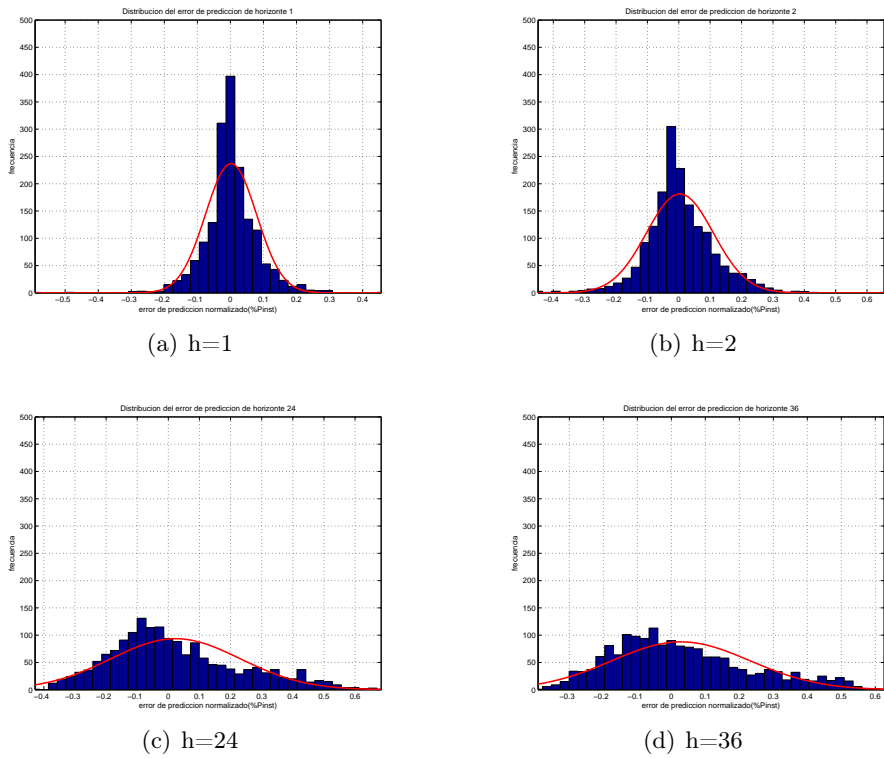


Figura 2.30: Histograma de los errores de predicción. Modelo M7.

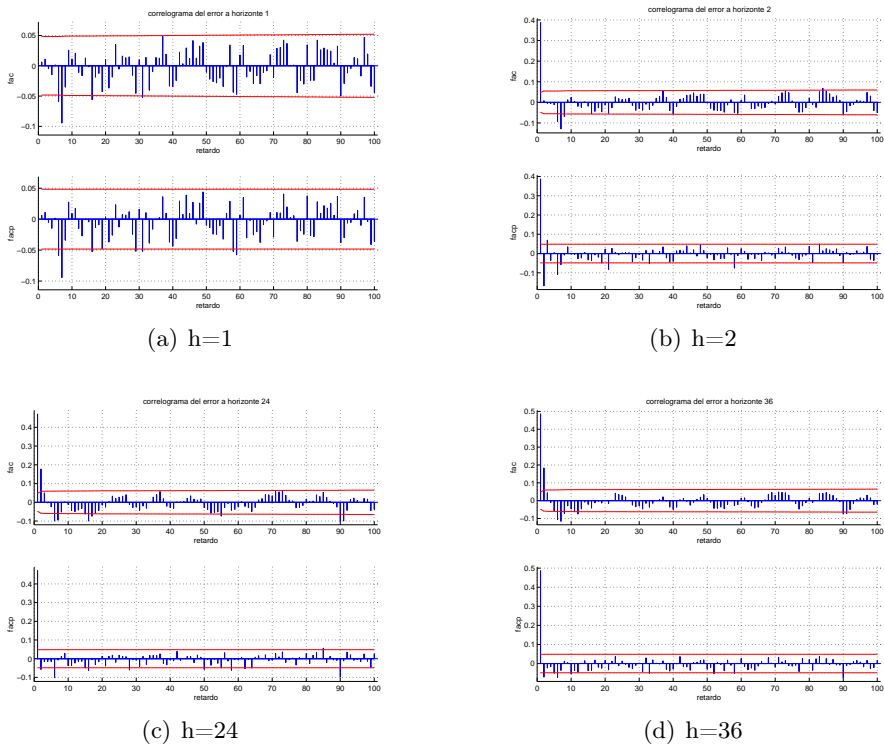


Figura 2.31: Correlograma de los errores de predicción. Modelo M7.

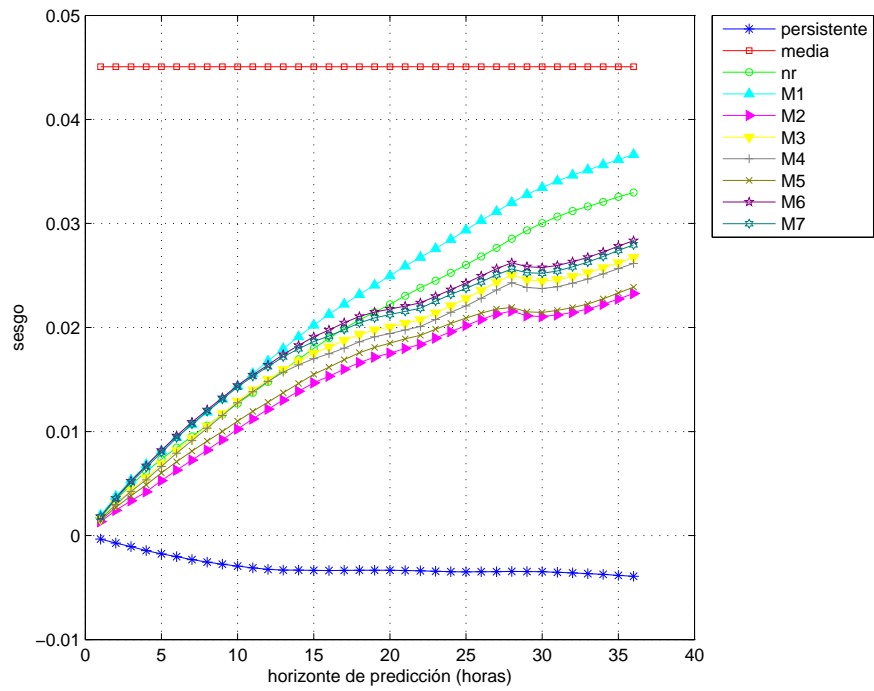


Figura 2.32: Sesgo de los modelos de referencia y los modelos de partida.

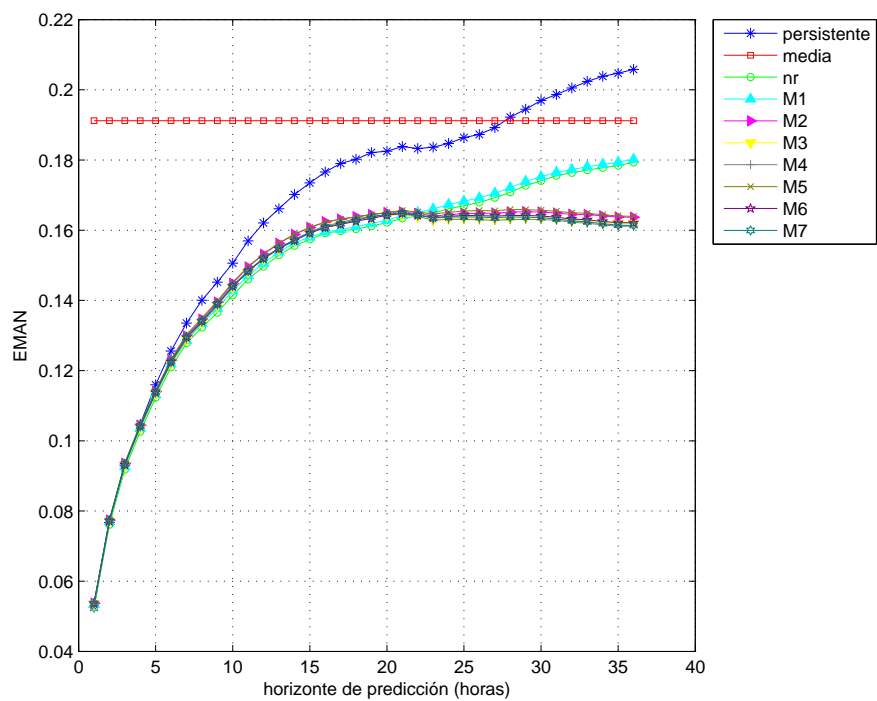


Figura 2.33: EMAN de los modelos de referencia y los modelos de partida.

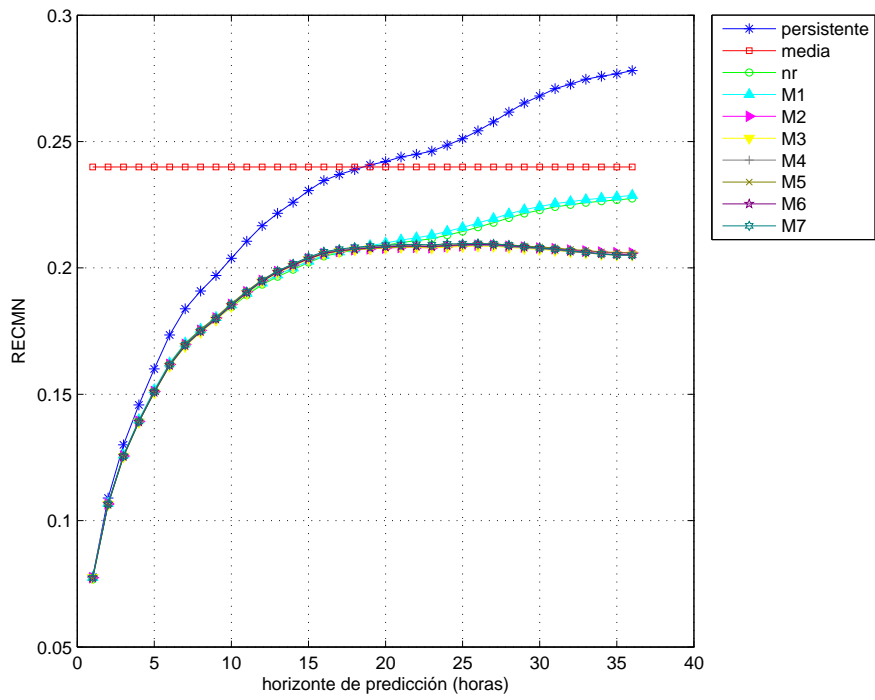


Figura 2.34: RECMN de los modelos de referencia y los modelos de partida.

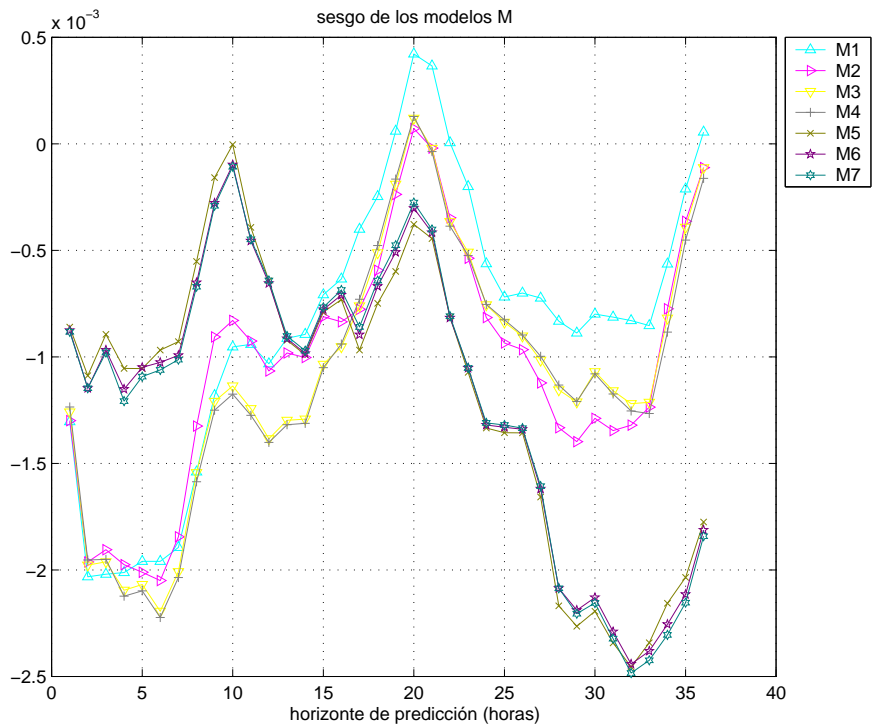


Figura 2.35: Sesgo de los modelos M.

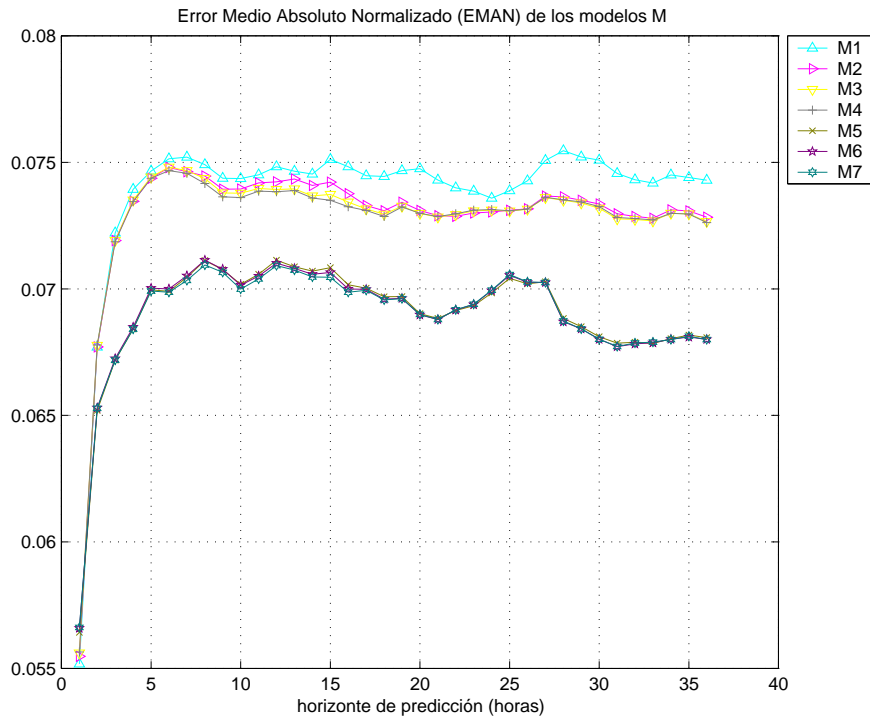


Figura 2.36: EMAN de los modelos M.

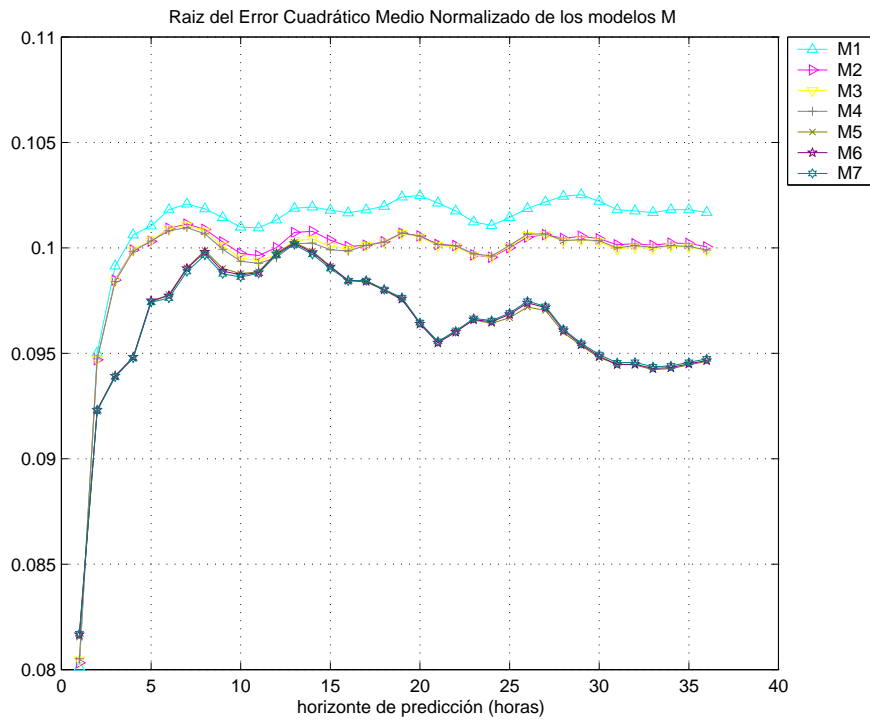


Figura 2.37: RECMN de los modelos M.

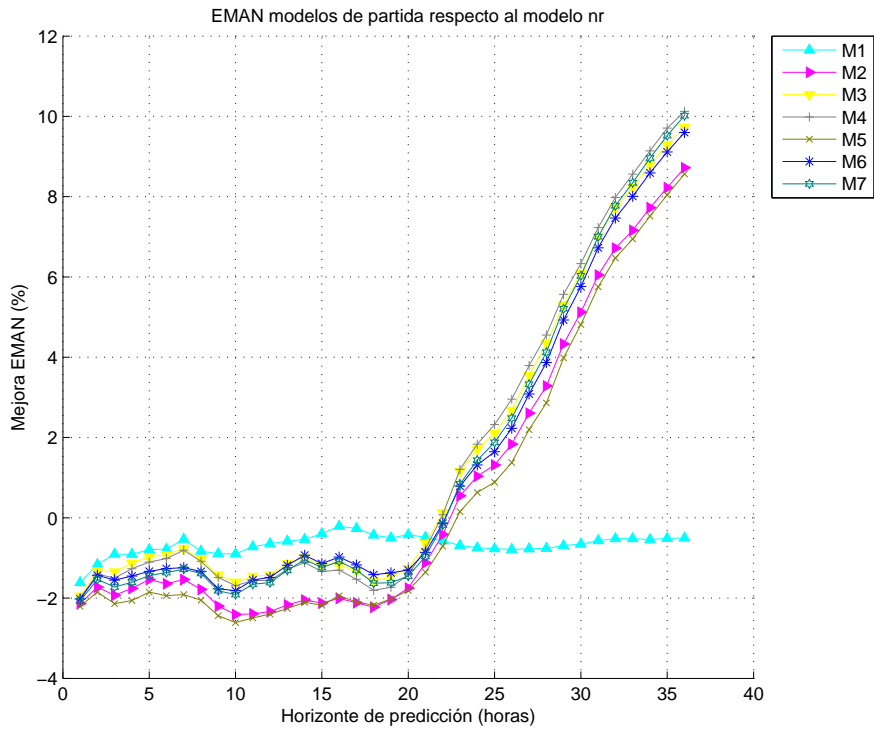


Figura 2.38: Coeficiente de mejora respecto al EMAN de los modelos de partida.

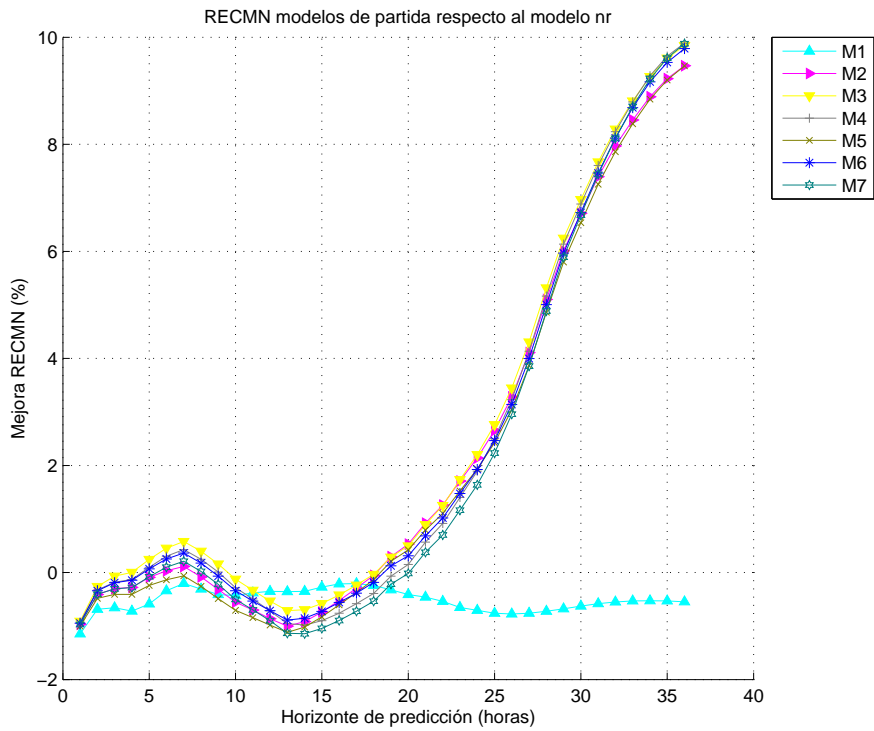


Figura 2.39: Coeficiente de mejora respecto a la RECMN de los modelos de partida.

Capítulo 3

Optimización de los modelos y propuesta de nuevos modelos

En este capítulo se estiman las matrices o hiperparámetros de los modelos planteados en Sipleólico mediante máxima verosimilitud. Posteriormente, se plantean nuevos modelos multivariantes que recogen de manera implícita, a través de las correlaciones entre parámetros, relaciones no recogidas explícitamente en los modelos. Se comparan los resultados obtenidos para las distintas estructuras matriciales planteadas, se analiza la estrategia de funcionamiento en tiempo real y, finalmente, se estima la incertidumbre en la estimación de los parámetros mediante técnicas de remuestreo.

3.1. Introducción

En SIPREOLICO se suponen conocidos los vectores y matrices que determinan el comportamiento dinámico del sistema que se está analizando. Además, los modelos propuestos para los parámetros cambiantes en el tiempo asociados a la ecuación de predicción de la potencia, suponen que estos evolucionan de manera aislada, sin tener en cuenta la posibilidad de que haya variables, no recogidas explícitamente en el modelo que puedan afectar el funcionamiento del sistema. Dichas variables pueden hacer variar varios parámetros del modelo a la vez, de manera que pueda existir algún grado de dependencia entre ellos. El objetivo de este capítulo es resolver las dos cuestiones anteriores. Las contribuciones del capítulo son:

1. Proponer la estimación por técnicas de máxima verosimilitud de los hiperparámetros del modelo. En particular, se usará el algoritmo EM.

2. Proponer nuevos modelos multivariantes para captar la evolución de las variables de estado. Estos modelos son de dos tipos:
 - a) Modelos en los que la estructura de dependencia de los parámetros o variables de estado tiene lugar únicamente a través de la matriz de varianzas-covarianzas del ruido.
 - b) Modelos en los que además puede haber una dependencia explícita a través de retardos del resto de variables de estado.
3. Comparar, en términos de predicción, todos los modelos propuestos entre sí, así como con los modelos de referencia y los modelos de partida expuestos en el capítulo anterior.
4. Plantear una estrategia adecuada para el funcionamiento en tiempo real.
5. Calcular mediante técnicas de remuestreo la incertidumbre asociada a los parámetros.

Este capítulo se ha ordenado de la siguiente manera. En la sección 2 se describe el algoritmo EM para la estimación de los parámetros de los distintos modelos y se presentan los resultados obtenidos. En la sección 3 se proponen modelos multivariantes para captar la evolución de las variables de estado y se analizan los resultados obtenidos en la estimación. En la sección 4, se comparan los resultados obtenidos en predicción de todos los modelos considerados. En la sección 5 se plantea la estrategia de funcionamiento en tiempo real y finalmente en la sección 6 se realiza el cálculo de la incertidumbre asociada a la estimación de los parámetros mediante técnicas de remuestreo.

3.2. Algoritmo EM

3.2.1. Introducción

El algoritmo EM, introducido por [7] y particularizado al ámbito de las series temporales por [51] y [54] es una técnica iterativa para realizar la estimación por máxima verosimilitud de un conjunto de hiperparámetros Θ en problemas en los que existen datos ocultos (en este caso, el vector de estado β). El algoritmo EM contempla tres conjuntos de datos:

1. Los datos observados que denotamos por Y .

2. Los datos faltantes que en nuestro caso son los parámetros cambiantes recogidos en el vector de estado β_t , que para todo el periodo de estimación denotamos por β .
3. El conjunto completo de datos (Y, β) .

Las iniciales EM corresponden a Expectation y Maximization, respectivamente. Cada iteración del algoritmo EM consiste en 2 etapas. En la primera etapa, E, se obtienen diversos momentos de las variables no observadas (el vector de estado β_t en nuestro caso). En la segunda etapa M o de maximización se obtienen los valores óptimos de los hiperparámetros (constantes) del modelo, que se denotan por Θ . Los distintos modelos de SIPREOLICO que se quieren estimar se pueden escribir en el espacio de los estados de la siguiente manera:

1. Ecuación de medida o de observación

$$y_t = A_t \beta_t + e_t \quad (3.1)$$

donde $y_t = p_{t+h}$ es la potencia eólica que se quiere predecir, A_t depende del modelo que estemos optimizando (ver ecuaciones 2.15 hasta 2.21) y e_t es el término de error que suponemos se distribuye siguiendo una normal de media 0 y varianza R . La dimensión del vector fila A_t es $m \times 1$. Para M1, $m = k + 2$, siendo k el parámetro definido en la sección 2.3.1. Para M2, M3 y M4 $m = k + d + 2$ con $d = 1, 2$ y 3 respectivamente y para M5, M6 y M7, $m = k + 4 + d$ con $d = 1, 2$ y 3 respectivamente.

2. Ecuación de estado para cada variable no observada del modelo, $i = 1, 2, \dots, m$

$$\beta_{it} = \beta_{it-1} + \eta_{it}$$

donde $\eta_{it} \sim N(0, \sigma_i^2)$, $cov(\eta_{it}, \eta_{j\tau}) = 0, \forall i \neq j$, y $t \neq \tau$. Sea $\beta_t = (\beta_{1t}, \dots, \beta_{mt})'$ y $\eta_t = (\eta_{1t}, \dots, \eta_{mt})'$. Se pueden recoger las ecuaciones de estado escalares en una única ecuación vectorial

$$\beta_t = \beta_{t-1} + \eta_t \quad (3.2)$$

donde η_t es ruido blanco multivariante con matriz de varianzas covarianzas diagonal dada por $Q = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$.

Sea $\beta_0 \sim (b_0, P_0)$ el estado inicial, donde $E(\beta_0) = b_0$ y $\text{var}(\beta_0) = P_0$. El conjunto de hiperparámetros constantes del modelo es en nuestro caso $\Theta = (b_0, P_0, Q, R)$. Nótese que

b_0 es un vector de dimensión $m \times 1$, R es un escalar y P_0 y Q son matrices de dimensión $m \times m$. Sea $Y = (p_1, p_2, \dots, p_T)$ la potencia eólica observada y $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_T)$ los parámetros no observados cambiantes en el tiempo. Se considera el conjunto completo de datos formado por las series de datos observados Y y los vectores de estado no observados $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_T)$ y se maximiza la esperanza de la densidad conjunta de todos los datos, condicionado a los datos observados, evaluada en la estimación disponible de los hiperparámetros desconocidos del modelo $\Theta^{(k-1)}$. La función de verosimilitud es

$$f_Y(Y; \Theta) = \frac{f_{Y,\beta}(Y, \beta; \Theta)}{f_{\beta|Y}(\beta|Y; \Theta)}.$$

El logaritmo de la verosimilitud es, por tanto,

$$\ln f_Y(Y; \Theta) = \ln f_{Y,\beta}(Y, \beta; \Theta) - \ln f_{\beta|Y}(\beta|Y; \Theta).$$

Tomando esperanzas con respecto a la densidad $f_{\beta|Y}$

$$\ln f_Y(Y; \Theta) = E_{\beta|Y}(\ln f_{Y,\beta}(Y, \beta; \Theta|Y)) - C$$

donde C es una constante, no involucrada en el proceso de maximización.

Por lo tanto, $\hat{\Theta} = \operatorname{argmax} \ln f_Y(Y; \Theta)$ puede calcularse como $\hat{\Theta} = \operatorname{argmax} E_{\beta|Y} \ln f_{Y,\beta}(Y, \beta; \Theta)$. La ventaja de optimizar $E_{\beta|Y} \ln f_{Y,\beta}(Y, \beta; \Theta)$ frente a $\ln f_Y(Y; \Theta)$ es que este segundo problema de optimización suele ser altamente no lineal y mucho más complejo de resolver que el primero. En nuestro caso particular

$$\ln f_{Y,\beta}(Y, \beta; \Theta) = \sum_{t=1}^T (\ln f(y_t|\beta_t; \Theta) + \ln f(\beta_t|\beta_{t-1}; \Theta)) + \ln f(\beta_0; \Theta). \quad (3.3)$$

Bajo el supuesto de gaussianidad, la esperanza de -2 veces la función de densidad conjunta $\ln f_{Y,\beta}(Y, \beta; \Theta)$ con respecto a la marginal de Y , puede escribirse como

$$\begin{aligned} E \{ \ln f_{Y,\beta}(Y, \beta; \Theta) \} &= \ln |P_0| + \operatorname{tr} \left\{ P_0^{-1} \left[P_0^T + (\beta_0^T - b_0) (\beta_0^T - b_0)' \right] \right\} + \\ &+ T \ln |Q| + \operatorname{tr} \left\{ Q^{-1} \left[S_{11} - S'_{10} - S'_{10} + S_{00} \right] \right\} + T \ln |R| + \\ &+ \operatorname{tr} \left\{ R^{-1} \sum_{t=1}^T \left[(p_t - A_t \beta_t^T) (p_t - A_t \beta_t^T)' + A_t P_t^T A_t' \right] \right\} \end{aligned} \quad (3.4)$$

donde a continuación se detallan todos los elementos en (3.4), no definidos hasta ahora:

$$\begin{aligned}\beta_{it}^T &= E(\beta_{it}|Y) \\ P_{it}^T &= E((\beta_{it} - \beta_{it}^T)^2|Y) \\ P_{it,t-1}^T &= E((\beta_{it} - \beta_{it}^T)(\beta_{it-1} - \beta_{it-1}^T)|Y)\end{aligned}$$

para $i = 1, \dots, m$. Y de manera vectorial sean

$$\beta_t^T = \begin{pmatrix} \beta_{1t}^T \\ \beta_{2t}^T \\ \vdots \\ \beta_{mt}^T \end{pmatrix} \quad (3.5)$$

$$P_t^T = \begin{pmatrix} P_{1t}^T & 0 & \cdots & 0 \\ 0 & P_{2t}^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{mt}^T \end{pmatrix} \quad (3.6)$$

$$P_{t,t-1}^T = \begin{pmatrix} P_{1t,t-1}^T & 0 & \cdots & 0 \\ 0 & P_{2t,t-1}^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{mt,t-1}^T \end{pmatrix}. \quad (3.7)$$

El procedimiento iterativo del algoritmo EM se puede resumir en las siguientes etapas:

1. Seleccionar el valor inicial de los hiperparámetros $\Theta = \{b_0, P_0, R, Q\}$.
2. Etapa E: En cada iteración j , ($j = 1, 2, \dots$), calcular los valores β_t^T, P_t^T , y $P_{t,t-1}^T$ para $t = 1, \dots, T$, mediante el suavizado de Kalman usando el valor de los hiperparámetros obtenidos en la iteración anterior $\Theta^{(j-1)}$. Emplear estos valores suavizados para calcular las matrices que se detallan a continuación:

$$\begin{aligned}S_{11} &= \sum_{t=1}^T (\text{diag}(\beta_t^T (\beta_t^T)') + P_t^T) \\ S_{10} &= \sum_{t=1}^T (\text{diag}(\beta_t^T (\beta_{t-1}^T)') + P_{t,t-1}^T) \\ S_{00} &= \sum_{t=1}^T (\text{diag}(\beta_{t-1}^T (\beta_{t-1}^T)') + P_{t-1}^T)\end{aligned}$$

3. Etapa M: En cada iteración j , ($j = 1, 2, \dots$), actualizar el valor de los hiperparámetros que se están optimizando. En nuestro caso:

$$Q^{(j)} = T^{-1} \left(S_{11} - S_{10} S_{00}^{-1} S'_{10} \right) \quad (3.8)$$

y

$$R^{(j)} = T^{-1} \sum_{t=1}^T \left[(y_t - A_t \beta_t^T)^2 + A_t P_t^T A_t' \right]. \quad (3.9)$$

Los valores de b_0 y P_0 se obtienen a partir de (3.5) y (3.6).

4. Repetir las etapas 2-3 hasta alcanzar la convergencia.

En cada iteración del algoritmo EM la verosimilitud aumenta, véase [56].

En nuestro caso, la variable y_t es la predicción un periodo hacia delante de la potencia. Optimizaremos los modelos utilizando los datos de estimación (8760 primeros datos) y con los valores obtenidos generaremos predicciones para los datos de predicción y para todos los horizontes de predicción de manera que sean verdaderas predicciones ex-ante. El algoritmo de estimación se detiene cuando hemos alcanzado 2000 iteraciones o la tolerancia es menor que 0.001, donde la tolerancia se define como

$$tol = \frac{|\Theta^{(k)}(i, j) - \Theta^{(k-1)}(i, j)|}{|\Theta^{(k)}(i, j)|}.$$

Como condiciones iniciales del algoritmo, se toman

- $R = 0,01$
- $Q = 0,001 \cdot I_{(m \times m)}$
- $b_0 = [0]_{m \times 1}$
- $P_0 = 100 \cdot I_{(m \times m)}$

Nótese que los valores iniciales propuestos difieren significativamente con los establecidos en Sipreólico dado que estos últimos estaban demasiado alejados del óptimo.

Por último, en vez de suponer que las variables de estado siguen un paseo aleatorio (ver ecuación 3.2), se plantea un modelo más general. De esta forma la ecuación de estado del modelo sería, para $i = 1, 2, \dots, m$

$$\beta_{it} = \phi_i \beta_{it-1} + \eta_{it},$$

donde $\eta_{it} \sim N(0, \sigma_i^2)$, $cov(\eta_{it}, \eta_{j\tau}) = 0, \forall i \neq j$, y $t \neq \tau$. Sea $\beta_t = (\beta_{1t}, \dots, \beta_{mt})'$, $\Phi = diag(\phi_1, \phi_2, \dots, \phi_m)$ y $\eta_t = (\eta_{1t}, \dots, \eta_{mt})'$. Al igual que en los casos anteriores, se pueden recoger las ecuaciones de estado escalares en una única ecuación vectorial

$$\beta_t = \Phi\beta_{t-1} + \eta_t, \quad (3.10)$$

donde η_t es ruido blanco multivariante con matriz de varianzas covarianzas diagonal dada por $Q = diag(\sigma_1^2, \dots, \sigma_m^2)$. En este caso, la matriz ϕ también se estima a través del algoritmo EM. Nótese que mientras que la optimización de las varianzas del modelo no tiene por qué suponer un cambio cualitativo en la dinámica de los parámetros, al optimizar los coeficientes ϕ_{ii} , $i = 1, 2, \dots, m$ podemos pasar a modelos estacionarios. En el caso que nos ocupa, en la etapa de maximización del algoritmo EM además de R y Q se obtiene:

$$\Phi^{(j)} = S_{10}S_{00}^{-1}. \quad (3.11)$$

3.2.2. Resultados obtenidos

En las tablas 3.1 y 3.2 del anexo pueden verse los resultados numéricos obtenidos en el proceso de optimización. Se han analizado dos casos; en el primero de ellos se optimizan únicamente las matrices de varianzas y covarianzas del modelo, es decir, se optimizan R y Q imponiendo estructura diagonal en dichas matrices. Se mantiene, por tanto, el modelo propuesto en Sipreólico pero se estima por máxima verosimilitud el valor numérico de las matrices en vez de imponerlo de una manera ad-hoc. En un segundo caso, se optimizan estas dos matrices y la matriz de transición del estado Φ pero, de igual modo, imponiendo estructura diagonal, es decir, seguimos manteniendo la formulación univariante del problema. En las gráficas 3.1 y 3.2 se muestran, para cada uno de los modelos M1 \dots M7, los resultados de predicción, en términos del sesgo y RECMN respectivamente, obtenidos mediante la optimización de las matrices del sistema.

Del análisis de los gráficos de la figura 3.2 se obtienen las siguientes conclusiones:

- Optimizar los parámetros en la muestra de estimación reduce la RECMN de predicción.
- La RECMN se estabiliza rápidamente con el horizonte de predicción aumentando solo ligeramente en los primeros horizontes.
- Las mayores ganancias relativas se obtienen para horizontes de predicción mayores.

3.2. Algoritmo EM

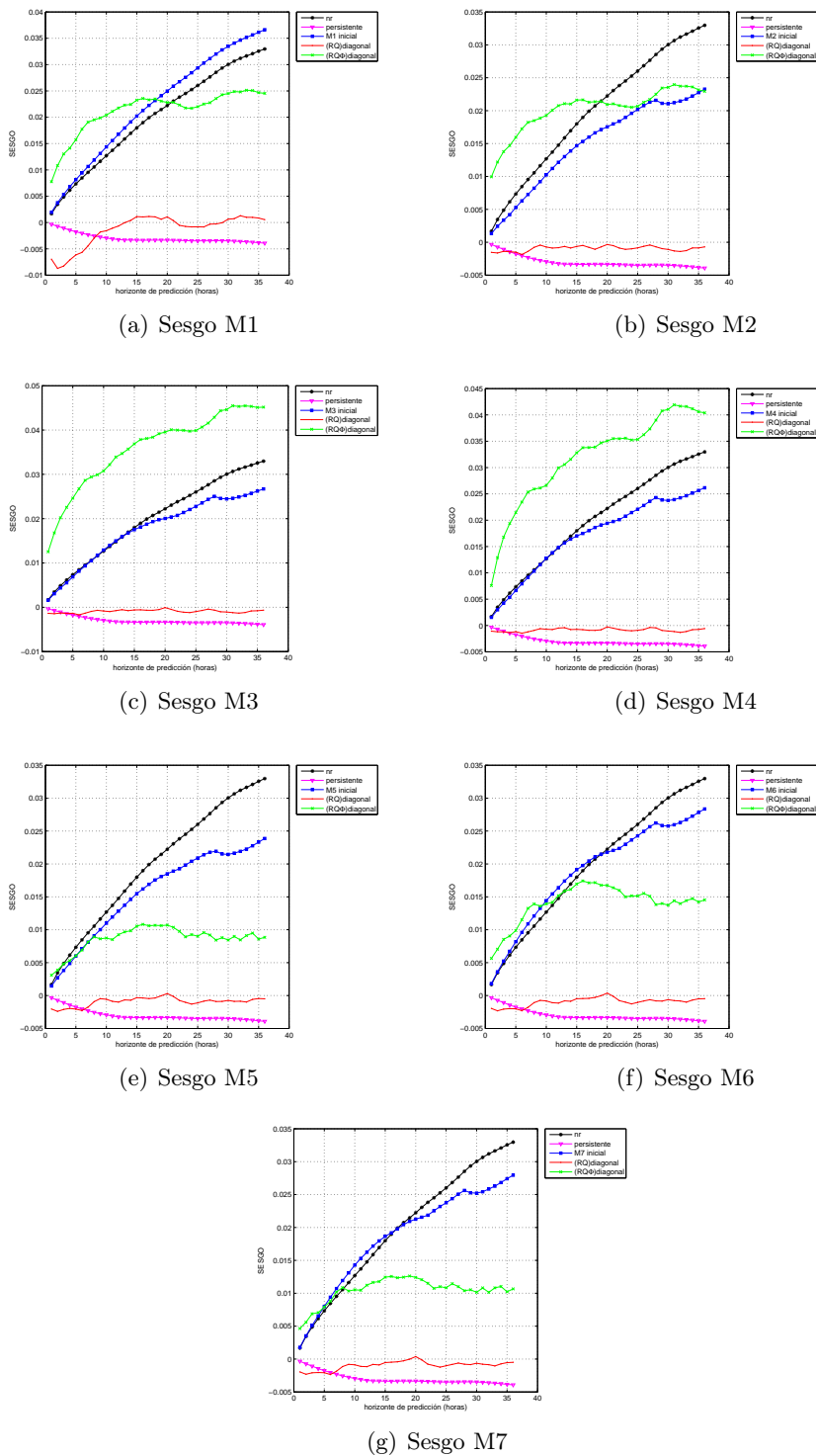


Figura 3.1: Sesgo de los errores de predicción para el mejor de los modelos de referencia, los modelos de partida (etiquetados como iniciales), y los modelos con las matrices estimadas por máxima verosimilitud. El modelo cuyos parámetros siguen un paseo aleatorio se ha etiquetado como RQdiagonal.

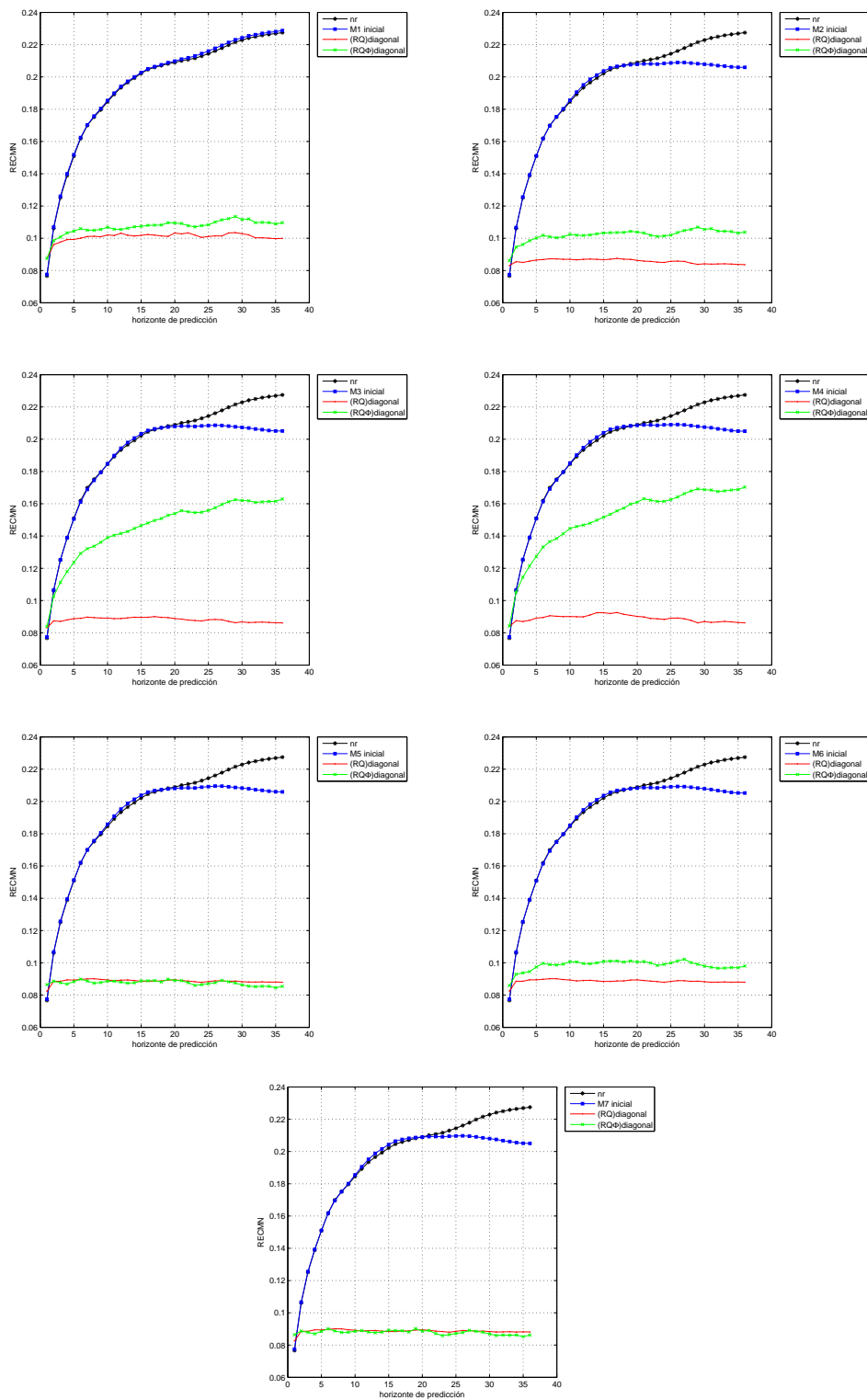


Figura 3.2: RECMN de los errores de predicción para el mejor de los modelos de referencia, los modelos de partida (etiquetados como iniciales), y los modelos univariantes con matrices estimadas por máxima verosimilitud. El modelo cuyos parámetros siguen un paseo aleatorio se ha etiquetado como RQ Φ diagonal.

En las gráficas de la figura 3.1 puede observarse cómo el sesgo en los modelos en los que se ha optimizado la matriz Φ (manteniendo estructura diagonal) aumenta respecto al valor obtenido en el modelo en el que se impone el paseo aleatorio para los parámetros, por lo que se realiza, con el fin de mejorar este estadístico, una nueva optimización imponiendo $\Phi(1, 1) = 1$ que, como se recordará, marca el nivel o valor de potencia media en torno al cual oscila la serie. Los resultados obtenidos se muestran en las gráficas de las figuras 3.3 y 3.4 para el sesgo y para la RECMN respectivamente. Se observa que el sesgo se corrige para todos los modelos (M1 a M7). La RECMN disminuye para los modelos M1 a M4 pero aumenta ligeramente en los modelos M5 a M7. En la tabla 3.3 del anexo se muestran los resultados numéricos. En dicha tabla puede comprobarse que los valores estimados de Φ están cercanos a la unidad en el modelo M1, autorregresivo puro, y para el resto de modelos en lo que respecta a los $\Phi_{i,i}$ correspondientes a la parte autorregresiva, pero no así para los $\Phi_{i,i}$ correspondientes a las predicciones de velocidad y dirección de viento meteorológicas por lo que, en este caso, el supuesto de paseo aleatorio para el vector de estado pudiera no ser adecuado.

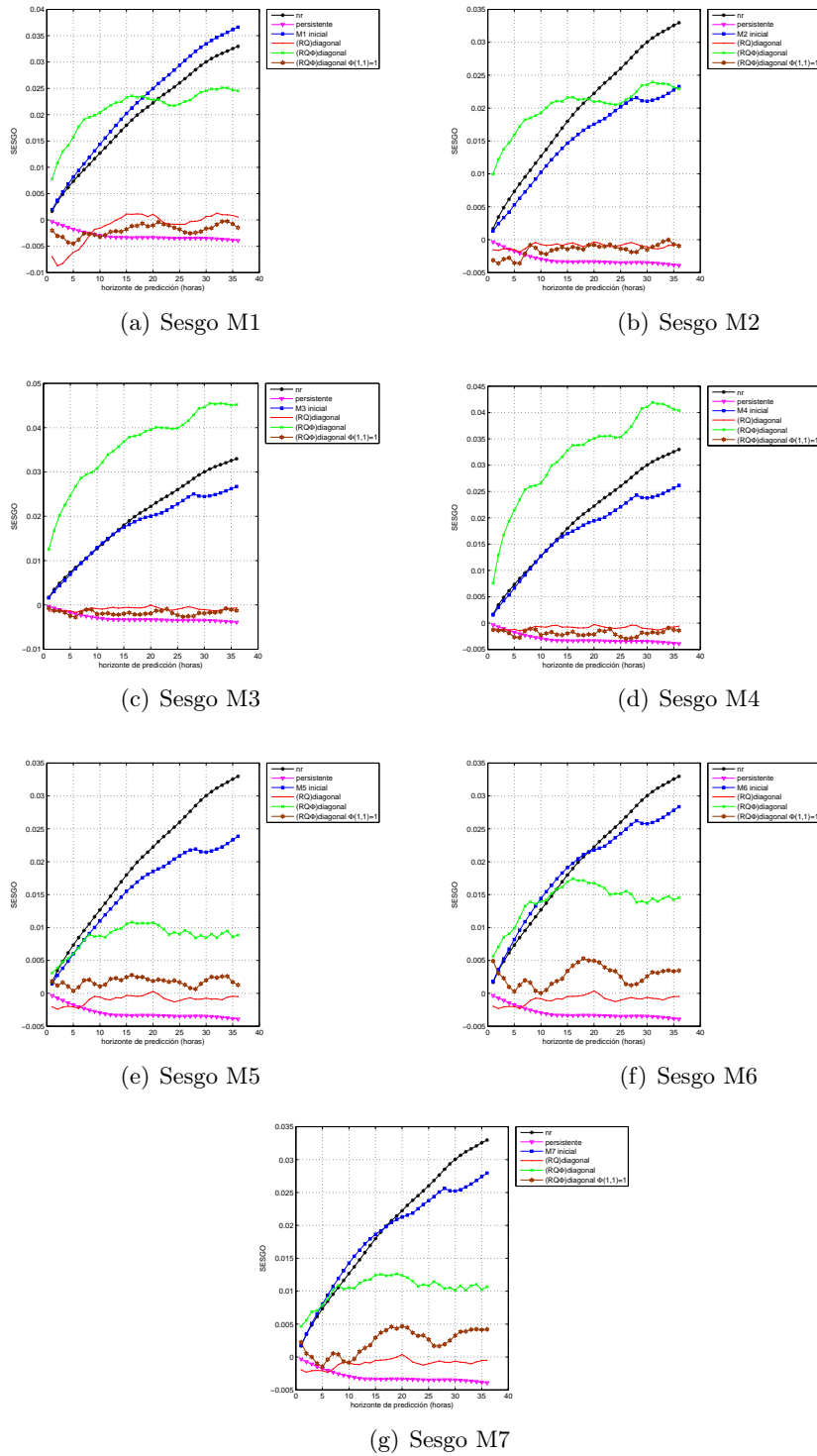


Figura 3.3: Sesgo de los errores de predicción para los modelos de referencia persistente y nueva referencia (etiquetado como nr), los modelos de partida (etiquetados como iniciales), y los modelos univariantes con matrices estimadas por máxima verosimilitud. El modelo con la restricción $\phi(1, 1) = 1$ se ha etiquetado como RQ ϕ diagonal $\phi(1, 1) = 1$.

3.2. Algoritmo EM

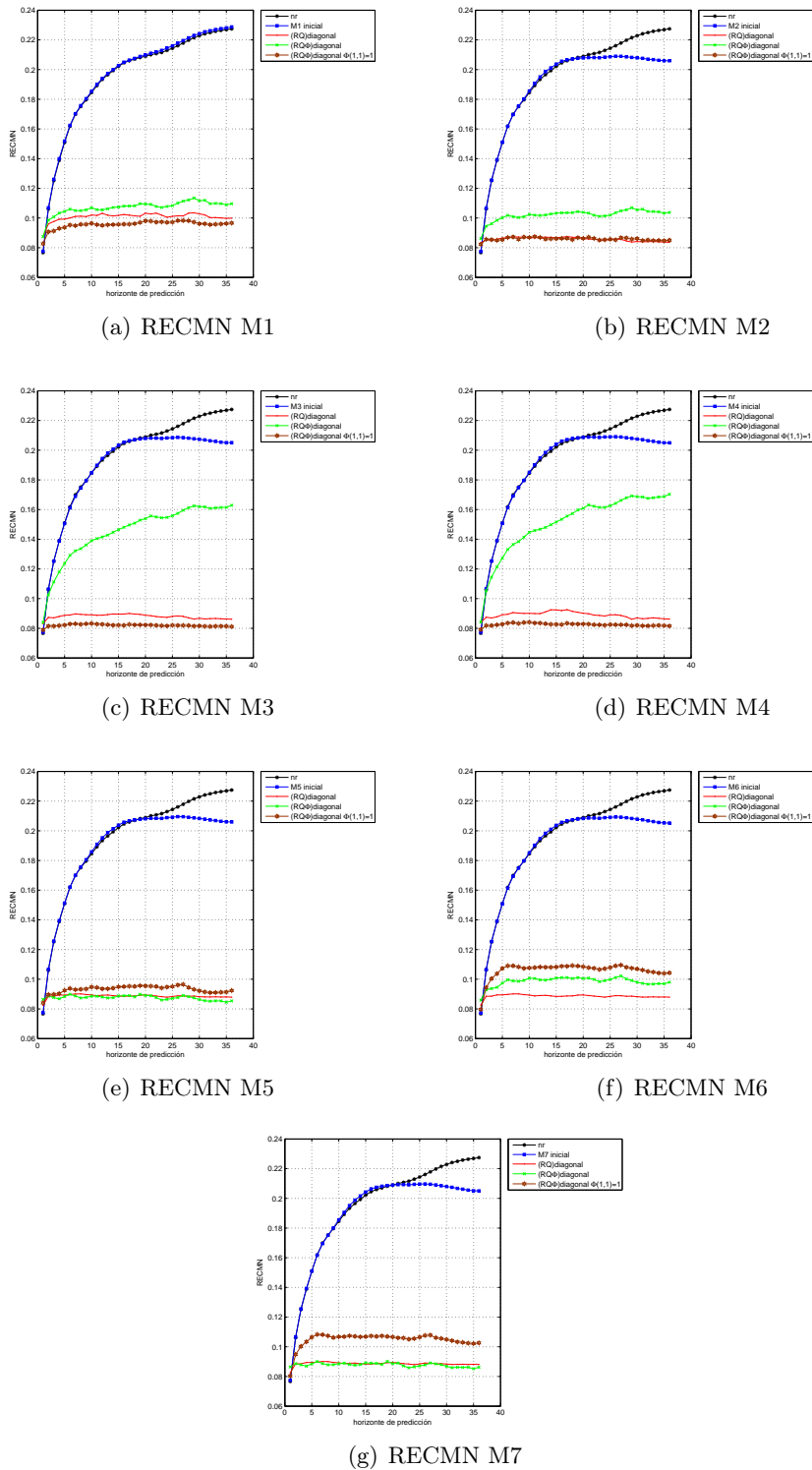


Figura 3.4: RECMN de los errores de predicción para el mejor de los modelos de referencia, los modelos de partida (etiquetados como iniciales), y los modelos univariantes con matrices estimadas por máxima verosimilitud. El modelo con la restricción $\phi(1, 1) = 1$ se ha etiquetado como RQ ϕ diagonal $\phi(1, 1) = 1$.

3.3. Nuevos modelos multivariantes para el estado

3.3.1. Introducción

En este apartado se generalizan los modelos autorregresivos de parámetros cambiantes contemplados en SIPREOLICO. Una de las razones para considerar en este análisis modelos de parámetros cambiantes es la naturaleza no lineal de la relación entre las variables. El carácter no lineal de esta relación puede ser debido, en parte, a la ausencia de regresores adicionales que pueden afectar a las variables de interés. (Véase, por ejemplo, [44]). Consideramos que la omisión de variables que pueden estar afectando a la potencia eólica podría producir parámetros que están correlados entre sí, ya que dichas variables omitidas pueden influir en varios parámetros simultáneamente. Para superar este problema se procede a generalizar los modelos contemplados para las variables de estado, es decir, los parámetros cambiantes en el tiempo del modelo, β_t , de tal forma que puedan estar correlados entre sí.

Se proponen dos tipos de modelos: en primer lugar, se propone un modelo donde los parámetros recogidos en el vector de estado β_t pueden estar correlados entre sí a través del ruido. Esto es, en vez de considerar que la matriz Q es diagonal, consideramos que esta matriz está llena. Esta idea fue introducida por [57] en el contexto de regresión y se denominó SURE ("Seemingly Unrelated Regression Equations") al modelo resultante. En este trabajo se adapta el modelo SURE al marco de predicción eólica de parámetros cambiantes con el tiempo y se denomina ahora el modelo como "de coeficientes aparentemente no relacionados" y lo denotamos por las siglas SUC ("Seemingly Unrelated Coefficients"). En el sistema compuesto por las ecuaciones (3.1) y (3.10), se mantiene la primera ecuación, mientras que los supuestos acerca del término de error en la ecuación (3.10) se sustituyen por $\eta_t \sim N(0, Q)$, donde ahora Q puede estar llena (las covarianzas entre los términos de error pueden ser distintas de cero). Nótese que, al igual que en la sección anterior, el término de error es $\eta_t = (\eta_{1t}, \dots, \eta_{mt})$, donde m depende del número de regresores incluidos en el modelo. Por supuesto, todavía se mantiene el supuesto de ausencia de correlación serial en el error, esto es, se sigue manteniendo el supuesto de que $cov(\eta_t, \eta_\tau) = 0, \forall t \neq \tau$. El vector de variables de estado $\beta_t = (\beta_{1t}, \dots, \beta_{mt})$ en el instante t evoluciona según el siguiente modelo

$$\beta_t = \beta_{t-1} + w_t; \quad \eta_t \sim N(0, Q).$$

Esto supone que las componentes del vector de estado del sistema pueden estar correladas (se puede suponer, por ejemplo, que la variación en los parámetros es debida a

causas comunes), aunque dicha correlación entra en el modelo a través del ruido de la ecuación de transición de estados.

La segunda generalización supone un modelo VAR(1) para el vector de estado $\beta_t = (\beta_{1t}, \dots, \beta_{mt})$. En este caso, la ecuación de transición (3.10) se sustituye por

$$\beta_t = \Phi\beta_{t-1} + \eta_t \quad (3.12)$$

donde Φ es una matriz ($m \times m$) que permite relacionar explícitamente no sólo el elemento i -ésimo de β_t con su propio pasado β_{it-1} , sino también con los valores pasados del resto de coeficientes β_{jt-1} . Además, se supone que la matriz de varianzas-covarianzas del ruido Q está llena, esto es, $\eta_t \sim N(0, Q)$, donde Q no tiene por qué ser diagonal. Se sigue manteniendo el supuesto de ausencia de correlación serial en el término de error ($cov(\eta_t, \eta_\tau) = 0, \forall t \neq \tau$).

Esto introduce algunas modificaciones en la estimación contemplada en la sección anterior. La esperanza de la función de densidad conjunta se escribe ahora como

$$\begin{aligned} E \{ \ln f_{Y,\beta}(Y, \beta; \Theta) \} &= \ln |P_0| + tr \left\{ P_0^{-1} \left[P_0^T + (\beta_0^T - b_0) (\beta_0^T - b_0)' \right] \right\} + \\ &+ T \ln |Q| + tr \left\{ Q^{-1} \left[S_{11} - S'_{10} \Phi - \Phi S'_{10} + \Phi S_{00} \Phi' \right] \right\} + T \ln |R| + \\ &+ tr \left\{ R^{-1} \sum_{t=1}^T \left[(p_t - A_t \beta_t^T) (p_t - A_t \beta_t^T)' + A_t P_t^T A_t' \right] \right\}. \end{aligned} \quad (3.13)$$

Además, mientras que cuando las variables de estado seguían paseos aleatorios univariantes, se podía procesar la ecuación del algoritmo EM para cada componente del vector de estado por separado, ahora, debido a las relaciones de dependencia que existen entre las variables de estado, hay que procesar cada etapa del algoritmo EM de manera multivariante. De ahí que ahora se calculan los vectores y matrices de primeros y segundos momentos suavizados del vector de estado de la etapa de E del algoritmo como

$$\begin{aligned} \beta_t^T &= E(\beta_t | Y) \\ P_t^T &= E((\beta_t - \beta_t^T)(\beta_t - \beta_t^T)' | Y) \\ P_{t,t-1}^T &= E((\beta_t - \beta_t^T)(\beta_{t-1} - \beta_{t-1}^T)' | Y). \end{aligned}$$

Las matrices $S_{ij}, i, j = 0, 1$ ya no tienen por qué ser diagonales y se calculan según

$$S_{11} = \sum_{t=1}^T (\beta_t^T (\beta_t^T)' + P_t^T),$$

$$S_{10} = \sum_{t=1}^T (\beta_t^T (\beta_{t-1}^T)' + P_{t,t-1}^T)$$

y

$$S_{00} = \sum_{t=1}^T (\beta_{t-1}^T (\beta_{t-1}^T)' + P_{t-1}^T).$$

En la etapa M de maximización, hay que añadir a las ecuaciones (3.9) y (3.8) para la estimación de R y Q , la siguiente ecuación

$$\Phi^{(j)} = S_{10} S_{00}^{-1}$$

para estimar Φ en la iteración j -ésima.

3.3.2. Resultados obtenidos

En las gráficas 3.5 y 3.6 se muestran los efectos que tienen en la predicción los nuevos modelos propuestos, y optimizados mediante el algoritmo EM, en términos de sesgo y RECMN respectivamente. Del análisis de los gráficos se obtienen las siguientes conclusiones:

- Los nuevos modelos propuestos reducen la RECMN en todos los casos (M1 a M7).
- La RECMN se estabiliza rápidamente con el horizonte de predicción aumentando solo ligeramente en los primeros horizontes.
- Las mayores ganancias relativas se obtienen para horizontes de predicción grandes.
- A diferencia de la optimización de los modelos de partida, el sesgo de los errores de predicción se mejora sin necesidad de imponer la restricción $\Phi(1, 1) = 1$, aunque es de destacar que en los modelos M3 y M4 esta mejora no se alcanza hasta los horizontes 17 y 8, respectivamente.

3.3. Nuevos modelos multivariantes para el estado

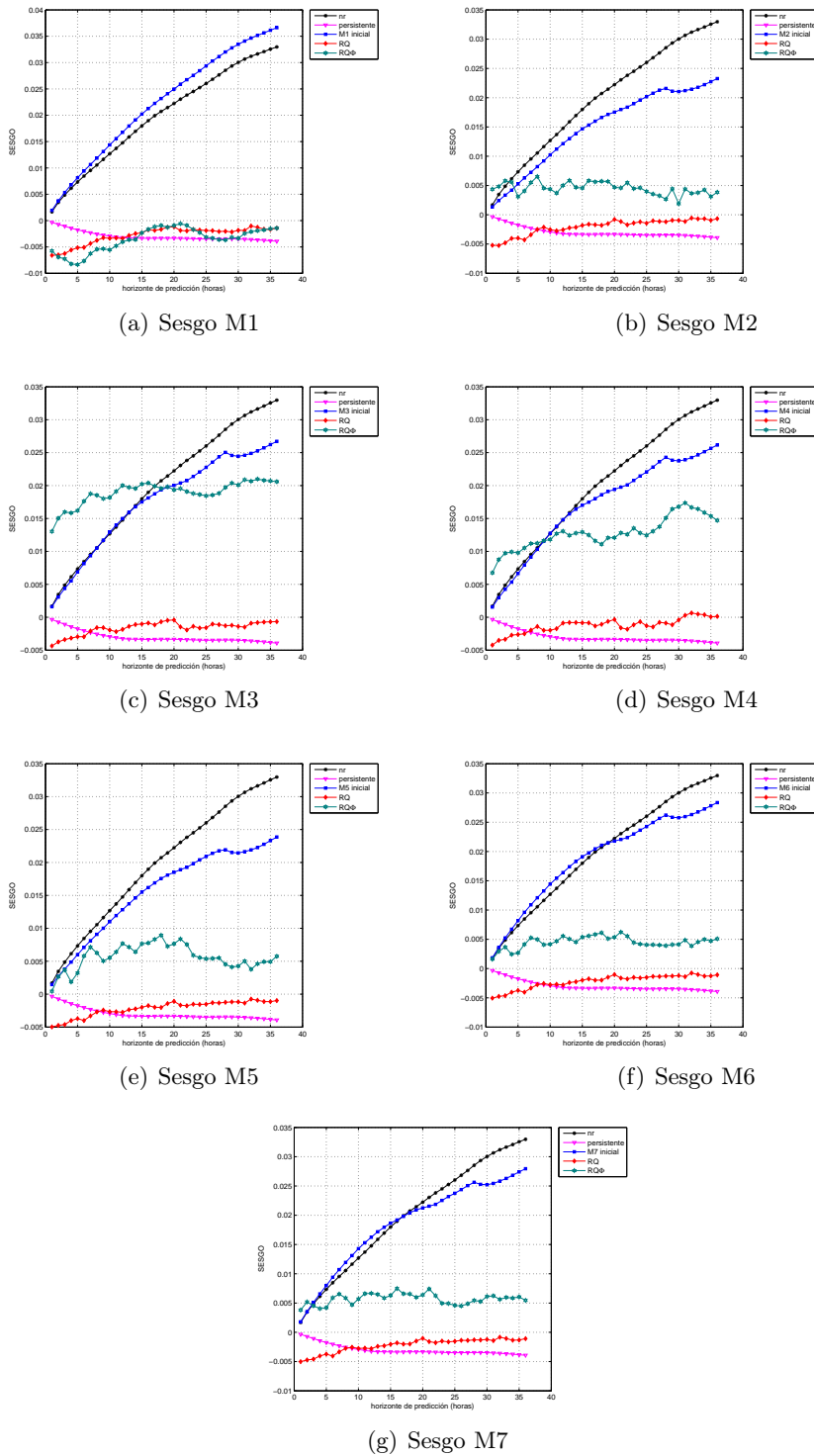


Figura 3.5: Sesgo de los nuevos modelos multivariantes propuestos. El modelo SUC se ha etiquetado como RQ y el modelo VAR(1) como $RQ\phi$. También se han incluido los modelos de referencia e iniciales.

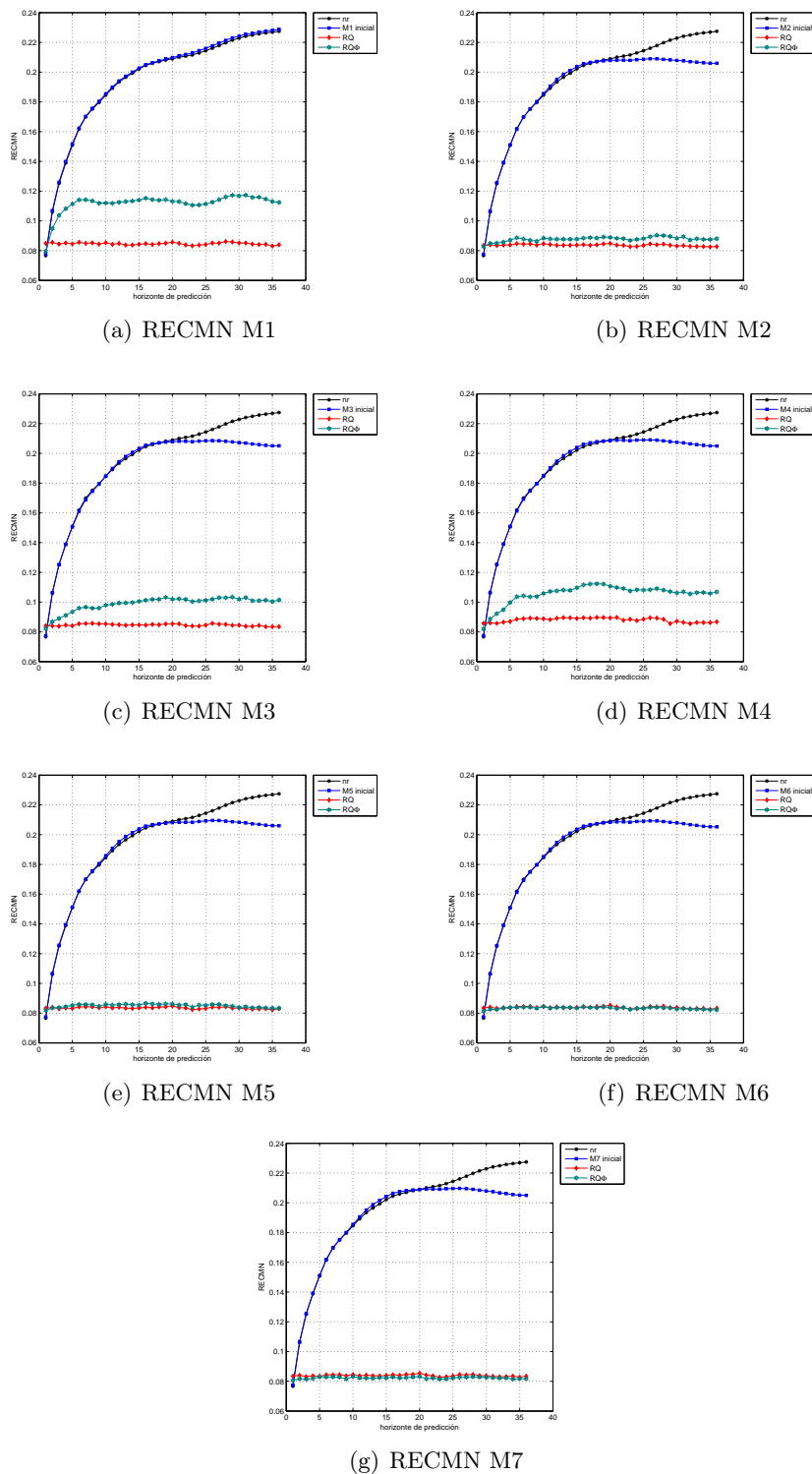


Figura 3.6: RECMN de los nuevos modelos multivariantes propuestos. El modelo SUC se ha etiquetado como RQ y el modelo VAR(1) como RQ ϕ . También se han incluido los modelos de referencia e iniciales.

3.4. Estrategia de funcionamiento en tiempo real.

Antes de proponer la estrategia de operación en tiempo real, resolveremos las dos cuestiones siguientes:

1. Una vez estimados los parámetros del modelo mediante máxima verosimilitud, ¿existe algún modelo mejor que los demás en términos de predicción?
2. ¿Sería adecuado optimizar los modelos para cada horizonte de predicción?

Para resolver la primera cuestión, en la figura 3.7 se muestra el sesgo de los errores de predicción de todos los casos propuestos y en la figura 3.8 se representa gráficamente el coeficiente de mejora de los modelos propuestos respecto de los modelos M1 a M7 de partida en lo que respecta a la RECMN. Analizando dichas figuras podemos observar:

- Respecto al sesgo, la estructura que mejores resultados proporciona es SUC (Seemingly Unrelated Coefficients) para todos los modelos M1,...,M7, incluso mejorando los resultados del modelo persistente.
- En términos de la RECMN, la mejor estructura es también SUC, aunque es de destacar que la propuesta VAR(1) también obtiene resultados similares para los modelos M5, M6 y M7 (aquellos modelos que tienen como regresor la dirección de viento). Para M2 los resultados de todas las estructuras consideradas son muy similares.

Imponer el mismo valor para las matrices para todos los horizontes de predicción significa un gran ahorro de coste computacional. Lo ideal sería optimizar los modelos para cada horizonte de predicción, pero si la ganancia, en términos de la RECMN es pequeña, podría ser preferible, de cara a una aplicación en tiempo real, usar las mismas matrices para todos los horizontes de predicción. En las secciones anteriores de este capítulo se ha demostrado que mejor que imponer el valor de las matrices de una manera ad-hoc es siempre más adecuado estimarlas, de manera que recojan, en cierto sentido, las características particulares de cada parque eólico. De esta manera, estimando el valor de las matrices para $h = 1$ hemos obtenido una reducción de la RECMN cercana al 60 % considerando todos los horizontes de predicción. Dado que el modelo SUC era el que mejores resultados globales proporcionaba, se han estimado las matrices de parámetros de estos modelos para todos los horizontes de predicción, obteniéndose una ganancia entre el 4 % y el 7 % respecto a la optimización de las matrices sólo a horizonte uno (véase la figura 3.9). Es decir, es importante estimar el modelo, al menos para un horizonte

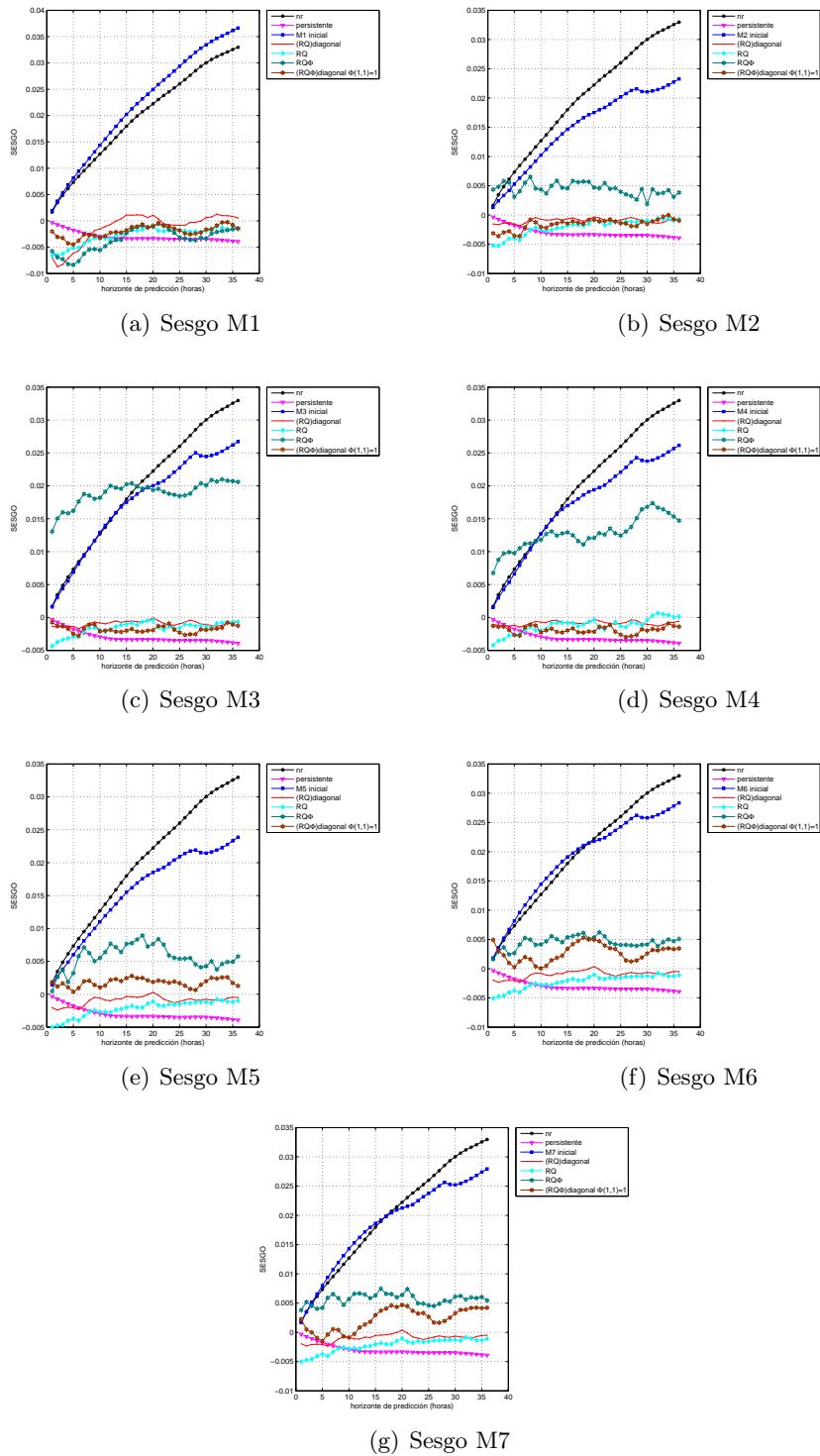
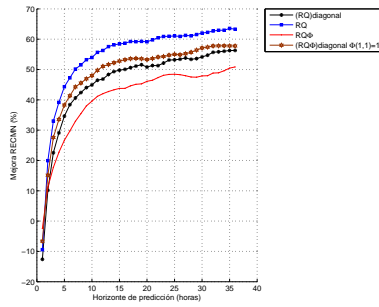
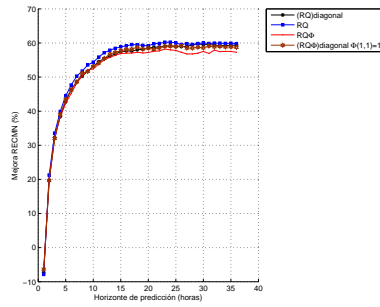


Figura 3.7: Comparación de resultados. Sesgo del error de predicción de todos los modelos propuestos.

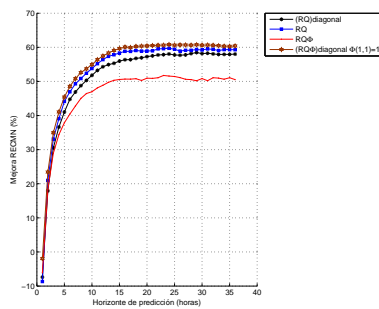
3.4. Estrategia de funcionamiento en tiempo real.



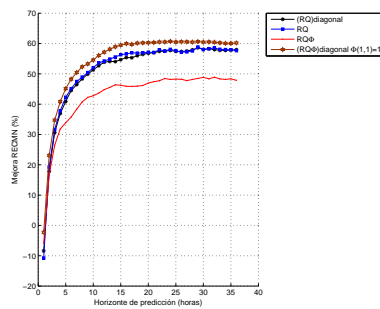
(a) Mejora RECMN M1



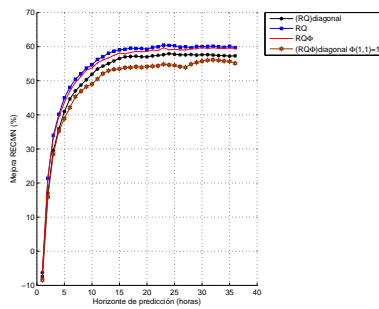
(b) Mejora RECMN M2



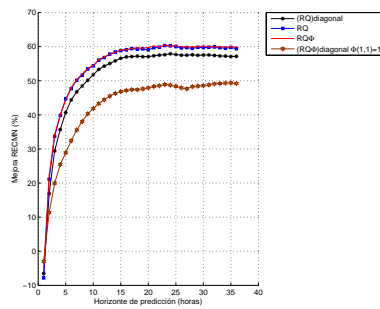
(c) Mejora RECMN M3



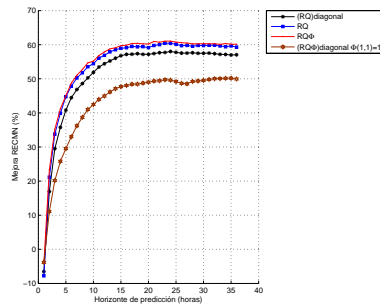
(d) Mejora RECMN M4



(e) Mejora RECMN M5



(f) Mejora RECMN M6



(g) Mejora RECMN M7

Figura 3.8: Comparación de resultados de los modelos propuestos. Coeficiente de mejora de la RECMN respecto a los modelos de partida.

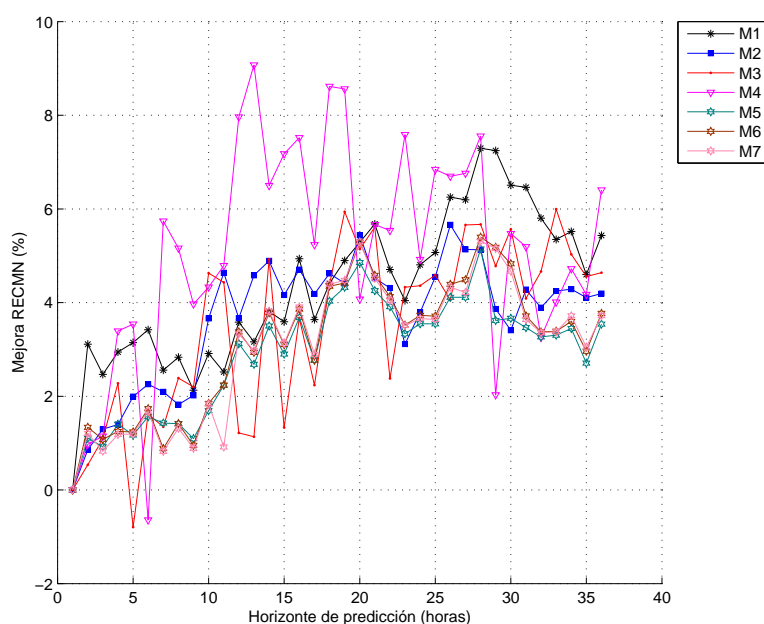


Figura 3.9: Mejora de la RECMN de los modelos optimizados para cada horizonte de predicción respecto al modelo con matrices estimadas unicamente para $h=1$.

de predicción, para reducir la RECMN. Haciendo esto, se captan, en cierto sentido, las características propias de cada parque eólico. Se podrían estimar las matrices del modelo para cada horizonte de predicción, si el tiempo de computación lo permite, pero en una herramienta en tiempo real, la ganancia adicional (en términos de la RECMN) es menor.

Dado que las estimaciones calculadas para $h = 1$ son óptimas, otra alternativa podría ser iterar las predicciones en vez de calcularlas directamente con un modelo para cada horizonte de predicción. Los resultados obtenidos son mucho peores y en la figura 3.10 únicamente se muestran los resultados para los primeros 10 horizontes de predicción dado el incremento exponencial de la RECMN (para el sesgo se obtienen resultados similares).

Teniendo en cuenta todas estas consideraciones, la estrategia de operación en tiempo real sería:

- Estimar las matrices de los modelos para un horizonte de predicción. En esta tesis se ha propuesto $h = 1$ dado que en este caso el término del error en la ecuación de medida sería ruido blanco y, por tanto, la estimación sería óptima.
- Aplicar las matrices obtenidas para $h = 1$ a todos los horizontes de predicción.
- No iterar las predicciones, sino que es más conveniente calcular predicciones di-

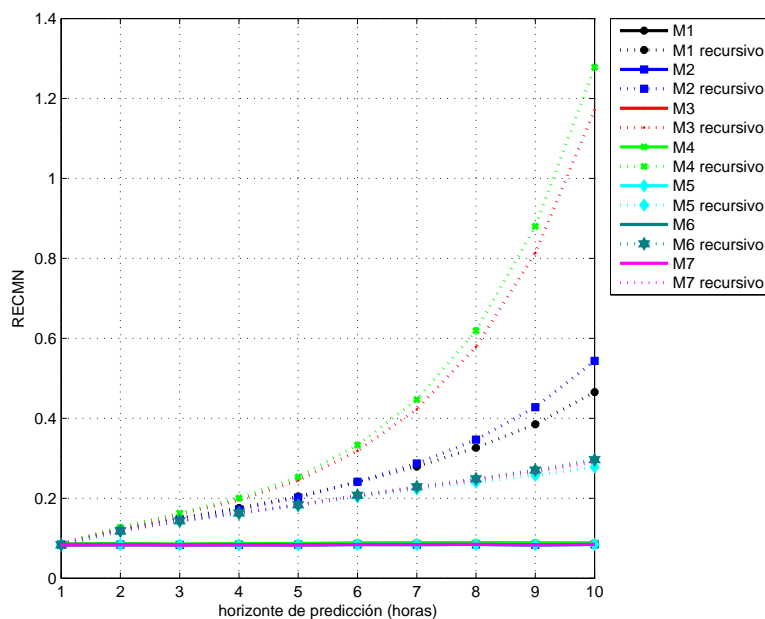


Figura 3.10: RECMN para los modelos recursivos.

rectamente para cada p_{t+h} .

3.5. Bootstrap de los parámetros

Una vez realizada la estimación de las matrices del sistema por máxima verosimilitud, es necesario proporcionar un valor de la incertidumbre asociada a dicho valor. En este trabajo este cálculo se ha realizado mediante técnicas de remuestreo bootstrap. El bootstrap descansa en la analogía entre la muestra y la población de la cual la muestra es extraída. De acuerdo a [9], dada una muestra con n observaciones, el estimador no paramétrico de máxima verosimilitud de la distribución poblacional es la función de densidad de probabilidad que asigna una masa de probabilidad de $1/n$ a cada una de las observaciones. La idea central es que muchas veces puede ser mejor extraer conclusiones sobre las características de la población estrictamente a partir de la muestra que se maneja, que estableciendo supuestos quizás poco realistas sobre la población. El bootstrap implica remuestreo de los datos obtenidos en una muestra con reemplazamiento para generar una estimación empírica de la distribución muestral completa de un estadístico. El bootstrap puede considerarse como un tipo especial de simulación denominada simulación basada en los datos.

Los métodos de remuestreo presentan varias ventajas. Primero, no es necesario realizar supuestos acerca de la población y permiten además estimar la distribución de un

estadístico de manera empírica. Segundo, no exigen tamaños de muestra excesivamente grandes para ser precisos y confiables (aunque sí es exigente en cuanto a número de réplicas bootstrap a realizar). Tercero, tienen magnitudes de sesgo significativamente más pequeñas que otros modelos de inferencia. Cuarto, son más fáciles de comprender y aplicar debido a su automaticidad, potencia y precisión.

Es decir, mediante bootstrap pretendemos, además de calcular la incertidumbre asociada a la estimación de los hiperparámetros, poder analizar su distribución. En particular, analizaremos la forma de la distribución bootstrap del estimador del hiperparámetro. Dado que la forma de la distribución bootstrap es una aproximación a la forma de la distribución muestral, analizaremos la distribución bootstrap para comprobar la normalidad de la distribución muestral. Por otro lado, un estimador de un hiperparámetro presenta sesgo si la distribución muestral no está centrada en el verdadero valor del parámetro. Podemos pues comprobar la existencia de sesgo analizando si la distribución bootstrap del estimador está centrado en el valor estimado para la muestra de partida. De manera más precisa podemos decir que el sesgo del estimador es la diferencia entre la media de su distribución muestral y el verdadero valor del parámetro. La estimación bootstrap del sesgo es la diferencia entre la media de la distribución bootstrap y el valor del estimador del hiperparámetro en la muestra de partida. La desviación típica bootstrap de un estimador se calcula para estimar la desviación típica de la distribución muestral del mismo.

En el presente trabajo seguiremos la metodología desarrollada en [49] y [50] para el cálculo de las réplicas bootstrap en series temporales. Dicha metodología parte de la representación del filtro de Kalman en función de las innovaciones:

$$\epsilon_t = y_t - A_t x_t^{t-1} \quad (3.14)$$

$$\Sigma_t = A_t P_t^{t-1} A_t' + R \quad (3.15)$$

$$x_{t+1}^t = \Phi x_t^{t-1} + \Phi K_t \epsilon_t \quad (3.16)$$

$$K_t = \left[\Phi P_t^{t-1} A_t' + S \right] \Sigma_t^{-1} \quad (3.17)$$

$$P_{t+1}^t = \Phi P_t^{t-1} \Phi' + Q - K_t \Sigma_t^{-1} K_t' \quad (3.18)$$

$$y_t = A_t x_t^{t-1} + \epsilon_t \quad (3.19)$$

En el algoritmo bootstrap se trabaja con los residuos estandarizados (e_t) ya que estos tienen, al menos, los dos primeros momentos idénticos.

$$e_t = \Sigma_t^{-1/2} \epsilon_t \quad (3.20)$$

siendo $\Sigma_t^{-1/2}$ la matriz definida por $\Sigma_t^{-1/2} \Sigma_t^{-1/2} = \Sigma_t^{-1}$.

Los coeficientes o matrices del sistema así como su estructura de correlación está definida de manera unívoca por el vector de hiperparámetros Θ_0 , es decir, $\Phi = \Phi(\Theta_0)$, $Q = Q(\Theta_0)$, $A_t = A_t(\Theta_0)$ y $R = R(\Theta_0)$. Ignorando una constante, la función de verosimilitud puede formularse en función del vector de innovaciones como

$$\begin{aligned} -2 \ln L_y(\Theta) &= \sum_{t=1}^n \left[\ln |\Sigma_t(\Theta)| + \epsilon_t(\Theta)' \Sigma_t(\Theta)^{-1} \epsilon_t(\Theta) \right] \\ &= \sum_{t=1}^n \left[\ln |\Sigma_t(\Theta)| + e_t(\Theta)' e_t(\Theta) \right]. \end{aligned} \quad (3.21)$$

Recordamos que no es necesario que el modelo sea gaussiano para considerar 3.21 como la función objetivo a emplear en la estimación de los hiperparámetros. Sea $\hat{\Theta}$ la estimación máximo verosímil de Θ_0 obtenida mediante el algoritmo EM. Sean $\epsilon_t(\hat{\Theta})$ y $\Sigma_t(\hat{\Theta})$ los valores del vector de innovaciones y su matriz de varianzas y covarianzas obtenidas ejecutando las ecuaciones (3.14) a (3.18) con el vector de hiperparámetros $\hat{\Theta}$. El procedimiento bootstrap se realiza siguiendo las siguientes etapas:

1. Se construye el vector de residuos estandarizados

$$e_t(\hat{\Theta}) = \Sigma_t^{-1/2}(\hat{\Theta}) \epsilon_t(\hat{\Theta})$$

2. Obtener $\{e_1^*(\hat{\Theta}), \dots, e_n^*(\hat{\Theta})\}$ remuestreando $\{e_1(\hat{\Theta}), \dots, e_n(\hat{\Theta})\}$ n veces con reemplazamiento.
3. Construir el conjunto de datos bootstrap $Y^* = \{y_1^*(\hat{\Theta}), \dots, y_n^*(\hat{\Theta})\}$ de la siguiente manera: Dado el vector $\xi_t = (x_{t+1}^t, y_t^t)$, uniendo las ecuaciones (3.16) y (3.19), se obtiene la ecuación vectorial de primer orden para ξ_t :

$$\xi_t = F_t \xi_{t-1} + H_t e_t \quad (3.22)$$

siendo,

$$F_t = \begin{bmatrix} \Phi & 0 \\ A_t & 0 \end{bmatrix} \quad H_t = \begin{bmatrix} K_t \Sigma_t^{-1/2} \\ \Sigma_t^{-1/2} \end{bmatrix}.$$

Para obtener $Y^* = \{y_1^*(\hat{\Theta}), \dots, y_n^*(\hat{\Theta})\}$ se resuelve la ecuación (3.22) con $e_t^*(\hat{\Theta})$ en vez de e_t .

4. A partir del conjunto de datos bootstrap $Y^* = \{y_t^*; t = 1, \dots, n\}$, calcular la verosimilitud $L_{Y^*}(\Theta)$ para obtener el estimador máximo verosímil de Θ , es decir, $\hat{\Theta}^*$.
5. Repetir las etapas 2 a 4 un número elevado de veces, B , con el fin de obtener un conjunto bootstrap de parámetros estimados $\{\hat{\Theta}_b^*; b = 1, \dots, B\}$. La distribución muestral de $\hat{\Theta} - \Theta_0$ puede entonces ser aproximada por la distribución de $\hat{\Theta}_b^* - \hat{\Theta}$ con $b = 1, \dots, B$.

Una vez calculadas las réplicas bootstrap estimaremos los intervalos de confianza mediante la aproximación normal y mediante el método de los percentiles (véase [8]). Si la distribución bootstrap de los hiperparámetros estimados presenta un perfil cercano a la distribución normal, los intervalos obtenidos por ambos métodos no diferirán significativamente.

En las figuras 3.11 y 3.12 se han representado gráficamente los histogramas bootstrap de los hiperparámetros estimados mediante máxima verosimilitud y los diagramas QQplot para el modelo M1 con estructura SUC. En la tabla 3.6 se muestran numéricamente los resultados obtenidos. Para el resto de modelos, M2, M3, ..., M7 los resultados obtenidos son similares. Como conclusiones podemos mencionar, atendiendo al QQplot, que los hiperparámetros podrían considerarse normales excepto para R, Q(3,3), Q(5,5), Q(2,5) y Q(3,5). En el gráfico de los histogramas puede comprobarse cómo el valor estimado para los hiperparámetros (indicado por la línea roja continua) presenta un sesgo importante respecto a la distribución bootstrap. También puede observarse la fuerte asimetría de los intervalos de confianza con colas izquierdas más pesadas. Además, los intervalos calculados por ambos métodos difieren significativamente. En este caso, en el que no se puede considerar que los hiperparámetros obtenidos presenten una distribución normal, los hiperparámetros obtenidos mediante el algoritmo EM son cuasimáximo verosímiles.

Existen aplicaciones recientes de la metodología bootstrap a series temporales en el espacio de los estados, como son [45] y [1]. En ambos casos se desarrollan algoritmos para la estimación de intervalos de predicción para sistemas con matrices invariantes

en el tiempo. Es de destacar que existen hasta el momento escasos desarrollos a nivel teórico y práctico de las técnicas bootstrap para su aplicación a series temporales con matrices del sistema variantes con el tiempo, por lo que la aplicación práctica realizada puede considerarse novedosa.

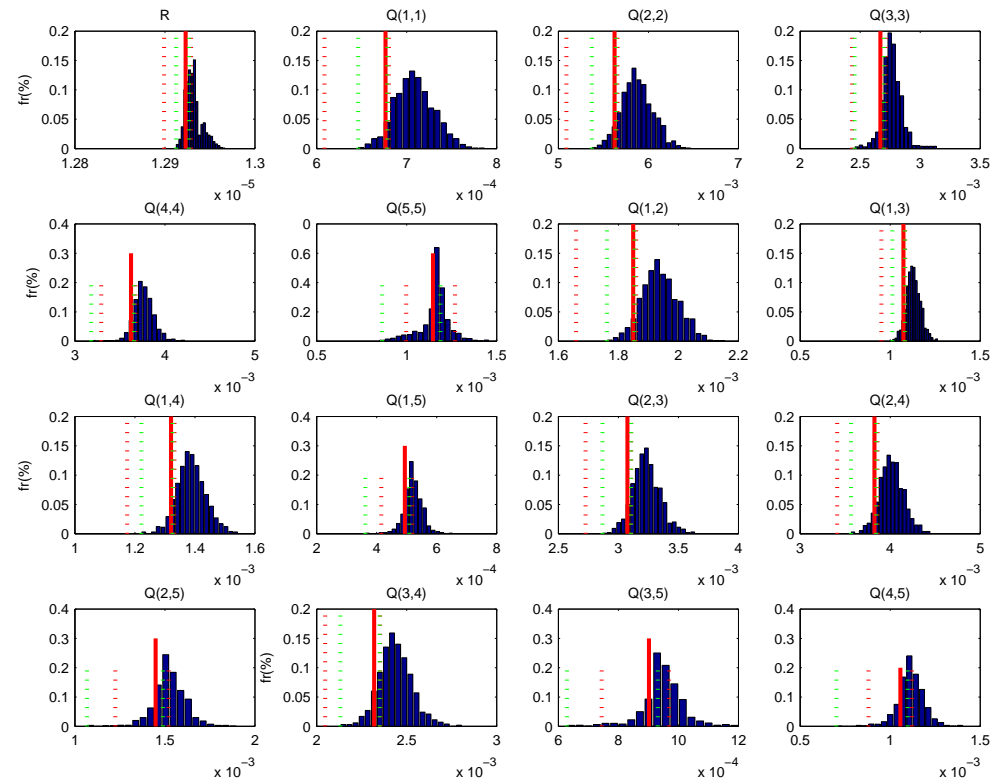


Figura 3.11: Histograma de la distribución bootstrap de los hiperparámetros del modelo SUC M1. Con línea roja continua se representa el valor óptimo del hiperparámetro obtenido mediante el algoritmo EM. Con línea roja discontinua se ha representado el intervalo de confianza obtenido mediante la aproximación normal y con línea verde discontinua el intervalo obtenido mediante el método de los percentiles.

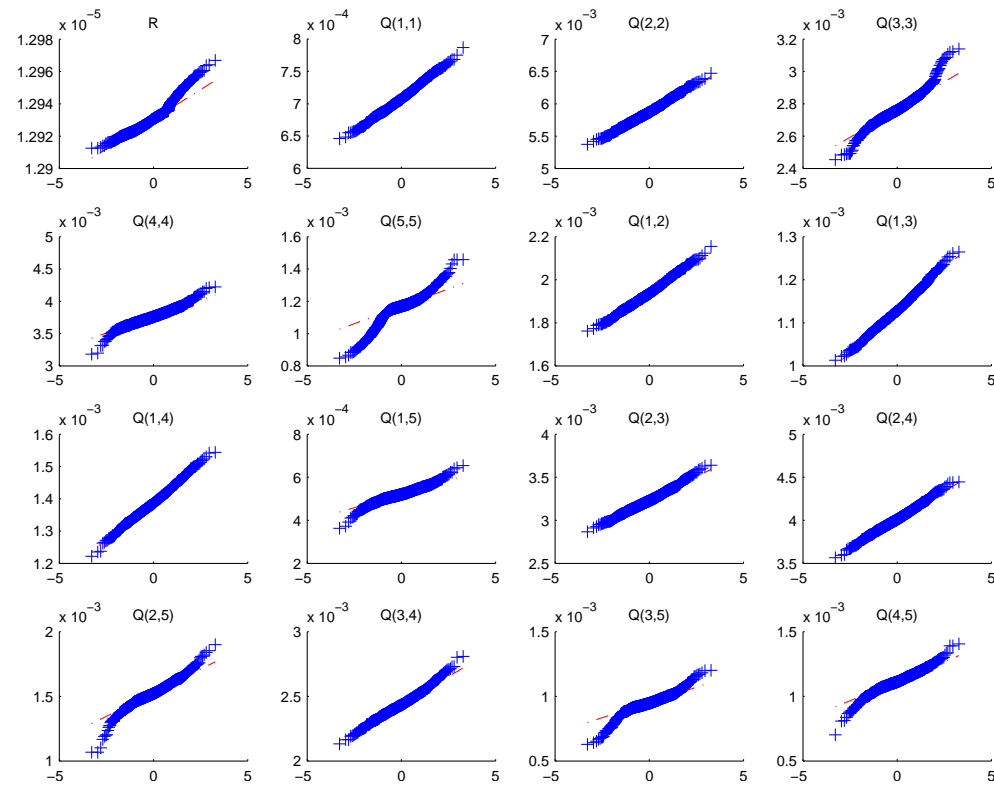


Figura 3.12: QQplot de la distribución bootstrap de los hiperparámetros del modelo SUC M1.

3.6. Apéndice: Resultados empíricos

3.6.1. Resultados de la optimización mediante el algoritmo EM

En las tablas 3.1, 3.2, 3.4 y 3.5 se muestran los resultados empíricos obtenidos mediante optimización máximo verosímil a través del algoritmo EM de los modelos M1 a M7 de Sipreólico en los siguientes casos:

- Estimación de las matrices R y Q , manteniendo la estructura univariante del problema.
- Estimación de las matrices R , Q y ϕ , manteniendo la estructura univariante del problema.
- Estimación de las matrices R y Q planteando la estructura SUC.
- Estimación de R , Q y ϕ .

Cuadro 3.1: Optimización de R y Q diagonales.

	M1	M2	M3	M4	M5	M6	M7
loglik	-1.8980e+04	-1.8842e+04	-1.8790e+04	-1.8700e+04	-1.8595e+04	-1.8605e+04	-1.8602e+04
R	2.2722e-03	1.0489e-03	1.5204e-03	1.3132e-03	1.5071e-03	1.4895e-03	1.4846e-03
q(1,1)	1.0171e-04	5.9508e-04	5.5524e-04	5.1668e-04	6.3586e-04	6.3224e-04	6.3276e-04
q(2,2)	1.7983e-02	9.4199e-04	9.1469e-04	8.9317e-04	9.6721e-04	9.6545e-04	9.6382e-04
q(3,3)	2.4271e-04	8.8738e-04	8.5970e-04	8.3593e-04	9.1559e-04	9.1534e-04	9.1491e-04
q(4,4)	2.8367e-04	8.7842e-04	8.3873e-04	8.1393e-04	9.1241e-04	9.1037e-04	9.0992e-04
q(5,5)	1.1682e-04	8.4160e-04	5.4020e-04	6.5205e-04	6.1589e-06	5.8391e-06	5.6788e-06
q(6,6)	-	2.7146e-05	1.0492e-05	1.1186e-05	9.5414e-04	9.4444e-04	9.4370e-04
q(7,7)	-	-	5.8690e-08	1.0481e-07	4.4070e-07	9.8578e-04	9.8825e-04
q(8,8)	-	-	-	5.3826e-10	3.2902e-07	4.3577e-07	9.9460e-04
q(9,9)	-	-	-	-	-	3.0008e-07	4.3775e-07
q(10,10)	-	-	-	-	-	-	3.0147e-07

Cuadro 3.2: Optimización de R, Q y Φ diagonales.

	M1	M2	M3	M4	M5	M6	M7
loglik	-1.8958e+04	-1.9102e+04	-1.8804e+04	-1.8735e+04	-1.9579e+04	-1.9395e+04	-1.9580e+04
R	2.4218e-03	1.1181e-03	1.9305e-03	9.9801e-04	2.6092e-04	1.4466e-03	3.1691e-04
q(1,1)	1.4777e-04	1.4692e-04	3.2483e-04	4.3476e-04	4.1166e-04	1.7324e-05	3.5494e-04
q(2,2)	9.8238e-04	9.6846e-04	9.5785e-04	9.2020e-04	2.2777e-03	9.7674e-04	9.7660e-04
q(3,3)	9.7594e-04	9.6807e-04	9.6564e-04	9.6332e-04	2.2966e-02	1.0473e-02	2.3127e-02
q(4,4)	1.2353e-02	1.0171e-02	9.6383e-04	9.6141e-04	9.2891e-03	7.6335e-03	1.0459e-02
q(5,5)	9.6397e-04	9.5134e-04	9.4525e-04	9.4189e-04	9.6097e-04	9.6070e-04	9.6084e-04
q(6,6)	-	2.9366e-05	1.8539e-05	2.3453e-05	9.7287e-04	9.6972e-04	9.6988e-04
q(7,7)	-	-	8.4604e-08	1.0973e-07	7.6155e-04	9.9172e-04	9.9172e-04
q(8,8)	-	-	-	1.0890e-010	2.6177e-04	7.5924e-04	9.9794e-04
q(9,9)	-	-	-	-	-	4.5479e-05	7.6074e-04
q(10,10)	-	-	-	-	-	-	2.5407e-04
$\phi(1,1)$	9.5023e-01	9.5874e-01	1.0174e-01	-2.6790e-01	9.5786e-01	9.5981e-01	9.5919e-01
$\phi(2,2)$	9.9794e-01	9.9504e-01	9.9583e-01	9.9905e-01	9.9651e-01	9.9645e-01	9.9636e-01
$\phi(3,3)$	9.1017e-01	9.6714e-01	9.7058e-01	9.6929e-01	9.5134e-01	9.5195e-01	9.5179e-01
$\phi(4,4)$	9.6238e-01	9.6998e-01	9.7591e-01	9.7814e-01	9.5578e-01	9.5291e-01	9.5301e-01
$\phi(5,5)$	9.4128e-01	9.3888e-01	9.5181e-01	9.5232e-01	9.4346e-01	9.4142e-01	9.4162e-01
$\phi(6,6)$	-	3.1774e-01	1.1656e-01	1.7281e-02	9.9316e-01	9.8509e-01	9.8500e-01
$\phi(7,7)$	-	-	2.3852e-01	1.9240e-01	5.8970e-01	9.9788e-01	9.9796e-01
$\phi(8,8)$	-	-	-	3.2400e-01	7.5980e-01	5.0346e-01	9.9906e-01
$\phi(9,9)$	-	-	-	-	-	7.4532e-01	5.9827e-01

continua en la siguiente página

<i>viene de la página anterior. Optimización de R, Q y Φ diagonales.</i>							
	M1	M2	M3	M4	M5	M6	M7
$\phi(10, 10)$	-	-	-	-	-	-	7.5335e-01

Cuadro 3.3: Optimización de R, Q y Φ diagonales con $\Phi(1, 1) = 1,0$.

	M1	M2	M3	M4	M5	M6	M7
loglik	-1.9153e+04	-1.9403e+04	-1.8990e+04	-1.8913e+04	-1.9413e+04	-1.9696e+04	-1.9705e+04
R	+2.2604e-03	+1.0014e-03	+8.6684e-04	+8.1091e-04	+1.4161e-03	+1.4112e-03	+1.4058e-03
q(1,1)	+1.1087e-04	+2.3976e-04	+4.9097e-04	+4.2847e-04	+1.6404e-06	+4.1114e-07	+4.1974e-07
q(2,2)	+9.8366e-04	+9.6800e-04	+9.5923e-04	+9.1433e-04	+9.7749e-04	+9.7699e-04	+2.0201e-03
q(3,3)	+9.7550e-04	+1.9425e-02	+9.6525e-04	+9.6322e-04	+6.9729e-03	+9.7305e-04	+9.7306e-04
q(4,4)	+1.3176e-02	+9.6562e-04	+9.6088e-04	+9.5914e-04	+1.1418e-02	+2.4283e-02	+2.3046e-02
q(5,5)	+2.0211e-03	+9.5349e-04	+9.4862e-04	+9.4616e-04	+9.6105e-04	+9.6098e-04	+9.6093e-04
q(6,6)	-	+1.3536e-05	+2.2274e-05	+2.0798e-05	+9.7372e-004	+9.7054e-04	+9.7052e-04
q(7,7)	-	-	+8.8034e-08	+9.7264e-08	+7.8220e-004	+9.9216e-04	+9.9224e-04
q(8,8)	-	-	-	+4.1238e-10	+5.9183e-05	+7.8024e-04	+9.9794e-04
q(9,9)	-	-	-	-	-	+2.5301e-05	+7.7826e-04
q(10,10)	-	-	-	-	-	-	+2.3411e-05
$\phi(1, 1)$	+1.0000e+00	+1.0000e+00	+1.0000e+00	+1.0000e+00	+1.0000e+00	+1.0000e+00	+1.0000e+00
$\phi(2, 2)$	+9.9778e-01	+9.9484e-01	+9.9476e-01	+9.9905e-01	+9.9567e-01	+9.9535e-01	+9.9616e-01
$\phi(3, 3)$	+9.4858e-01	+9.6462e-01	+9.6736e-01	+9.6707e-01	+9.5388e-01	+9.5686e-01	+9.5688e-01
$\phi(4, 4)$	+9.4503e-01	+9.5328e-01	+9.5521e-01	+9.5734e-01	+9.3474e-01	+1.4252e-01	+1.2775e-01
$\phi(5, 5)$	+9.2804e-01	+9.3817e-01	+9.4128e-01	+9.4217e-01	+9.3531e-01	+9.3656e-01	+9.3636e-01
$\phi(6, 6)$	-	+7.7976e-01	+7.8028e-01	+7.7192e-01	+9.9381e-01	+9.8032e-01	+9.8041e-01
$\phi(7, 7)$	-	-	+7.4342e-01	+7.2634e-01	+6.3196e-01	+9.9788e-01	+9.9798e-01
$\phi(8, 8)$	-	-	-	+7.4302e-01	+7.8206e-01	+4.7798e-01	+9.9906e-01
$\phi(9, 9)$	-	-	-	-	-	+7.8902e-01	+4.4514e-01

continua en la siguiente página

viene de la página anterior. Optimización de R , Q y Φ diagonales con $\Phi(1,1) = 1,0$.							
	M1	M2	M3	M4	M5	M6	M7
$\phi(10,10)$	-	-	-	-	-	-	+7.8619e-01

Cuadro 3.4: Optimización de R y Q.

	M1	M2	M3	M4	M5	M6	M7
loglik	-1.9921e+04	-1.9907e+04	-1.9831e+04	-1.9544e+04	-1.9882e+04	-1.9889e+04	-1.9883e+04
q(1,1)	+1.2923e-05	+9.3142e-06	+5.1517e-05	+1.0020e-04	+1.0950e-05	+1.0931e-05	+1.0958e-05
q(2,2)	+6.7656e-04	+7.1844e-04	+1.7373e-04	+2.3884e-04	+6.7399e-04	+6.7241e-04	+6.7146e-04
q(3,3)	+5.6250e-03	+2.5397e-03	+2.1290e-03	+2.5112e-03	+2.5316e-03	+2.9919e-03	+2.8667e-03
q(4,4)	+2.6693e-03	+2.3246e-03	+3.6471e-03	+3.3879e-03	+1.9187e-03	+1.9683e-03	+1.9245e-03
q(5,5)	+3.6236e-03	+2.7485e-03	+2.2146e-03	+3.4964e-03	+2.2627e-03	+2.2547e-03	+2.2583e-03
q(6,6)	+1.1465e-03	+1.2142e-03	+4.6830e-04	+1.1285e-03	+9.6175e-04	+9.7976e-04	+9.5823e-04
q(7,7)	-	+8.1560e-06	+5.3019e-05	+1.3014e-04	+1.8831e-03	+1.5362e-03	+1.6513e-03
q(8,8)	-	-	+4.1062e-07	+3.1212e-06	+9.3733e-06	+1.1105e-03	+1.0694e-03
q(9,9)	-	-	-	+1.3069e-08	+1.2788e-05	+1.0368e-05	+1.0125e-03
q(10,10)	-	-	-	-	-	+1.3103e-05	+1.0189e-05
q(2,1)	-	-	-	-	-	-	+1.2617e-05
q(3,1)	+1.8493e-03	+1.0692e-03	+4.5553e-04	+6.1808e-04	+1.0822e-03	+1.1670e-03	+1.1300e-03
q(3,2)	+1.0736e-03	+8.7335e-04	+5.4050e-04	+6.5130e-04	+7.7996e-04	+7.9417e-04	+7.7181e-04
q(4,1)	+3.0750e-03	+1.8850e-03	+2.4223e-03	+2.5113e-03	+1.5322e-03	+1.7086e-03	+1.6394e-03
q(4,2)	+1.3208e-03	+9.9003e-04	-1.6343e-06	+5.5034e-04	+8.9240e-04	+8.9264e-04	+8.8346e-04
q(4,3)	+3.8260e-03	+2.0800e-03	+1.0381e-03	+2.4784e-03	+1.7190e-03	+1.8611e-03	+1.8260e-03
q(8,8)	+2.3194e-03	+1.8749e-03	+1.5038e-03	+2.8171e-03	+1.3735e-03	+1.3859e-03	+1.3745e-03
q(5,1)	+4.9418e-04	+4.3363e-04	+1.6459e-04	+2.7859e-04	+3.4688e-04	+3.5799e-04	+3.4558e-04
q(5,2)	+1.4489e-03	+9.8888e-04	+8.4032e-04	+1.1845e-03	+7.2677e-04	+7.9575e-04	+7.5808e-04
q(5,3)	+9.0213e-04	+8.6081e-04	+1.0676e-03	+1.3375e-03	+5.7063e-04	+5.9324e-04	+5.7203e-04

continua en la siguiente página.

<i>viene de la página anterior. Optimización de R y Q.</i>							
	M1	M2	M3	M4	M5	M6	M7
q(5,4)	+1.0572e-03	+9.3917e-04	+5.9163e-04	+1.2591e-03	+6.2569e-04	+6.3976e-04	+6.2732e-04
q(6,1)	-	+6.2820e-07	+5.4441e-05	-9.5669e-05	+6.3398e-04	+4.9275e-04	+5.0863e-04
q(6,2)	-	-1.6446e-05	+2.8940e-04	-8.8883e-05	+1.4291e-03	+1.2674e-03	+1.3174e-03
q(6,3)	-	+4.5976e-05	+3.7128e-04	-4.9296e-05	+1.1589e-03	+9.5941e-04	+1.0075e-03
q(6,4)	-	+5.4505e-05	+2.3917e-04	+5.6196e-05	+1.2360e-03	+8.3983e-04	+1.0717e-03
q(6,5)	-	+1.6559e-05	+1.3332e-04	-2.5785e-05	+4.3217e-04	+3.6239e-04	+3.9227e-04
q(7,1)	-	-	-3.0535e-06	+1.0274e-05	+5.4041e-05	+2.6323e-04	+2.0653e-04
q(7,2)	-	-	-2.2808e-05	+2.9029e-05	+1.1059e-04	+5.3928e-04	+4.2518e-04
q(7,3)	-	-	-2.6902e-05	+3.2487e-05	+8.3389e-05	+4.2017e-04	+3.4062e-04
q(7,4)	-	-	-2.0088e-05	+2.2029e-05	+9.7598e-05	+4.4944e-04	+3.5909e-04
q(7,5)	-	-	-1.0676e-05	+1.3180e-05	+6.2266e-05	+1.6734e-04	+1.3421e-04
q(7,6)	-	-	-4.2456e-06	-1.4836e-05	+6.9713e-05	+1.8532e-04	+1.6779e-04
q(8,1)	-	-	-	-7.6931e-07	-7.5857e-05	+5.7912e-05	+1.0129e-04
q(8,2)	-	-	-	-2.9760e-06	-1.1164e-04	+1.2758e-04	+1.8790e-04
q(8,3)	-	-	-	-3.6343e-06	-8.8915e-05	+9.1137e-05	+1.5413e-04
q(8,4)	-	-	-	-3.9155e-06	-9.5744e-05	+1.0303e-04	+1.7582e-04
q(8,5)	-	-	-	-1.5469e-06	-4.4018e-05	+6.6777e-05	+5.7972e-05
q(8,6)	-	-	-	+5.7079e-07	-6.0781e-05	+6.1169e-05	+8.6733e-05
q(8,7)	-	-	-	-1.7518e-07	-5.8934e-06	+2.5467e-05	+2.3874e-05
q(9,1)	-	-	-	-	-	-7.6946e-05	+5.7100e-05
q(9,2)	-	-	-	-	-	-1.2171e-04	+1.2311e-04
q(9,3)	-	-	-	-	-	-9.1596e-05	+8.9340e-05

continua en la siguiente página.

viene de la página anterior. Optimización de R y Q.

	M1	M2	M3	M4	M5	M6	M7
q(9,4)	-	-	-	-	-	-9.9390e-05	+1.0231e-04
q(9,5)	-	-	-	-	-	-4.6332e-05	+6.4920e-05
q(9,6)	-	-	-	-	-	-5.1200e-05	+6.6296e-05
q(9,7)	-	-	-	-	-	-2.0661e-05	+2.0228e-05
q(9,8)	-	-	-	-	-	-6.4914e-06	+8.6528e-06
q(10,1)	-	-	-	-	-	-	-7.5162e-05
q(10,2)	-	-	-	-	-	-	-1.1419e-04
q(10,3)	-	-	-	-	-	-	-8.6533e-05
q(10,4)	-	-	-	-	-	-	-9.3788e-05
q(10,5)	-	-	-	-	-	-	-4.4152e-05
q(10,6)	-	-	-	-	-	-	-5.0027e-05
q(10,7)	-	-	-	-	-	-	-1.5922e-05
q(10,8)	-	-	-	-	-	-	-8.7664e-06
q(10,9)	-	-	-	-	-	-	-6.2537e-06

Cuadro 3.5: Optimización de R y Q y Φ .

Optimización de R, Q y Φ							
	M1	M2	M3	M4	M5	M6	M7
loglik	-2.0311e+04	-1.9855e+04	-1.9224e+04	-1.8780e+04	-1.9783e+04	-1.9759e+04	-1.9970e+04
R	+7.5083e-05	+3.2757e-04	+7.7638e-04	+5.1131e-04	+3.7668e-04	+4.4599e-04	+2.3065e-04
q(1,1)	+7.1610e-04	+7.8717e-04	+7.5746e-04	+7.3986e-04	+6.9262e-04	+6.9022e-04	+6.9329e-04
q(2,2)	+2.6098e-03	+2.2935e-03	+9.3523e-04	+8.8560e-04	+1.8808e-03	+1.8036e-03	+2.8247e-03
q(3,3)	+2.7826e-03	+1.5617e-03	+1.0190e-03	+1.0254e-03	+1.7693e-03	+1.6033e-03	+1.7718e-03
q(4,4)	+3.4381e-03	+1.5464e-03	+1.0413e-03	+1.0250e-04	+2.1729e-03	+1.9566e-03	+2.5162e-03
q(5,5)	+1.3178e-03	+1.1868e-03	+9.9441e-04	+1.0034e-03	+1.2040e-03	+1.1610e-03	+1.0880e-03
q(6,6)	-	+2.4634e-05	+2.5405e-04	+5.0151e-04	+1.1211e-03	+1.0882e-03	+1.1301e-03
q(7,7)	-	-	+2.3720e-06	+2.8507e-05	+4.8833e-04	+1.0140e-03	+1.0294e-03
q(8,8)	-	-	-	+1.1420e-07	+4.3001e-04	+5.2520e-04	+1.0037e-03
q(9,9)	-	-	-	-	-	+4.6331e-04	+3.8662e-04
q(10,10)	-	-	-	-	-	-	+3.4252e-04
q(2,1)	+1.0483e-03	+6.1206e-04	-1.5743e-05	+6.9433e-05	+7.0751e-04	+6.3879e-04	+8.9005e-04
q(3,1)	+1.0805e-03	+4.9792e-04	-7.0817e-06	-4.7215e-06	+6.6380e-04	+5.8675e-04	+6.6555e-04
q(3,2)	+2.0172e-03	+1.0969e-03	+2.1248e-04	+2.2822e-04	+1.1088e-03	+9.4245e-04	+1.3733e-03
q(4,1)	+1.1935e-03	+4.3924e-04	-1.8663e-05	+8.0133e-06	+7.2013e-04	+6.2750e-04	+7.8801e-04
q(4,2)	+2.1141e-03	+9.4479e-04	+1.7652e-04	+8.4696e-05	+1.1297e-03	+9.5035e-04	+1.6230e-03
q(4,3)	+2.0877e-03	+8.7054e-04	+2.3828e-04	+2.4037e-05	+8.2634e-04	+6.4485e-04	+1.2638e-03
q(5,1)	+4.2787e-04	+2.8689e-04	+3.9704e-06	+5.8459e-05	+3.1880e-04	+2.8026e-04	-2.0326e-05
q(5,2)	+8.7929e-04	+5.9212e-04	+1.2638e-04	+1.5710e-04	+5.9005e-04	+5.1295e-04	+3.5595e-04

continua en la siguiente página.

viene de la página anterior. Optimización de R , Q y ϕ .

	M1	M2	M3	M4	M5	M6	M7
q(5,3)	+9.0354e-04	+5.2008e-04	+1.4177e-04	+1.6654e-04	+5.4277e-04	+4.5601e-04	+2.5229e-04
q(5,4)	+9.5808e-04	+4.4472e-04	+1.4433e-04	+4.3299e-05	+5.6978e-04	+4.7429e-04	+3.0991e-04
q(6,1)	-	-5.4624e-05	-1.9679e-04	-1.6572e-04	+1.6637e-04	+1.3407e-04	+1.3227e-04
q(6,2)	-	+6.3706e-05	+1.5450e-04	+1.2237e-04	+3.9446e-04	+3.6239e-04	+5.4244e-04
q(6,3)	-	+5.1934e-05	+1.6312e-04	+1.4390e-04	+4.3717e-04	+3.4795e-04	+4.1962e-04
q(6,4)	-	+5.7624e-05	+1.6441e-04	+1.2924e-05	+4.3316e-04	+3.5587e-04	+5.0358e-04
q(6,5)	-	+1.6754e-05	+8.5233e-05	+1.0082e-04	+2.0351e-04	+1.6466e-04	+1.1941e-04
q(7,1)	-	-	+1.0988e-05	+7.3507e-06	+6.3335e-05	+8.0184e-05	+9.5689e-05
q(7,2)	-	-	-1.2911e-05	-8.2376e-06	+1.0163e-04	+1.3843e-04	+2.5027e-04
q(7,3)	-	-	-1.3792e-05	+2.0313e-06	+7.3932e-05	+1.5986e-04	+2.2791e-04
q(7,4)	-	-	-1.3456e-05	+5.3470e-06	+1.0819e-04	+1.5347e-04	+2.3063e-04
q(7,5)	-	-	-8.4280e-06	-1.1765e-05	+6.6011e-05	+6.9783e-05	+5.9309e-05
q(7,6)	-	-	-2.2472e-05	-1.0497e-04	+1.1520e-05	+5.1144e-05	+1.0893e-04
q(8,1)	-	-	-	+1.4765e-07	-9.3807e-05	+6.6076e-05	+4.4029e-05
q(8,2)	-	-	-	-7.0636e-07	-9.8393e-05	+1.0807e-04	+1.1053e-04
q(8,3)	-	-	-	-8.6572e-07	-8.9018e-05	+8.9815e-05	+7.7835e-05
q(8,4)	-	-	-	+3.5265e-08	-9.8422e-05	+1.0336e-04	+1.0614e-04
q(8,5)	-	-	-	+1.9998e-07	-4.8631e-05	+6.9999e-05	+1.5177e-05
q(8,6)	-	-	-	+5.4652e-06	-3.1273e-05	+2.3130e-05	+2.2298e-05
q(8,7)	-	-	-	-1.6969e-06	-2.1107e-05	+5.5481e-06	+2.0043e-05
q(9,1)	-	-	-	-	-	-9.4910e-05	+7.0054e-05
q(9,2)	-	-	-	-	-	-9.3305e-05	+1.3581e-04

continua en la siguiente página.

<i>viene de la página anterior. Optimización de R, Q y ϕ.</i>							
	M1	M2	M3	M4	M5	M6	M7
q(9,3)	-	-	-	-	-	-7.9695e-05	+9.9043e-05
q(9,4)	-	-	-	-	-	-8.7353e-05	+1.2086e-04
q(9,5)	-	-	-	-	-	-4.9367e-05	+4.7384e-05
q(9,6)	-	-	-	-	-	-3.0636e-05	+2.7177e-05
q(9,7)	-	-	-	-	-	-7.6358e-06	+2.8191e-06
q(9,8)	-	-	-	-	-	-1.7592e-05	+4.6864e-06
q(10,1)	-	-	-	-	-	-	-9.6828e-05
q(10,2)	-	-	-	-	-	-	-1.2421e-04
q(10,3)	-	-	-	-	-	-	-8.8938e-05
q(10,4)	-	-	-	-	-	-	-1.0697e-04
q(10,5)	-	-	-	-	-	-	-1.2691e-05
q(10,6)	-	-	-	-	-	-	-2.8452e-05
q(10,7)	-	-	-	-	-	-	-1.1537e-05
q(10,8)	-	-	-	-	-	-	-2.9102e-07
q(10,9)	-	-	-	-	-	-	-2.1875e-05
$\phi(1,1)$	+9.7483e-01	+8.7735e-01	+9.5662e-01	+8.5537e-01	+9.6749e-01	+9.6876e-01	+9.6605e-01
$\phi(2,2)$	+9.9734e-01	+9.9524e-01	+9.9673e-01	+9.9905e-01	+9.9220e-01	+9.8814e-01	+8.5921e-01
$\phi(3,3)$	+5.7461e-01	+9.4244e-01	+9.6236e-01	+9.6161e-01	+8.5884e-01	+8.8132e-01	+9.5992e-01
$\phi(4,4)$	+5.9703e-01	+9.6838e-01	+9.6802e-01	+9.6966e-01	+8.4482e-01	+8.6346e-01	+8.3137e-01
$\phi(5,5)$	+7.8791e-01	+9.4995e-01	+9.4181e-01	+9.3683e-01	+9.0306e-01	+9.0605e-01	+9.3664e-01
$\phi(6,6)$	-	+6.4325e-01	+2.9064e-01	+6.5807e-01	+9.9396e-01	+9.8251e-01	+9.8360e-01
$\phi(7,7)$	-	-	+7.2111e-01	+1.3032e-01	+8.0209e-01	+9.9808e-01	+9.9824e-01

continua en la siguiente página.

<i>viene de la página anterior. Optimización de R, Q y ϕ.</i>							
	M1	M2	M3	M4	M5	M6	M7
$\phi(8, 8)$	-	-	-	+1.0106e+00	+8.2838e-01	+8.1554e-01	+9.9906e-01
$\phi(9, 9)$	-	-	-	-	-	+8.3106e-01	+8.5202e-01
$\phi(10, 10)$	-	-	-	-	-	-	+8.2767e-01
$\phi(2, 1)$	+6.1726e-02	-1.9757e-01	-9.6650e-02	-1.2548e-01	+3.1932e-02	+5.7997e-03	-1.8180e-02
$\phi(3, 1)$	+2.4394e-02	-1.8601e-01	-1.2613e-01	-1.4612e-01	+1.2139e-02	-7.5507e-03	+1.2854e-02
$\phi(3, 2)$	-4.9898e-03	+2.0578e-02	+1.9390e-02	+1.4758e-02	-5.1518e-03	-1.4732e-02	-2.5577e-02
$\phi(4, 1)$	+9.3337e-03	-8.2342e-02	-7.9041e-02	-7.8679e-02	+5.4094e-02	+1.4946e-02	+3.0862e-02
$\phi(4, 2)$	+7.3596e-03	+1.1220e-02	+1.8027e-02	+9.6358e-03	-3.0948e-02	-3.4386e-02	-5.4960e-02
$\phi(4, 3)$	-3.9621e-01	-5.0111e-02	+4.7245e-03	+6.9142e-03	-8.2514e-02	-7.0153e-02	-8.4394e-02
$\phi(5, 1)$	+3.6900e-02	-5.4333e-02	-4.1193e-02	-6.3907e-02	+2.0533e-02	-6.1228e-04	-5.1401e-03
$\phi(5, 2)$	+3.4812e-03	+4.7651e-03	+3.4757e-03	-1.8809e-03	-9.2130e-03	-1.2090e-02	-6.9003e-03
$\phi(5, 3)$	-1.2929e-01	+2.8931e-03	+7.0281e-03	+9.1002e-03	-4.7763e-02	-4.5157e-02	-4.0129e-02
$\phi(5, 4)$	-1.0482e-01	+3.1943e-02	+1.9520e-02	+6.3997e-02	-3.9841e-02	-3.3621e-02	-1.4213e-02
$\phi(6, 1)$	-	-2.9325e-02	-1.1893e-01	-7.6221e-02	+6.5344e-03	+1.8205e-02	-1.4698e-02
$\phi(6, 2)$	-	+7.1807e-03	+1.0352e-02	+3.3243e-03	+5.6074e-03	+1.2364e-03	-8.6334e-03
$\phi(6, 3)$	-	-5.0161e-03	-3.0744e-03	-3.5839e-03	-2.5343e-03	-1.5325e-02	-1.3622e-02
$\phi(6, 4)$	-	+1.0067e-02	+5.5583e-03	+2.0999e-02	-4.8871e-02	-4.4433e-02	-4.2235e-02
$\phi(6, 5)$	-	+7.9973e-03	+1.5103e-03	+7.8373e-03	-6.7618e-03	-1.2330e-02	+1.7326e-03
$\phi(7, 1)$	-	-	+8.0405e-03	-4.1553e-03	+1.2038e-02	+7.7325e-03	-1.2531e-03
$\phi(7, 2)$	-	-	-3.0122e-04	+8.6608e-04	-5.3041e-03	-4.4862e-03	+2.4926e-03
$\phi(7, 3)$	-	-	-2.6484e-04	+6.5204e-04	-2.7673e-02	+6.6994e-03	+4.7721e-03
$\phi(7, 4)$	-	-	-4.0827e-05	-2.7164e-04	-2.2273e-02	-2.1634e-02	-3.6147e-02

continua en la siguiente página.

<i>viene de la página anterior. Optimización de R, Q y ϕ.</i>							
	M1	M2	M3	M4	M5	M6	M7
$\phi(7, 5)$	-	-	-5.0216e-05	-1.5666e-03	-9.5787e-03	-7.3376e-03	-3.9364e-03
$\phi(7, 6)$	-	-	+2.3737e-02	-8.9532e-02	+2.8460e-03	+3.8690e-03	-4.6120e-03
$\phi(8, 1)$	-	-	-	+7.5933e-04	-4.3804e-03	+8.4397e-03	-2.8922e-03
$\phi(8, 2)$	-	-	-	-9.3214e-05	+1.5723e-03	-5.5176e-03	-1.1582e-02
$\phi(8, 3)$	-	-	-	-2.1902e-05	-1.5719e-03	-2.1713e-02	-3.0827e-03
$\phi(8, 4)$	-	-	-	+3.2169e-05	+6.0393e-03	-1.9451e-02	-8.3665e-03
$\phi(8, 5)$	-	-	-	+1.2265e-04	+2.4074e-03	-7.7643e-03	-2.1897e-03
$\phi(8, 6)$	-	-	-	+6.8736e-03	-5.4394e-03	+4.7090e-03	-3.1785e-03
$\phi(8, 7)$	-	-	-	+4.3511e-02	-2.6164e-02	-9.8722e-04	+3.2855e-03
$\phi(9, 1)$	-	-	-	-	-	-6.7623e-03	+1.4564e-02
$\phi(9, 2)$	-	-	-	-	-	+3.8048e-03	-1.1451e-02
$\phi(9, 3)$	-	-	-	-	-	+4.4904e-03	-1.0318e-02
$\phi(9, 4)$	-	-	-	-	-	+1.0420e-02	-1.9234e-02
$\phi(9, 5)$	-	-	-	-	-	+4.9959e-03	+1.6513e-03
$\phi(9, 6)$	-	-	-	-	-	-5.9579e-03	+3.5904e-03
$\phi(9, 7)$	-	-	-	-	-	-5.4878e-04	-1.0135e-03
$\phi(9, 8)$	-	-	-	-	-	-2.6308e-02	-1.4723e-03
$\phi(10, 1)$	-	-	-	-	-	-	-5.1985e-03
$\phi(10, 2)$	-	-	-	-	-	-	+6.2722e-03
$\phi(10, 3)$	-	-	-	-	-	-	+7.3878e-03
$\phi(10, 4)$	-	-	-	-	-	-	+8.0034e-03
$\phi(10, 5)$	-	-	-	-	-	-	+8.7786e-04

continua en la siguiente página.

<i>viene de la página anterior. Optimización de R, Q y ϕ.</i>							
	M1	M2	M3	M4	M5	M6	M7
$\phi(10, 6)$	-	-	-	-	-	-	-9.9368e-03
$\phi(10, 7)$	-	-	-	-	-	-	-6.0513e-03
$\phi(10, 8)$	-	-	-	-	-	-	-6.2779e-04
$\phi(10, 9)$	-	-	-	-	-	-	-2.5580e-02

3.6.2. Resultados numéricos del remuestreo mediante bootstrap.

Cuadro 3.6: Resultados bootstrap modelo SUC M1.

	R	q11	q12	q13	q14	q15	q22	q23	q24	q25	q33	q34	q35	q44	q45	q55
optimo	1,292e-5	6,766e-4	1,849e-3	1,074e-3	1,321e-3	4,942e-4	5,625e-3	3,075e-3	3,826e-3	1,449e-3	2,669e-3	2,319e-3	9,021e-4	3,624e-3	1,057e-3	1,146e-3
Bootstrap																
min	1,291e-5	6,463e-4	1,762e-3	1,013e-3	1,222e-3	3,625e-4	5,372e-3	2,867e-3	3,565e-3	1,067e-3	2,455e-3	2,132e-3	6,283e-4	3,181e-3	7,021e-4	8,475e-4
max	1,297e-5	7,869e-4	2,155e-3	1,265e-3	1,544e-3	6,545e-4	6,473e-3	3,639e-3	4,446e-3	1,900e-3	3,140e-3	2,808e-3	1,201e-3	4,224e-3	1,406e-3	1,458e-3
media	1,293e-5	7,085e-4	1,939e-3	1,130e-3	1,388e-3	5,215e-4	5,877e-3	3,233e-3	4,016e-3	1,525e-3	2,766e-3	2,438e-3	9,477e-4	3,765e-3	1,114e-3	1,160e-3
mediana	1,293e-5	7,070e-4	1,935e-3	1,127e-3	1,386e-3	5,200e-4	5,867e-3	3,227e-3	4,010e-3	1,521e-3	2,761e-3	2,433e-3	9,462e-4	3,761e-3	1,112e-3	1,170e-3
sesgo	8,901e-9	3,198e-5	9,003e-5	5,616e-5	6,764e-5	2,731e-5	2,524e-4	1,580e-4	1,901e-4	7,612e-5	9,637e-5	1,190e-4	4,554e-5	1,413e-4	5,691e-5	1,330e-5
σ	9,132e-9	2,175e-5	6,083e-5	4,012e-5	4,799e-5	3,113e-5	1,721e-4	1,154e-4	1,364e-4	9,007e-5	8,527e-5	9,281e-5	6,827e-5	1,156e-4	7,329e-5	8,236e-5
curtosis	0,6725	-0,0629	-0,0279	0,1914	0,2487	2,5904	-0,0092	0,3661	0,2376	2,9233	2,5074	0,7074	3,5408	2,3587	2,5262	2,1934
c asim	0,8958	0,1903	0,1813	0,2233	0,1222	-0,1567	0,1803	0,2555	0,1317	-0,3059	0,4330	0,2755	-0,5574	-0,0055	-0,3513	-0,7095
Intervalos de Confianza calculados con el método del percentil																
lower	1,292e-5	6,747e-4	1,844e-3	1,066e-3	1,313e-3	4,739e-4	5,609e-3	3,054e-3	3,801e-3	1,386e-3	2,641e-3	2,295e-3	8,369e-4	3,597e-3	9,981e-4	9,864e-4
upper	1,295e-5	7,465e-4	2,043e-3	1,198e-3	1,470e-3	5,696e-4	6,171e-3	3,431e-3	4,241e-3	1,664e-3	2,905e-3	2,598e-3	1,051e-3	3,948e-3	1,227e-3	1,279e-3
length	3,026e-8	7,176e-5	1,987e-4	1,314e-4	1,576e-4	9,570e-5	5,612e-4	3,777e-4	4,400e-4	2,775e-4	2,637e-4	3,032e-4	2,144e-4	3,506e-4	2,290e-4	2,927e-4
Intervalos de Confianza calculados con el método del percentil corregido por sesgo																
pct1	30,30	6,70	6,4750	11,12	8,82	30,93	7,15	12,20	10,33	32,49	17,63	16,15	36,23	14,14	38,69	68,20
pct2	7,076e-3	8,405e-5	7,707e-5	3,244e-4	1,726e-4	7,605e-3	9,952e-5	4,195e-4	2,651e-4	9,069e-3	1,225e-3	9,434e-4	1,353e-2	6,392e-4	1,738e-2	2,428e-1
lower	1,291e-5	6,463e-4	1,762e-3	1,013e-3	1,222e-3	3,633e-4	5,372e-3	2,868e-3	3,565e-3	1,067e-3	2,455e-3	2,132e-3	6,300e-4	3,181e-3	7,201e-4	8,726e-4
upper	1,293e-5	6,784e-4	1,853e-3	1,083e-3	1,327e-3	5,093e-4	5,633e-3	3,108e-3	3,849e-3	1,491e-3	2,701e-3	2,354e-3	9,319e-4	3,658e-3	1,098e-3	1,189e-3
length	1,407e-8	3,217e-5	9,085e-5	7,017e-5	1,049e-4	1,460e-4	2,614e-4	2,402e-4	2,838e-4	4,235e-4	2,457e-4	2,221e-4	3,019e-4	4,769e-4	3,775e-4	3,159e-4
Intervalos de Confianza normal																
lower	1,290e-5	6,088e-4	1,659e-3	9,514e-4	1,174e-3	4,157e-4	5,090e-3	2,727e-3	3,412e-3	1,225e-3	2,433e-3	2,048e-3	7,443e-4	3,292e-3	8,798e-4	9,977e-4
upper	1,293e-5	6,804e-4	1,859e-3	1,083e-3	1,332e-3	5,181e-4	5,656e-3	3,107e-3	3,860e-3	1,521e-3	2,713e-3	2,353e-3	9,689e-4	3,673e-3	1,121e-3	1,269e-3
length	3,004e-8	7,154e-5	2,001e-4	1,320e-4	1,579e-4	1,024e-4	5,663e-4	3,797e-4	4,487e-4	2,963e-4	2,805e-4	3,053e-4	2,246e-4	3,804e-4	2,411e-4	2,71e-4

Capítulo 4

Combinación de predicciones

En este capítulo se presenta una estrategia novedosa de combinación de predicciones en el campo de la predicción eólica, empleando técnicas multivariantes de reducción de la dimensión del problema. Se analizan distintas alternativas de combinación y se comparan entre sí.

4.1. Introducción

La estrategia de funcionamiento de la herramienta de predicción desarrollada hasta ahora es disponer de un conjunto de predictores operando en paralelo de manera que se necesita una segunda etapa para obtener una única predicción final. Esta estrategia es muy común por parte de los desarrolladores de herramientas de predicción eólica, por dos motivos principalmente: por un lado, porque la precisión relativa de los diferentes modelos cambia a lo largo del tiempo (dependiendo de su capacidad para captar la relación no lineal entre viento y potencia generada en cada momento) y en función del horizonte de predicción. Por otro lado, esta estrategia permite que la herramienta de predicción siga funcionando en caso de que alguno de los modelos no pueda ejecutarse adecuadamente (por ejemplo, en caso de que no lleguen las predicciones meteorológicas, siempre podrán ejecutarse los modelos de referencia y el modelo M1). Dado el descenso del precio del servicio de predicción, los operadores de mercado suelen disponer de varios proveedores de este servicio e incluso disponer de herramienta de predicción propia. Como lo que se desea al final del proceso es una única predicción eólica, se realiza una combinación de las predicciones alternativas que fueron obtenidas de modos o fuentes diferentes. El objetivo último del proceso de combinación será que la predicción final resultante sea mejor o al menos igual de precisa que las predicciones que se combinan.

El origen de la combinación de predicciones se remonta a [2] donde se demuestra que se podría mejorar la exactitud de las predicciones mediante la combinación de un conjunto de predictores. Desde entonces, numerosos avances y desarrollos han tenido lugar en este campo. Para examinar los avances y conocimientos más recientes sobre combinación de predicciones pueden consultarse, entre otros, [4], [6], [11], [35], y más recientemente [52]. En las revisiones más recientes (véase por ejemplo [19] o [52]) se ha puesto de manifiesto la carencia de optimalidad de los predictores individuales debido, entre otras razones, a errores de modelado, errores de estimación, no estacionaridades, no linealidades, no disponer del conjunto completo de información de entrada, o simplemente a las diferentes funciones objetivo asociadas a cada predictor. Una discusión sobre la aplicación de las diferentes estrategias de combinación de predicciones al problema que nos ocupa puede encontrarse en [48].

Es decir, partimos de un conjunto de predictores, $y_{i,t+h|t}$ con $i = 1, \dots, N$ y el objetivo es llegar a una única predicción final $\hat{y}_{c,t+h|t}$ que, en cada instante, sea mejor o igual de precisa que cada uno de los predictores individuales. Se trata pues de realizar una combinación de predicciones a partir de los N predictores de partida, de la siguiente manera:

$$\hat{y}_{c,t+h|t} = \mathbf{w}'_t(h) \mathbf{y}_{t+h|t} \quad (4.1)$$

donde $\mathbf{y}_{t+h|t} = (y_{1,t+h|t}, \dots, y_{N,t+h|t})'$ es el conjunto de predictores de partida y $\mathbf{w}_t(h) = (w_1(h), \dots, w_N(h))'$ es el vector de ponderación. El subíndice t indica que este vector es variante con el tiempo y se calcula para cada instante t . También ha querido denotarse la dependencia con el horizonte de predicción h . En caso de ser necesario puede añadirse una constante a la ecuación (4.1) para corregir un posible sesgo en la combinación final. Es decir, la clave de la de combinación de predicciones es cómo reducir la dimensión del problema desde N predictores iniciales a un único predictor final, $\hat{y}_{c,t+h|t}$.

Existen dos estrategias posibles: combinar todos los predictores de partida o bien seleccionar únicamente los d mejores que serán los que se combinen. En el primer caso, como inconveniente, puede citarse que la inclusión de modelos con pobres resultados, a los que se asigna un coeficiente de ponderación pequeño aunque no nulo, va en detrimento del peso asignado a los mejores predictores y, por tanto, puede perjudicar al resultado o combinación final que se obtenga. En la segunda clase de métodos de combinación, lo que se busca es seleccionar el mejor de los predictores de partida. En el caso de seleccionar un único modelo (en este caso $d = 1$), la selección del mejor predictor no suele ser la mejor opción, ya que los posibles cambios en la precisión de los predictores de partida harían que se estuviese cambiando continuamente de modelo seleccionado. Para evitar esta variabilidad, se seleccionan los d mejores predictores que se combinarán

posteriormente. En la estrategia clásica que se realiza en combinación de predicciones, en el instante t se desea obtener $\hat{y}_{c,t+h|t}$ seleccionando los d mejores predictores. La selección de estos de predictores es basa en evaluar el funcionamiento de cada predictor. Pero en el instante t sólo podemos evaluar el error cometido en $t-h$. Es decir, en t , podemos evaluar $e_{t|t-h} = p_t - \hat{y}_{c,t|t-h}$. A medida que h aumenta, la relación entre lo que sucede en t y en $t+h$ es cada vez menor, por lo que estas estrategias pudieran no ser adecuadas.

La estrategia que se propone en esta tesis, basada en [43], aunque aplicada a distinto problema, es realizar la combinación de predicciones mediante técnicas multivariantes de reducción de datos. En este caso, la ponderación de cada predictor no se obtiene a partir del error cometido hasta el instante t , sino que se obtiene a partir de la cantidad de información, respecto al resto de predictores, que puede aportar cada predictor en la combinación final.

En predicción eólica, como en otros campos, es usual que la predicción final se construya sobre una o más combinaciones lineales de los predictores (véase, por ejemplo, [48]) denominándose este procedimiento combinación en dos etapas. En la metodología que se propone en esta tesis, en la primera etapa se construyen r combinaciones lineales o factores $f_{j,t+h|t} = \mathbf{w}'_{jt}(h)\mathbf{y}_{t+h|t}$. En la segunda etapa se obtiene la predicción final como

$$\hat{y}_{c,t+h|t} = \hat{\beta}_{0t} + \hat{\beta}_{1t}f_{1,t+h|t} + \cdots + \hat{\beta}_{rt}f_{rt}, \quad (4.2)$$

donde $\hat{\boldsymbol{\beta}}_t = (\hat{\beta}_{0t}, \dots, \hat{\beta}_{rt})'$ es la estimación por mínimos cuadrados ordinarios de la regresión,

$$p_t = \beta_{0t} + \beta_{1t}f_{1,t|t-h} + \cdots + \beta_{rt}f_{r,t|t-h} + u_t, \quad (4.3)$$

siendo p_t la variable observada en el instante t . Nótese que $\hat{y}_{c,t+h|t}$ es una verdadera predicción a priori ya que los coeficientes β_{jt} se estiman únicamente con información hasta el instante t y no incluyen ninguna información del instante a predecir. En resumen, el procedimiento de combinación consiste en calcular dos conjuntos de vectores de ponderación: el primero de ellos, $\mathbf{w}_t(h)$ para calcular las combinaciones lineales f_{jt} de los predictores de partida y que denominaremos factores, y el segundo, el vector $\boldsymbol{\beta}_t(h)$ para combinar los factores y obtener la predicción final.

Se han considerado dos técnicas multivariantes de reducción de datos para producir las r combinaciones lineales (f_{1t}, \dots, f_{rt}), mínimos cuadrados parciales (PLS de ahora en adelante) y componentes principales (CCPP). Ambas metodologías se diferencian en que al predecir p_{t+h} , el método de componentes principales no emplea la información disponible sobre la variable a predecir, es decir, no utiliza (p_t) para el cálculo de los

factores (f_{1t}, \dots, f_{rt}) , mientras que la regresión por mínimos cuadrados parciales sí emplea la información contenida en la variable a predecir, p_t , para calcular el vector de ponderación $\mathbf{w}_{jt}(h)$.

Una particularidad a destacar del problema que nos ocupa es que estamos aplicando las técnicas multivariantes de reducción de datos a predictores de la misma variable, por lo que podrán presentar cierto grado de correlación entre ellas a diferencia de cuando se trata de combinar variables heterogéneas.

Este capítulo se ha organizado de la siguiente manera: la sección 2 describe brevemente las distintas técnicas de reducción de dimensión; la sección 3 aplica las técnicas presentadas a los modelos desarrollados en el capítulo 3 y, finalmente, la sección 4 concluye.

4.2. Técnicas multivariantes de reducción de datos

Como se ha comentado anteriormente, clasificamos las técnicas multivariantes de reducción de datos en dos grupos: aquellos que tienen en cuenta la variable a predecir en el proceso de reducción de la dimensión o formación de factores, y aquellas técnicas que no tienen en cuenta esta variable. En el primero de los grupos emplearemos mínimos cuadrados parciales mientras que en el segundo analizaremos componentes principales.

Las técnicas mencionadas consideran como factores en el instante t , las combinaciones lineales $f_{jt} = \mathbf{w}'_{jt}(h)\mathbf{y}_{t+h|t}$. La diferencia entre los métodos analizados, es el conjunto de datos y la metodología empleados para calcular el vector de ponderaciones $\mathbf{w}_{jt}(h)$: PLS emplea la información disponible de la variable a predecir, p_t , por lo que el vector de ponderaciones se estima a partir del conjunto de datos $(p_\tau \quad \mathbf{y}_{\tau|\tau-h})$ con $\tau = 1, \dots, t$ para generar verdaderas predicciones a priori. La técnica de CCP no emplea la información contenida en la variable a predecir p_t para calcular el vector de ponderaciones $\mathbf{w}_{jt}(h)$ y, por lo tanto, se utilizan los predictores disponibles en el instante t , $\mathbf{y}_{t+h|t}$. Es decir, en el primer caso, $\widehat{\mathbf{w}}_{jt}(h) = \widehat{\mathbf{w}}_j(p_h \mathbf{y}_{h|0}, p_{h+1} \mathbf{y}_{h+1|1}, \dots, p_t \mathbf{y}_{t|t-h})$, mientras que en el segundo caso, $\widehat{\mathbf{w}}_{jt}(h) = \widehat{\mathbf{w}}_j(\mathbf{y}_{h|0}, \mathbf{y}_{h+1|1}, \dots, \mathbf{y}_{t|t-h}, \mathbf{y}_{t+h|t})$. A continuación revisaremos brevemente cada una de las técnicas que se van a aplicar.

4.2.1. Componentes principales

El método de componentes principales fue inicialmente desarrollado por Pearson a finales del siglo XIX y posteriormente estudiado por Hotelling en los años 30 del siglo XX. Sin embargo, hasta la aparición de los ordenadores no se empezaron a popularizar.

Para una revisión más en profundidad de esta técnica puede consultarse [42]. Para estudiar las relaciones que se presentan entre N variables correladas (que miden información común) se puede transformar el conjunto original de variables en otro conjunto de nuevas variables incorreladas entre sí (que no tenga repetición o redundancia en la información) llamado conjunto de componentes principales. Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra. Es decir, sea X_1, \dots, X_N un conjunto de variables y sea Σ_X su matriz de covarianzas. Las componentes principales ordenadas se obtienen de las sucesivas combinaciones lineales $\mathbf{f}_i = \mathbf{w}_i' \mathbf{X}$, $i = 1, \dots, N$ donde \mathbf{w}_i es el autovector de módulo 1 de Σ_X asociado con el i -ésimo mayor valor propio de dicha matriz. En nuestro caso la matriz X contiene el conjunto de predictores de la potencia eólica $\mathbf{y}_{t+h|t} = (y_{1,t+h|t}, \dots, y_{N,t+h|t})'$ con $t = 1, \dots, t$. Las sucesivas componentes principales son ortogonales entre sí, es decir, están incorreladas. Por tanto, cada una de ellas recoge variabilidad o información no contenida en las anteriores. Para obtener la predicción final, seleccionamos las r primeras componentes principales $\mathbf{f}_t = (f_{1t}, \dots, f_{rt})'$ y calculamos la regresión a través de las ecuaciones (4.2) y (4.3) para obtener la predicción final $\hat{y}_{c,t+h|t}$.

4.2.2. Mínimos cuadrados parciales

Mediante la técnica conocida como PLS, se busca un conjunto de variables latentes o factores que ajusten una descomposición simultánea de las matrices X e Y bajo la condición de que expliquen tanto como sea posible la covarianza entre X e Y y que la citada descomposición sea eficaz para ajustar Y . Esta técnica data de los años 60 y fue propuesta inicialmente por [55]. Es un proceso secuencial donde la correlación de la variable a predecir (hasta el instante t , para obtener verdaderas predicciones a priori) y los predictores se emplean para formar los factores. Es decir, sea $x = \mathbf{y}_{t|t-h} = (y_{1,t|t-h}, \dots, y_{N,t|t-h})'$ el vector de dimensión $N \times 1$ de predicciones a horizonte h de la variable observada, p_t , en el instante t . Nótese que estas predicciones se han calculado en el instante $t - h$.

La primera componente o factor, h_{1t} , se obtiene proyectando la covarianza entre la variable a predecir y las predicciones de su valor, en la dirección de las predicciones, es decir,

$$h_{1t} = \alpha \sum_{i=1}^N cov(y_{i,t|t-h}, p_t) y_{i,t|t-h},$$

siendo α la constante de proporcionalidad que depende de la normalización escogida. Se realiza una regresión para cada componente y con los residuos se determinará el siguiente factor. Es decir, se realiza la regresión de p_t con cada uno de los predictores, $p_{i,t|t-h}$ con $i = 1, \dots, N$ y se mantienen los residuos que contendrán toda la información ortogonal al primer componente h_{1t} . Sean $M_{1t}^{p_t}$ y $M_{1t}^{y_{i,t|t-h}}$, $i = 1, \dots, N$ los residuos de estas regresiones. Para construir el segundo factor empleamos esta información, ortogonal al primer factor, y construimos la nueva combinación lineal

$$h_{2t} \propto \sum_{i=1}^N \text{cov}(M_{1t}^{y_{i,t|t-h}}, M_{1t}^{p_t}) M_{1t}^{y_{i,t|t-h}}.$$

El siguiente componente o factor emplea información ortogonal a los dos primeros. Se realiza la regresión de p_t y cada uno de los predictores, $x_{it} = y_{i,t|t-h}$, $i = 1, \dots, N$ sobre una constante, h_{1t} y h_{2t} y se almacenan los residuos $M_{2t}^{p_t}$ y $M_{2t}^{y_{i,t|t-h}}$. Se emplean dichos residuos para calcular el tercer factor. El procedimiento se repite hasta que hayamos calculado los r factores. Como se ha mencionado anteriormente, se emplea solo información disponible hasta el instante t para calcular los vectores de ponderación $\mathbf{w}_{1t}(h), \dots, \mathbf{w}_{rt}(h)$ y con ellos se evalúa la predicción para $t+h$, $h_{it+h} = f_{it} = \mathbf{w}'_{it}(h) \mathbf{y}_{t+h|t}$, $i = 1, \dots, r$ en (4.2) y (4.3). La diferencia entre CCP y PLS es que mientras que el primero trata de encontrar combinaciones lineales que expliquen $\text{var}(X)$, PLS calcula combinaciones lineales en la dirección de $\text{cov}(X, p_t)$. En nuestra aplicación particular, las direcciones de las CCP tratan de resumir la variabilidad entre predictores, mientras que los factores obtenidos mediante PLS son las combinaciones lineales de X que son más útiles para predecir la variable de salida p_t .

4.3. Resultados obtenidos

En esta sección mostramos los resultados obtenidos con las técnicas explicadas en el apartado anterior. Para la realización de la combinación a través de factores, la selección del número de componentes o factores no lo estimamos a priori, sino que realizamos el análisis de 1 a 4 componentes y a posteriori estimamos el número óptimo en términos de la RECMN. Mantenemos un número bajo de factores (de 1 a 4) dado que el número de predictores individuales de partida no es demasiado elevado y debido a que un factor adicional aporta cada vez menos información y la incertidumbre asociada a la estimación de un mayor número de valores de los vectores de ponderación puede anular el efecto positivo de la introducción de un nuevo factor.

El conjunto de predictores de partida serán el modelo de referencia definido en la ecuación 2.3 así como los modelos M1 a M7 con estructura SUC (es decir, con las matrices R y Q llenas).

De los 1711 datos disponibles que no hemos empleado para la estimación de las matrices de los modelos, seleccionamos los 100 primeros para el cálculo inicial de las matrices de varianzas y covarianzas y el resto de datos, 1611, se destinarán para la validación de los resultados obtenidos. Nótese que en cada instante de tiempo t se reestiman los vectores de ponderación $\mathbf{w}_t(h)$ y $\beta_t(h)$ dado que la matriz de datos se actualiza con la última información disponible.

4.3.1. Modelo de referencia

Como modelos de referencia compararemos los resultados obtenidos con la media de las predicciones en cada instante t , con la mediana de los predictores y con una regresión simple de los predictores individuales calculada mediante mínimos cuadrados ordinarios. En este último caso se han analizado dos regresiones, una realizada sin tener en cuenta la posibilidad de tener término independiente, es decir, $y_c(t+h|t) = X * \beta(t)$ y una segunda incluyendo término independiente, es decir, $y_c(t+h|t) = \beta_{0t} + X * \beta(t)$. De manera análoga al capítulo 2, los modelos de referencia son esquemas sencillos pero que en la práctica proporcionan muy buenos resultados en la mayoría de las aplicaciones, y a los que las técnicas multivariantes que planteamos deberían ser capaces de mejorar.

En las figura 4.1 se comparan los resultados obtenidos en términos del sesgo. Analizando dicha figura puede comprobarse como el mejor modelo de referencia es la mediana de los modelos de partida. El sesgo de la media aumenta significativamente con el horizonte de predicción debido al valor elevado del sesgo del modelo de referencia que se ha combinado, no representado en la gráfica para mayor claridad (véase no obstante la figura 2.11(a)). La combinación de predicciones realizada mediante mínimos cuadrados ordinarios tiene un comportamiento diferenciado entre el modelo que incluye término independiente del que no lo incluye. Así, el modelo sin término independiente presenta un comportamiento peor. Ambos modelos presentan mayor sesgo que los predictores de partida puesto que como parte de la combinación, al igual que ocurría con la media, está incluido el modelo nr que presenta predicciones muy sesgadas, aunque el modelo con término independiente presenta menor sesgo en los menores horizontes de predicción.

En la figura 4.2 se comparan los distintos modelos de referencia en términos de la RECMN. Puede comprobarse como las dos alternativas de ponderación de las predicciones mediante mínimos cuadrados ordinarios mejoran el comportamiento de los

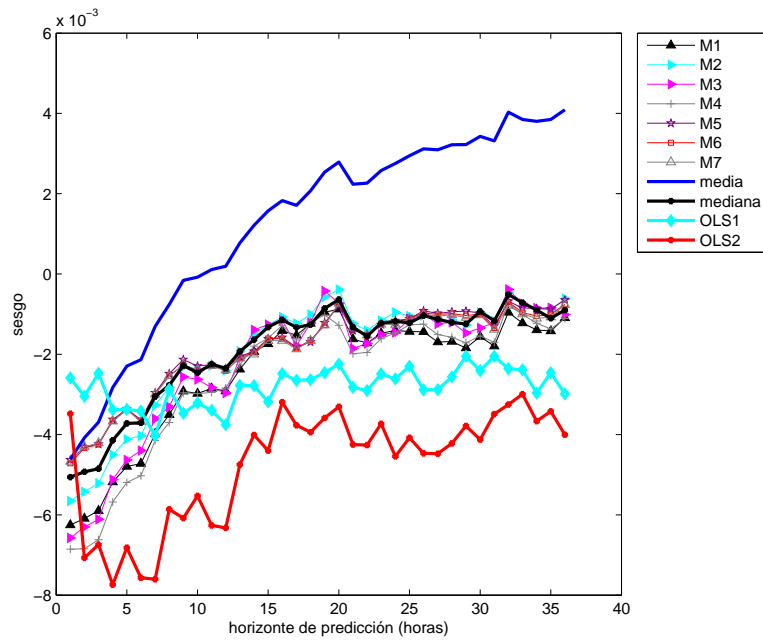


Figura 4.1: Sesgo de los modelos de partida y distintas alternativas de referencia de combinación de predicciones. OLS1 y OLS2 denotan la combinación de predicciones realizada mediante mínimos cuadrados ordinarios con y sin término independiente respectivamente.

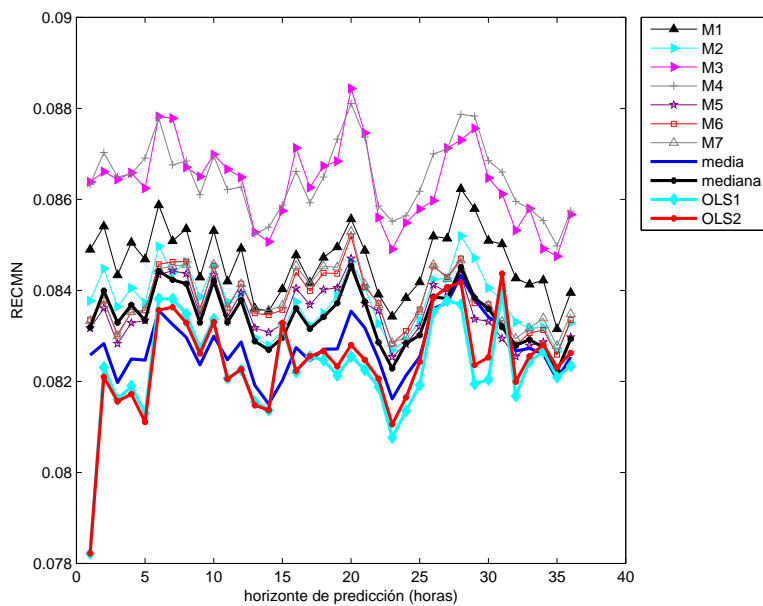


Figura 4.2: RECMN de los modelos de partida y distintas alternativas de referencia de combinación de predicciones. OLS1 y OLS2 denotan la combinación de predicciones realizada mediante mínimos cuadrados ordinarios con y sin término independiente respectivamente.

predictores individuales para todos los horizontes de predicción. La media mejora el comportamiento de los predictores de partida, pero en menor medida que los algoritmos de mínimos cuadrados ordinarios, no pudiendo superar a alguno de los predictores individuales a horizontes elevados ($h > 30$). La mediana presenta un comportamiento muy próximo a los mejores predictores de partida.

A la vista de los resultados, seleccionamos como método de referencia para la combinación de predicciones la ponderación realizada por mínimos cuadrados ordinarios con término independiente, ya que es el que presenta menor RECMN.

4.3.2. Componentes principales

En las figuras 4.3 y 4.4 se muestran los resultados de la combinación de predicciones realizada mediante la regresión de los componentes principales en términos del sesgo y de la RECMN del error de predicción, respectivamente. Como puede observarse, el sesgo de todas las alternativas es muy parecido y similar al sesgo de la combinación mediante mínimos cuadrados ordinarios. Respecto a la RECMN puede observarse que este esquema de combinación sólo supera ligeramente los resultados obtenidos mediante mínimos cuadrados ordinarios. La combinación de 4 factores es la de menor RECMN en 18 de los 36 horizontes de predicción. El resto de alternativas (la combinación de una, dos o tres componentes principales) sólo consigue mejorar los resultados obtenidos mediante el modelo de referencia en ciertos horizontes puntuales. A medida que aumenta el horizonte de predicción (a partir de horizonte 25) se diferencia claramente el comportamiento de la combinación de uno y dos factores de la combinación de 3 ó 4 factores. Es decir, a medida que aumenta el horizonte de predicción, dado que la incertidumbre es mayor, se obtienen mejores resultados aumentando el número de factores (recordemos que cada factor es ortogonal a los anteriores y explica, en orden decreciente, cierto porcentaje de la varianza total de los predictores de partida).

Es interesante interpretar el significado de cada factor o componente principal. Este análisis además se deberá realizar para cada horizonte de predicción. Así, por ejemplo, en la figura 4.5 se muestra el peso que se asigna a cada predictor individual en el caso de horizonte 1. Puede observarse cómo este factor representa un promedio ponderado de todos los predictores de partida, asignando menor peso al modelo de referencia (denotado como nr en la gráfica). En la figura 4.6 se muestra el peso asignado a cada predictor individual para calcular el primer componente principal a horizonte 36. Como puede observarse, al modelo de referencia se le asigna un coeficiente prácticamente nulo, consecuencia lógica del alto error de este predictor para dicho horizonte de predicción.

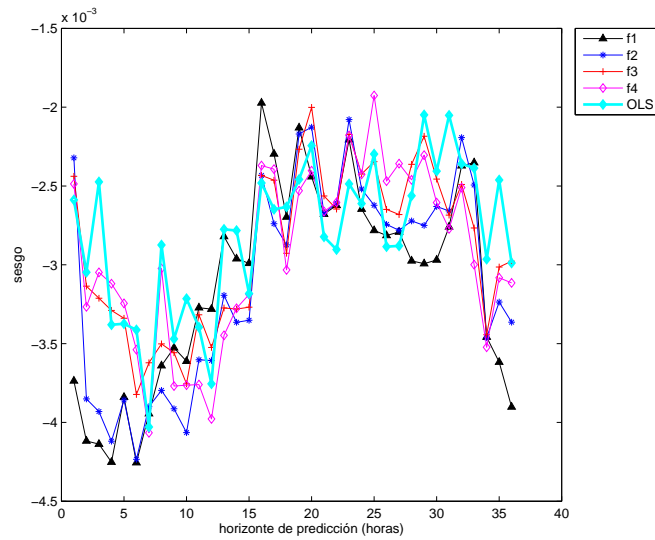


Figura 4.3: Sesgo de la combinación de predicciones mediante componentes principales en función del número de factores seleccionados; f1, f2, f3, y f4 denotan uno, dos, tres y cuatro componentes principales respectivamente.

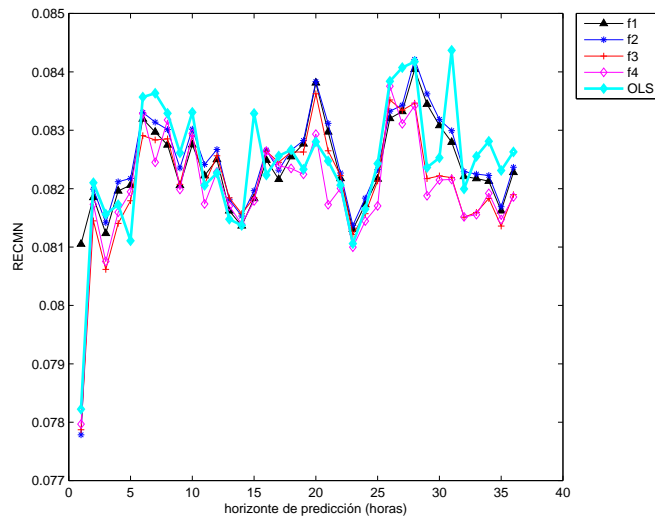


Figura 4.4: RECMN de la combinación de predicciones mediante componentes principales en función del número de factores seleccionados; f1, f2, f3, y f4 denotan uno, dos, tres y cuatro componentes principales respectivamente.

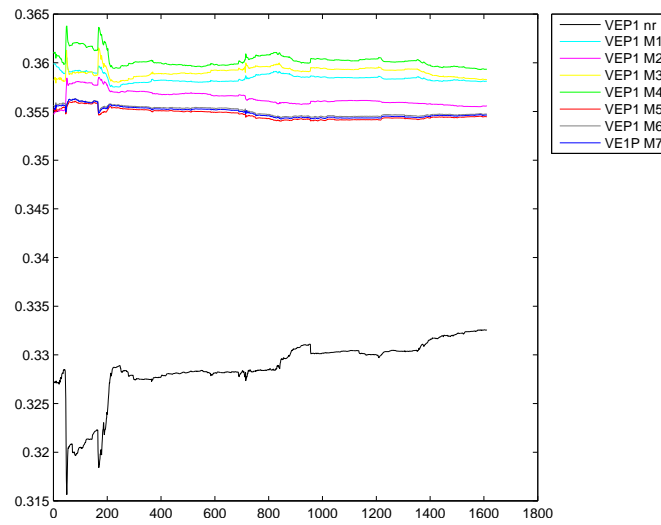


Figura 4.5: Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la primera componente principal a horizonte 1.

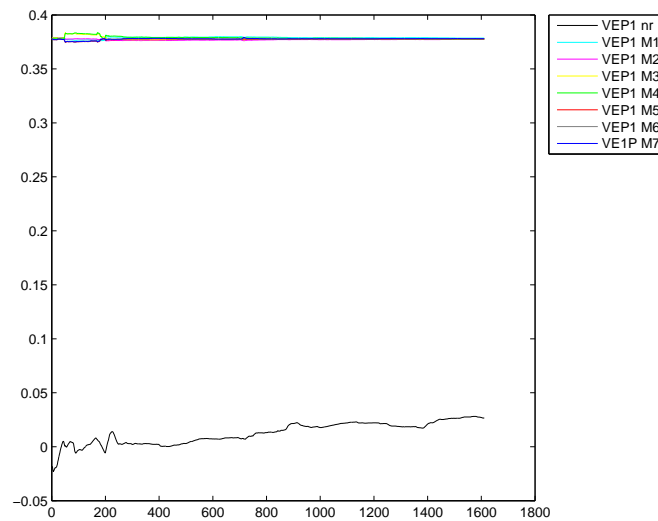


Figura 4.6: Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la primera componente principal a horizonte 36.

En la figura 4.7 se muestra la misma gráfica excluyendo este predictor. Como puede observarse, este factor vuelve a ser una media ponderada de los predictores individuales excluyendo el modelo de referencia.

Analicemos ahora el segundo componente principal. Para ello, como en el caso anterior, en la figura 4.8 se muestra la evolución temporal de los coeficiente asignados a cada predictor individual para formar este segundo componente principal a horizonte 1. Como puede observarse, este es un factor de forma. El coeficiente asociado al modelo de referencia toma valores muy elevados alternando de signo según el instante temporal.

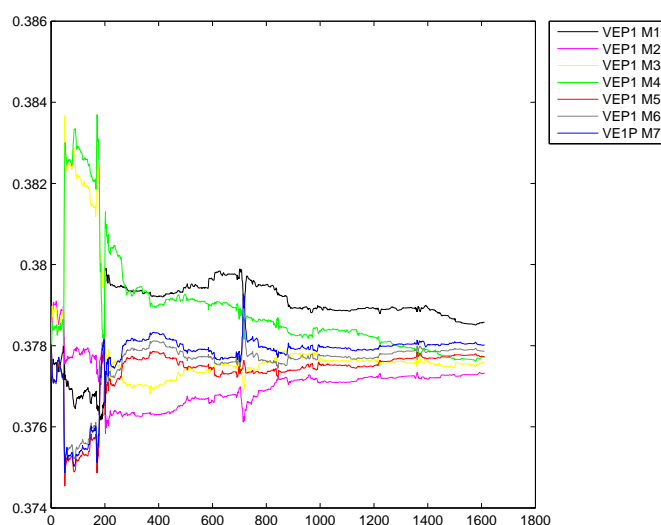


Figura 4.7: Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la primera componente principal a horizonte 36 para los modelos M1 a M7.

Los modelos M1, M2, M3 y M4 tienen ponderaciones del mismo signo y contrarias al signo del modelo nr, como si trataran de compensar el efecto de este predictor en la formación del componente. Los modelos M5, M6 y M7 tienen un coeficiente de ponderación prácticamente nulo en la formación de este componente. Como el valor de los predictores está comprendido entre cero y uno, el valor del factor, dados los signos de los coeficientes de ponderación, es muy pequeño comparado con el primer componente principal donde todos los coeficientes de ponderación son positivos. En la figura 4.9 se muestran los coeficientes de ponderación para la formación de este segundo factor para horizonte 36. Como en el caso anterior, el valor del coeficiente de ponderación asignado al modelo de referencia toma valores extremos entre -1 y 1. Para analizar mejor el resto de coeficientes, en la figura 4.10, se muestran los coeficientes asignados a los modelos M1 a M7. En este caso el valor de los coeficientes asociados a M5, M6 y M7 no son nulos, y existe cierta diferencia, en cuanto a valor asignado, entre M5 con M6 y M7. En este caso, el valor del coeficiente asociado a los modelos M1 a M7 sirven para contrarrestar el valor del coeficiente asociado al modelo nr de referencia. Como en el caso de horizonte 1, al tener los coeficientes de ponderación distinto signo, el factor resultante será de una magnitud mucho menor que el primer factor.

Analicemos ahora el tercer factor. Para ello, en la figura 4.11 se muestra la evolución temporal de los coeficientes asignados a cada predictor individual para la formación de este tercer factor. Este vuelve a ser un factor de forma, ya que asigna valores negativos ligeramente menores que $-0,4$ a los modelos M5, M6 y M7 (que son los modelos que contienen la predicción de la dirección de viento). A los modelos nr, M1, M3 y M4

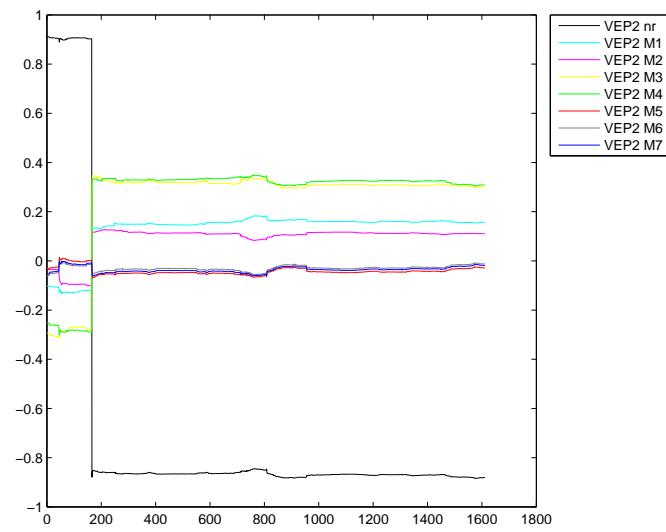


Figura 4.8: Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la segunda componente principal a horizonte 1.

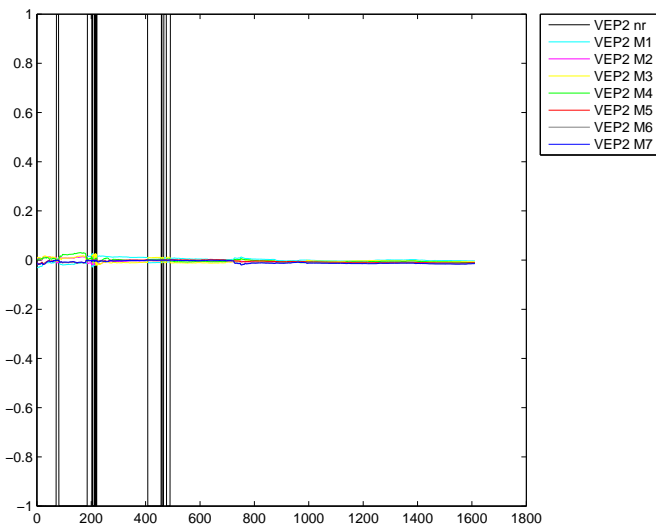


Figura 4.9: Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la segunda componente principal a horizonte 36.

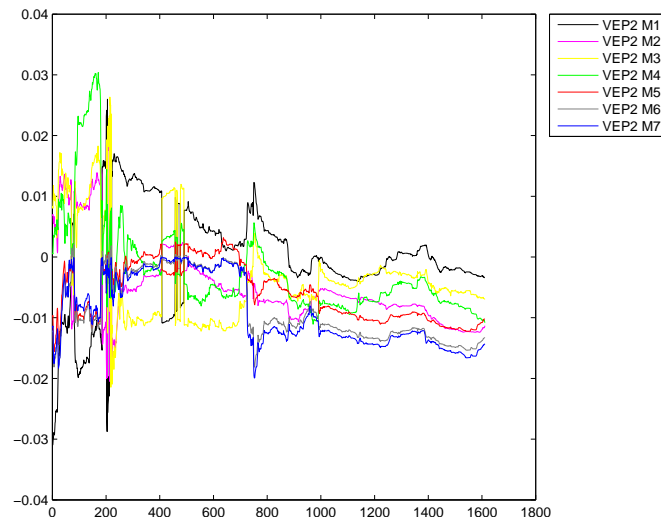


Figura 4.10: Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la segunda componente principal a horizonte 36 para los modelos M1 a M7.

les asigna un coeficiente de ponderación positivo, mientras que al modelo M2 le asigna un coeficiente prácticamente nulo (este modelo es el que incluye además de la parte autorregresiva, la predicción de la velocidad de viento, y como se vió en el capítulo anterior, es el que mejores resultados obtenía junto con M5, M6 y M7).

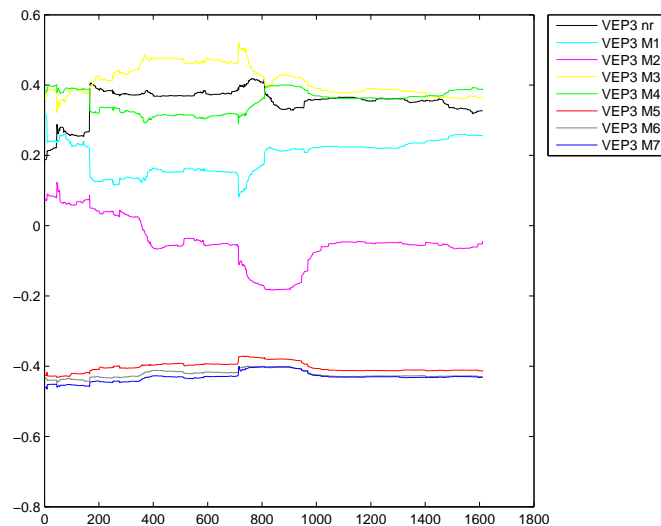


Figura 4.11: Evolución temporal del coeficiente de ponderación de cada predictor en la formación de la tercera componente principal a horizonte 1.

En el caso de horizonte 36, vemos cómo los coeficientes de ponderación de los modelos nr, M1 y M2 son prácticamente nulos, el valor de los coeficientes para M5, M6 y M7 son negativos y próximos a $-0,4$ mientras que el valor del peso asignado a los modelos

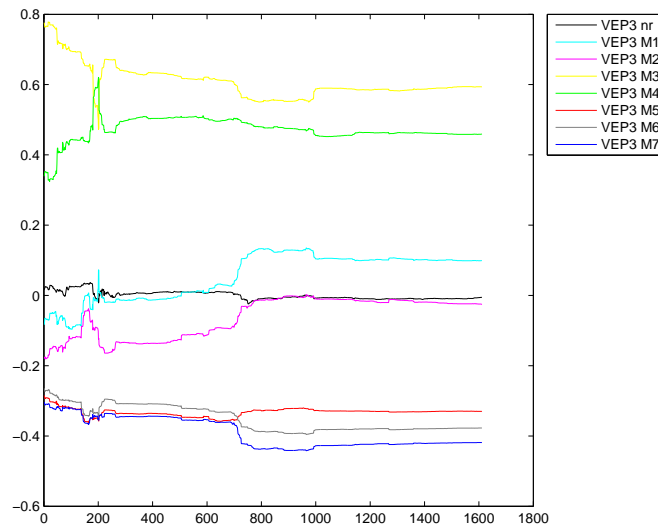


Figura 4.12: Evolución temporal del coeficiente de ponderación de cada predictor en la formación del tercer componente principal a horizonte 36.

M3 y M4 (M3 incluye la predicción de la velocidad de viento y el cuadrado de este valor y M4 además el cubo de la predicción de viento) es positivo y con valores próximos a $+0,6$. Es decir, en este caso, también es un factor de forma.

4.3.3. Mínimos cuadrados parciales

Los resultados se muestran gráficamente en las figuras 4.13 y 4.14 para el sesgo y la RECMN respectivamente. Atendiendo al sesgo, podemos observar que no existen grandes diferencias entre los diferentes esquemas y la combinación mediante OLS. Respecto a la RECMN, se puede comprobar que los resultados son similares a los obtenidos mediante la combinación de componentes principales. La combinación de factores obtenidos mediante mínimos cuadrados parciales mejora ligeramente los resultados obtenidos mediante mínimos cuadrados ordinarios, y a medida que aumenta el horizonte de predicción, se puede comprobar un mejor comportamiento de la combinación de 3 y 4 factores respecto a la combinación de 1 o 2 factores.

4.3. Resultados obtenidos

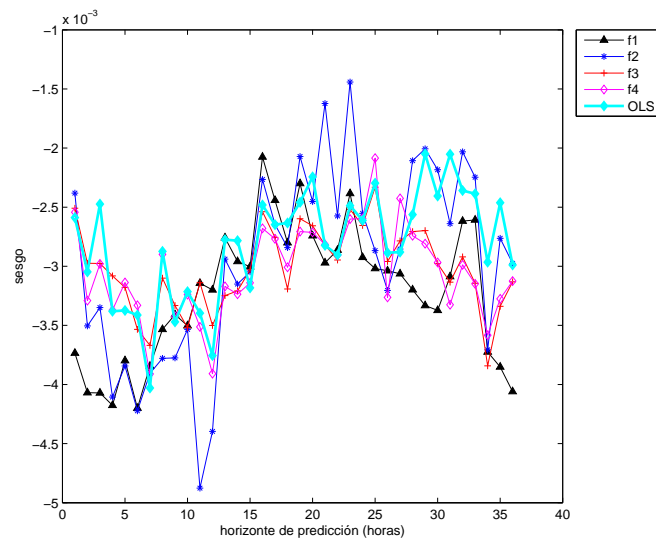


Figura 4.13: Sesgo de la combinación de predicciones mediante mínimos cuadrados parciales en función del número de factores seleccionados.

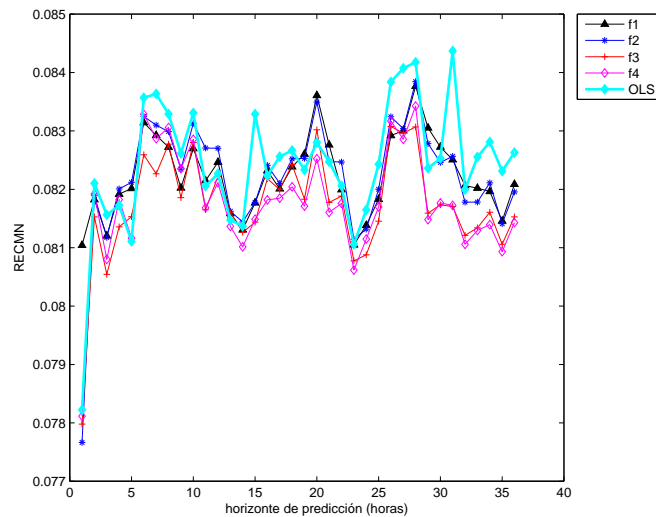


Figura 4.14: RECMN de la combinación de predicciones mediante mínimos cuadrados parciales en función del número de factores seleccionados.

4.4. Conclusiones

En este capítulo se han aplicado, de una manera novedosa en el área de predicción eólica, las técnicas multivariantes de reducción de datos para la realización de la combinación de las predicciones disponibles, con el objetivo de calcular una única predicción final. La ventaja de estas técnicas de combinación frente a las técnicas clásicas, es que éstas últimas calculan los pesos o coeficientes de ponderación en función del error de predicción. Es decir, en el instante t , se calcula $\mathbf{w}_t(h)$ como una función de $e(t|t-h)$, $e(t-1|t-h-1)$, \dots , $e(h|0)$. A medida que h se va haciendo mayor, la relación entre lo que sucede en t y en $t+h$ es menor, por lo que esta estrategia para horizontes de predicción grandes puede no ser adecuada (cuando el horizonte de predicción es grande, la dinámica del sistema puede haber cambiado, por ejemplo, el sistema puede estar en otro régimen, y si asignamos en $t+h$ la ponderación de cada predictor según lo ocurrido en t puede no ser adecuado). En las técnicas propuestas, la ponderación de cada predictor no se obtiene del error cometido hasta ese momento, sino de la cantidad de información, respecto al resto de predictores, que puede aportar. Como principales conclusiones podemos destacar que la técnica propuesta para el cálculo de la predicción final se ha demostrado que es capaz de obtener una predicción final tan buena o incluso mejor que los predictores de partida, objetivo inicial de este capítulo. Las técnicas analizadas han sido componentes principales y mínimos cuadrados parciales, demostrándose que los resultados obtenidos con ambas técnicas son similares.

La combinación de predicciones mediante técnicas multivariantes de reducción de datos mejoran ligeramente los resultados obtenidos mediante técnicas clásicas cómo puede ser mínimos cuadrados ordinarios. La razón principal es que los modelos de partida están anidados. Es decir, M2 es el resultado de añadir a M1 el término de la predicción meteorológica de la velocidad de viento. M3 es el resultado de añadir a M2 un término con la predicción de velocidad de viento al cuadrado y así sucesivamente. Este hecho afecta a los resultados de la combinación de predicciones. Sin embargo, esto no implica que no sea interesante realizar combinación de predicciones. Como se ha comentado a lo largo de esta tesis, la relación entre la velocidad de viento y la potencia eólica es altamente no lineal por lo que, en cada instante de tiempo, unos modelos aproximarán mejor que otros el verdadero valor de la potencia, por lo que siempre es mejor combinar que seleccionar un único predictor. Aplicando estas técnicas hemos visto como siempre se obtienen resultados al menos similares al mejor predictor a combinar.

A medida que aumenta el horizonte de predicción es más interesante aumentar el número de factores, aunque es de destacar que prácticamente no hay diferencia entre los

resultados obtenidos con 3 y 4 factores. El método de mínimos cuadrados parciales, al igual que el de componentes principales, realiza una reducción del espacio de entrada, como etapa previa a la regresión lineal realizada para obtener una única predicción final. En el caso de mínimos cuadrados parciales los factores maximizan la covarianza entre los datos del espacio de entrada (en nuestro caso el conjunto de predictores de partida) y la variable a predecir, mientras que componentes principales maximizan la varianza explicada de la matriz de datos de partida X sin tener en cuenta la variable a predecir. Sin embargo, no se obtienen diferencias apreciables entre los resultados obtenidos con uno y otro método. Este hecho viene a corroborar que ninguna de las dos soluciones es mejor que la otra para todos los problemas, aunque a priori podría parecer que el método de mínimos cuadrados parciales podría obtener combinaciones con mayor capacidad predictiva.

En otros problemas, para obtener el mismos resultados, son necesarios menor número de factores con la técnica de mínimos cuadrados parciales que con componentes principales. No es nuestro caso, y obtenemos resultados similares con ambas técnicas atendiendo al número de factores.

Capítulo 5

Conclusiones

5.1. Conclusiones

Partiendo de la utilidad que presenta para el operador del sistema u otros agentes conocer con antelación suficiente la potencia eólica que será generada por un parque eólico, en esta tesis doctoral se han propuesto técnicas para mejorar los modelos paramétricos que están actualmente operativos en herramientas de predicción eólica. Del análisis inicial de estos modelos, el primer desarrollo realizado ha sido la estimación por técnicas de máxima verosimilitud de las matrices de los modelos o hiperparámetros en vez de imponerlos de una manera ad-hoc. Se ha demostrado que esto permite reducir la RECMN en todos los horizontes de predicción y en todos los modelos planteados, alcanzándose reducciones de hasta el 60% en un parque eólico en terreno complejo como es el caso del parque eólico de Sotavento. Otra aportación realizada en el capítulo 3 de la tesis ha sido el planteamiento de modelos multivariantes que permitan recoger a través de las correlaciones de los parámetros del modelo o bien a través de las correlaciones en el ruido de la ecuación de estado, relaciones entre variables no recogidas explícitamente en los modelos. Como en el caso anterior, las matrices de los modelos se han estimado mediante máxima verosimilitud, obteniéndose también una mejora significativa en la reducción del error cuadrático medio de predicción. Como conclusiones a este capítulo, se ha analizado el funcionamiento en tiempo real del procedimiento de estimación de las matrices y del cálculo de las predicciones en tiempo real, demostrándose que siempre es adecuado estimar el valor de las matrices, al menos para un horizonte de predicción. La estimación para todos los horizontes, si el tiempo de ejecución del algoritmo lo hace posible, permite reducciones adicionales del error de predicción, pero mucho menores que las que se obtienen estimando al menos las matrices para un horizonte de

predicción. Además, se ha corroborado que incluso con matrices estimadas de manera óptima, se obtienen mejores resultados calculando la predicción para cada horizonte h directamente, que calculándola mediante un proceso iterativo.

Por último, se ha analizado, mediante técnicas de remuestreo "bootstrap", la distribución muestral de los hiperparámetros calculados mediante máxima verosimilitud. Los histogramas "bootstrap" de los hiperparámetros presentan un exceso de curtosis y asimetría. Estas dos características no hubieran podido ser detectadas mediante la metodología asintótica lo que podría provocar la subestimación de los intervalos de confianza calculados para los hiperparámetros. Otra conclusión derivada de este análisis es que los parámetros estimados son cuasimáximo verosímiles.

En el capítulo 4 se han aplicado, de una manera novedosa en el área de predicción eólica, las técnicas multivariantes de reducción de datos para la realización de la combinación de las predicciones disponibles, con el objetivo de calcular una única predicción final. La ventaja de estas técnicas de combinación frente a las técnicas clásicas, es que estas últimas calculan los pesos o coeficientes de ponderación en función del error de predicción. Es decir, en el instante t , se calcula $w_t(h)$ como una función de $e(t|t-h)$, $e(t-1|t-h-1)$, \dots , $e(h|0)$. A medida que h se va haciendo mayor, la relación entre lo que sucede en t y en $t+h$ es menor, por lo que esta estrategia para horizontes de predicción grandes puede no ser adecuada. En las técnicas propuestas, la ponderación de cada predictor no se obtiene del error cometido hasta ese momento, sino en la cantidad de información, respecto al resto de predictores, que puede aportar. Las técnicas analizadas han sido componentes principales y mínimos cuadrados parciales, demostrándose que los resultados obtenidos con ambas técnicas son similares. El método en dos etapas propuesto permite, para cada instante t , obtener una predicción mejor que los predictores de partida, demostrándose la utilidad de la metodología de combinación planteada en dicho capítulo y la ventaja de disponer de varios modelos de predicción operando en paralelo en problemas donde la naturaleza no lineal del mismo hace que los modelos sean aproximaciones más o menos válidas para cada rango de datos y valores de las entradas al sistema.

5.2. Desarrollos futuros

La estrategia desarrollada de estimación óptima de las matrices del sistema puede implementarse en tiempo real, para un único horizonte de predicción, las veces que se estime necesario. La técnica planteada en esta tesis, emplearía, en un instante t , todos los datos disponibles. Otra alternativa a explorar es realizar la estimación de las matrices

con una ventana de datos con el objetivo de recoger únicamente las últimas dinámicas presentes. El esquema de trabajo sería ejecutar ambos modelos en paralelo (el modelo de corto plazo y el que contiene todas las dinámicas) para combinarlos posteriormente.

Desde el punto de vista teórico, la principal dificultad del algoritmo EM para la estimación de los hiperparámetros de los modelos estriba en la obtención de estimadores máximo verosímiles restringidos a determinadas condiciones sobre los hiperparámetros, tales como estacionariedad o la existencia de momentos marginales de orden superior a dos.

Desde el punto de vista teórico es necesario además una revisión de las técnicas de remuestreo bootstrap para su aplicación en series temporales con matrices del sistema variantes con el tiempo. Existen escasas aplicaciones prácticas de estas técnicas a problemas de dicha naturaleza, dadas las cuestiones teóricas aún pendientes de resolver. En un segundo paso debería analizarse la viabilidad de la aplicación de estas técnicas en problemas en tiempo real, pues dado el número de remuestreos o simulaciones a realizar para calcular intervalos de confianza, son claramente técnicas off-line, al menos para el problema que nos ocupa. Este trabajo es interesante porque una vez obtenida la incertidumbre asociada a la estimación de los hiperparámetros el siguiente paso sería incorporar ésta en los intervalos de predicción. El filtro de Kalman proporciona estimaciones del estado (β_t^{t+1}) y la covarianza del instante $t+1$, P_t^{t+1} , a partir de la información disponible hasta el instante t , es decir, dado el conjunto de datos $\{y_1, \dots, y_t\}$. Si se asume que los errores son normales, β_t^{t+1} y P_t^{t+1} son la media condicionada y errores cuadráticos medios de predicción condicionados, respectivamente. Es decir, en este caso se cumple:

$$\begin{pmatrix} \beta_t \\ y_t \end{pmatrix} \Big| Y_{t-1} \sim N \left[\begin{pmatrix} \beta_{t-1}^t \\ A_t \beta_{t-1}^t \end{pmatrix}, \begin{pmatrix} P_{t-1}^t & P_{t-1}^t A_t' \\ A_t P_{t-1}^t & \Sigma_t \end{pmatrix} \right] \quad (5.1)$$

Cuando el modelo no es condicionalmente gaussiano, el filtro de Kalman es todavía óptimo en el sentido de proporcionar el estimador lineal de menor error cuadrático medio. En los casos prácticos como el que nos ocupa, el valor de las matrices del sistema es desconocido y se sustituyen por sus estimadores consistentes. En esta tesis se han calculado los estimadores mediante la maximización de la función de verosimilitud, obteniéndose los estimadores cuasi máximo verosímiles. Sean $\hat{\Phi}$, \hat{R} y \hat{Q} las matrices del sistema donde los verdaderos parámetros se han sustituido por sus estimaciones cuasi máximo verosímiles. Al ejecutar el filtro de Kalman con estos valores se obtienen $\hat{\beta}_t^{t+1}$ y \hat{P}_t^{t+1} . $\hat{\beta}_t^{t+1}$ es un estimador de la media condicionada β_t^{t+1} . Sin embargo, \hat{P}_t^{t+1} no es un estimador del error cuadrático medio de predicción de β_t^{t+1} porque no tiene en cuenta

la incertidumbre existente en la estimación de los parámetros.

$$\begin{aligned}
ECMP_{t+1|t} &= \mathbf{E}_t \left[(\hat{\beta}_t^{t+1} - \beta_{t+1})(\hat{\beta}_t^{t+1} - \beta_{t+1})' \right] = \\
&= \mathbf{E}_t \left[(\hat{\beta}_t^{t+1} - \beta_t^{t+1})(\hat{\beta}_t^{t+1} - \beta_t^{t+1})' \right] + \\
&+ \mathbf{E}_t \left[(\beta_t^{t+1} - \beta_{t+1})(\beta_t^{t+1} - \beta_{t+1})' \right] = \\
&= \mathbf{E}_t \left[(\hat{\beta}_t^{t+1} - \beta_t^{t+1})(\hat{\beta}_t^{t+1} - \beta_t^{t+1})' \right] + P_t^{t+1} \quad (5.2)
\end{aligned}$$

donde \mathbf{E}_t denota que la esperanza es condicionada a $\{Y_1, \dots, Y_t\}$.

$\mathbf{E}_t \left[(\hat{\beta}_t^{t+1} - \beta_t^{t+1})(\beta_t^{t+1} - \beta_{t+1})' \right]$ es cero bajo la suposición de normalidad condicionada. El segundo término de 5.2 se denomina incertidumbre del filtro (véase [10]) y representa lo alejado que está el verdadero valor del estado de su estimación y es inherente al filtro de Kalman. Sin embargo, el primer término de 5.2 es la incertidumbre debida a los parámetros y representa la discrepancia entre el valor del estado estimado con los verdaderos parámetros y el valor calculado mediante el filtro de Kalman con los parámetros estimados. Es decir, \hat{P}_t^{t+1} no tiene en cuenta la incertidumbre asociada a la estimación de los parámetros que se refleja en el primer término de 5.2 por lo que subestima el verdadero error cuadrático medio de predicción condicionado al conjunto de datos Y_t . Los intervalos de predicción calculados mediante la aproximación asintótica no tienen en cuenta esta nueva fuente de incertidumbre y en la práctica resultan bastante inexactos.

Para dar mayor utilidad a las predicciones de potencia es necesario que vayan acompañadas de una medida de incertidumbre. En el contexto de la predicción probabilista, el algoritmo de predicción de potencia propuesto se podría utilizar para calcular intervalos de predicción, utilizando conjuntos de predicciones de viento “ensemble”. Las predicciones “ensemble” o por conjuntos permiten informar sobre el nivel de significación del pronóstico operacional del modelo meteorológico. La precisión de una predicción no sólo depende del periodo del pronóstico, sino que también está influido por la estabilidad en la situación atmosférica. La atmósfera es un sistema caótico, lo que significa que variaciones muy pequeñas pueden traer grandes consecuencias. En una situación atmosférica inestable, por ejemplo, unas décimas de grado de diferencia en la temperatura inicial sobre el Atlántico, puede producir situaciones meteorológicas completamente diferentes en el tercer día del pronóstico sobre la península ibérica. El pronóstico “ensemble” o por conjuntos se calcula provocando n perturbaciones muy pequeñas (en la escala temporal y espacial) en las condiciones iniciales o de contorno del modelo operacional. El resultado mostrará n curvas adicionales para cada parámetro predicho. Un nivel de

correlación alto entre los n pronósticos implica una situación atmosférica estable y un pronóstico muy cercano a la realidad. Si los resultados presentan una alta dispersión, podemos estar ante situaciones atmosféricas inestables y pronósticos de baja fiabilidad. La idea es trasladar dicha metodología a la predicción de potencia eólica. De esta manera pueden además calcularse los intervalos de predicción. El conjunto de predicciones de potencia resultante daría lugar a una estimación probabilista de la predicción de potencia. Esta alternativa resulta más viable para su desarrollo a corto plazo, que la citada previamente mediante técnicas de remuestreo, aunque depende de la disponibilidad de las predicciones meteorológicas por conjuntos.

Debido a que los modelos son aproximaciones al verdadero comportamiento del sistema real y que cada uno de ellos funciona mejor en unos instantes que en otros, y atendiendo a la gráfica potencia velocidad, podría analizarse la realización de combinación de predicciones por regímenes. Otra opción a explorar siguiendo con las técnicas multivariantes de reducción de datos sería el análisis factorial dinámico dado que todas las predicciones quieren predecir la misma variable y, por tanto, tendrán un alto componente común.

5.3. Difusión de resultados

- Poncela, M. Saludes, S. García, R. Sanz, S. Predicción eólica. Reoltec: Grupo de trabajo de Recurso y Emplazamiento: Presentación de Predictores, Madrid, 17 de Marzo de 2006.
- Penabad, E. Balseiro, C. F. Poncela, M. Hermida, A. Dorado, A. Empleo de modelos deterministas y estadísticos para la predicción de energía eólica. XXIX Jornadas Científicas de la Asociación Meteorológica Española. 24 al 26 de abril de 2006.
- Poncela, M. Poncela, P. Forecasting wind energy through time-varying parameter models. Tipo de participación: Ponencia Congreso: 27th international Symposium on forecasting. International Institute of Forecasting, Julio de 2007, New York City (EEUU).
- Poncela, M. Poncela, P. Perán, JR. Forecasting wind energy through time-varying parameter models. Applied Energy, editorial Elsevier, enviado.

Bibliografía

- [1] Alonso, A. García-Martos, C. Rodríguez, J. y Sánchez, M.J. (2011) Seasonal dynamic factor analysis and bootstrap inference: application to electricity market forecasting. *Technometrics*, 53(2):137–15.
- [2] Bates, J. M. and Granger, C. W. J. (1969) The combination of forecasts. *Operational Research Quarterly*, 20:451–68.
- [3] Beyrer, H. G. Heinemann, D. Mellinhoff, Mönnich, K. Waldl, H. P. (1999) Forecast of regional power output of wind turbines. *Proceedings of the European Wind Energy Conference*. Nice France págs. 1070-1073. ISBN 1 902916 00 X.
- [4] Clemen RT (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5:559-583.
- [5] Costa, A., Crespo, A. Navarro, J., Lizcano, G., Madsen, H., Feitosa, E. (2008) A review on the young history of the wind power short-term prediction. *Renewable & Sustainable Energy Reviews*, 12:1725-1744.
- [6] Diebold FX y López J.A. (1996) Forecast evaluation and combination. *Handbook of Statistics*, 14:241-268 (North-Holland, Amsterdam).
- [7] Dempster, A. P., Laird, N. M. y Rubin, D. B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 39:1-38.
- [8] DiCiccio, T. J., y Efron, B. (1996) Bootstrap confidence intervals. *Journal of Statistical Science* 11(3):189-228.
- [9] Efron, B. Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1(1):54-75.
- [10] Durbin J., and Siem Jan Koopman (2001). *Time series analysis by state space methods*. Oxford Statistical Science Series 24. ISBN 978-0-19-852354-3.

- [11] de Menezes, L. Bunn, D. y Taylor, J. (2000) Review of guidelines for the use of combined forecasts. *European Journal of Operational Research* 120:190-204.
- [12] Fente, L. F. De Castro, M. Zubiaur, R. (2003) Gamesa Energía Cassandra Project. *Proceedings of the European Wind Energy Conference*. Madrid Spain
- [13] Focken, U., Lange, M. and Waldl, H.-P. (2001) Previento - A wind power prediction system with an innovative upscaling algorithm. Helm, P. and Zervos, A.(eds.), *Wind Energy for the new millennium*. *Proceedings of the European Wind Energy Conference*, Copenhagen, Denmark, 02.-06.07.2001, 826-829.
- [14] Gallardo, C. Gaertner M A., Tejada, C. Martínez, N. Calabria, S. Martínez, N. Fernández B. (2003) Casandra project results at wind power 72h range daily operational forecast in Spain. *Proceedings of the European Wind Energy Conference*. Madrid Spain.
- [15] Giebel, G. Landberg, L. (2001) ZEPHYR-Project. The next generation prediction system. *Proceedings of the European Wind Energy Conference* Copenhagen, Denmark págs. 777-780.
- [16] Giebel, Landberg, L., Nielsen T. S., and Madsen H. (2002) The zephyr project. The next generation prediction system. In *Proc. Global Windpower Conf. Exhib.*, Paris, France, Apr. 2-5, 2002.
- [17] Giebel, G. Brownsword, R. Kariniotakis G. (2003) The state of the art in short-term prediction of wind power. A literature overview. [En red]. Disponible: <http://anemos.cma.fr> (deliverable D1.1 of ANEMOS project).
- [18] Giebel G., Landberg L., Kariniotakis G., and Brownsword R. (2003) State-of-the-art on methods and software tools for short-term prediction of wind energy production. *Proceedings of the European Wind Energy Conference*, Madrid, Spain.
- [19] Hendry, D. and Clements, M. (2004) Pooling of forecasts. *Econometrics Journal* 7:1-31.
- [20] Joensen, A. Giebel, G. Landberg, L. Madsen, H and Nielsen, H. Aa. (1999) Model output statistics applied to wind power prediction. *Proceedings of the European Wind Energy Conference* Nice France págs. 1177-1180. ISBN 1 902916 00 X.
- [21] Joensen, A. Madsen, H. Nielsen, T.S. (1997) Non-parametric statistical methods for wind power prediction. *Proceedings of the European Wind Energy Conference* Dublin Ireland págs. 788-792. ISBN 0 9533922 0 1.

-
- [22] Kariniotakis, G. Nogaret, E. Dutton, A. G. Halliday, J. A. Androutsos (1999) Evaluation of advanced wind power and load forecasting methods for the optimal management of isolated power systems. Proceedings of the European Wind Energy Conference Nice France págs. 1082-1085. ISBN 1 902916 00 X.
- [23] Landberg, L. (1998) A mathematical look at a Physical Power Prediction Model. *Wind Energy* 1:23-28.
- [24] Landberg, L. (1999) Operational results form a physical power prediction model. Proceedings of the European Wind Energy Conference Nice France págs. 1086-1089. ISBN 1 902916 00 X.
- [25] Landberg, L. Giebel, G. Madsen, H. Jorgensen, J.U. Laursen, L. Tofting, J. Christensen, H.S. Bjerge, C. Nielsen and Emsholm. (2002) Wind Farm production prediction-the Zephyr model. Technical report. Riso National Laboratory, Roskilde, Denmark. ENS-1363/99-0017.
- [26] Lei, Ma. Shiyan, Luan, Chuanwen, Jiang. et al. (2009) A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews* 13:915-920.
- [27] Loureiro, Y.(2004) El modelo de predicción de Cener. Predicción y Mercado de Energía Eólica. Reoletos Conferencias. Madrid, 30 de Noviembre de 2004.
- [28] Louka, P. Galanis, G. Siebert, N. Kariniotakis, G. (2008) Improvements in wind speed forecasts for wind power prediction purposes using Kalman Filtering. *Journal of Wind Engineering and Industrial Aerodynamics* 96:2348-2362.
- [29] Lozano, T. Orellana, P. Navarro, E. Abascal A., Piernagorda, E. Prediction and planning of wind energy production (2003) Proceedings of the European Wind Energy Conference. Madrid, España.
- [30] Madsen, H. Kariotakis G., Nielsen H Aa. Nielsen T. S. Pinson P. (2004) A protocol for standardizing the performance evaluation of short-term wind power prediction models. Technical Report of Anemos Project.
- [31] Martí I.,Nielsen T. S., Madsen H., Navarro J., Roldán A., Cabezón D., Barquero C. G. (2001) Prediction models in complex terrain. Proceedings of the EWEC Copenhagen, Denmark.
- [32] Martí I.,Nielsen T. S., Madsen H., Roldán A., Pérez S. (2002) Improving prediction models in complex terrain.Proceedings of the 2002 Global Windpower Conference. Paris, France.

- [33] Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., Conzelmann, G. (2009) Wind power forecasting: State-of-the-art 2009. Report ANL/DIS-10-1, Argonne National Laboratory, November 2009.
- [34] Muñoz, L. (2004) Experiencias de un modelo de predicción e implicaciones en el ámbito español. Predicción y Mercado de Energía Eólica. Recoletos Conferencias. Madrid, 30 de Noviembre de 2004.
- [35] Newbold, P. y D.I. Harvey (2002) Forecasting combination and encompassing. In M. Clements and D. Hendry (Eds.). A companion to economic forecasting. Oxford: Blackwells, 268-283.
- [36] Nielsen, H. Aa., Nielsen T. S. y Madsen H. (1998) Conditional parametric arx-models Journal of Time Series Analysis.
- [37] Nielsen, H. Aa., Nielsen T. S. Joensen A. y Madsen H. (2000) Tracking time-varying coefficient functions. International Journal of adaptive Control and Signal Processing 14:813-828.
- [38] Nielsen T. S., y Madsen, H. and Tofting (2002) Experiences with statistical methods for predicting wind power in Proc. GlobalWindpower Conf. Exhib., Paris, France, Apr. 2-5, 2002.
- [39] Nielsen T. S. (2002) Online prediction and control in nonlinear stochastic systems. Ph. D. Thesis Department of Mathematical Modelling. Technical University of Denmark.
- [40] Nielsen T. S., Nielsen H. Aa., Madsen H., (2002) Prediction of wind power using time-varying coefficient-functions. 15th Triennial World Congress. IFAC 2002.
- [41] Penabad, E. Balseiro, C. F. Poncela, M. Hermida, A. Dorado, A. (2006) Empleo de modelos deterministas y estadísticos para la predicción de energía eólica. XXIX Jornadas Científicas de la Asociación Meteorológica Española. 24 al 26 de abril de 2006.
- [42] Peña D. (2002) Análisis de datos multivariantes. McGrawHill ISBN: 8448136101.
- [43] Poncela P., Rodríguez J., Sánchez-Mangas R. y Senra E. (2011) Forecast combination through dimension reduction techniques. International Journal of Forecasting, 27(2):224-237.
- [44] Raj, B. and Ullah, A. (1981) Econometrics, A varying coefficients approach. Croom Helm, Ltd 2-10, London. ISBN 0-7099-0313-8.

-
- [45] Rodríguez, A. y Ruiz, E. (2009). Bootstrap prediction interval in space state models. *Journal of time series analysis*, 30:167-178.
- [46] Sánchez I., Usaola J., Ravelo O., Velasco C., Domínguez J., Lobo M. G., González G., Soto F., (2002) SIPREOLICO-A wind power prediction system based on flexible combination of dynamic models. Application to spanish power system. *Proceedings of the World Wind Energy Conference and Exhibition 2002*.
- [47] Sánchez I. (2006) Short term prediction of wind energy production. *International Journal of Forecasting* 22:43–56.
- [48] Sánchez I. (2008) Adaptive combination of forecasts with application to wind energy. *International Journal of Forecasting* 24:679–693.
- [49] Stoffer, D.S. y Wall K. (1991) Bootstrapping state space models: Gaussian maximum likelihood estimation and the Kalman Filter. *Journal of the American Statistical Association* 86:1024-1033.
- [50] Stoffer, D. S. y Wall K. (2004) Resampling in State Space Models. En *State Space and Unobserved Component Models Theory and Applications*, Capítulo 9, pp 227-258. Andrew Harvey, Siem Jan Koopman and Neil Shephard, eds. Cambridge: Cambridge University Press.
- [51] Shumway RH y Stoffer DS. (1982) An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis* 3:253-64.
- [52] Timmerman, A. (2006). Forecast combinations. In G. Elliot, C.W.J. Granger and A. Timmerman (Eds.), *Handbook of Economic Forecasting*. Amsterdam: North-Holland, pp. 135-196.
- [53] Thor S-E, Weis-Taylor P. (2002) Long term research and development needs wind energy for the time frame 2000-2020. *Wind Energy* 5:73-75.
- [54] Watson MW y Engle RF. (1983) Alternative algorithms for the estimation of dynamic, mimic and varying coefficient regression models. *Journal of Econometrics* 23:385-400.
- [55] Wold, H. (1966) Estimation of principal components and related models by iterative least squares. In P. R. Krishnaiah (Ed). *Multivariate Analysis*. New York: Academic Press, pp. 391-420.
- [56] Wu C. F. (1983) On the convergence properties of the EM algorithm. *Annals of Statistics*, 11:95-103.

- [57] Zellner A.(1962) An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. *Journal of the American Statistical Association* 57:348-368.

