

TESIS DOCTORAL

Modelos avanzados para
la predicción a corto plazo
de la producción eléctrica
en parques eólicos

Luis Alfredo Fernández Jiménez



UNIVERSIDAD DE LA RIOJA

TESIS DOCTORAL

Modelos avanzados para
la predicción a corto plazo
de la producción eléctrica
en parques eólicos

Luis Alfredo Fernández Jiménez

Universidad de La Rioja
Servicio de Publicaciones
2008

Esta tesis doctoral, dirigida por el doctor D. Ignacio J. Ramírez Rosado, fue leída el 6 de septiembre de 2007, y obtuvo la calificación de Sobresaliente Cum Laude Unanimidad

© Luis Alfredo Fernández Jiménez

Edita: Universidad de La Rioja
Servicio de Publicaciones

ISBN 978-84-691-8299-4

MODELOS AVANZADOS PARA LA PREDICCIÓN A CORTO PLAZO DE LA PRODUCCIÓN ELÉCTRICA EN PARQUES EÓLICOS

TESIS DOCTORAL

Autor:

D. Luis Alfredo Fernández Jiménez

Director:

Dr. D. Ignacio J. Ramírez Rosado



UNIVERSIDAD DE LA RIOJA

DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

Logroño, Junio de 2007

AGRADECIMIENTOS

Quiero aprovechar estas líneas para expresar mi agradecimiento a quienes han hecho posible que esta tesis doctoral haya llegado a buen término.

Al Director de la misma, profesor Ignacio J. Ramírez Rosado, por su continuo apoyo, orientación y consejo en los trabajos realizados.

A los investigadores del Instituto de Engenharia de Sistemas e Computadores, INESC, de Oporto, por las atenciones y recomendaciones recibidas en mis estancias en sus instalaciones. En concreto, debo agradecer muy especialmente al profesor Vladimiro Miranda, por sus sugerencias en el tema de la tesis, y al profesor Claudio Monteiro, por su apoyo continuo estos últimos años.

A los técnicos de Red Eléctrica de España S.A., en especial a D. Gerardo González, D. José Luís Mata y Dña. Cristina Gómez, por su colaboración en la obtención de datos con los que desarrollar los modelos.

A los técnicos de Iberdrola S.A., en especial a D. Jesús M^a Olartecochea, por su colaboración en la obtención de los datos.

A los técnicos de Energía Hidroeléctrica de Navarra S.A.; y a los de Eólicas de La Rioja S.A., en especial al profesor Félix Sáenz, por sus comentarios y su colaboración en la obtención de datos.

A los compañeros del Departamento de Ingeniería Eléctrica de la Universidad de La Rioja, principalmente a los integrantes del Grupo de Investigación de “Planificación, Operación y Control de Sistemas de Energía Eléctrica” de la Universidad de La Rioja: Montse, Eduardo, Enrique y los dos Pedros. Y, por supuesto, a Charo, por su ayuda en la resolución de todos los trámites administrativos.

A las niñeras tan especiales que he tenido estos últimos meses, que me han permitido eludir en parte mis obligaciones de padre: mis suegros y mis padres.

A quienes más paciencia y comprensión han mostrado por mis ausencias, por mis despistes, por todo... a Nuria y a Elisa.

A todos los que me educaron.

A mis padres y hermanos.

A Nuria y a Elisa.

| | |
|---|----|
| Capítulo 1 INTRODUCCIÓN Y OBJETIVOS DE LA TESIS | 1 |
| 1.1. INTRODUCCIÓN | 1 |
| 1.2. OBJETIVOS DE LA TESIS | 3 |
| 1.3. ESTRUCTURA Y CONTENIDO DE LA TESIS | 6 |
| | |
| Capítulo 2 GENERACIÓN DE ENERGÍA ELÉCTRICA EN PARQUES EÓLICOS | 11 |
| 2.1. LA ENERGÍA EÓLICA | 11 |
| 2.1.1. El viento | 11 |
| 2.1.2. Evolución del aprovechamiento de la energía eólica. | 14 |
| 2.2. INTEGRACIÓN EN LA RED ELÉCTRICA DE LA PRODUCCIÓN ELÉCTRICA DE ORIGEN EÓLICO | 16 |
| 2.2.1. Repercusiones en la planificación de la Red Eléctrica..... | 17 |
| 2.2.2. Repercusiones en la operación de la Red Eléctrica..... | 18 |
| 2.3. NECESIDAD DE PREDICCIÓN A CORTO PLAZO DE LA PRODUCCIÓN ELÉCTRICA EN PARQUES EÓLICOS | 22 |
| 2.4. VALOR ECONÓMICO DE LA PREDICCIÓN | 27 |
| 2.5. HORIZONTES DE PREDICCIÓN | 30 |
| 2.6. INCERTIDUMBRE EN LA PREDICCIÓN | 32 |
| | |
| Capítulo 3 REVISIÓN BIBLIOGRÁFICA DE LOS MODELOS DE PREDICCIÓN A CORTO PLAZO DE LA PRODUCCIÓN ELÉCTRICA DE ORIGEN EÓLICO | 35 |
| 3.1. EVALUACIÓN DE MODELOS DE PREDICCIÓN. CRITERIOS | 35 |
| 3.2. FAMILIAS DE MODELOS DE PREDICCIÓN DE LA PRODUCCIÓN ELÉCTRICA DE ORIGEN EÓLICO EN EL CORTO PLAZO | 39 |
| 3.2.1. Modelos de predicción basados en series temporales de viento o en series de energía eléctrica generada. | 40 |
| 3.2.2. Modelos de predicción basados en modelos atmosféricos..... | 47 |
| 3.3. CONCLUSIONES | 57 |
| | |
| Capítulo 4 MODELOS DE PREDICCIÓN METEOROLÓGICA | 61 |
| 4.1. CONCEPTOS BÁSICOS DE LOS MODELOS DE PREDICCIÓN METEOROLÓGICA | 61 |
| 4.2. MODELOS ATMOSFÉRICOS | 64 |
| 4.2.1. Modelos atmosféricos globales..... | 65 |
| 4.2.2. Modelos atmosféricos locales..... | 68 |
| 4.3. EL MODELO ATMOSFÉRICO MM5 | 70 |
| 4.4. CONCLUSIONES | 74 |
| | |
| Capítulo 5 MODELOS PARA LA PREDICCIÓN DE PRODUCCIÓN ELÉCTRICA EN PARQUES EÓLICOS | 77 |
| 5.1. SERIES TEMPORALES | 77 |

| | |
|--|-----|
| 5.2. MODELOS ARIMA | 79 |
| 5.2.1. Modelos auto-regresivos | 80 |
| 5.2.2. Modelos de media móvil | 81 |
| 5.2.3. Modelos ARMA | 82 |
| 5.2.4. Modelos ARIMA | 82 |
| 5.2.5. Modelos SARIMA | 83 |
| 5.2.6. Identificación de modelos ARIMA | 84 |
| 5.2.7. Modelos ARIMA multivariantes | 85 |
| 5.2.8. Modelos con variables exógenas | 85 |
| 5.2.9. Ventajas e inconvenientes de los modelos ARIMA | 87 |
| 5.2.10. Aplicación de los modelos ARIMA para la predicción de la producción eléctrica en parques eólicos en el corto plazo. | 87 |
| 5.3. REDES NEURONALES ARTIFICIALES | 88 |
| 5.3.1. Perceptrones multicapa | 91 |
| 5.3.2. Red neuronal unidireccional multicapa generalizada | 94 |
| 5.3.3. Red neuronal modular | 96 |
| 5.3.4. Red neuronal de función de base radial. | 97 |
| 5.3.5. Redes neuronales de análisis de componentes principales | 98 |
| 5.3.6. Redes neuronales recurrentes | 101 |
| 5.3.7. Mapas auto-organizados. | 107 |
| 5.3.8. Ventajas e inconvenientes de los modelos con redes neuronales | 109 |
| 5.3.9. Aplicación de los modelos de redes neuronales en la predicción de la producción eléctrica en parques eólicos en el corto plazo | 110 |
| 5.4. SISTEMAS DE INFERENCIA DIFUSA | 111 |
| 5.4.1. Estructura de un sistema de inferencia difusa. | 113 |
| 5.4.2. Sistemas de inferencia difusa de Takagi-Sugeno | 114 |
| 5.4.3. Partición del espacio de entrada | 115 |
| 5.4.4. Sistemas neuro-difusos. | 118 |
| 5.4.5. Ventajas e inconvenientes de los modelos basados en sistemas de inferencia difusa. | 120 |
| 5.4.6. Aplicación de modelos basados en sistemas de inferencia difusa en la predicción de la producción eléctrica en parques eólicos en el corto plazo. | 121 |
| 5.5. CONCLUSIONES | 122 |
| | |
| Capítulo 6 OPTIMIZACIÓN DE MODELOS DE PREDICCIÓN CON ALGORITMOS GENÉTICOS | 125 |
| 6.1. ALGORITMOS GENÉTICOS | 125 |
| 6.2. OPTIMIZACIÓN DE MODELOS DE PREDICCIÓN BASADOS EN REDES NEURONALES | 131 |
| 6.3. OPTIMIZACIÓN DE MODELOS DE PREDICCIÓN BASADOS EN SISTEMAS DE INFERENCIA DIFUSA | 131 |
| 6.3.1. Sistema de optimización propuesto | 134 |
| 6.3.2. Resultados del sistema de optimización propuesto | 138 |
| 6.4. CONCLUSIONES | 139 |
| | |
| Capítulo 7 NUEVOS MODELOS AVANZADOS PARA LA PREDICCIÓN A CORTO PLAZO DE LA PRODUCCIÓN ELÉCTRICA EN PARQUES EÓLICOS | 141 |
| 7.1. CARACTERÍSTICAS DE LOS PARQUES EÓLICOS. DATOS PARA LA PREDICCIÓN | 141 |
| 7.2. SELECCIÓN Y NORMALIZACIÓN DE VARIABLES PARA LOS MODELOS DE PREDICCIÓN | 144 |

| | |
|--|------------|
| 7.2.1. Selección de variables de entrada para el modelo..... | 144 |
| 7.2.2. Normalización de variables de entrada..... | 146 |
| 7.2.3. Técnicas de selección y normalización utilizadas..... | 148 |
| 7.3. NUEVOS MODELOS DE PREDICCIÓN DE LA PRODUCCIÓN ELÉCTRICA BASADOS EN SERIES TEMPORALES..... | 149 |
| 7.3.1. Modelo de predicción basado en descomposición de Fourier..... | 150 |
| 7.3.2. Modelo de predicción basado series temporales fuzzy..... | 153 |
| 7.3.3. Modelo de predicción basado en regresión lineal..... | 158 |
| 7.4. NUEVOS MODELOS DE PREDICCIÓN DE LA PRODUCCIÓN ELÉCTRICA BASADOS EN PREDICIONES DE MODELOS ATMOSFÉRICOS..... | 161 |
| 7.4.1. Nuevo modelo de predicción con información limitada del modelo atmosférico..... | 161 |
| 7.4.2. Nuevo modelo de predicción con información completa del modelo atmosférico..... | 164 |
| 7.5. CONCLUSIONES..... | 166 |
| | |
| Capítulo 8 RESULTADOS COMPUTACIONALES..... | 169 |
| 8.1. PREDICCIÓN DE LA VELOCIDAD MEDIA DEL VIENTO A MUY CORTO PLAZO..... | 169 |
| 8.1.1. Datos disponibles..... | 169 |
| 8.1.2. Modelos de predicción analizados..... | 171 |
| 8.1.3. Desarrollo de los modelos de predicción..... | 176 |
| 8.1.4. Resultados obtenidos..... | 179 |
| 8.2. PREDICCIÓN DE LA ENERGÍA ELÉCTRICA GENERADA EN UN PARQUE EÓLICO A CORTO PLAZO..... | 187 |
| 8.2.1. Datos disponibles..... | 188 |
| 8.2.2. Modelos de predicción analizados..... | 189 |
| 8.2.3. Desarrollo de los modelos de predicción..... | 190 |
| 8.2.4. Resultados obtenidos..... | 196 |
| 8.3. PREDICCIÓN DE LA POTENCIA MEDIA HORARIA GENERADA EN UN PARQUE EÓLICO A CORTO PLAZO..... | 196 |
| 8.3.1. Datos disponibles..... | 197 |
| 8.3.2. Modelos de predicción analizados..... | 198 |
| 8.3.3. Resultados obtenidos..... | 202 |
| 8.4. PREDICCIÓN A CORTO PLAZO DE LA POTENCIA MEDIA HORARIA GENERADA EN UN PARQUE EÓLICO CON INFORMACIÓN LIMITADA DE UN MODELO ATMOSFÉRICO..... | 207 |
| 8.4.1. Datos disponibles..... | 208 |
| 8.4.2. Modelos de predicción analizados..... | 210 |
| 8.4.3. Resultados obtenidos..... | 215 |
| 8.5. PREDICCIÓN A CORTO PLAZO DE LA POTENCIA ELÉCTRICA GENERADA EN UN PARQUE EÓLICO CON INFORMACIÓN COMPLETA DE UN MODELO ATMOSFÉRICO..... | 219 |
| 8.5.1. Datos disponibles..... | 220 |
| 8.5.2. Modelos de predicción analizados..... | 223 |
| 8.5.3. Resultados preliminares obtenidos..... | 228 |
| 8.5.4. Mejora de las predicciones de la potencia eléctrica a muy corto plazo..... | 238 |
| 8.5.5. Nuevo sistema global para la predicción..... | 251 |
| 8.5.6. Aplicación del nuevo sistema global de predicción..... | 262 |
| 8.6. CONCLUSIONES..... | 269 |
| | |
| Capítulo 9 CONCLUSIONES..... | 273 |
| 9.1. RESUMEN DE LA INVESTIGACIÓN Y CONCLUSIONES..... | 273 |

| | |
|---|-----|
| 9.2. PRINCIPALES APORTACIONES DE LA TESIS | 276 |
| 9.3. FUTURAS LÍNEAS DE INVESTIGACIÓN | 279 |
| BIBLIOGRAFÍA | 283 |
| ANEXO. IMPLEMENTACIÓN DE UN NUEVO FILTRO KALMAN | 301 |
| I. El filtro Kalman. | 301 |
| II. Aplicación del nuevo filtro Kalman en la mejora de la predicción de velocidad media del viento..... | 304 |

1. INTRODUCCIÓN Y OBJETIVOS DE LA TESIS

En este primer capítulo se presenta una primera aproximación a la importancia del problema de la predicción a corto plazo de la producción eléctrica en parques eólicos, se indican los objetivos propuestos en la realización de los trabajos de investigación de la tesis doctoral, y finalmente se señalan de forma resumida los contenidos de sus capítulos.

1.1. INTRODUCCIÓN

La predicción de sucesos futuros ha fascinado al hombre desde los tiempos más remotos. En los primeros relatos escritos, las referencias a la predicción del futuro son constantes. La predicción requiere un estudio meticuloso de los sucesos pasados y de las relaciones existentes entre ellos, para a partir de ahí, intentar extrapolar un suceso futuro.

Según historiadores clásicos, el primer hombre que fue capaz de realizar una predicción fundamentada sobre una base científica fue Thales de Mileto, uno de los siete sabios de Grecia. Thales fue capaz de predecir un eclipse total de sol, ocurrido el 28 de mayo del 585 a. C. Evidentemente, la capacidad de predicción de Thales le venía de su estudio de las centenarias series cronológicas de eclipses que habían registrado los astrónomos egipcios, en las que descubrió un patrón que se repetía en el tiempo.

Pero si en tiempos de Thales de Mileto la predicción de un suceso natural constituía causa de admiración general, en el momento actual, con la tecnología disponible, la predicción de valores futuros de cualquier variable entra dentro de lo que podríamos catalogar como técnica cotidiana. El desarrollo de modelos de predicción ha sido, durante los últimos 50 años, un campo científico relativamente activo y en el que se han sucedido continuos avances. La búsqueda de la relación existente entre el valor actual de una variable y los valores pasados de la misma, o de otras variables que puedan influir sobre el valor de la primera, constituye una de las bases del desarrollo de modelos de predicción. La suposición de que entendiendo las relaciones entre los datos del pasado, podremos extrapolar el comportamiento futuro, es la idea que subyace en todos los modelos de predicción.

Muchos son los ejemplos que se pueden dar de variables, cuyo conocimiento previo de su valor futuro, nos resultaría especialmente (económicamente) interesante. Pero bajo un punto de vista global, la predicción de valores futuros de variables puede ser importante para las tareas indicadas a continuación.

- Planificar: el mantenimiento de unas condiciones de funcionamiento de un sistema requiere una anticipación a los valores futuros que afectan a tal sistema.
- Adquirir recursos: los tiempos necesarios en la adquisición y transporte de productos, la preparación de personal cualificado o la puesta en marcha de nuevas unidades de producción, obliga a tomar las decisiones de su adquisición o formación de tal forma que estén disponibles en el momento que se les precise. Para esa toma de decisiones, una predicción ajustada es de gran importancia económica y estratégica.
- Rentabilizar esfuerzos: la predicción de valores futuros puede facilitar la toma de decisiones financieras o laborales que permitan obtener la máxima rentabilidad en el corto o medio plazo.

Son muchas las actividades humanas, especialmente las económicas, en las que la predicción juega un papel de primordial importancia.

El sector de actividad de la energía eléctrica no es ajeno a la necesidad de predicción de las variables que le afectan. Pero no sólo para las grandes decisiones de trascendencia futura (política de expansión, política de inversiones, etc.), sino también para su funcionamiento diario. El carácter diferenciador de la energía eléctrica de otro tipo de suministros vitales en nuestra sociedad tecnológica, es el hecho de que actualmente no se puede almacenar a gran escala. Desde un punto de vista práctico, se debe generar tanta energía eléctrica como se vaya a consumir. Esto obliga a una planificación de la producción eléctrica día tras día, y obliga a decidir que centrales eléctricas deben estar en funcionamiento en cada momento.

En los últimos 20 años, dentro del sector de la energía eléctrica, el problema de predicción que más interés ha suscitado ha sido el de la previsión de la demanda eléctrica en el corto plazo (entre 1 y 7 días). En las publicaciones especializadas es posible encontrar decenas de artículos describiendo nuevos modelos de predicción de la demanda de energía eléctrica, modelos en los que el desarrollo de nuevas técnicas meta-

heurísticas (próximas a la inteligencia artificial, en el campo del “soft computing”), se ha dejado sentir. El objetivo de todos estos modelos de predicción es proporcionar un valor de predicción de energía demandada lo más aproximado posible al futuro valor real.

Pero, ante el espectacular incremento de potencia instalada en instalaciones generadoras de energía eléctrica abastecida por recursos o fuentes de energía renovable en los últimos años, ha surgido un nuevo problema de predicción en el corto plazo en el sector de la energía eléctrica: el de la predicción de la energía eléctrica que estas instalaciones van a generar. Dentro de las fuentes renovables utilizadas en estas instalaciones, la eólica es la más importante. En España, por ejemplo, la potencia instalada en parques eólicos a finales del año 2006 era de 11615 MW, y la energía generada en ellos representó el 9.03% de toda la consumida en España en el año 2006 (fuente: Red Eléctrica de España, balance diario del 31 de diciembre de 2006).

Hace una decena de años, cuando la potencia instalada en parques eólicos representaba un porcentaje casi despreciable de la potencia instalada en todo el sistema eléctrico nacional, se empezaba a considerar que la energía generada en los mismos representaba una “demanda negativa”. Pero es ahora, cuando la potencia instalada en parques eólicos supone un porcentaje significativo del total del sistema eléctrico de generación (además con clara tendencia a aumentar en los próximos tres años), cuando más necesaria se hace la predicción de la producción de energía eléctrica en parques eólicos en el corto plazo. Esta predicción tiene, como se verá en el siguiente capítulo, importantes repercusiones técnicas y económicas.

1.2. OBJETIVOS DE LA TESIS

En la realización de los trabajos de investigación, cuyos resultados se recogen en este documento, el objetivo fundamental perseguido ha sido el de desarrollar un conjunto de nuevos modelos avanzados para la predicción a corto plazo de la producción eléctrica en parques eólicos, capaces de lograr los mejores valores de predicción posibles con los datos disponibles.

El grado de amplitud de los datos disponibles es, frecuentemente, un elemento básico que determina un reto importante a la hora de crear los modelos de predicción de la producción eléctrica. Obviamente la calidad de las predicciones depende, en buena parte, de la cantidad de datos históricos de la producción eléctrica y de la existencia de

predicciones de las variables meteorológicas. Por ello, los originales modelos de predicción de la generación eléctrica que se han creado han sido desarrollados específicamente para contemplar la mayor o menor cantidad de datos utilizables para dicha predicción. El resultado final ha sido la creación de cinco nuevos modelos o sistemas de predicción de la producción eléctrica en parques eólicos.

Seguidamente indicamos los objetivos específicos de los trabajos de investigación de la tesis recogidos en este documento.

1. Creación de nuevos modelos de predicción, a corto plazo, de la producción eléctrica en parques eólicos.

Los nuevos modelos deben alcanzar buenos valores de predicción con relación a la información (datos) utilizados: serie temporal de valores de la velocidad media del viento, serie temporal de valores de la potencia eléctrica media horaria y/o predicciones de condiciones o de valores de variables meteorológicas. Además, se deben determinar los horizontes de predicción que razonablemente se adapten a las características de la información disponible, manteniendo un nivel de calidad aceptable de las predicciones.

2. Desarrollo de un sistema de predicción con horizontes de predicción de hasta 72 horas, para los casos en que se disponga de información relativamente más amplia (suficientes valores históricos de producción eléctrica y predicciones meteorológicas adecuadas).

Por un lado, los modelos publicados en las revistas especializadas tienen como horizonte máximo las 48 horas. Por otro lado, de acuerdo a la legislación española, los propietarios de parques eólicos acogidos a uno de los sistemas regulados de venta de la energía eléctrica producida, pueden estar obligados a suministrar previsiones de generación de energía eléctrica, sobre la base de una programación horaria, con una anticipación mínima de 30 horas al inicio del día correspondiente a las previsiones (es decir, previsiones con antelación de hasta 54 horas). Por ello, el desarrollo de un sistema de predicción cuyo horizonte alcance las 72h, superará el de los modelos descritos en la literatura internacional, al tiempo que podrá satisfacer las necesidades actuales de predicción en el contexto español.

3. Análisis de bondad de las predicciones de producción eléctrica logradas con los

nuevos modelos creados.

Para cada uno de los problemas planteados (casos o situaciones de predicción específicas en parques eólicos), diferentes en cuanto a la información disponible (datos) para la creación de los modelos de predicción, se analizarán los resultados obtenidos con los nuevos modelos que se desarrollen, y también con los modelos presentados en los trabajos publicados en la literatura internacional, utilizando el mismo tipo de información (datos disponibles) para las predicciones. Evidentemente, los resultados de predicción logrados con los nuevos modelos creados deberán superar a los del resto de modelos publicados.

4. Desarrollo de sistemas de predicción de la producción eléctrica más económicos.

Se optará por soluciones de menor coste. Por ello se elegirán aquellas soluciones (sistemas) cuya implementación, en las instalaciones del propietario o gestor del parque eólico, eviten la necesidad de compra diaria de cualquier tipo de datos. Todos los datos necesarios para la obtención de los valores de predicción de los nuevos modelos a desarrollar, deben ser propios del parque eólico (valores de variables medidas en el mismo parque) o públicos y fácilmente accesibles sin coste económico de adquisición.

5. Aplicación y desarrollo de herramientas para el diseño de los nuevos modelos de predicción de la producción eléctrica.

Se utilizarán las herramientas necesarias que permitan construir y optimizar los nuevos modelos (y simular los modelos publicados en la literatura internacional), tanto de predicción de la producción eléctrica en parques eólicos, como de otras variables (predicción de variables meteorológicas) que sean de interés en la predicción de la generación eléctrica en parques eólicos.

Además de los objetivos específicos, como característica general para el desarrollo de los modelos de predicción se procurará, cuando sea posible, utilizar variables de entrada que presenten los mínimos inconvenientes en cuanto a su disponibilidad y que, de acuerdo a los gestores de los parques eólicos, sean suficientemente fiables.

También, para el sistema de predicción más complejo, a desarrollar para el

mayor horizonte de predicción (hasta 72 horas), se hará una evaluación aproximada del valor económico de sus predicciones, en términos del beneficio que se obtendría como consecuencia de utilizar las mismas para realizar ofertas de venta de la energía eléctrica producida en el contexto español.

1.3. ESTRUCTURA Y CONTENIDO DE LA TESIS

El documento está dividido en nueve capítulos. Este primer capítulo presenta el resto del documento, especificando los objetivos de partida en la realización de los trabajos de investigación cuyos resultados se recogen en otros capítulos.

En el Capítulo 2 se describe la naturaleza de la energía eólica, la evolución histórica del aprovechamiento de la misma por el hombre, y la situación actual y perspectivas para un futuro próximo. Se analizan las repercusiones que tiene, en el contexto español, la integración, en la Red Eléctrica, de la energía eléctrica producida en parques eólicos, incidiendo en los aspectos que hacen de la predicción de la producción eléctrica eólica una necesidad para un adecuado funcionamiento del Sistema Eléctrico de forma económica y fiable. Se describen también las penalizaciones económicas que afectan a los desvíos (diferencia entre la potencia media horaria programada y la real) en la producción de energía eléctrica en los parques eólicos.

El Capítulo 3 corresponde al estado del arte de los modelos de predicción de la producción eléctrica en parques eólicos. Se comentan los diferentes modelos descritos en artículos publicados en la literatura internacional, con los resultados obtenidos en cada caso y las limitaciones de dichos modelos de predicción. Así, se realiza una clasificación de los modelos de predicción publicados en dos familias. La primera familia de modelos de predicción corresponde a aquellos que utilizan como variable explicativa las series temporales de velocidad media del viento o de producción eléctrica media en el parque eólico. La segunda familia de modelos incluye aquellos que utilizan como variables explicativas valores de predicción de variables meteorológicas obtenidos con modelos de predicción numérica atmosférica (modelos atmosféricos). El capítulo contiene además una descripción de los criterios para la evaluación de los modelos de predicción, esto es, criterios para la medida de la bondad de los resultados de predicción ofrecidos por los modelos.

En el Capítulo 4 se describen los modelos atmosféricos. Estos modelos conocidos dentro del campo meteorológico como modelos numéricos de predicción, son

los utilizados para obtener los valores de predicción de las variables meteorológicas que se usan como variables explicativas en los modelos de predicción de la producción eléctrica en parques eólicos (modelos de la segunda familia). En este capítulo se muestra la clasificación de los modelos atmosféricos en función de la escala espacio-temporal para la que están diseñados. Finalmente se presenta el modelo numérico de predicción MM5, que además de reunir las características que permite definirlo como uno de los más avanzados en la predicción de variables meteorológicas (entre ellas el viento) en el corto plazo, ha sido implementado y está en funcionamiento en el laboratorio de investigación de Planificación, Operación y Control de Sistemas de Energía Eléctrica del área de Ingeniería Eléctrica de la Universidad de La Rioja.

En el Capítulo 5 se hace un repaso por las técnicas más utilizadas en el desarrollo de modelos de predicción de producción eléctrica en parques eólicos, presentadas en los trabajos publicados en la literatura internacional. Así, se describen los modelos basados en la metodología de Box-Jenkins (modelos ARIMA), los modelos basados en varias familias de redes neuronales artificiales, y los modelos basados en los sistemas de inferencia difusa (lógica fuzzy). Para cada uno de estas familias de modelos se presentan las ventajas e inconvenientes que su desarrollo conlleva, así como la aplicación que de ellos se ha hecho en la predicción de la producción eléctrica en parques eólicos.

El Capítulo 6 describe la técnica de optimización utilizada para la mejora de los nuevos modelos de predicción basados en redes neuronales artificiales y en sistemas de inferencia difusa. Se enumeran los parámetros de los modelos de predicción basados en redes neuronales artificiales que han sido objeto de optimización. En el caso de los modelos de predicción basados en sistemas de inferencia difusa se describen las características de un método original desarrollado para realizar la optimización de los parámetros que lo definen.

El Capítulo 7 presenta los nuevos modelos creados en los trabajos de investigación. Inicialmente se presentan los datos disponibles para la realización de estos trabajos, con sus características más relevantes. A continuación se describen los métodos utilizados para la selección y normalización de las variables de entrada a los modelos. Finalmente, se describen uno a uno los nuevos modelos desarrollados para acometer la predicción, a corto plazo, de la producción eléctrica en parques eólicos. Cada modelo de predicción descrito está diseñado para utilizar diferentes variables de entrada: las disponibles en cada caso. Los horizontes de predicción de los diferentes modelos están comprendidos entre 0.5 y 72 horas.

El Capítulo 8 corresponde a los resultados computaciones logrados utilizando los nuevos modelos descritos en el capítulo anterior. Se presentan cinco casos distintos, diferenciados por los datos disponibles para el desarrollo de cada uno de los correspondientes modelos. Todos los datos utilizados son datos reales obtenidos de estaciones meteorológicas y de parques eólicos. Para cada uno de los casos, se utiliza uno de los modelos creados, y se obtienen las predicciones, así como los errores de predicción con datos de testeo. La bondad de los modelos propuestos se ilustra comparando sus resultados respecto a los de otros modelos de predicción de producción eléctrica en parques eólicos descritos en la literatura internacional. Se describe un avanzado sistema global de predicción, para el caso con mayor número de datos disponibles, consistente en la selección óptima, en función de la hora y del horizonte de predicción, de las predicciones obtenidas por diferentes modelos de predicción. Las predicciones generadas por este sistema global de predicción para el muy corto plazo hacen uso de un nuevo filtro Kalman, cuya misión es mejorar las predicciones de velocidad media del viento proporcionadas por el modelo atmosférico. El horizonte máximo de predicción del sistema global es de 72 horas. Por último, para este sistema global de predicción se presenta un análisis de la estimación del valor económico asociado a sus predicciones, en el caso de haber sido utilizadas en la presentación de ofertas de venta (o previsiones de generación) en el contexto de la legislación española.

El capítulo 9 recoge un resumen de los trabajos de investigación realizados, con las correspondientes conclusiones y las aportaciones originales. Se incluye también una relación de los trabajos de investigación futuros que se pretenden emprender, continuación natural de los recogidos en este documento.

Conviene destacar que los trabajos de investigación recogidos en esta tesis doctoral han dado lugar a varios artículos científico-técnicos, el primero de ellos presentado en la 20ª International Conference on Modelling, Identification and Control, (Austria, 2001); el segundo presentado en la 22ª International Conference on Modelling, Identification and Control, (Austria, 2003); el tercero publicado como artículo invitado en la revista Control and Intelligent Systems (Canadá, 2004); el cuarto presentado en la 23ª International Conference on Modelling, Identification and Control, (Suiza, 2004); y el quinto presentado en la 25ª International Conference on Modelling, Identification and Control, (Austria, 2006). Todos estos artículos están indicados en la bibliografía del presente documento.

Así mismo, los trabajos desarrollados en la tesis doctoral han estado

encuadrados en el proyecto de investigación DPI2005-07494 del Plan Nacional de I+D+I financiado por el Ministerio de Educación y Ciencia y fondos FEDER; en el proyecto de investigación ANGI2004/02 del Plan Riojano de I+D+i del Gobierno de La Rioja financiado por la Comunidad Autónoma de La Rioja; y en el proyecto de investigación HP2001-0037 con financiación del Ministerio de Ciencia y Tecnología, dentro del Programa Bilateral de Acciones Integradas de Investigación Científica y Técnica entre España y Portugal, en colaboración con el Departamento de Engenharia Electrotécnica e Computadores de la Universidad de Porto y con el prestigioso instituto europeo de investigación INESC, Instituto de Engenharia de Sistemas e Computadores en Porto (Oporto).

2. GENERACIÓN DE ENERGÍA ELÉCTRICA EN PARQUES EÓLICOS

En este capítulo se presenta el panorama actual de la energía eléctrica de origen eólico, tanto a nivel nacional como internacional, su integración en los sistemas de energía eléctrica (sistemas eléctricos de potencia) y su importancia en la planificación y operación de los mismos. También se analizan en este capítulo los horizontes de predicción más importantes bajo los puntos de vista del funcionamiento de los sistemas eléctricos de potencia y de su interés económico; y se estudia el valor que tiene una predicción adecuada de la producción eléctrica de origen eólico, en el contexto del mercado eléctrico español, así como las repercusiones en la operación de los sistemas de energía eléctrica. Igualmente se describe una de las principales causas de incertidumbre en la predicción de la producción eléctrica en parques eólicos: errores de predicción en los modelos atmosféricos.

2.1. LA ENERGÍA EÓLICA

En los siguientes apartados se describe la naturaleza física del fenómeno meteorológico que conocemos por viento y la evolución de su aprovechamiento por el hombre desde el principio de la historia hasta nuestros días.

2.1.1. El viento.

La energía eólica tiene su origen, como el de la mayor parte de las fuentes energéticas del planeta, en la energía solar. El calentamiento de masas de aire como consecuencia de la radiación solar contribuye a la aparición de corrientes ascendentes, siendo el espacio que liberan ocupado por otras masas adyacentes de aire más frío. El movimiento de las masas de aire como consecuencia del calentamiento ocasiona también la aparición de diferencias de presión, que a su vez contribuyen al propio movimiento del aire en forma de fuerzas de gradiente de presión.

El movimiento del aire, o viento, a alturas por encima de 1000 m de la superficie terrestre viene condicionado, fundamentalmente, por las diferencias de presión entre las distintas zonas y por la fuerza de Coriolis, consecuencia del giro de la Tierra. Una masa

de aire inicialmente en reposo se moverá a una zona de menor presión que la que ocupa. Cuando esa masa de aire comienza su movimiento, puede ser desviada por la fuerza de Coriolis (en función de su dirección). La “desviación” del viento es proporcional a la velocidad del mismo: a mayor velocidad mayor desviación con respecto a la dirección inicial. Conforme el viento gana velocidad, la “desviación” aumenta hasta que la fuerza de Coriolis iguala a la fuerza de gradiente de presión. Llegado a esta situación el viento fluirá paralelo a las líneas isobaras. Cuando esto sucede, denominamos al viento como geostrófico. Los vientos geostróficos (también llamados globales) no vienen influenciados por la superficie terrestre, ya que aparecen a alturas superiores a los 1000 metros.

La superficie terrestre representa una fricción para el viento que fluye sobre la misma. Esta fricción puede contribuir a cambiar la dirección del viento y disminuir su velocidad. La composición o estructura del terreno afecta directamente al valor de la fricción. Una superficie muy “rugosa”, como puede ser una zona boscosa, afectará significativamente a la velocidad del viento en la superficie, reduciendo su velocidad con respecto a otras áreas adyacentes pero menos “rugosas”, como puede ser la superficie de un lago [Troen91].

Conforme vamos ganando altura sobre la superficie del terreno, las características superficiales (rugosidad, obstáculos) tienen cada vez menos influencia en el valor de la velocidad o dirección del viento. Cuando alcanzamos una altura para la cual la influencia de las características del terreno es nula, el viento tiene las propiedades de viento geostrófico. Esa altura es la que constituye la altura de la capa límite, que depende de factores como el tipo de terreno, el valor de la velocidad del viento, el perfil vertical de temperaturas, la hora del día, etc. Habitualmente esta capa límite se encuentra en alturas entre 1000 y 2000 m.

Desde el punto de vista del aprovechamiento de la energía del viento, con las tecnologías actuales, el viento a nivel de superficie de la tierra es mucho más importante que el viento geostrófico. En este sentido, ciertas características del terreno pueden provocar la aparición de unos vientos locales que tienen una escasa dependencia con los vientos globales. Una muestra clara de esto son las zonas costeras, donde la característica local de cercanía de una gran masa de agua provoca la formación de un viento que fluye hacia el interior durante el día y hacia el mar durante la noche. Este viento, conocido como brisa, es consecuencia de las diferencias de temperatura entre las superficies del agua del mar y la tierra. Otro viento local es el que se produce en zonas montañosas, en las que el calentamiento del aire en laderas soleadas hace que éste

ascienda por las mismas por el día y descienda por las noches.

La utilización del viento, del movimiento del aire, como fuente energética implica transformar su energía cinética en otra forma de energía aprovechable por el hombre. Para el caso concreto de transformación de la energía del viento (energía eólica) en energía eléctrica, realizamos primero una transformación de energía eólica a energía cinética de rotación que, actuando sobre el eje de un generador eléctrico, se transformará en energía eléctrica. La transformación de la energía eólica en energía cinética de rotación se realiza por medio de la acción del viento sobre unas palas unidas a un eje de giro (eje a su vez acoplado mecánicamente con el eje del generador eléctrico). La fuerza del viento sobre las palas produce un par de giro en el eje. La potencia mecánica transferida al eje por el viento depende de la densidad del aire, del área de barrido de las palas y de la velocidad del viento, tal como muestra la ecuación (2.1).

$$p_m(t) = \frac{1}{2} \rho(t) A v^3(t) \quad (2.1)$$

En la ecuación (2.1) $p_m(t)$ representa la potencia mecánica en el instante t , A representa el área de barrido de las palas, $\rho(t)$ la densidad del aire en el instante t , y $v(t)$ la velocidad del viento en el instante t . Se ha considerado que el área de barrido de las palas no cambia con el tiempo, cosa que si puede ocurrir con la densidad del aire y con la velocidad del viento.

La ecuación (2.1) corresponde a un límite teórico de potencia extraíble del viento. En realidad, la potencia extraída es menor ya que en el proceso hay pérdidas: el viento no transfiere toda su energía cinética al pasar por el área barrida por las palas, hay pérdidas por fricciones, pérdidas en forma de calor, etc. Es preciso introducir un coeficiente de rendimiento en la anterior ecuación para obtener la potencia eléctrica, $p(t)$, que podemos extraer del viento, tal como muestra la ecuación (2.2). En dicha ecuación C_p representa el rendimiento de la máquina que realiza la conversión de energía eólica a eléctrica (aerogenerador).

$$p(t) = \frac{1}{2} C_p \rho(t) A v^3(t) \quad (2.2)$$

El coeficiente C_p depende del tipo de máquina e incluso de las condiciones de funcionamiento, pero tiene un límite teórico, el conocido como límite de Betz. Este

límite, formulado por el físico Albert Betz en 1919, está establecido en un valor de 59.3%. Para un aerogenerador moderno el valor máximo para el coeficiente de rendimiento C_p está en torno al 45%.

2.1.2. Evolución del aprovechamiento de la energía eólica.

El progreso humano, tanto económico como tecnológico, ha estado desde el primer momento condicionado por la existencia de unos recursos energéticos suficientes para satisfacer las necesidades que el propio progreso plantea. El viento representó, desde los mismos albores de la historia, una fuente de energía fácilmente aprovechable, fundamentalmente como sistema de propulsión en barcos. Fue precisamente este aprovechamiento de la energía eólica, la principal clave en el desarrollo del comercio y en la expansión de los grandes imperios de la antigüedad, cuna de nuestra actual civilización.

Como primera aplicación de la energía eólica, diferente a la de la navegación a vela, se puede citar un posible proyecto de irrigación emprendido por el rey babilónico Hammurabi en el siglo XVII a. C. por medio de molinos de viento [Eur05]. Posteriormente, bajo el imperio persa (siglo VII a. C.) se desarrollaron los primeros molinos de viento de eje vertical, tanto en aplicaciones de molienda como de irrigación.

Las primeras aplicaciones de energía eólica en Europa, aparte de la navegación a vela, se demoraron hasta la edad media, donde como consecuencia de la invasión musulmana de la península ibérica, se transfirió la tecnología que permitió su implantación en el viejo continente. Ya para finales del siglo XIII los molinos de viento europeos (los de La Mancha en España, o los holandeses) presentaban un aspecto similar al que mantuvieron hasta bien entrado el siglo XVIII, y con un uso exclusivo en molienda e irrigación. La difusión de los molinos de viento en el viejo continente fue enorme, llegándose a estimar en 30000 el número de los molinos de viento en funcionamiento en Europa a finales del siglo XIX [Practical05].

La primera instalación de producción de energía eléctrica a partir de energía eólica fue la realizada en Cleveland, Ohio, por Charles F. Brush en 1888. Como características más interesantes de esta instalación se pueden citar que la energía eléctrica producida era en forma de corriente continua, que la potencia nominal era de 12 kW, a pesar de contar con un diámetro de aspas de 17 metros de longitud (con 144 aspas), y que usaba un multiplicador mecánico entre el eje de las aspas y el eje del

generador. Se mantuvo en funcionamiento satisfactorio durante unos 20 años [Neg02].

Uno de los más importantes pioneros en el desarrollo de la aplicación de la energía eólica en la generación de electricidad fue Dane Poul la Cour, quien en 1902 construyó un aerogenerador con sólo cuatro palas, capaz de entregar una potencia de 73 kW, en Askov, Dinamarca. También es reseñable su construcción de un túnel de viento y sus estudios de aerodinámica, que lo convierten en el precursor de una tecnología.

El aumento de potencia de los aerogeneradores recibe un impulso en 1931, cuando se pone en servicio uno de 100 kW en Balaclava, costa del mar Caspio, Rusia. Este aerogenerador llegó a producir más de 200 MWh durante algo más de dos años. Pocos años después, en 1941, en Vermont, USA, se instala un aerogenerador de 1.25 MW de potencia (marca Smith-Putman), con dos palas de casi 30 metros de longitud. Esta potente máquina estuvo operando de forma intermitente durante unos pocos días, para acabar finalmente rompiéndose una de las palas cerca del eje, aparentemente por fatiga del material. Todavía quedaban años de estudio, desarrollo y sobre todo elaboración de nuevos diseños y materiales, que hicieran realidad un funcionamiento fiable y duradero de estas máquinas [Dodge06].

Ya en la década de los 50, un discípulo de la Cour, el danés Johannes Juul, estableció las bases de lo que son los aerogeneradores modernos, al desarrollar un modelo de 200 kW con tres palas, con regulación pasiva por pérdida aerodinámica (stall control), y un generador de corriente alterna conectado a la red eléctrica. Los trabajos de Juul son continuados en la década de los 70 por un autodidacta, Christian Riisager, quien construye un aerogenerador de 22 kW, utilizando componentes convencionales (un motor comercial actuando como generador, elementos de automóviles, etc.), cosechando un importante éxito social e impulsando el desarrollo de la industria del sector en Dinamarca.

La década de los 80 del pasado siglo significó para la industria eólica el despegue definitivo, con un importante abaratamiento de los costes, consecuencia del uso de nuevos materiales y nuevas técnicas de fabricación. Este abaratamiento de costes permite que la energía eólica sea competitiva frente a otras tecnologías energéticas. La construcción de grandes parques eólicos con aerogeneradores de 55 kW es una realidad a principios de la década: en Palm Springs, California, se construye un parque con 1000 aerogeneradores Micon de 55 kW. A partir de entonces parques eólicos con decenas o centenares de aerogeneradores se construyen por todo el mundo. Con el desarrollo de máquinas cada vez más potentes (ya existen prototipos de hasta 5 MW), la potencia

media de los nuevos parques eólicos es de decenas o centenares de MW.

En los últimos 20 años el mercado de energía eólica ha estado en continua expansión, fundamentalmente en Europa, que acapara más de la mitad de la potencia instalada total. Así, por ejemplo, en España la potencia eólica instalada se ha multiplicado por factor cercano a 340 en los últimos 15 años. La potencia instalada a nivel mundial en parques eólicos ha pasado de 7500 MW al final de 1997 a cerca de 74000 MW a finales de 2006, esperándose alcanzar los 160 GW a finales de 2010 [World07]. La lista de los 10 países del mundo con mayor potencia instalada en parques eólicos a finales de 2006 es la reflejada en la Tabla 2.1.

| País | Potencia instalada MW |
|-------------|----------------------------------|
| Alemania | 20622 |
| España | 11615 |
| USA | 11603 |
| India | 6270 |
| Dinamarca | 3136 |
| China | 2405 |
| Italia | 2123 |
| Reino Unido | 1963 |
| Portugal | 1650 |
| Francia | 1567 |

Tabla 2.1. Países con mayor potencia instalada en parques eólicos a finales de 2006.

Y el crecimiento de nuevos parques eólicos no ha cesado: por poner el ejemplo sólo de España, los objetivos del Plan de Energías Renovables 2005-2010 marcan para el final del periodo una potencia instalada en nuestro país de 20155 MW, lo que representa un 73% más de lo ya instalado a finales de 2006 [Idae05].

2.2. INTEGRACIÓN EN LA RED ELÉCTRICA DE LA PRODUCCIÓN ELÉCTRICA DE ORIGEN EÓLICO

En los dos siguientes subapartados se describen las repercusiones que tiene la integración de potencia eléctrica de parques eólicos en la planificación y operación de la Red Eléctrica, en el contexto español.

2.2.1. Repercusiones en la planificación de la Red Eléctrica.

En los primeros años de explotación comercial de la energía eólica en forma de energía eléctrica (década de los 80), la potencia instalada en parques eólicos representaba un porcentaje insignificante de la capacidad total del sistema eléctrico al que se conectaba. Generalmente la conexión se realizaba en los niveles de media tensión (unas decenas de kV). Los parques eólicos actuaban como una fuente generadora más, con factores de potencia cercanos a la unidad y cumpliendo las normas de seguridad específicas de cada red eléctrica en cuanto a derivas de tensión, desvío de frecuencia, etc. Los operadores de redes eléctricas también solían poner condiciones, en cuanto a posibles distorsiones en la forma de onda como flickers o distorsiones armónicas.

Pero conforme iba aumentando la potencia instalada en parques eólicos, el porcentaje que esta potencia instalada representaba frente a la capacidad total de la Red Eléctrica fue haciéndose mayor. El desarrollo de nuevos aerogeneradores cada vez más potentes, con la consiguiente construcción de parques de mayor capacidad, trajo también otra consecuencia: la conexión se hace en subestaciones conectadas directamente en niveles de tensión más elevados (redes de reparto o redes de transporte). Por otra parte, las zonas de mayor demanda de energía eléctrica no siempre coinciden con las de mayor potencial eólico (de hecho lo normal es la situación contraria: las zonas con mayor potencial eólico suelen corresponder a zonas casi despobladas y, por tanto, con escasa demanda energética). Así surge una necesidad de acuerdo entre los promotores de parques eólicos y los operadores de la Red Eléctrica, acuerdo que ha venido establecido en la legislación vigente.

El objetivo subyacente en la legislación vigente (nuevas reglas) es el de hacer pasar a los parques eólicos de generadores “pasivos o neutros”, a asumir un papel más activo en las redes eléctricas. El nuevo papel de los parques eólicos implica un aumento (o, por lo menos, no deterioro) de la calidad de suministro de las propias redes eléctricas.

Las nuevas reglas de conexión de parques eólicos a la Red Eléctrica imponen limitaciones de potencia máxima que se puede conectar, y condiciones mínimas de seguridad y calidad de suministro. Por particularizar para el caso de España, la Administración ya ha establecido normas generales que posibilitan estas limitaciones o condiciones. Así, en la Resolución de 11 de febrero de 2005, de la Secretaría General de la Energía (Ministerio de Industria, Turismo y Comercio), por la que se aprueba un

conjunto de procedimientos de carácter técnico e instrumental necesarios para realizar la adecuada gestión técnica del Sistema Eléctrico (BOE de 1 de marzo de 2005), aprueba los procedimientos de operación siguientes:

- P.O. 12.1. Solicitudes de acceso para la conexión de nuevas instalaciones a la red de transporte.
- P.O. 12.2. Instalaciones conectadas a la red de transporte: requisitos mínimos de diseño, equipamiento, funcionamiento y seguridad, y puesta en servicio.
- P.O. 12.3. Requisitos de respuesta frente a huecos de tensión de las instalaciones eólicas.

Así, el artículo 4.2 del P.O. 12.1. en su apartado de acceso para generación establece:

“El Operador del Sistema y Gestor de la Red de Transporte establecerá la capacidad de acceso en un punto de la red como la producción total simultánea máxima que puede inyectarse en dicho punto.”

Este último párrafo del referido artículo posibilita que el operador del sistema (en este caso Red Eléctrica de España, S.A.) pueda limitar la potencia máxima de un parque eólico en función de la capacidad de la red para el punto de conexión. El resto del articulado, y sobre todo los P.O. 12.2 y 12.3, imponen las condiciones técnicas que tienen que cumplir las instalaciones generadoras (que afectan también a los parques eólicos), en aspectos relacionados con la seguridad de las mismas. En este sentido están regulados los valores máximos de tiempos críticos de eliminación de faltas y de potencia de cortocircuito, la normativa aplicable a la coordinación de aislamiento, los valores máximos de incrementos de tensión o la capacidad de soportar huecos de tensión.

La estructura administrativa de nuestro país queda mejor reflejada en la Orden ITC/2675/2006, de 1 de agosto, por la que se inicia el procedimiento para efectuar propuestas de desarrollo de la red de transporte de energía eléctrica, de la red básica de gas natural y de las instalaciones de almacenamiento de reservas estratégicas de productos petrolíferos (BOE de 17 de agosto de 2006), en la que establece como agentes en la planificación de la red eléctrica de transporte y referido a la integración de nuevos parques eólicos, a las autoridades autonómicas. Esta orden obliga a todos los sujetos del

sistema eléctrico, Comunidades Autónomas y promotores de nuevos proyectos de generación eléctrica, a suministrar la información necesaria para la planificación de la expansión de la red de transporte. El tratamiento dado a los nuevos parques eólicos incluye, como novedad, que en el proceso de determinación de la capacidad máxima de conexión intervengan el operador del Sistema, la Administración Central y las Comunidades Autónomas:

“La información aportada servirá de base y tendrá carácter de complementaria a la ya existente a la hora del establecimiento de los límites de la capacidad máxima de conexión de parques eólicos a la red eléctrica compatibles con la seguridad del sistema, desagregados en su caso por nudos y regiones, en cuyo proceso de determinación participarán conjuntamente con el operador del sistema la Administración General del Estado y las Comunidades Autónomas.”

2.2.2. Repercusiones en la operación de la Red Eléctrica.

Una de las características actuales más peculiares de la energía eléctrica producida en parques eólicos es su carácter no gestionable. Entendemos como carácter gestionable de la generación de energía eléctrica, a la capacidad de afrontar lo desconocido en términos de producción de potencia eléctrica. La energía eléctrica producida en parques eólicos actuales es no gestionable por las características indicadas a continuación.

- Alta vulnerabilidad ante contingencias (por ejemplo, ante los huecos de tensión).
- Régimen de generación eléctrica con alta intermitencia (en función de la evolución del viento, que puede presentar importantes cambios en intervalos cortos de tiempo).
- Imposibilidad de control de la producción eléctrica (se genera toda la energía eléctrica que el viento “aporte”, sin que sea posible aumentar esa energía eléctrica, ni sea económicamente rentable reducirla).
- Escasa fiabilidad en los valores futuros de la producción de energía eléctrica (frente a la programación de la generación eléctrica de las centrales clásicas, en los parques eólicos sólo se puede llegar a unas predicciones de potencia

eléctrica generada afectadas por un error de predicción).

Entre los problemas técnicos actuales más importantes, originados por los parques eólicos de mayor dimensión, en cuanto a la operación de la Red Eléctrica (en un contexto de una creciente integración de energía de origen eólico), se encuentran los derivados de la estabilidad del Sistema Eléctrico y la variabilidad de la energía eléctrica (de origen eólico) producida. La estabilidad del Sistema viene afectada, entre otros aspectos, por la mayor sensibilidad de los parques eólicos a los huecos de tensión, lo que hace que desconecten a la menor anomalía. La variabilidad de la potencia eléctrica producida, consecuencia del carácter no continuo del viento, obliga a compensarla (en forma generación sí gestionable) con una reserva de generación eléctrica suficiente en magnitud y en velocidad de respuesta.

Centrándonos una vez más en España, la Administración ha regulado ciertos aspectos que pretenden ofrecer soluciones a los problemas de operación de la red eléctrica. Una primera solución es la de posibilitar la gestión de los parques eólicos. Así, el RD 661/2007 de 25 de mayo (BOE del 26 de mayo de 2007) por el que se regula la actividad de producción de energía eléctrica en régimen especial, en su artículo 18 dice:

“Todas las instalaciones de régimen especial con potencia superior a 10 MW deberán estar adscritas a un centro de control de generación, que actuará como interlocutor con el operador del sistema, remitiéndole la información en tiempo real de las instalaciones y haciendo que sus instrucciones sean ejecutadas con objeto de garantizar en todo momento la fiabilidad del sistema eléctrico.”

Esta obligación de conexión a un centro de control viene desarrollada en el procedimiento de operación P.O. 3.7 de programación de la generación renovable no gestionable. En este P.O. 3.7 se faculta al operador del sistema a dar las instrucciones oportunas de modificación de la producción de las unidades de generación renovable no gestionable (caso de los parques eólicos) por medio de sus respectivos centros de control. También se identifican en el citado procedimiento cinco causas que obligan a la reducción de la producción, y son las reseñadas seguidamente.

1. *Congestión en la evacuación de la generación:* la congestión responde a la aparición de sobrecargas no admisibles en elementos de la red de transporte, como consecuencia de un exceso de generación de energía en una zona geográfica por encima de la capacidad de evacuación de la misma por la red.

2. *Estabilidad*: asociada a los huecos de tensión responsables de la desconexión de parques eólicos, facultando al operador a que determine la máxima potencia eólica que se puede integrar en el Sistema sin comprometer su seguridad.
3. *Potencia de cortocircuito*: cuando en algún nudo de la red de transporte el valor de la potencia de cortocircuito ponga en riesgo la seguridad del sistema, puede obligarse a la reducción de la producción.
4. *Viabilidad de los balances de potencia*: el operador del sistema debe garantizar la viabilidad de los balances de potencia activa y reactiva teniendo en cuenta las circunstancias singulares de operación y los límites técnicos de las unidades gestionables, para poder cubrir la demanda en los periodos horarios siguientes. Esto puede significar restricciones técnicas para las unidades no gestionables (reducciones de producción).
5. *Excedentes de generación no integrables en el sistema*: en aquellos casos en que se presente una demanda inferior a la prevista o una generación, de las unidades no gestionables, superior a la prevista, el operador puede reducir la producción de estas últimas.

La segunda solución regulatoria ofrecida por la Administración ha sido el procedimiento de operación P.O. 12.3 (BOE de 24 de octubre de 2006) sobre requisitos de respuesta frente a huecos de tensión de las instalaciones eólicas. Este procedimiento regula, fundamentalmente, la no desconexión de la Red Eléctrica de parques eólicos por huecos de tensión producidos por faltas correctamente despejadas.

Queda todavía camino por andar para resolver los problemas de operación de la Red Eléctrica derivados de la conexión de parques eólicos. Como soluciones a contemplar en el futuro inmediato se pueden citar las indicadas a continuación.

1. Mejora de la tecnología de los parques eólicos para minimizar el impacto de su conexión en la Red Eléctrica.
2. Mejora de las infraestructuras de intercambio de energía eléctrica con el exterior.

3. Planificación simultánea de la expansión de la Red Eléctrica con la de construcción de nuevos parques eólicos.

2.3. NECESIDAD DE PREDICCIÓN A CORTO PLAZO DE LA PRODUCCIÓN ELÉCTRICA EN PARQUES EÓLICOS

Los sistemas eléctricos (sistemas de energía eléctrica o sistemas eléctricos de potencia) presentan como problema característico, frente a otros sistemas de energía, el de la imposibilidad actual (desde un punto de vista práctico) de poder almacenar energía eléctrica a gran escala. Por otra parte, un sistema eléctrico fiable debe asegurar que la demanda de energía eléctrica, por parte de los usuarios, va a estar suficientemente cubierta por la generación eléctrica disponible en cada momento. Como consecuencia de esto, se hizo evidente desde los primeros momentos de puesta en funcionamiento de los sistemas de generación, transporte, reparto y distribución de energía eléctrica, la necesidad de disponer de herramientas de predicción de valores futuros de las variables que intervienen en el funcionamiento del Sistema Eléctrico.

La primera variable que fue objeto de atención fue la demanda de energía eléctrica. La previsión de la demanda de energía eléctrica se utiliza para la planificación de la energía eléctrica producida, en la planificación de las compras de combustible para las centrales eléctricas que lo precisen, para la secuenciación de los periodos de mantenimiento de los elementos del Sistema Eléctrico, para la planificación de las redes eléctricas, e incluso para la determinación de las tarifas eléctricas [Bunn85].

El significativo aumento de la potencia instalada en parques eólicos en los últimos años, con respecto a la potencia total instalada en los sistemas eléctricos nacionales en países como Dinamarca, Alemania y España, entre otros, ha obligado a desarrollar también sistemas de predicción de la potencia eléctrica generada en parques eólicos. En los primeros momentos la energía eléctrica producida en parques eólicos era vista como una demanda negativa de cara a los operadores del Sistema, pero además como una demanda prácticamente imprevisible [Hartnell00].

La dependencia que tiene la energía eléctrica generada en un parque eólico de las condiciones meteorológicas, y la alta variabilidad de éstas, hacen que el problema de la predicción de la energía generada en el mismo sea una labor compleja. La naturaleza intermitente del viento origina que, incluso para horizontes temporales pequeños, sea

difícil asegurar cual va a ser exactamente el valor de su velocidad o dirección.

Un funcionamiento seguro del sistema de energía eléctrica no sólo debe cubrir la demanda, sino que debe dar respuesta inmediata a los problemas que en ella puedan surgir, garantizando el suministro al mayor número de usuarios y, obviamente, respetar todas las restricciones técnicas que pueda presentar la Red Eléctrica. Sólo bajo ese aspecto ya es importante disponer de herramientas para la predicción de la producción eléctrica en parques eólicos, por las repercusiones económicas que un conocimiento de antemano de los valores de energía generada puede tener en la determinación de la programación de las centrales eléctricas clásicas, y en la determinación de la reserva de generación y de los cambios en el despacho de las unidades generadoras convencionales.

Pero si el operador del Sistema Eléctrico tiene la necesidad de conocer de antemano cuáles van a ser los valores de la energía eléctrica producida en los parques eólicos, que vierten su generación en la Red Eléctrica que controla, la necesidad por parte de los propietarios de los parques no es menor, y más si venden su energía en el mercado eléctrico liberalizado.

El mercado de producción de energía eléctrica persigue coordinar la libre competencia en la generación de la misma, con unos criterios preestablecidos (y regulados) de calidad y seguridad en el suministro eléctrico. Las ofertas de energía eléctrica que los agentes negocian en el mercado de producción eléctrica responden a sus previsiones de demanda eléctrica, a sus capacidades de generación y a la disponibilidad de las redes eléctricas.

El papel de un productor de energía eléctrica, actuando en el mercado liberalizado, es la de comunicar al operador del mismo, en los periodos habilitados para ello, sus ofertas de generación eléctrica “por unidad de producción” (central térmica, unidad de gestión de centrales hidroeléctricas, unidad de gestión de parques eólicos), con expresión de un precio y cantidad de energía eléctrica, para cada una de las horas correspondientes al horizonte de programación.

La Figura 2.1 representa los intervalos de tiempo de actuación en el mercado eléctrico español (fuente: Resolución de 24 de mayo de 2006 de la Secretaría general de energía, por la que se aprueban las reglas de funcionamiento del mercado diario e intradiario de producción de energía eléctrica, BOE de 30 de mayo de 2006). Hay una única sesión de mercado diario y seis sesiones de mercado intradiario.

El mercado diario tiene por objeto programar las transacciones de energía eléctrica para el día siguiente sobre una base horaria. De cara a los productores, como es el caso de los propietarios de parques eólicos, las ofertas de producción eléctrica se deben presentar ante el Operador del Mercado (OMEL) antes de las 10 horas del día anterior al correspondiente a la programación que se va a realizar. Otros agentes del mercado (comercializadores, distribuidores y consumidores cualificados) presentan sus ofertas de compra de energía eléctrica. El Operador realiza la casación de las ofertas, determinándose el precio marginal y el volumen de energía eléctrica para cada unidad de compra y venta en cada periodo horario. El resultado de este proceso de casación es el llamado programa base de casación. El programa diario definitivo se obtiene, tras el análisis por el Operador de la Red de Transporte (Red Eléctrica de España), del programa base de casación, bajo el punto de vista de la seguridad de la red, y con la resolución de las posibles restricciones técnicas, lo que puede dar lugar a una reasignación de la generación de las centrales eléctricas (unidades generadoras que deben suministrar la energía eléctrica demandada).

El mercado intradiario tiene como misión ajustar los desvíos en generación eléctrica y/o en demanda eléctrica que se puedan producir con posterioridad a haberse fijado el programa diario definitivo. El mercado intradiario consta de seis sesiones en las que pueden presentar ofertas de compra o venta de energía eléctrica aquellos agentes que hayan participado en la sesión del mercado diario. En la Figura 2.1 está representada por una flecha roja el límite de presentación de ofertas para cada una de las sesiones del mercado eléctrico español, así como el periodo de programación correspondiente a cada sesión (en color azul).

Los productores de energía eléctrica de origen eólico no tienen obligación de participar en el mercado eléctrico español. El R.D. 2818/1998 de producción de energía eléctrica por instalaciones abastecidas por recursos o fuentes de energía renovables, residuos o cogeneración, establece el régimen económico de las instalaciones acogidas a lo que denomina régimen especial (las instalaciones que únicamente utilicen como energía primaria energía eólica están incluidas como tales). Según el citado R.D., el precio por la energía eléctrica entregada por estas instalaciones es calculado sobre la base del precio medio horario, en el mercado eléctrico, más una prima de producción. El precio medio horario es calculado mes a mes y una vez transcurrido el mes que se vaya a facturar. Bajo estas condiciones económicamente ventajosas han operado todos los parques eólicos en España hasta finales de 2003.

La publicación del R.D. 436/2004 (BOE 27 de marzo de 2004) marca el inicio de la irrupción en el mercado eléctrico de los productores propietarios de parques eólicos. Este R.D. obliga al titular de la instalación, si está por debajo de los 50 MW instalados, a elegir entre dos sistemas de venta de la energía eléctrica producida. Puede optar por venderla a una empresa distribuidora de energía eléctrica, percibiendo por ello una retribución en forma de tarifa regulada, única para todos los períodos de programación, que se define como un porcentaje de la tarifa eléctrica media o de referencia regulada en el R.D. 1432/2002. O puede optar por vender la energía eléctrica producida directamente en el mercado diario, percibiendo en este caso el precio que resulte en el mercado, más un incentivo por participar en él, más un complemento por energía reactiva y una prima, si la instalación concreta tiene derecho a percibirla.

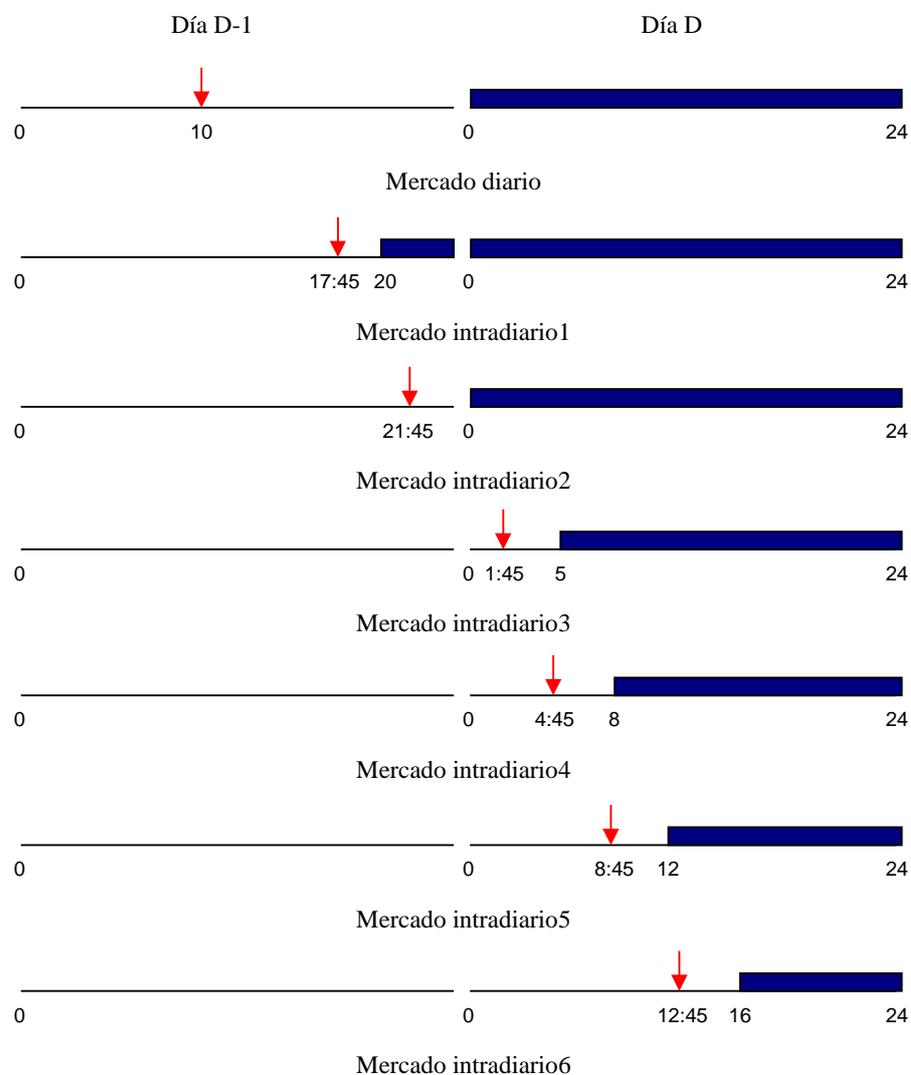


Figura 2.1. Límites temporales de entrega de ofertas de producción y horizonte de aplicación en el mercado eléctrico español.

Las ventajosas condiciones de acceso al mercado han propiciado que el porcentaje de energía eléctrica de origen eólico retribuida en forma de tarifa regulada sea, a finales de 2006, de sólo el 4.1%, cuando 3 años antes era del 100%. Obviamente el resto de la energía eléctrica de origen eólico, el 95.9%, ha sido vendida en el mercado eléctrico. Pero, tanto si el productor eólico ha optado por vender su energía producida a una empresa distribuidora, como si ha decidido acudir a realizar la venta en el mercado eléctrico, viene obligado a realizar una previsión de su generación.

- Para instalaciones con potencias instaladas superiores a 10 MW que opten por vender la energía producida a la empresa distribuidora, deben comunicar a la misma una previsión de la energía eléctrica a ceder a la red, en cada uno de los períodos de programación del mercado de producción de energía eléctrica, con una antelación mínima de 30 horas al inicio del día objeto de programación (artículo 19 del citado R.D.).
- Para instalaciones que opten por vender la energía eléctrica producida en el mercado eléctrico, deben realizar sus ofertas de venta en las mismas condiciones que el resto de productores de energía eléctrica que participan en el mismo, esto es, con expresión de un precio y cantidad de energía, para cada una de las horas correspondientes al horizonte de programación. Es decir, deben realizar unas previsiones de energía eléctrica generada para cada una de las horas que componen el periodo de programación (las 24 horas del día siguiente para el mercado diario, y entre 8 y 28 horas si participan en el mercado intradiario).

La reciente publicación del R.D. 661/2007, de 25 de mayo, por el que se regula la actividad de producción de energía eléctrica en régimen especial (BOE de 26 de mayo de 2007), mantiene las dos opciones de venta, aunque con ligeros cambios para la primera opción: ya no es necesario vender la energía eléctrica producida a una empresa distribuidora y sólo se mantiene la obligación de presentar la previsión de la energía eléctrica a ceder a la red, en cada uno de los períodos de programación del mercado de producción de energía eléctrica, con una antelación mínima de 30 horas al inicio del día objeto de programación, para las instalaciones conectadas a la red de transporte.

Así, nos encontramos con una doble necesidad de predicción de la energía eléctrica producida en los parques eólicos. Por una parte, el Operador del Sistema Eléctrico necesita conocer de antemano la potencia eléctrica que se va a inyectar en el Sistema para un periodo de programación (el más importante es el día siguiente), sobre

una base horaria; y, además, el Operador necesita conocer el nudo en el que se va a inyectar esa potencia eléctrica para poder realizar los análisis del funcionamiento previsto del Sistema. Por otra parte, los propietarios de los parques eólicos, tanto si optan por vender su energía a tarifa regulada, como si optan por realizar ofertas de venta en el mercado eléctrico, están obligados a realizar una previsión de la energía vertida en la Red Eléctrica para cada una de las 24 horas del día siguiente, o para el posterior (para la venta de la energía eléctrica generada a tarifa regulada para parques eólicos conectados a la red de transporte). Los desvíos con respecto a las previsiones tienen un coste económico según veremos en el apartado siguiente.

2.4. VALOR ECONÓMICO DE LA PREDICCIÓN

Sobre la importancia de la previsión de la generación eléctrica de origen eólico, y sus importantes repercusiones económicas sobre las compañías productoras, se pueden consultar diversos trabajos publicados en los últimos años. Estos trabajos están orientados bajo dos perspectivas: la reducción de costes de operación originada por la reducción de la reserva (consecuencia a su vez de una adecuada predicción de la producción eléctrica en parques eólicos), o la reducción de las penalizaciones económicas debida a una predicción ajustada para un productor de energía eléctrica (con parques eólicos) que realice la venta de la energía eléctrica en el mercado.

Así, en [Chan83] se recoge un primer modelo probabilístico para la determinación de la reserva rodante óptima para los siguientes minutos (hasta 30) en el caso de una compañía generadora con significativa generación de origen eólico.

[Schwartz98] expone el resultado de un estudio realizado mediante entrevistas a siete compañías eléctricas norteamericanas, propietarias de parques eólicos, en las que éstas reflejan su interés en el desarrollo de herramientas de predicción de la generación eléctrica sobre una base horaria y con un horizonte de hasta 48 horas.

[Milligan95] expone una simulación de la disminución del beneficio de una compañía generadora de energía eléctrica, con un porcentaje significativo de potencia eléctrica instalada en forma de parques eólicos, en función de los errores en la predicción de la generación eólica. Estos errores de predicción de la potencia generada en los parques eólicos se traducen en una reducción de los beneficios, como consecuencia de la necesidad de aumentar la reserva rodante o por la necesidad de comprar energía eléctrica a otras compañías.

[Watson99] presenta un trabajo parecido al anterior, pero evaluando el ahorro de combustible convencional para una compañía productora de energía eléctrica que dispusiera de predicciones de velocidad de viento para sus parques eólicos. El estudio está realizado con las predicciones de dos modelos: el modelo persistente y un modelo de predicción numérica meteorológica (modelo atmosférico). El gasto de combustible, con ambos modelos, es comparado con el gasto realizado en el caso de conocer de antemano y con exactitud la generación en el parque eólico. Los gastos realizados con las predicciones del modelo atmosférico son inferiores a los gastos con las predicciones del modelo persistente, lo que justifica económicamente el uso de modelos atmosféricos en la predicción de la energía eléctrica generada en parques eólicos.

Más recientemente [Usaola04] presenta un caso referido al mercado eléctrico español, utilizando el modelo de predicción Sipreólico (que será comentado en el Capítulo 3), en la predicción de la potencia eléctrica media horaria de un parque eólico, con la que se presenta la oferta de venta de la energía eléctrica producida al mercado eléctrico. La reducción de los ingresos debida a desvíos en la producción ofertada, como consecuencia de los errores del modelo de predicción, son del 7.5% del total de ingresos; pero esa reducción es del 9.5% si las predicciones son realizadas con el modelo persistente, o del 10% si no se realiza ninguna predicción.

Para el caso español, la legislación aplicable contempla la penalización por desvío en la producción de energía eléctrica entre el valor programado y el valor real. Estas penalizaciones son distintas en función de la opción seleccionada por el productor. Seguidamente comentamos las opciones de venta.

- Opción de venta de la energía producida a una empresa distribuidora: a estas instalaciones se les repercutirá un coste de desvío por cada periodo de programación en el que la producción se desvíe más allá de la tolerancia permitida respecto a su previsión. El desvío en cada uno de estos periodos de programación se calculará como el valor absoluto de la diferencia entre la previsión de generación y la medida correspondiente.

Hasta el 31 de mayo de 2007 la tolerancia permitida era del 20% con respecto a la previsión, y el coste de los desvíos de cada mes era un 10 por ciento del resultado de multiplicar la tarifa eléctrica media o de referencia de cada año (definida en el artículo 2 del Real Decreto 1432/2002, de 27 de diciembre, y publicada en el real decreto por el que se establece la tarifa

eléctrica), por la suma de todos los desvíos de dicho mes que hubieran excedido esa tolerancia establecida.

Desde el 1 de junio de 2007 y hasta la entrada en vigor de la figura del comercializador de último recurso (prevista para el 1 de enero de 2009), y siempre que la empresa distribuidora actúe como representante del parque eólico, la tolerancia permitida es del 5% con respecto a la previsión. A los desvíos (diferencia absoluta entre la energía prevista y la real) que superen la tolerancia permitida, hora a hora, se les aplica el coste de desvío fijado en el mercado eléctrico organizado por cada período de programación.

- Opción de venta en el mercado eléctrico: la penalización económica de los desvíos entre la producción prevista y la real se rige bajo las reglas del mercado eléctrico. Los ingresos de la unidad de producción de energía eléctrica para un día, I , se pueden calcular por medio de la ecuación (2.3) y el coste o ingreso adicional por el desvío (coste si es negativo, ingreso si es positivo) por medio de la ecuación (2.4).

$$I = \sum_{t=1}^{24} P_{d,t} \cdot PM_t + \sum_{t=1}^{24} PM_{i,t} (P_{i,t} - P_{d,t}) + \text{desvío} \quad (2.3)$$

$$\text{desvío} = \begin{cases} + PM_t^{\text{venta}} (P_{gen,t} - P_{ult,t}) & \text{si } P_{gen,t} > P_{ult,t} \\ - PM_t^{\text{compra}} (P_{ult,t} - P_{gen,t}) & \text{si } P_{gen,t} < P_{ult,t} \end{cases} \quad (2.4)$$

En estas dos últimas ecuaciones $P_{d,t}$ representa la potencia eléctrica horaria programada en el mercado para esa unidad (parque eólico o conjunto de los mismos inyectando potencia eléctrica horaria en el mismo nudo) para la hora t ; PM_t representa el precio marginal del mercado para la potencia eléctrica horaria en la hora t y fijada en el mercado diario; $PM_{i,t}$ representa el precio marginal de la potencia eléctrica horaria fijado en el mercado intradiario para la hora t ; $P_{gen,t}$ representa la potencia eléctrica horaria realmente entregada por la unidad generadora a la Red Eléctrica en la hora t ; PM_t^{venta} representa el precio de mercado para la venta de la potencia eléctrica horaria en la hora t ; PM_t^{compra} representa el precio de mercado para la compra de la potencia eléctrica horaria en la hora t ; $P_{ult,t}$ representa el último valor de la potencia eléctrica horaria programada para la unidad de producción en la última

sesión del mercado (diario o intradiario) para la hora t .

Así, tanto en el caso de venta de la energía eléctrica producida a una empresa distribuidora, como en el caso de venta de energía en el mercado, existe una penalización en el caso de que la energía programada (la prevista por el productor) no coincida con la realmente vertida en la red eléctrica. Además, todos los desvíos producidos, independientemente de su signo, deberán “pagar” los sobrecostos de la energía eléctrica de regulación secundaria y terciaria y, en su caso, de la parte del procedimiento de gestión de desvíos correspondientes, en proporción al valor absoluto de los desvíos de potencia eléctrica horaria.

Estas penalizaciones por desvíos entre la potencia eléctrica horaria programada y la finalmente generada han propiciado la elaboración de ofertas conjuntas de producción hidroeléctrica y eólica [Márquez07] para maximizar los ingresos, cuando se opta por vender la energía eléctrica producida en el mercado (la energía hidroeléctrica cubre los posibles desvíos en la producción del parque eólico). También se han propuesto estrategias para minimización de penalizaciones, independientemente de las previsiones de generación en el parque eólico, sobre la base de predicción de un modelo persistente [Bathurst02].

2.5. HORIZONTES DE PREDICCIÓN

El desarrollo de cualquier modelo de predicción debe contemplar, desde el primer momento de diseño, el horizonte de predicción, esto es, el periodo de tiempo (contabilizado desde un instante determinado en que realizamos la predicción), que determina el momento futuro para el que realizamos las predicciones. En otros problemas de predicción relacionados con el Sistema Eléctrico, como es el caso de la predicción de la demanda eléctrica, según el valor del horizonte la predicción puede clasificarse como de muy corto plazo, corto plazo, medio plazo o largo plazo.

Para el caso de la predicción de la potencia eléctrica producida en un parque eólico no existe unanimidad en cuanto cuales son las fronteras que separan una clasificación de otra; es más, la mayoría de los autores de artículos relacionados con el tema rehuyen cuantificar los términos y los que si lo hacen dan valores diferentes, en especial para la predicción a muy corto plazo. El horizonte de predicción y el interés de la predicción para cada grupo es el indicado a continuación.

- *Predicciones a muy corto plazo*: el horizonte de predicción es de unas horas, no existiendo unanimidad en su valor. Así, como valores extremos [Usaola04] propone 4 horas y [Möhrlen04] propone 9 horas. Para un productor las predicciones en este periodo son de interés con objeto de poder actuar en el mercado intradiario. Para el operador de la Red Eléctrica estas predicciones son de interés para el mantenimiento de la seguridad del Sistema Eléctrico. Si se reduce el horizonte de predicción, también disminuyen los errores de predicción.
- *Predicciones a corto plazo*: el horizonte de predicción abarca desde el límite con el muy corto plazo hasta las 48 ó 72 horas. En la mayoría de los trabajos publicados se limita a 48 horas, o incluso se limita a 36 horas. Esta predicción es la de mayor interés para poder actuar en el mercado diario eléctrico. Para el caso español, dado que el plazo de presentación de ofertas de venta en el mercado diario se cierra a las 10 horas, un horizonte de predicción de 38 horas cubriría todo el día siguiente. En otros países el momento de cierre de recepción de ofertas puede diferir del caso español, con lo que el número de horas de horizonte para cubrir las predicciones del día siguiente puede variar. Así, en el Capítulo 3 se citan modelos, descritos en publicaciones internacionales, con horizontes de predicción de 36 horas (han sido desarrollados para países en los que el plazo de recepción de ofertas de venta en el mercado se cierra a las 12 horas, como es el caso de Dinamarca).

Las predicciones de producción eléctrica correspondientes no al día siguiente, sino al posterior, actualmente carecen de interés económico en relación al mercado, ya que no se pueden realizar ofertas en el mercado eléctrico español. Sin embargo, son estas las predicciones necesarias para aquellos parques eólicos que no venden la energía producida en el mercado eléctrico. Para los parques eólicos que venden la energía producida en el mercado, pueden resultar también interesantes, para poder programar labores de su mantenimiento durante las horas para las que las predicciones apuntan un valor muy bajo de potencia eléctrica generada.

- *Predicciones a “más largo plazo”*: corresponden a un horizonte por encima de las 48 ó 72 horas llegando a un límite de 7 días. Cuanto mayor sea el horizonte, mayores son los errores de predicción, resultando a día de hoy difícil realizar predicciones fiables de variables atmosféricas por encima de

esos 7 días. Estas predicciones (hasta 7 días) pueden ser de interés para la previsión del despacho de centrales eléctricas convencionales, y/o la programación de las labores de mantenimiento en algunas de esas centrales (en la programación del sistema de generación): en periodos con previsiones de alto régimen de viento durante varios días, los parques eólicos pueden suministrar unos valores de potencia eléctrica media cercanos a su potencia instalada, posibilitando que otro tipo de centrales puedan reducir su generación eléctrica o programar labores de mantenimiento que podrían afectar a su producción normal.

En todo caso, dada la importante trascendencia económica del tema, la mayor parte de los trabajos publicados describiendo modelos de predicción se han centrado en la predicción de la producción eléctrica en un corto plazo, considerado éste con un horizonte de predicción entre 36 y 48 horas.

2.6. INCERTIDUMBRE EN LA PREDICCIÓN

Los modelos usados para las predicciones de la producción eléctrica a corto plazo (hasta las 48 ó 72 horas de horizonte de predicción) pueden usar, como entradas a los mismos, las predicciones (meteorológicas) generadas por otros modelos, los modelos atmosféricos. Estos son, generalmente, propiedad de los institutos meteorológicos nacionales, que proporcionan las predicciones de los valores numéricos de las variables meteorológicas: velocidad y dirección del viento, temperatura, presión atmosférica, etc. A partir de estos valores numéricos, fruto de la predicción de los modelos atmosféricos, se calcula la predicción de la potencia eléctrica generada en un parque eólico por aproximaciones estadísticas o físicas (utilizando modelizaciones de leyes físicas para transformar las predicciones de velocidad y dirección del viento en potencia eléctrica generada). Pero la predicción de la producción eléctrica que realizan estos últimos modelos tiene una gran limitación: los modelos de predicción de la producción eléctrica, en principio, no pueden mejorar las predicciones (meteorológicas) de los modelos atmosféricos.

Los modelos atmosféricos, atendidos por los institutos meteorológicos, debido a los altos requerimientos computacionales y a la inmediatez relativa de obtención de resultados, no han sido diseñados para realizar unas predicciones precisas del viento en la superficie del terreno, y menos aún para el rango de velocidades entre 5 y 15 m/s (no son vientos ni muy flojos ni muy fuertes). Para la predicción de la potencia eléctrica

generada en un parque eólico la variable meteorológica más importante es el viento. Y es, justamente, en el rango de velocidades del viento antes citado, donde actualmente aparece la mayor variabilidad en la potencia eléctrica generada en un aerogenerador.

La Figura 2.2 presenta la curva de potencia (potencia generada en función de la velocidad del viento, con condiciones de viento constante) para un aerogenerador de última generación. Como se puede apreciar la mayor variación de potencia eléctrica generada se da en el rango de velocidades de viento de 4 a 12 m/s, aproximadamente. Por ejemplo, para una velocidad del viento de 10 m/s, la potencia generada es del orden de 8 veces más que para una velocidad del viento de la mitad de ese valor, es decir, 5 m/s. Un ligero error en la predicción de la velocidad del viento por parte del modelo atmosférico puede significar un alto error de predicción en la potencia eléctrica generada.

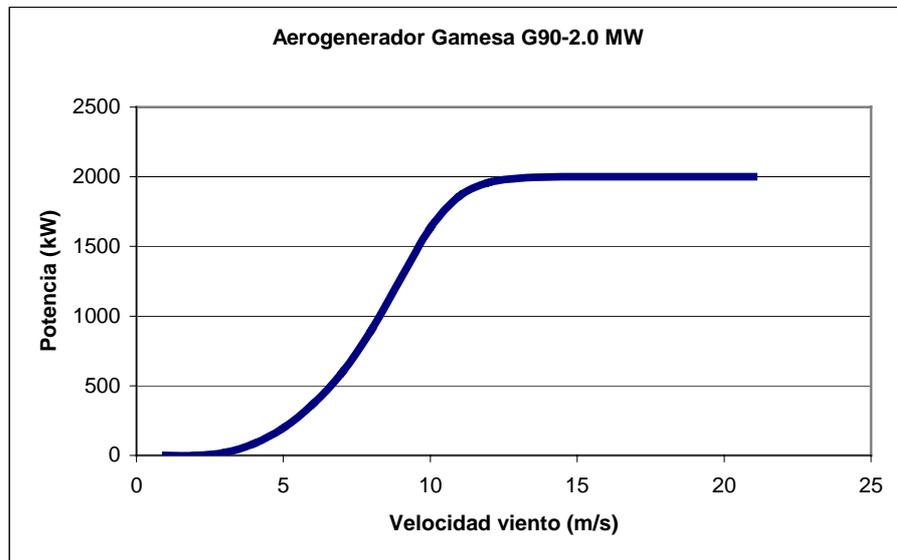


Figura 2.2. Curva de potencia del aerogenerador G90-2.0 MW fabricado por Gamesa Corporación Tecnológica.

Las fuentes de error en los modelos atmosféricos pueden ser varias: desde un deficiente modelado de la superficie del terreno, hasta condiciones iniciales incorrectas, consecuencia de errores en el proceso de medida o interpretación de las variables meteorológicas que definen el estado de partida de la atmósfera. El primer tipo de error, el causado por un deficiente modelado de la superficie del terreno, puede ser corregido estadísticamente: sólo es un sesgo en el valor de la predicción de la velocidad del viento (aunque con un valor dependiente de la dirección). El segundo tipo de error (condiciones iniciales incorrectas) no puede ser corregido de antemano y es responsable de diferencias importantes entre los valores de predicción y los reales, tanto en

magnitud como en el instante de tiempo en que ocurren (el llamado “error de fase”, que se refiere a las variaciones o cambios previstos en el valor de la variable meteorológica que ocurren antes o después de lo previsto).

Con todo, la predicción de la potencia eléctrica generada en el corto plazo por un parque eólico constituye una tarea difícil, no exenta de importantes errores. Pero a pesar de su dificultad, resulta ser actualmente imprescindible, tanto desde un punto de vista técnico como económico. Precisamente para suministrar a los usuarios (con interés económico en las mismas) estas predicciones de producción eléctrica, una información valiosa acerca del grado de incertidumbre o fiabilidad que tienen dichas predicciones, se han elaborado en los últimos cinco años modelos para la medida de la incertidumbre asociada a la de la predicción de la potencia eléctrica generada en parques eólicos.

3. REVISIÓN BIBLIOGRÁFICA DE LOS MODELOS DE PREDICCIÓN A CORTO PLAZO DE LA PRODUCCIÓN ELÉCTRICA DE ORIGEN EÓLICO

Este capítulo presenta el estado actual de los modelos de predicción de producción eléctrica, a corto plazo, en parques eólicos que han sido publicados en la literatura internacional. Para cada uno de ellos se describen las características más relevantes y sus limitaciones, así como los resultados de predicciones de producción eléctrica (si así lo han publicado los autores). Se ha incluido, además, un apartado de criterios de evaluación de modelos de predicción, señalando los indicadores estadísticos escogidos para la evaluación de los modelos propuestos en la literatura especializada.

3.1. EVALUACIÓN DE MODELOS DE PREDICCIÓN. CRITERIOS

Como se ha indicado anteriormente, en este apartado se comentan los indicadores estadísticos más usuales en la evaluación de modelos de predicción de producción eléctrica en parques eólicos. Para cualquier tarea de predicción con datos cuantitativos son varias las etapas que se deben seguir ([Makridakis98], [DeLurgio98]). De forma resumida estas etapas son las indicadas a continuación.

1. *Definición del problema.*

Incluye la comprensión de cómo se usarán las predicciones, para qué se usarán, y quién las usará. En definitiva, en este primer paso se pretende recabar la máxima información de los usuarios, suministradores de datos y expertos en el tema, para poder ofrecer a los usuarios las predicciones con las características que mejor se adapten a sus necesidades.

2. *Recogida de información.*

Es necesario recoger datos históricos de las variables de interés, incluyendo la variable a predecir, así como de todas las posibles variables explicativas accesibles. La información puede ser cuantitativa (datos numéricos) o cualitativa (opiniones de expertos).

3. *Análisis preliminar.*

El estudio de los valores estadísticos descriptivos (media, varianza,

correlaciones, etc.) puede ayudar en la selección de los modelos más adecuados, reconocer outliers, seleccionar las variables explicativas de mayor peso, etc.

4. *Elección y ajuste de modelos.*

Un modelo de predicción esencialmente consiste en un sistema de ecuaciones (lineales o no) que relacionan los valores futuros de la variable a predecir con los valores presentes y pasados de la propia variable y de las variables explicativas. El ajuste del modelo se basa en la adecuación de las características o parámetros del modelo, de forma que se minimizan los errores de predicción para el conjunto de datos con los que se ha construido. Este error será tanto menor cuanto mejor se haya realizado la elección del modelo, es decir, cuanto más idóneo sea el modelo a la hora de representar las relaciones reales entre la variable a predecir y las variables explicativas.

5. *Evaluación de modelos.*

El funcionamiento de un modelo debe ser evaluado, obviamente no con los datos utilizados en su ajuste (para algunos modelos, como los basados en las redes neuronales artificiales, con los datos usados en su aprendizaje o entrenamiento), sino con otro conjunto de valores (valores de testeo) de las variables explicativas no utilizado en el ajuste del modelo. Esta prueba permitirá la comparación de los resultados (valores de predicción) entre diferentes modelos y la selección del modelo que arroje los menores errores de predicción.

6. *Uso del modelo.*

Una vez comprobado su funcionamiento, el modelo puede ser usado para realizar las predicciones.

La quinta etapa es la más importante cuando tratamos de seleccionar definitivamente un modelo de predicción entre varios posibles. Se puede facilitar esta selección de modelos mediante un análisis de los valores alcanzados por unos indicadores estadísticos sobre las predicciones del modelo, estudiando así la bondad de las mismas. Son muchos los indicadores que se han propuesto en la literatura especializada para la evaluación de modelos de predicción. En los siguientes párrafos se recogen los indicadores estadísticos más relevantes.

Se define el error de predicción, ecuación (3.1), como la diferencia entre el valor real y el valor de predicción.

$$e(t) = y(t) - \hat{y}(t) \quad (3.1)$$

En la ecuación (3.1), $e(t)$ representa el error de predicción para el instante t , $y(t)$ el verdadero valor de la variable, e $\hat{y}(t)$ el valor previsto para esa variable en el instante t .

Cuando se dispone de varios valores de predicción para la variable en el instante t , $\hat{y}(t)$, correspondientes a previsiones realizadas con diferentes horizontes de predicción, esto es, en diferentes instantes de tiempo, el error de predicción para la predicción realizada hace k unidades de tiempo viene expresado en la ecuación (3.2).

$$e(t+k|t) = y(t+k) - \hat{y}(t+k|t) \quad (3.2)$$

En la ecuación (3.2), el término $e(t+k|t)$ corresponde al error de la predicción realizada en t para el instante $t+k$.

El error de predicción puede descomponerse en dos componentes: el error sistemático, $e_s(t)$, y el error aleatorio, $e_a(t)$, tal y como indica la ecuación (3.3). El error sistemático en un proceso puede estar originado por medidas erróneas (por ejemplo anemómetros mal calibrados) o por factores humanos. En algunos casos, el error sistemático es considerado como si fuese una constante, siempre que no se produzca un cambio en los procesos de toma de datos o medidas.

$$e(t) = e_s(t) + e_a(t) \quad (3.3)$$

El error sistemático se conoce también como sesgo. El sesgo de un modelo se puede determinar calculando el error medio para conjunto de datos de evaluación (los datos correspondientes a un periodo de tiempo). Cuando el modelo de predicción permite varios horizontes de predicción distintos, puede calcularse el sesgo para cada uno de los horizontes de predicción según la ecuación (3.4).

$$\text{sesgo}(k) = \frac{1}{N} \sum_{t=1}^N e(t+k|t) \quad (3.4)$$

En la ecuación (3.4) N representa el número total de datos del conjunto de predicciones para el horizonte k .

Se define el error medio, ME (mean error), como el valor medio del error de predicción. Se calcula, para cada horizonte de predicción, por medio de la ecuación (3.5).

$$ME(k) = \frac{1}{N} \sum_{t=1}^N e(t+k|t) \quad (3.5)$$

Los indicadores estadísticos más usuales para evaluar el funcionamiento de un modelo de predicción son el error absoluto medio, MAE, y la raíz del error cuadrático medio, RMSE. En la expresión del error absoluto medio intervienen las dos componentes que hemos definido para el error (en el caso del sesgo únicamente interviene el error sistemático, al suponer que el error aleatorio presenta una distribución estadística normal con valor medio nulo). El indicador MAE se calcula según la ecuación (3.6).

$$MAE(k) = \frac{1}{N} \sum_{t=1}^N |e(t+k|t)| \quad (3.6)$$

El indicador RMSE se calcula, tal como expresa la ecuación (3.7), como la raíz del error medio cuadrático (raíz del valor medio del cuadrado del error). El indicador RMSE tiene las mismas unidades que la serie temporal que predecimos: por ejemplo, en la predicción de la velocidad media del viento, el indicador RMSE tiene unidades de velocidad (las mismas que presente la serie utilizada en la predicción).

$$RMSE(k) = \sqrt{\frac{1}{N} \sum_{t=1}^N (e(t+k|t))^2} \quad (3.7)$$

En la predicción de la potencia eléctrica generada en un parque eólico, tanto el indicador MAE como el indicador RMSE se suelen referir a la máxima potencia que el parque eólico puede producir (potencia instalada). En ese caso, se denominan estos errores como errores normalizados. Así el indicador MAE normalizado se obtiene mediante la ecuación (3.8), y el indicador RMSE normalizado según la ecuación (3.9), siendo P_{inst} la potencia instalada del parque eólico.

$$MAE_n(k) = \frac{1}{P_{inst}} \frac{1}{N} \sum_{t=1}^N |e(t+k|t)| \quad (3.8)$$

$$RMSE_n(k) = \frac{1}{P_{inst}} \sqrt{\frac{1}{N} \sum_{t=1}^N (e(t+k|t))^2} \quad (3.9)$$

En muchos casos para cuantificar la mejora en el error de predicción que presenta un modelo con respecto a otro, que se toma como referencia, se realiza una comparación relativa entre los valores de alguno de los indicadores de error definidos

anteriormente. De esta forma, se define dicha mejora mediante la ecuación (3.10).

$$\text{mejora}_{\text{ref,CE}}(k) = \frac{\text{CE}_{\text{ref}}(k) - \text{CE}(k)}{\text{CE}_{\text{ref}}(k)} \quad (3.10)$$

En la ecuación (3.10), $\text{mejora}_{\text{ref,CE}}(k)$ representa la reducción porcentual del indicador o criterio de error escogido, CE, en el modelo evaluado, con respecto al mismo valor del indicador en el modelo de referencia. Los criterios de error escogidos suelen ser el MAE y el RMSE, especialmente este último.

En la literatura especializada sobre técnicas de predicción de series temporales ([Wei90], [Makridakis98], [DeLurgio98], [Armstrong01]) pueden encontrarse otros criterios o indicadores de error, como pueden ser el coeficiente de correlación, la desviación estándar del error, el criterio de información de Akaike, criterio de información Bayesiano de Schwarz, etc., que han sido utilizadas puntualmente en alguno de los trabajos publicados. El criterio más utilizado es el indicador RMSE normalizado a la potencia instalada en el parque eólico.

En [Madsen05], los autores sugieren utilizar como criterios de medida del error los siguientes: sesgo, MAE y RMSE, todos ellos normalizados (divididos por la potencia instalada en el parque eólico). Además sugieren también utilizar el indicador “mejora” para comparar la eficiencia de los modelos.

3.2. FAMILIAS DE MODELOS DE PREDICCIÓN DE LA PRODUCCIÓN ELÉCTRICA DE ORIGEN EÓLICO EN EL CORTO PLAZO

La necesidad, por parte de generadores y operadores de los sistemas eléctricos, de disponer de herramientas de predicción de la energía eléctrica producida de origen eólico que proporcionen el mínimo error posible, se recoge ya en los primeros trabajos publicados. La necesidad de modelos de predicción de la generación se justifica, en estos primeros trabajos, sobre los beneficios económicos que aporta un sistema de predicción, especialmente para los casos de una importante penetración de los sistemas de generación eléctrica de origen eólico. Se entiende esta importante penetración como un porcentaje significativo de potencia instalada en parques eólicos frente a centrales eléctricas convencionales.

El primer modelo auténtico de predicción de velocidad del viento, específico para ser usado en la estimación de la producción de energía eléctrica, es el presentado

en [Bossanyi85], donde se propone el uso de un filtro Kalman para predicciones de la velocidad media del viento para los próximos minutos (hasta una hora). Los resultados son relativamente buenos para horizontes de predicción por debajo de los 10 minutos, pero ya no tanto para un horizonte de una hora, donde claramente el autor apunta que el método no resulta ser el más adecuado para la predicción de la velocidad media horaria.

Tras estos primeros trabajos publicados, son dos los tipos de modelos de predicción propuestos en la literatura internacional. Por una parte se desarrollan modelos que utilizan como datos de entrada valores pasados, bien de la propia serie temporal de la variable a predecir (velocidad del viento o energía producida), o bien de otras variables explicativas que pueden contribuir a reducir los errores de predicción. Denominaremos a estos modelos como pertenecientes a la familia de series temporales. Frente a estos, se desarrollan otros modelos que incorporan como entradas, además de las comentadas para modelos de series temporales, los valores de predicción para variables meteorológicas obtenidos por los programas de predicción numérica existentes en institutos meteorológicos (modelos atmosféricos).

Algunos de los trabajos publicados se limitan a la predicción de la velocidad media horaria del viento. En otros trabajos se ofrece también la predicción de la potencia eléctrica media horaria producida. El cálculo de esta última se realiza por medio de la relación entre potencia eléctrica producida y velocidad del viento, obtenida de forma empírica para el propio parque para el que se realiza la predicción, o utilizando la ofrecida por el fabricante del aerogenerador (curva de potencia del fabricante).

En la literatura internacional pueden encontrarse un par de interesantes revisiones bibliográficas sobre el tema. Una buena fuente sobre estado actual de los modelos de previsión a corto plazo la constituyen [Landberg03] y [Giebel03].

3.2.1. Modelos de predicción basados en series temporales de viento o en series de energía eléctrica generada.

Los modelos basados en la metodología de Box-Jenkins (conocidos también como modelos ARIMA) [Box76], fueron los primeros en ser utilizados, y aún siguen usándose si bien con horizontes de predicción limitados a unas pocas horas. Así, en [Contaxis91] se utiliza un modelo auto-regresivo (concretamente un modelo AR(3)) para la predicción de la velocidad media del viento, utilizando este valor de predicción para el control de un sistema generador diesel/eólico en un sistema aislado y en la programación a muy corto plazo de la unidad que va a suministrar energía a la carga del

sistema (también utiliza un modelo auto-regresivo de media móvil, ARMA, para la predicción de la demanda). El horizonte de predicción es de 30 minutos y llega a encadenar 10 periodos de tiempo, alcanzando así un horizonte total de predicción de 5 horas. Sólo realiza una prueba de predicción del modelo para 5 horas, demostrando la viabilidad del modelo para la programación de unidades generadoras en el muy corto plazo.

Otros modelos de tipo ARIMA para la predicción de la velocidad del viento pueden encontrarse en [Kamal97], con el desarrollo de un modelo ARMA para la predicción de la velocidad media horaria, incluyendo el cálculo de intervalos de confianza para la predicción; en [Schlink98], con una aplicación para la predicción de la velocidad media del viento para los próximos 10 minutos en un aeropuerto; en [Poggi03], donde se desarrollan modelos auto-regresivos para cada mes con intención de predecir la velocidad media horaria para las próximas 3 horas; en [Torres05], donde se describen cinco modelos ARMA para la predicción de la velocidad media horaria del viento en otros tantos lugares situados en Navarra, con un horizonte de predicción de 10 horas, y obteniendo una reducción en el error de predicción de hasta un 20% con respecto al modelo persistente.

La evaluación de los resultados obtenidos con modelos de predicción de potencia eléctrica generada es tratada en [Nielsen98], donde concretamente se propone un modelo alternativo al clásico de persistencia, con horizontes de hasta 48 horas, como modelo de referencia. Este modelo, que también puede ser utilizado para realizar predicciones, se comporta como el modelo persistente con horizontes de predicción muy cortos, mientras que para horizontes mayores (varias horas) su salida se va adaptando progresivamente al valor medio de la serie temporal.

La previsión de la generación de energía eléctrica a corto plazo en sistemas aislados, con especial interés en la predicción de la energía de origen eólico, vuelve a ser tratada en [Kariniotakis96]. En este caso se propone la predicción de la producción eléctrica en parques eólicos por medio del uso de modelos basados en redes neuronales artificiales. Esta técnica de predicción había sido ya utilizada en la predicción a corto plazo de otras variables de interés para los sistemas eléctricos, fundamentalmente en la predicción de la demanda de energía eléctrica ([Hsu91], [Park91], [Lee92], [Ho92], [Peng92], por citar sólo los primeros trabajos). En [Kariniotakis96], los autores proponen tres tipos de redes neuronales recurrentes para la predicción de la potencia eléctrica media producida en un parque eólico: una primera red neuronal con 12 salidas y que realiza la predicción de la potencia eléctrica media generada en las dos próximas

horas en intervalos de 10 minutos; una segunda red neuronal con una salida (predicción a 10 minutos) pero que se utiliza recursivamente para completar la predicción hasta el horizonte de 2 horas, con selección óptima del número de entradas, considerando las correspondientes 12 etapas del proceso recursivo; y una tercera red neuronal similar a la anterior, pero en la que la selección del número de entradas se ha realizado teniendo en cuenta exclusivamente el error de la primera etapa. Los resultados muestran que para horizontes de predicción pequeños (menos de 40 minutos) el comportamiento de la primera red neuronal es peor que el de las otras dos (cuyos comportamientos son casi similares); mientras que con un horizonte de predicción superior a esos 40 minutos, la red neuronal con 12 salidas tiene un mejor comportamiento al ofrecer errores de predicción menores.

Los modelos de predicción basados en redes neuronales, en la forma de perceptrones multicapa (MLPs), empiezan a mostrar mejores resultados frente a los modelos clásicos ARIMA en la predicción de valores medios diarios y mensuales de la velocidad del viento. A este resultado se llega en [Mohandes98] con su estudio de una serie temporal de velocidades medias de viento en la ciudad de Jeddah (Arabia Saudí), al comparar las predicciones obtenidas con un modelo ARMA y las obtenidas con una red neuronal (en concreto, una red neuronal MLP con una única entrada: la velocidad media en el día anterior).

La disponibilidad de muchos puntos de medida de la velocidad del viento propicia la realización de trabajos en los que se incluye el estudio de la posible correlación entre los valores medidos de velocidad del viento en diferentes estaciones meteorológicas. [Alexiadis98] propone un modelo basado en una red neuronal para la predicción de la velocidad del viento en la isla de Syros (Grecia) utilizando como entradas, además de los últimos valores conocidos de la serie temporal en la propia isla, los valores medidos en dos islas próximas. Los resultados muestran una mejora máxima del 32% con respecto al modelo persistente con un horizonte de predicción de una hora. El mismo método lo vuelve a aplicar en la bahía de Tesalónica (Grecia) [Alexiadis99] con una mejora de hasta el 27%, con respecto al modelo persistente, para un horizonte de predicción de 2 horas.

Las bases del modelo de correlación de velocidades del viento se encuentra muy detalladas en [Salmon99], donde se realiza un estudio intensivo de la correlación existente entre las series temporales de velocidad y dirección del viento en cinco estaciones meteorológicas de Canadá, si bien su estudio no está orientado a la predicción, sino al estudio estadístico de esta correlación. La principal aplicación de

esta correlación entre velocidades de viento se encuentra en la estimación de los recursos eólicos de un lugar, conocida una serie temporal limitada de velocidad del viento medido en ese lugar, pero disponiendo de una serie mucho mayor para otro lugar relativamente cercano. Así, en [Joensen99] se realiza una estimación de los parámetros de la distribución estadística de la velocidad media del viento de un lugar a otro; y en [Denison01] se realiza una estimación del número medio de días de “calma” (días con velocidades medias horarias por debajo de un umbral). Una aproximación completamente diferente se presenta en [Hunt01], donde se realiza la predicción de la velocidad media del viento en un punto, conocido su valor en otro punto cercano, por medio de la aplicación de wavelets. Para el desarrollo del modelo de predicción estudia la contribución de diferentes componentes de la descomposición wavelet para cada uno de los 12 sectores en que divide el espacio, de acuerdo a la dirección del viento. Los resultados que presenta, para un horizonte de predicción de 10 minutos, son mejores que los obtenidos con un modelo de regresión lineal.

El primer trabajo con una comparación amplia entre diferente modelos, para la predicción de la velocidad media del viento, aparece en [Sfetsos00], que compara el modelo persistente, modelos ARIMA, modelos basados en redes neuronales y modelos basados en sistemas neuro-difusos. Excepto para los modelos ARIMA, en los demás modelos considera como entradas los últimos seis valores conocidos de la velocidad media horaria del viento. Es precisamente uno de los modelos basados en redes neuronales el que obtiene los mejores resultados, aunque sólo consigue un 4.19% de mejora sobre el modelo persistente para la predicción de la velocidad media del viento con un horizonte de una hora. La serie temporal que utiliza está obtenida en la isla de Creta y corresponde a una serie de velocidades medias horarias. En un trabajo previo [Sfetsos99], aunque más limitado en el número de modelos a comparar, se obtuvieron los mejores resultados con una red neuronal del tipo MLP, incluyendo otras variables de entrada adicionales: radiación solar, temperatura y presión atmosférica.

Como consecuencia del trabajo desarrollado en un proyecto de investigación europeo (el llamado MORE CARE), [Hatzigiorgiou01] presenta los componentes principales de este proyecto que pretende desarrollar un sistema de control para redes eléctricas aisladas o con interconexión débil y con alta penetración de energía eólica. Para la predicción de la generación eólica utilizan dos modelos distintos: el primero para la predicción a corto plazo (10 horas en concreto) basado en un sistema difuso auto-regresivo, que utiliza como variable exógena para la predicción de la velocidad la propia dirección del viento. Justifica este modelo como el que mejores resultados ha ofrecido en comparación con el modelo persistente. El segundo modelo utiliza el

resultado obtenido por un modelo de predicción numérica atmosférica (perteneciente, por tanto, a la segunda familia de modelos que será comentada en las próximas páginas). Este segundo modelo permite extender el horizonte de predicción hasta las 48 horas. Para la transformación de velocidad del viento a potencia eléctrica utilizan la curva de potencia del aerogenerador, con correcciones en función de la dirección para tener en cuenta el posible efecto de “sombra” que puede aparecer entre aerogeneradores próximos.

Los modelos de correlación, ya presentados en años anteriores, vuelven con una nueva aplicación en un modelo de predicción de velocidad del viento de tipo difuso en [Damousis01] y en [Damousis04]. Utilizan un algoritmo genético (GA) para la optimización de los parámetros de un sistema de inferencia difusa de tipo Takagi-Sugeno, con una participación del espacio de entrada (valores previos de la velocidad del viento en el propio lugar de prueba y en otras estaciones cercanas) en rejilla. Este modelo llega a mejorar los resultados obtenidos con el modelo persistente entre un 9.5% y un 28.4% dependiendo del horizonte de predicción (mayor mejora cuanto mayor es el horizonte de predicción). El horizonte de predicción está comprendido entre 30 y 240 minutos. Realiza las pruebas en dos zonas, una llana (golfo de Tesalónica en la Grecia continental) y otra montañosa (la región cercana a Sitia, en la isla de Creta).

Las dos primeras publicaciones derivadas de los trabajos de investigación realizados en esta tesis, [Ramirez01] y [Ramirez04a], corresponden también a un modelo de esta familia de series temporales. En ellas se utiliza un sofisticado modelo de tres etapas: en la primera etapa se calculan los coeficientes de la transformada de Fourier correspondientes a los 24 últimos datos conocidos de velocidad media horaria del viento; en la segunda etapa, mediante 23 sistemas de inferencia difusa del tipo Takagi-Sugeno (TSK) se realiza la predicción de los coeficientes para la transformada de Fourier de la velocidad media del viento para la próxima hora; y en la última etapa se calcula la velocidad media del viento para la próxima hora a partir de las predicciones de los coeficientes obtenidas en la etapa anterior. Los resultados mejoran en poco más de un 9% al modelo persistente y superan a los de otros modelos basados en redes neuronales, ARIMA y sistemas de inferencia difusa.

Un nuevo trabajo de Sfetsos, [Sfetsos02], presenta dos aproximaciones distintas a la predicción de la velocidad del viento con un horizonte de predicción de una hora. En la primera desarrolla un modelo basado en redes neuronales, que utiliza como entradas, los últimos valores conocidos de la serie temporal de velocidad media horaria. Los resultados del modelo basado en redes neuronales son ligeramente mejores que los

del modelo persistente (en torno a un 3% de mejora) para los dos ejemplos que presenta. En la segunda aproximación utiliza de nuevo redes neuronales, si bien ahora las entradas corresponden a la serie temporal de velocidad de viento, pero con un espaciado entre datos de 10 minutos (es decir, toma como entradas la velocidad del viento 10 minutos antes, 20 minutos antes, etc.), y utiliza iterativamente la salida de esta red neuronal para obtener la predicción para los siguientes 60 minutos. Los resultados presentan ahora una mejora de, al menos, el 10% sobre el modelo persistente para un horizonte de predicción de una hora.

En [More03] se presenta un modelo basado en redes neuronales MLP para la predicción de la velocidad media mensual, semanal y diaria del viento, demostrando su superioridad sobre los modelos ARIMA que son aplicados a las mismas series temporales. [Maqsood05] presenta la idea de desarrollar modelos basados en cuatro tipos de redes neuronales, perceptrones multicapa, redes neuronales recurrentes de Elman, redes neuronales de función de base radial (RBF) y redes neuronales de Hopfield, en la predicción, para las próximas 24 horas, de tres variables meteorológicas, entre ellas la velocidad media horaria del viento. Se construye una red neuronal de cada tipo para cada una de las estaciones del año. Los resultados más satisfactorios son los obtenidos por la red neuronal de función de base radial, aunque sólo ligeramente mejores a los de las otras redes neuronales. Los resultados mejoran, a su vez, usando una combinación (“ensemble”) de los resultados de las cuatro redes neuronales, que consiste en realizar una media ponderada con los resultados de predicción de las cuatro redes. En [Maqsood04] los autores demuestran que el ensemble de redes neuronales permite reducir la desviación estándar de las predicciones del conjunto con respecto a las de cualquiera de las redes neuronales consideradas individualmente. El principal inconveniente en [Maqsood05] es lo limitado de los grupos de testeo, ya que sólo considera 1 día por estación del año.

El tercero de los artículos publicados como consecuencia de los trabajos de investigación de esta tesis, [Ramirez03], presenta un original sistema de predicción de la energía producida en un parque eólico con horizonte de predicción de 1 día. Se basa en el uso de series temporales fuzzy, que ya había sido utilizado, aunque con peores resultados que el modelo persistente, en la predicción de temperaturas medias diarias [Chen00]. En [Ramirez03] el uso de esta técnica, junto con la información difusa y de tipo lingüístico relativa al viento, tal como “viento suave”, “viento fuerte”, etc. (que, por ejemplo, puede aportar un experto, o bien simplemente puede recogerse en un diario de prensa), permite que el sistema de predicción desarrollado sea capaz de mejorar en un

14.3% la predicción de la energía total producida por un parque eólico durante un día (mejora de la predicción con respecto al modelo persistente).

En el cuarto artículo publicado derivado de los trabajos de investigación realizados, [Ramirez04b], se presenta otro modelo de predicción a corto plazo de la potencia eléctrica media horaria generada en un parque eólico. Este modelo está basado en el agrupamiento de datos históricos de características similares (por medio del agrupamiento substractivo, “subtractive clustering” [Chiu94]), en el desarrollo de un sistema de regresión lineal para cada uno de los grupos de datos y en la obtención del valor de predicción como media ponderada de los valores obtenidos por los modelos de regresión para cada grupo. Los resultados obtenidos con el modelo de predicción, para un horizonte de predicción de una hora, son mejores que los obtenidos mediante otros ocho modelos (modelo persistente, modelos basados en redes neuronales y en sistemas de inferencia difusa). El modelo propuesto se aplica a la predicción de la potencia eléctrica media horaria generada en un parque eólico con un horizonte de predicción de hasta 6 horas. La mejora con respecto al modelo persistente es del 4.7% para el horizonte de una hora; y del 14,1% para el horizonte de predicción de 6 horas (sobre la base del error RMSE).

En [Flores05] se desarrolla un modelo basado en una red neuronal MLP para la predicción de la velocidad media horaria del viento con horizonte 1 hora, para su aplicación en el control de la generación de un parque eólico. Los datos de entrada corresponden a los últimos tres valores conocidos de la velocidad media horaria.

El modelo presentado en [Potter06], perteneciente a esta primera familia de modelos basados en series temporales, está basado en un sistema neuro-difuso ANFIS ([Jang93]) para la predicción de la velocidad media del viento para un horizonte de 2.5 minutos. El ajuste de los datos de entrada (valores medios de velocidad) se realiza por medio de splines, lo que permite reducir considerablemente el error de predicción con respecto al modelo persistente, si bien los autores sólo estudian una de las dos posibles componentes de la velocidad del viento (por ejemplo la componente de dirección este-oeste, si bien no la especifican).

Por último, en [El-Foury06] se presenta un modelo de predicción con horizonte una hora basado en modelos “grises” [Deng89]. Los modelos grises constituyen una nueva teoría, todavía casi incipiente, desarrollada fundamentalmente en China, y de la que comienzan a aparecer las primeras aplicaciones. Los resultados obtenidos con el modelo mejoran en un 12% al modelo persistente sobre el criterio del error medio.

En la Tabla 3.1 se presenta resumidamente las características más relevantes de los modelos de predicción, pertenecientes a esta primera familia basada en series temporales, propuestos en la literatura internacional. En dicha tabla se indica la variable que predicen (velocidad media horaria o potencia eléctrica media horaria), el tipo de modelo, el horizonte de predicción máximo, los criterios de evaluación y la mejora obtenida con respecto al modelo persistente (si lo indican).

| Referencia | Variable de predicción | Tipo modelo | Horizonte | Criterio evaluación | Mejora respecto a persistente |
|-----------------|------------------------|-----------------------------------|-----------|---------------------|-------------------------------|
| Kariniotakis96 | potencia eléctrica | Red neuronal | 2 horas | RMSE | 13.5% |
| Kamal97 | velocidad del viento | ARIMA | 6 horas | --- | --- |
| Schlink98 | velocidad del viento | ARIMA | 30 min. | --- | --- |
| Alexiadis99 | velocidad del viento | Red neuronal | 2 horas | MAE | 27% |
| Sfetsos00 | velocidad del viento | ARIMA, Red neuronal, neuro-difuso | 1 hora | RMSE | 4.9% |
| Hatziargyriou01 | potencia eléctrica | neuro-difuso | 10 horas | --- | --- |
| Ramirez01 | velocidad del viento | neuro-difuso + Fourier | 1 hora | RMSE | 9% |
| Sfetsos02 | velocidad del viento | ARIMA | 1 hora | RMSE | 2 - 5% |
| Poggi03 | velocidad del viento | ARIMA | 3 horas | --- | --- |
| Ramirez03 | energía eléctrica | Fuzzy | 24 horas | RMSE | 14.3% |
| Ramirez04b | potencia eléctrica | Regresión | 6 horas | RMSE | 14.1% |
| Damousis04 | velocidad del viento | Fuzzy + GA | 4 horas | RMSE | 18.2% |
| Maqsood05 | velocidad del viento | Red neuronal | 24 horas | MAE | --- |
| Torres05 | velocidad del viento | ARIMA | 10 horas | RMSE | 12 - 20% |
| Flores05 | velocidad del viento | Red MLP | 1 hora | --- | --- |
| Potter06 | velocidad del viento | neuro-difuso | 2.5 min. | MAE | 87% |
| El-Foury06 | potencia eléctrica | “gris” | 1 hora | ME | 12% |

Tabla 3.1. Resumen de características de modelos de predicción de la primera familia.

3.2.2. Modelos de predicción basados en modelos atmosféricos.

La introducción de predicciones de variables meteorológicas (obtenidas de modelos de previsión atmosférica, “modelos atmosféricos”, desarrollados fundamentalmente por los institutos meteorológicos nacionales) en los modelos de predicción de producción de energía eléctrica de origen eólico, ha permitido mejorar sustancialmente la precisión en los resultados de éstos últimos, sobre todo en horizontes de predicción por encima de las 6 horas. Existen varios modelos de predicción meteorológica siendo el más usado el conocido como HIRLAM, que ha sido utilizado por varios institutos meteorológicos, entre ellos el de España. Estos modelos atmosféricos se van perfeccionando año tras año, permitiendo en la actualidad la

previsión aceptable de las principales variables meteorológicas, entre ellas el viento, con una anticipación de hasta 72 horas.

El primer trabajo donde utilizan predicciones de variables meteorológicas de los modelos atmosféricos, para la predicción de la producción de energía eléctrica de origen eólico, es anglo-danés [Watson94]. Utilizan los resultados del modelo atmosférico británico MESO y del modelo europeo HIRLAM, ambos con una corrección estadística en las previsiones, para tratar de determinar el ahorro de combustible convencional que hubiese existido en los años 1989 y 1990 en los sistemas eléctricos inglés y galés, en el caso de haber reducido al mínimo la reserva rodante como consecuencia de la disponibilidad de predicciones a corto plazo de generación de energía eléctrica de origen eólico. Es evidente que los ahorros son tanto mayores cuanto mayor es la potencia instalada en parques eólicos, así como la precisión en el modelo de predicción de la generación (los ahorros son mucho mayores que los correspondientes al uso del modelo persistente). Los horizontes de predicción que utilizan son 18 horas con el modelo MESO/MOS y 36 horas con el modelo HIRLAM/WASP. La técnica MOS (Model Output Statistics) constituye una de las primeras herramientas desarrolladas para la reducción de errores en modelos atmosféricos [Glahn72]. El citado modelo anglo-danés de predicción de la generación, es un modelo de regresión lineal que utiliza como variables explicativas salidas del modelo atmosférico (componentes este-oeste y norte-sur del viento, temperaturas, etc., a diferentes alturas o niveles de presión) y como variables a ajustar las correspondientes al viento (velocidad y dirección) para un punto representativo del parque eólico. Una descripción más en detalle de las características del modelo HIRLAM/WASP (predicción atmosférica numérica con el modelo HIRLAM y ajuste a las condiciones locales donde esté situado el parque eólico con el modelo físico WASP) se encuentra en [Landberg94].

El modelo atmosférico de predicción numérica de variables meteorológicas HIRLAM se revela como el más utilizado en el desarrollo de modelos de predicción a corto plazo de energía eléctrica producida en un parque eólico. El modelo HIRLAM (High Resolution Limited Area Modelling) es el fruto de un proyecto de desarrollo y mantenimiento de un sistema de previsión numérica atmosférica a corto plazo para su uso por parte de los institutos meteorológicos de Dinamarca, Finlandia, Islandia, Irlanda, Holanda, Noruega, Suecia y España. El desarrollo del proyecto comenzó en 1985 con un conjunto de “rutinas” programadas en Fortran y ejecutadas en los más potentes ordenadores disponibles en los organismos participantes. Las predicciones de las variables atmosféricas (temperatura, velocidad y dirección del viento, humedad, etc.) se realizan, a mediados de los 90, para un conjunto de puntos en el espacio que forman

una “rejilla” tridimensional, siendo constante la separación (resolución) entre los puntos de su proyección en el plano horizontal (superficie del terreno), y siendo variable la separación vertical entre los mismos, la cual viene determinada por el valor de la presión atmosférica [Cats96]. Las resoluciones utilizadas varían entre los 5 y los 50 km. La última versión del modelo HIRLAM mantiene las características esenciales de las versiones anteriores, pero mejorando diversos aspectos, entre ellos la resolución [Unden02].

A partir de los datos numéricos para las próximas 48 horas proporcionados por el modelo HIRLAM, se han desarrollado otros modelos específicos para la predicción de la potencia eléctrica media horaria producida en un parque eólico. El primero de ellos aparece en [Joensen97] con un modelo estadístico basado en regresión local. Las variables explicativas corresponden a la propia potencia eléctrica media horaria producida, a los pronósticos (salidas del modelo HIRLAM) de velocidad y dirección del viento, y a la medida de la velocidad del viento. Como principales novedades incorpora una curva empírica para la potencia eléctrica producida por el parque dependiendo de la dirección del viento, y la modelización de la componente diurna del viento. Esta componente refleja los vientos que aparecen en la costa marítima como consecuencia de las diferencias de temperaturas respecto a las del territorio interior y es modelizada de acuerdo a la hora del día y la estación del año.

En [Landberg97a], [Landberg97b] y con más detalle en [Landberg99] se describe el modelo de predicción desarrollado en el Risø National Laboratory de Dinamarca. Este modelo utiliza las predicciones obtenidas con el modelo HIRLAM (con una resolución de 26x26 km) para calcular los valores del viento en superficie utilizando las ecuaciones matemáticas que relacionan los distintos “perfiles” de velocidad de viento (relación entre velocidad del viento y altura sobre el terreno). Los valores de viento en superficie son ajustados de acuerdo a las características locales por medio de las herramientas del software estadístico WASP, también desarrollado por Risø [Troen91], que modeliza el efecto de la rugosidad del terreno circundante y los posibles obstáculos cercanos. Con el uso de técnicas de regresión lineal (MOS) y de evaluación de efectos de sombra entre aerogeneradores (software PARK, también parte del software WASP) se realiza la predicción de la potencia eléctrica media horaria con un horizonte de predicción de 36 horas. Con el desarrollo de las tecnologías de la información y comunicación, se publica en [Landberg01] una propuesta consistente en la implementación del modelo de predicción on-line, es decir, como programa informático capaz de tomar los datos más recientes de forma automática (últimos

valores medidos de variables meteorológicas, nuevas predicciones numéricas, etc.) y de suministrar predicciones de generación de energía eléctrica a través de Internet.

Tras las primeras publicaciones describiendo modelos de predicción de generación de energía eléctrica basados, a su vez, en predicciones atmosféricas del modelo HIRLAM, en [Beyer99] se describe el primer modelo basado en predicciones, no del modelo HIRLAM, sino del modelo Deutschlandmodell, desarrollado por el instituto meteorológico nacional alemán con una resolución de 14x14 km. El modelo de predicción de generación descrito es muy similar al de [Landberg99] en el sentido de que en aquél se calcula el viento en superficie a partir de las leyes físicas que lo rigen, se reducen los errores por medio de una regresión lineal (MOS) y se determina la potencia eléctrica media horaria producida por medio de la curva de potencia para un horizonte de predicción de 48 horas. La principal novedad en [Beyer99] radica en las pruebas que se realizan en la predicción de la producción de energía eléctrica para toda una región, donde los errores RMSE se reducen a la mitad en comparación con los de la media de los errores RMSE para los parques de la región considerados individualmente. Una versión actualizada de la modelización a partir de las leyes físicas, específica para parques eólicos sobre la superficie del mar, aparece en [Tambke05], donde se llega a la conclusión de que los modelos de predicción atmosférica conducen a errores relativos menores para parques eólicos sobre el mar que sobre tierra.

La idea de aprovechar el efecto de agregación (de producción de energía eléctrica de los parques eólicos), para mejorar los resultados de las predicciones a nivel nacional es el utilizado en el modelo operativo alemán “previento” [Focken01], [Focken02]. El modelo de predicción es el ya presentado en [Beyer99], si bien en este caso lo utilizan para realizar la predicción, con un horizonte de hasta 48 horas, de la producción media horaria de un conjunto de 30 parques eólicos situados en la mitad norte (la más llana) de Alemania. Después de analizar la correlación existente entre los datos de generación eléctrica en varios parques eólicos, se llega a la conclusión de que es posible realizar la predicción de la generación en todos los parques eólicos de una región, considerando sólo las predicciones de un conjunto reducido de parques eólicos. Los límites a la reducción del error de predicción, como consecuencia de la agregación de parques eólicos, es tratado en [Gastón06], donde se estudia la reducción del error de predicción para un conjunto de parques eólicos, llegando a la conclusión de que, en general, el error de predicción se reduce conforme aumenta el nivel de agregación (número de parques eólicos considerados), si bien por encima de tres parques las reducciones no son tan significativas. También en el caso de parques eólicos con el mismo régimen de viento (misma región y características orográficas similares), apenas

hay reducción del error de predicción como consecuencia de la agregación (de producción de energía eléctrica de los parques eólicos).

El desarrollo de nuevos modelos de predicción numérica atmosférica hace que se publiquen trabajos acerca de las bondades de alguno de los últimos modelos atmosféricos en aparecer. En [Roulston01] y [Roulston03] se analiza la utilidad de las predicciones de los modelos atmosféricos, concretamente las predicciones “ensemble” realizadas por el Centro Europeo de Predicción Atmosférica a Medio Plazo (ECMWF). Estas predicciones, con un horizonte de hasta 10 días, presentan 51 escenarios distintos para las variables meteorológicas, es decir, un conjunto de 51 predicciones diferentes para cada punto, para cada variable y cada horizonte de predicción. Los 51 escenarios distintos son creados a partir de pequeñas perturbaciones (pequeñas variaciones) sobre los valores de las condiciones iniciales establecidas con los últimos datos registrados (asimilados). En [Roulston03] se presenta un breve estudio de simulación sobre el beneficio económico obtenido en un parque eólico, en el caso de vender en el mercado la energía producida, en función de los resultados de seis predicciones probabilísticas de variables meteorológicas (seis escenarios).

La propuesta apuntada en [Landberg01] sobre sistemas de predicción on-line sirve de base para el sistema de predicción danés Zephyr desarrollado para su aplicación por parte de las compañías eléctricas de ese país. La descripción del sistema, en cuanto a los módulos constituyentes y características de acceso y uso, se puede encontrar en [Giebel01] y [Nielsen02a]. Una descripción más detallada del modelo de predicción aparece en [Nielsen01] y en [Nielsen02b]. Se trata de un modelo basado en las predicciones del modelo atmosférico HIRLAM, que a diferencia del modelo de [Landberg99] ya no precisa de la determinación del viento local, puesto que incorpora una nueva mejora que consiste en la predicción de la velocidad y dirección del viento para una altura de 10 m sobre el terreno. El horizonte de predicción del sistema Zephyr es de 48 horas, con valores medios de producción cada media hora. El modelo de predicción está basado en regresión local, ya descrito en [Joensen97], en una estimación paramétrica condicional y en una estimación de la curva de potencia de forma empírica (a partir de las medidas de velocidad y dirección del viento, y de potencia media generada). Realizan la predicción de la producción de energía eléctrica para todo el parque eólico, incluyendo efectos diurnos (patrón de viento repetitivo día a día, como es el caso de las brisas en zonas costeras) y auto-regresivos.

Al inicio de esta última década, el desarrollo de mejoras en los modelos atmosféricos de predicción numérica, centra los trabajos de colaboración realizados

entre el instituto danés de meteorología y la Universidad de Cork en Irlanda. En [Moehrlen02] se presentan los resultados obtenidos con el modelo atmosférico HIRLAM en la predicción, con horizontes entre 6 y 24 horas, de la potencia eléctrica media horaria para un parque eólico situado en terreno montañoso utilizando la curva de potencia del parque. La principal novedad se centra en que han variado los parámetros del modelo HIRLAM, entre ellos la resolución horizontal (dejándola en un valor entre 30 y 1.4 km), siendo la conclusión más relevante que el error sistemático en la predicción se reduce aumentando la resolución del modelo HIRLAM. En [Jørgensen02a] se enumeran las mejoras introducidas en el modelo HIRLAM que incluyen las numéricas (mejora en la resolución de las ecuaciones), y un mejor tratamiento a la rugosidad superficial, tanto de la superficie terrestre como de la marina. En [Jørgensen02b] se describe el sistema completo con la integración de los modelos HIRLAM, WAM (modelado de la superficie marina), e HIRPOM (módulo para la obtención de la potencia eléctrica media horaria a partir de las predicciones de velocidad y dirección del viento). La propuesta más interesante que aparece en estos tres trabajos es la de la medida de la incertidumbre en la predicción utilizando un “ensemble” de predicciones del modelo HIRLAM. En [Roulston01] se apunta la idea de utilizar predicciones “ensemble”, pero para la realización de predicciones probabilísticas. En este caso, [Jørgensen02a], el “ensemble” corresponde a un conjunto de resultados de predicción de variables meteorológicas con el modelo HIRLAM, pero con distintas condiciones iniciales de los parámetros del modelo atmosférico. Si la dispersión de los resultados de predicción de variables meteorológicas para el “ensemble” es baja, la incertidumbre en la predicción de la potencia eléctrica media generada también será baja. Por el contrario, una alta dispersión en los resultados (valores de predicción de variables meteorológicas para un instante determinado) se traducirá en una alta incertidumbre en la predicción de la potencia eléctrica media generada. Para más información sobre las características de los “ensemble” ofrecidos por servicios meteorológicos internacionales puede consultarse [Giebel03b].

La mejora de la predicción de los valores de variables meteorológicas por medio de “ensemble”, pero bajo un punto de vista diferente, aparece en [Landberg02]. En este caso el conjunto de datos que forman el “ensemble” son las predicciones de variables meteorológicas realizadas para un instante concreto, pero con distintos horizontes, por el mismo modelo atmosférico de predicción numérica y con un único escenario (sin perturbaciones). Si tenemos en cuenta que la mayoría de los modelos atmosféricos se actualizan cada 6 ó 12 horas, podemos disponer de varios valores de predicción para la velocidad y dirección del viento en las próximas 48 ó 72 horas. Cuando los valores para un instante determinado presentan pequeñas diferencias entre sí, significa que la

situación atmosférica es estable, por lo que el error cometido por el propio modelo atmosférico será pequeño. Por el contrario una gran diferencia entre los valores de predicción del “ensemble” significará que la situación atmosférica es inestable, lo que conducirá a mayores errores en la predicción de las variables meteorológicas. Evidentemente, el uso de esos valores de predicción de las variables meteorológicas (situación atmosférica inestable) ocasionará mayores errores en la predicción de la producción eléctrica, que en el caso de usar valores de predicción de variables meteorológicas correspondientes a una situación atmosférica estable.

El problema de la incertidumbre en la predicción vuelve a ser tratado en [Pinson03] y con más detalle en [Pinson04]. En el primero, [Pinson03], se describe un modelo de predicción con un horizonte de hasta 10 horas, basado en sistemas neuro-difusos ANFIS que utiliza como entradas valores pasados de la potencia eléctrica media horaria generada en el parque eólico y predicciones de dirección y velocidad del viento suministradas por el modelo HIRLAM. La selección de entradas del sistema ANFIS, así como su estructura (número de reglas) es optimizada por medio de un algoritmo “simplex” junto con restricciones no lineales. A pesar de que se describe dicho sistema, no se aporta ningún dato de su funcionamiento, centrándose más en recoger una medida de la incertidumbre de la predicción por medio de intervalos de confianza. Esta forma de estimar la incertidumbre en la predicción se encuentra más desarrollada en [Pinson04], donde se define un “índice de riesgo meteorológico” en función de la dispersión de los valores de predicción en un “ensemble” de resultados obtenidos con el modelo HIRLAM, y realizando después el ajuste de los intervalos de confianza de acuerdo a este índice.

En [Bremnes04] se regresa otra vez a las predicciones probabilísticas como forma de cuantificar la incertidumbre en la predicción. Esta predicción probabilística se basa en el uso de la regresión local lineal de “cuantiles” (percentiles si están expresados en tantos por ciento) [Yu98]. Para predicción a corto plazo de la potencia eléctrica media horaria en un parque eólico, el “cuantil” z es el valor de potencia eléctrica media para el cual la probabilidad de producir menos que dicho valor es, justamente, z (siendo z , evidentemente, un número entre 0 y 1). A partir de las predicciones de variables meteorológicas realizadas con el modelo HIRLAM y utilizando como variables las predicciones de velocidad y dirección del viento para 10 metros de altura y el número del mes correspondiente al instante para el que se hace la predicción, se calcula la potencia eléctrica media horaria producida mediante regresión polinómica de segundo grado. En [Bremnes04] no se presentan los resultados obtenidos en la predicción, sino el estudio estadístico de estos “cuantiles”. En [Bremnes06] se presentan otros dos métodos

para el cálculo de los “cuantiles” aparte de la regresión lineal local. En [Nielsen06a] se desarrolla otro modelo para la obtención de los percentiles del 25% y del 75% para la predicción de la potencia eléctrica media realizada con la última versión del sistema danés Zephyr. El modelo de [Nielsen06a] se basa en la regresión lineal de “cuantiles”, el uso de las predicciones de variables meteorológicas con el modelo HIRLAM (densidad del aire y dirección del viento a 10 m) y aproximación de funciones por medio de splines.

Un nuevo modelo atmosférico no europeo, el modelo MM5, presenta su primera aplicación en la predicción de la potencia eléctrica media horaria en [Larson06]. En este trabajo se utilizan las predicciones medias horarias de valores de variables meteorológicas obtenidas con este modelo atmosférico para, por medio de redes neuronales artificiales, realizar la predicción de la velocidad media del viento. Para cada mes se desarrolla un modelo de predicción distinto. La inclusión de datos medidos en estaciones meteorológicas situadas hasta 147 km de distancia, permiten mejorar las predicciones para horizontes de unas pocas horas. El trabajo sólo aporta datos de mejora sobre el modelo persistente para un horizonte de predicción de 2 horas, donde las redes neuronales obtienen mejoras entre un 8 y un 36 %, dependiendo del mes. Otro modelo atmosférico nuevo, el COAMPS (Coupled Ocean/Atmosphere Mesoscale Prediction System) desarrollado en el laboratorio de investigación de la Armada norteamericana se presenta en [Perivolaris06], con la particularidad de realizar una predicción de velocidad de viento para cada uno de los aerogeneradores de un parque eólico.

Las predicciones sobre velocidad y dirección del viento obtenidas con el modelo MM5 son utilizadas en el quinto artículo publicado [Ramirez06], fruto de los trabajos de investigación de la presente tesis. En él se desarrollan modelos de predicción de la potencia eléctrica media horaria generada en un parque eólico basados en cuatro familias distintas de redes neuronales. Se desarrolla un modelo distinto para cada una de las cuatro familias (perceptrones multicapa, redes neuronales multicapa generalizadas, redes neuronales de función de base radial y redes neuronales con análisis de componentes principales) para cada uno de los horizontes de predicción (de 1 a 24 horas). Los 96 modelos de redes neuronales son optimizados por medio de la aplicación de un algoritmo genético, que permite automáticamente la selección idónea de entradas, del número de neuronas en capas ocultas y del valor de los parámetros del algoritmo de entrenamiento (“back-propagation”). Los resultados de los modelos basados en las redes neuronales son comparados con los obtenidos mediante el modelo persistente y mediante un modelo auto-regresivo, siendo los resultados de los modelos de redes neuronales superiores. La mejora de las predicciones de dichos modelos basados en

redes neuronales con respecto a las del modelo persistente, es del 35.3% para un horizonte de predicción de 24 horas y utilizando como criterio el indicador RMSE.

La primera publicación en revista sobre un modelo de predicción numérica a corto plazo desarrollado en España aparece en [Sánchez06]. El modelo recibe el nombre de “Sipreólico” y hace uso de las predicciones de variables meteorológicas obtenidas con el modelo HIRLAM del Instituto Nacional de Meteorología, siendo el horizonte de predicción de 36 horas. El sistema de predicción se basa en una combinación adaptativa lineal de las predicciones obtenidas con nueve modelos dinámicos lineales y no lineales. Estos modelos alternativos incluyen un modelo auto-regresivo puro, modelos auto-regresivos con variable exógena o externa, y modelos auto-regresivos con componentes de regresión lineal no paramétrica. Los parámetros de los nueve modelos son estimados recursivamente. La predicción final se realiza combinando linealmente los resultados obtenidos por los modelos dinámicos que mejores resultados han ofrecido en las últimas predicciones. Lamentablemente, el autor no presenta los resultados de ninguna prueba de predicción de la producción eléctrica en un parque eólico. Entre las limitaciones de este sistema de predicción se encuentran la del horizonte máximo de predicción (36 horas); y la falta de información sobre resultados de los modelos propuestos y la bondad de los mismos (con relación a los resultados de otros modelos publicados en la literatura internacional). Igualmente un inconveniente de estos modelos es la utilización de las predicciones de un modelo atmosférico cuyos valores no pueden adquirirse gratuitamente (exigen contrato con el Instituto Nacional de Meteorología).

En [Sideratos07] se describe un sistema de predicción de la potencia eléctrica media generada en un parque eólico con un horizonte de predicción de 48 horas con variables de entrada obtenidas del modelo HIRLAM. Se trata de un sistema compuesto por 12 redes neuronales RBF y tres sistemas de inferencia difusa. Los sistemas de inferencia difusa son utilizados para determinar la calidad de las predicciones del modelo HIRLAM. Los autores prueban el sistema propuesto con los datos del parque eólico de Klim situado junto a la costa marítima danesa logrando mejoras en torno al 43% sobre el modelo persistente para un horizonte de 24 horas, y del 45% para 36 horas (sobre la base del error RMSE normalizado). Como limitaciones significativas del sistema de predicción se puede resaltar que los resultados presentados corresponden a un parque eólico en un terreno llano (con predicciones de producción eléctrica comparativamente más sencillas que para un terreno de interior con orografía compleja), y para el que se han evaluado hasta 12 modelos de predicción distintos con apenas diferencias de valores de predicción entre ellos ([Martí06], [Karionatakis06]); la estructura de dicho sistema (número de neuronas de cada red, por ejemplo) se ha

establecido sólo mediante un proceso de prueba y error, limitado escasamente a tres pruebas. Adicionalmente no hay evidencia, ni análisis posterior, de que el índice de calidad de las predicciones del modelo HIRLAM determinado por los sistemas de inferencia difusa, verdaderamente responda a la exactitud o inexactitud en la predicción de los valores de velocidad media suministrados por dicho modelo atmosférico.

Un buen resumen de las técnicas estadísticas utilizadas en los distintos modelos de predicción de esta familia (basada en modelos atmosféricos) puede encontrarse en [Nielsen06b], donde se describen las técnicas, pero lamentablemente sin ningún tipo de resultado obtenido para algún ejemplo de aplicación.

En la Tabla 3.2 se presentan resumidamente las características más relevantes de los modelos de predicción pertenecientes a esta segunda familia (basada en predicciones de modelos atmosféricos) ya comentados en los párrafos anteriores. En dicha tabla se recogen los modelos utilizados, el modelo atmosférico que proporciona las predicciones de variables meteorológicas como la velocidad y dirección de viento, temperatura, etc., y las características esenciales de las técnicas usadas en el modelo de predicción de la producción eléctrica en un parque eólico. También figuran en la tabla los horizontes máximos de predicción de cada modelo, el criterio de evaluación seleccionado para la determinación de los errores de predicción, así como la mejora con respecto al modelo persistente (mejora correspondiente al horizonte máximo considerado, en caso de que el modelo estudiado haya presentado algún ejemplo de aplicación).

| Referencia | Modelo de predicción | Modelo atmosférico | Horizonte | Criterio evaluación | Mejora respecto a persistente |
|-------------------|-----------------------------|---------------------------|------------------|----------------------------|--------------------------------------|
| Watson94 | MOS - WASP | HIRLAM | 36 horas | --- | --- |
| Joensen97 | Estadístico | HIRLAM | 24 horas | RMSE | 30% |
| Landberg99 | Físico+MOS+WASP | HIRLAM | 36 horas | RMSE | 48% |
| Beyer99 | Físico+MOS | Deutschland | 48 horas | RMSE | --- |
| Focken01 | Físico+MOS+correlación | Lokalmodel | 48 horas | RMSE | --- |
| Giebel01 | Estadístico | HIRLAM | 48 horas | --- | --- |
| Ramirez06 | Neuronal | MM5 | 24 horas | RMSE | 35.3% |
| Larson06 | Neuronal | MM5 | 2 horas | RMSE | 14 – 37% |
| Sanchez06 | Estadístico | HIRLAM | 36 horas | MSE | --- |
| Sideratos07 | Neuronal + fuzzy | HIRLAM | 41 horas | RMSE | 42% |

Tabla 3.2. Resumen de características de modelos de predicción de la segunda familia.

3.3. CONCLUSIONES

Son varias las conclusiones que se pueden extraer de la revisión bibliográfica sobre la predicción de la producción eléctrica en un parque eólico y que son comentadas a continuación.

1. Los modelos de la primera familia (modelos basados en series temporales) presentan mejores resultados que los modelos de la segunda familia (los basados en valores de predicción de variables meteorológicas obtenidos de modelos atmosféricos) para horizontes de predicción relativamente bajos, con un límite entre las 3 y las 6 horas. Así, en [Ramirez06] un modelo auto-regresivo (por tanto de la primera familia) presenta menores errores de predicción que los modelos de la segunda familia basados en redes neuronales para horizontes de predicción por debajo de las 6 horas.

En general, los modelos publicados parecen “focalizados” para un intervalo específico de horizonte de predicción. Así, los modelos de la primera familia están diseñados para realizar las predicciones de potencias eléctricas en parques eólicos en el muy corto plazo, mientras que los modelos de la segunda familia están más orientados a realizar la predicción para el día siguiente (marcando su aplicación para la realización de ofertas de venta en el mercado eléctrico, así como para otras necesidades relativas a la operación del Sistema Eléctrico). Ningún sistema de predicción de la producción eléctrica combina ambas características.

2. En los modelos de predicción de la primera familia no existe un criterio de evaluación de error común, puesto que utilizan tanto el error medio (ME), como el error absoluto medio (MAE) o la raíz del error medio cuadrático (RMSE). Por el contrario los modelos de predicción publicados y pertenecientes a la segunda familia, utilizan de forma generalizada el indicador RMSE, como criterio de evaluación del error en la predicción.
3. A pesar de haberse propuesto un modelo de referencia en [Nielsen98], para obtener los resultados de predicción de potencias eléctricas con los que comparar los obtenidos con cualquier otro modelo propuesto, se utiliza como modelo de referencia exclusivamente el modelo persistente, independientemente del horizonte de predicción.
4. Los errores de predicción dependen del horizonte de predicción: cuanto

mayor es el horizonte de predicción mayor es el error. Sin embargo, en la comparación de modelos propuestos con el modelo persistente, la mejora de la predicción con cualquier modelo propuesto crece con el horizonte de predicción.

5. El nivel de detalle que aparece en las publicaciones sobre modelos de la primera familia es muy superior al de los modelos de la segunda familia. Es más, mientras que para los modelos de la primera familia se presentan de forma generalizada los resultados y errores obtenidos con la aplicación del modelo, para los modelos de la segunda familia son casi excepcionales las publicaciones que lo hacen, quedándose la mayoría en una mera descripción del modelo propuesto.
6. Los resultados publicados y obtenidos con los distintos modelos son muy dispares. Igualmente, puede afirmarse que no es posible alcanzar resultados (de predicción de producción eléctrica) equiparables de un lugar geográfico (para el que se realiza la predicción) a otro. Ello es debido esencialmente a que las condiciones locales de viento dependen altamente de la orografía del lugar donde se sitúa el parque eólico, y también de la situación de cada uno de los aerogeneradores que lo componen [Landberg99], de tal forma que resulta casi imposible alcanzar un nivel de error similar con el mismo modelo de predicción para todos los parques eólicos. Así, no es posible obtener un nivel de error de predicción similar para un parque eólico situado en un terreno llano, que para otro parque situado en un terreno significativamente montañoso.
7. Los errores de predicción se pueden reducir como consecuencia de la agregación (predicción para un conjunto de parques eólicos): se obtiene menor error en la predicción a nivel regional o nacional para un conjunto de parques eólicos que en la predicción de un único parque [Gastón06].
8. Para el desarrollo de modelos de predicción (de producción eléctrica) de la segunda familia, se han utilizado diferentes modelos atmosféricos que casi siempre han sido desarrollados, mantenidos y ejecutados por institutos meteorológicos nacionales. En este sentido, una buena parte de los diseñadores de nuevos modelos de predicción de la producción eléctrica en un parque eólico vienen a ser, en cierta forma, “clientes” de los institutos meteorológicos (o de empresas que a su vez son clientes de dichos institutos).

9. La reducción de los errores de predicción de producción eléctrica para modelos de la segunda familia está ligada a la reducción de los errores de predicción de las variables meteorológicas de los modelos atmosféricos. Esta reducción de errores (mejora de las predicciones de potencia eléctrica) se consigue fundamentalmente aumentando la resolución de dichos modelos atmosféricos.
10. Las predicciones de variables atmosféricas de interés para la predicción de la producción eléctrica en un parque eólico son, esencialmente, la velocidad y dirección del viento, utilizándose excepcionalmente otras (temperatura, densidad del aire, humedad, etc.).
11. La adaptación (si ésta es posible) de la mayor parte de los modelos publicados, a otros parques eólicos, debe ser realizada por un experto: la correcta modelización física, la adecuación de los parámetros, el ajuste de “ventanas temporales”, etc., requiere de la participación de un experto.

4. MODELOS DE PREDICCIÓN METEOROLÓGICA

En este capítulo se presentan las bases sobre las que se desarrollan los modelos de predicción numérica de variables meteorológicas en el corto plazo. Se presentan las características de resolución espacial y temporal de estos modelos de predicción numérica, describiendo los aspectos más relevantes de los modelos globales y meso-escalares. Se analiza más en detalle un modelo meso-escalar de predicción numérica (modelo atmosférico MM5), bien conocido a nivel internacional, que suministra predicciones de las variables meteorológicas de mayor interés en nuestros trabajos de investigación sobre la predicción de la producción eléctrica en parques eólicos. También se expondrá, de forma resumida, las experiencias realizadas para la puesta en marcha de dicho modelo MM5 en el laboratorio de investigación de Planificación, Operación y Control de Sistemas de Energía Eléctrica, perteneciente al Área de Ingeniería Eléctrica de la Universidad de La Rioja.

4.1. CONCEPTOS BÁSICOS DE LOS MODELOS DE PREDICCIÓN METEOROLÓGICA

La predicción de la evolución de las variables meteorológicas en el corto plazo ha sido un problema científico de indudable interés social desde la antigüedad. El conocer de antemano las temperaturas de mañana, o si van a caer precipitaciones en los dos próximos días, puede cambiar decisiones con una gran influencia socio-económica. Por poner un ejemplo directamente relacionado con la región de La Rioja, unas predicciones de importantes lluvias pronosticadas con unos días de anticipación, pueden hacer adelantar las fechas de vendimia, ya que el exceso de agua puede afectar a la calidad final del vino que se obtenga de la uva recogida tras las lluvias.

La atmósfera es, esencialmente, una mezcla de gases y agua en sus distintos estados. Esta mezcla se encuentra en un continuo movimiento, sometida a la acción del campo gravitatorio de la propia Tierra. Además sobre esta mezcla está actuando la radiación solar, responsable del calentamiento de la misma. Pero como en cualquier sistema real, se cumplen un conjunto de leyes físicas: leyes de conservación de energía, masa y momento, además de la ecuación de estado de los gases. Estas leyes físicas se pueden representar mediante ecuaciones diferenciales que relacionan las derivadas totales de ciertas magnitudes físicas respecto a la variable tiempo.

Las ecuaciones diferenciales que corresponden a la modelización matemática de las leyes físicas que rigen el comportamiento dinámico de la atmósfera, reciben la denominación de ecuaciones primitivas. La resolución de dichas ecuaciones proporcionará unos resultados numéricos para las variables que intervienen en las mismas. Esta es la base de la predicción numérica meteorológica.

El objetivo de la predicción numérica meteorológica es determinar el estado de la atmósfera para un instante futuro, a partir de la condición inicial actual. Para ese objetivo se debe disponer de un modelo matemático capaz de resolver las ecuaciones primitivas y que incluya los posibles intercambios energéticos (radiación, turbulencia, calor latente, etc.). Puesto que las ecuaciones primitivas son no lineales, en la práctica se suele recurrir a distintas aproximaciones que simplifican la resolución numérica [Gutierrez04].

La resolución de las ecuaciones diferenciales se realiza con varias técnicas numéricas, siendo las más usuales la de diferencias finitas, la de elementos finitos y la de esquemas de interpolación. Con las soluciones numéricas para las distintas variables meteorológicas se obtiene el estado futuro de la atmósfera, lo que en términos coloquiales no es más que la predicción del tiempo.

Pero, para la obtención de las soluciones numéricas, los modelos atmosféricos discretizan el espacio y el tiempo. Básicamente, en un análisis con coordenadas cartesianas, se considera una rejilla 4D (tres dimensiones espaciales y tiempo) sobre la que se aplican las técnicas de integración de las ecuaciones diferenciales. La resolución espacial de esa rejilla (tamaño de los prismas en una rejilla tridimensional, ya que la componente vertical no es igual a las otras dos componentes) influye directamente en la precisión del modelo atmosférico. Así, una disminución del tamaño de la rejilla significará un aumento de la precisión, si bien, también significará un aumento en el tiempo computacional necesario para la resolución de las ecuaciones: a la par que reducimos el tamaño de la rejilla, debe reducirse el paso de integración utilizado en las técnicas numéricas de resolución, para poder mantener la estabilidad de las soluciones. Una importante área de investigación en el campo de la predicción meteorológica, hoy en día, lo constituye el aumento de la resolución en los modelos, sin que este aumento de resolución signifique importantes incrementos en el tiempo computacional para la resolución de los mismos.

En la rejilla de análisis el significado de la coordenada vertical (altura) es

diferente según el modelo que se utilice. Existen modelos que trabajan con coordenadas verticales z definidas sobre un nivel de referencia (por ejemplo, nivel medio del mar en algún punto). En otros modelos atmosféricos se utiliza una coordenada conocida como σ , que es una coordenada vertical que “sigue” el terreno. En la Figura 4.1 se puede apreciar la diferencia entre una discretización vertical con una coordenada z o con una coordenada σ . Evidentemente, un modelo atmosférico que trabaje con coordenadas σ necesita una definición precisa de la orografía del terreno.

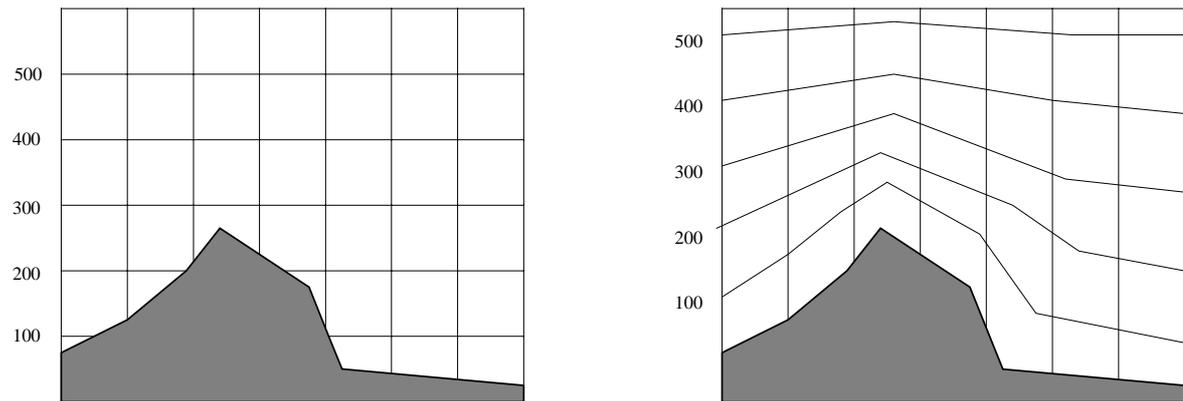


Figura 4.1. Coordenadas verticales z (a la izquierda) y coordenadas verticales σ (a la derecha).

Los modelos de predicción numérica atmosférica (modelos atmosféricos) suelen tener tres fases diferenciadas: Preproceso, Proceso y Postproceso. En la fase de Preproceso se realizan varias tareas, entre ellas las de definición del dominio de análisis (determinación de la zona de estudio), resolución de la rejilla, características del suelo y “asimilación” de los datos atmosféricos disponibles [Daley99]. Esta asimilación contempla el establecimiento del estado inicial o punto de partida de la atmósfera y consiste en la interpolación sobre los puntos de la rejilla (vértices de los prismas) de las observaciones de las variables meteorológicas disponibles. La interpolación va seguida de una inicialización dinámica de los datos interpolados para comprobar su consistencia dinámica (análisis de la conservación de masa). El proceso de asimilación es muy delicado, ya que los errores en el mismo tendrán como consecuencia una incertidumbre en el plazo de predicción (hasta cuando son válidas las predicciones), debido a la no linealidad de los modelos. Tras el Preproceso, el resultado es una rejilla con los valores de presión, velocidad y dirección del viento, temperatura y humedad para cada uno de los prismas que la forman.

La fase de Proceso del modelo atmosférico consiste en la resolución de las ecuaciones diferenciales para cada uno de los elementos de la rejilla. Es decir, para cada uno de los prismas delimitados por la rejilla de análisis se resuelven las ecuaciones que

rigen el comportamiento de la atmósfera. La resolución se realiza de forma iterativa con un intervalo de tiempo (paso o paso de integración) definido (del orden de segundos).

En la fase de Postproceso se determina, a partir de las variables básicas (presión, velocidad y dirección del viento, temperatura y humedad), otras variables de interés meteorológico, y se realizan las representaciones gráficas como las que aparecen en los programas televisivos con información meteorológica.

4.2. MODELOS ATMOSFÉRICOS

No todos los modelos atmosféricos trabajan con la misma escala espacio-temporal. Cada modelo atmosférico trata de seguir la evolución de la atmósfera para la escala que lo define. Por ejemplo, para la predicción del viento en la superficie del terreno la escala del modelo atmosférico debe ser pequeña: pequeña en lo espacial por la alta influencia que tiene la orografía sobre el viento, y pequeña en lo temporal por la alta variabilidad del viento (ráfagas, por ejemplo). Así, en función de la escala espacio-temporal, se distinguen tres clases de modelos: macro-escala, meso-escala y micro-escala.

En [Orlanski75] se plantea una normalización para las distintas escalas, proponiendo 8 subdivisiones contenidas en las tres escalas. La escala correspondiente a cada subdivisión se adapta mejor para el estudio de algún fenómeno atmosférico concreto. Estas subdivisiones, con su escala espacial y temporal, y algunos de los posibles fenómenos que podrían estudiar, figuran en la Tabla 4.1. En dicha tabla se da el límite inferior para cada una de las subdivisiones. En general, un modelo atmosférico con alta resolución espacial (escala espacial pequeña) tendrá un bajo alcance (validez) de las predicciones que realice (escala temporal pequeña). Por el contrario los modelos con una baja resolución espacial (escala espacial grande) tendrán un alcance temporal mucho mayor.

Tanto los modelos atmosféricos de macro-escala como los de micro-escala estudian fenómenos cuyas escalas escapan de las variables de interés en la predicción de la producción eléctrica en parques eólicos. Para esta tarea, los modelos atmosféricos que pueden aportar la información con mayor interés (viento en superficie) son los modelos de meso-escala, al menos en su escala espacial. La escala temporal de un modelo de meso-escala puede aumentarse si sus condiciones iniciales se establecen a partir de un modelo de mayor escala. Por ejemplo, es posible realizar predicciones con un modelo

meso- β con una escala temporal de varios días, sobre la base de encadenar predicciones para intervalos de varias horas, tomándose las condiciones iniciales de la atmósfera, para cada intervalo, a partir de las predicciones de un modelo de macro-escala. Por ello, se va a describir con mayor detalle los modelos de macro-escala y de meso-escala.

| Nombre | Escala espacial | Escala temporal | Fenómenos |
|-----------------|-----------------|-----------------|----------------------------|
| Macro- α | 10000 km | 1 mes | Circulación general |
| Macro- β | 2000 km | 1 semana | Depresiones y anticiclones |
| Macro- γ | 200 km | 1 día | Frentes, huracanes |
| Meso- β | 20 km | 6 horas | Brisas, vientos de montaña |
| Meso- γ | 2 km | 1 hora | Tormentas |
| Meso- δ | 200 m | 30 minutos | Tornados |
| Micro- β | 20 m | 1 minuto | Trombas, estelas |
| Micro- γ | 2 m | 1 segundo | Turbulencias |

Tabla 4.1. Escalas espacio-temporales para modelos atmosféricos.

4.2.1. Modelos atmosféricos globales.

Los modelos atmosféricos de macro-escala son los que tienen un mayor alcance en las predicciones meteorológicas, esto es, tienen un mayor horizonte de validez de las mismas. Como ya se ha comentado anteriormente, un aumento de la resolución espacial implica la necesidad de reducción del paso de integración de las ecuaciones diferenciales que definen el modelo, lo que, debido a no linealidad, conduce a valores que resultan ser no válidos en un menor tiempo de predicción, es decir, el alcance de la validez de la predicción se reduce (disminución de la escala temporal).

Los llamados modelos atmosféricos globales son modelos de macro-escala, que simulan el flujo general atmosférico de todo el planeta y, por tanto, son modelos que realizan la integración de las ecuaciones primitivas sobre una rejilla 4D que cubre toda la Tierra. En lugar de utilizar coordenadas cartesianas utilizan coordenadas esféricas, indicándose la resolución espacial del modelo atmosférico por medio de la separación entre puntos de la misma en grados. La Figura 4.2 representa la cobertura de puntos de la rejilla de un modelo global con una resolución de 2.5° , lo que corresponde a un total de 144×73 puntos sobre la proyección en la superficie terrestre. Los modelos atmosféricos globales pueden tener diferente número de niveles para la altura sobre la superficie de la Tierra. Algunos modelos atmosféricos expresan la componente vertical en función de la presión; así se refieren a la altura de presión de x milibares (utilizan las

coordenadas σ); otros modelos atmosféricos utilizan un sistema de niveles verticales fijos. El mayor valor de la componente vertical, expresado en unidades de longitud, es aproximadamente 64 km. La superficie de la Tierra está parametrizada a la misma resolución que el modelo atmosférico, por lo que su representación es un tanto “grosera”, al quedar la orografía del terreno muy poco detallada.

Los modelos atmosféricos globales reciben también la denominación de modelos hidrostáticos ya que usan esta aproximación (de comportamiento hidrostático) para analizar la evolución de la atmósfera. El comportamiento hidrostático equivale a suponer que existe un equilibrio perfecto entre la fuerza del gradiente de presión y la fuerza gravitatoria, lo que en el fondo significa la “no evolución” de la atmósfera en la componente vertical.

Dada la baja resolución espacial de este tipo de modelos atmosféricos (los de mayor resolución la tienen establecida en 0.5°), pueden simular la evolución de la atmósfera con alcances temporales de semanas. Estos modelos se utilizan para estudios de variaciones estacionales e, incluso, para estudios de cambio climático.

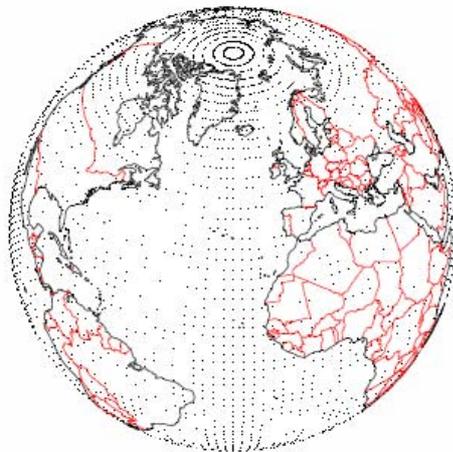


Figura 4.2. Puntos de la rejilla para un modelo atmosférico global de resolución 2.5° .

Los grandes organismos dedicados a estudios meteorológicos son los que dedican sus recursos informáticos a la resolución de los modelos atmosféricos, que se resuelven varias veces por día (2 ó 4 veces, a intervalos de 12 ó 6 horas). Como ejemplos de estos organismos, pueden citarse el NCEP americano (National Centre for Environmental Prediction, perteneciente al National Weather Service del gobierno norteamericano) y el ECMWF europeo (European Centre for Medium-Range Weather Forecasts). El NCEP americano pone a disposición pública, por medio de varios

servidores ftp, los resultados de varios de los modelos atmosféricos globales de que dispone. Los más interesantes son los modelos atmosféricos GFS (Global Forecast System, anteriormente conocidos como AVN). Se trata de dos modelos atmosféricos con resoluciones espaciales de 2.5° y 1°, aunque recientemente se ha puesto en ejecución el denominado modelo atmosférico “GFS master” con una resolución espacial de 0.5°.

El modelo atmosférico GFS, con resolución espacial de 1°, es resuelto 4 veces al día, concretamente con la asimilación de los datos de la atmósfera a las 0, 6, 12 y 18 horas GMT. Produce predicciones de las variables meteorológicas para los puntos de la rejilla con horizontes de 3 a 180 horas a intervalos de 3 horas. Los resultados, en ficheros con un formato específico llamado GRIB, están disponibles desde unas 3 horas y media después de la asimilación de datos (los primeros datos con un menor horizonte de predicción), hasta unas 4 horas y media después de la asimilación de datos (los últimos). La Figura 4.3 representa la proyección de los puntos de la rejilla sobre la superficie terrestre. El número de niveles verticales, determinados por coordenadas σ , es de 26. El número total de datos por proyección horizontal de la rejilla, es decir, sin tener en cuenta la componente vertical, es de 329. Para todos los puntos de la rejilla (considerando uno distinto en cada uno de los 26 niveles verticales) aporta los valores de la componente este-oeste y norte-sur del viento, con lo que determina su velocidad y dirección.

Para la zona de España la resolución horizontal del modelo GFS 1° es de unos 100 km (distancia entre dos puntos contiguos). Evidentemente esta resolución no es suficiente para poder predecir la dinámica del viento en la superficie terrestre (a unas decenas de metros de la superficie terrestre) para cualquier lugar no coincidente con los puntos de análisis, lo que obliga a utilizar un modelo atmosférico con mayor resolución espacial.

4.2.2. Modelos atmosféricos locales.

Los modelos atmosféricos globales aportan información de la evolución de la atmósfera para todo el planeta, pero en ocasiones sólo se necesita información para las zonas continentales (la información para las zonas oceánicas no tiene tanto interés). Esto dio lugar al nacimiento de los modelos atmosféricos llamados regionales, que abarcan en su análisis grandes zonas: Europa, Norteamérica, etc. Siguen siendo modelos atmosféricos de macro-escala pero con la desventaja frente a los modelos atmosféricos

globales de necesitar, además de las condiciones iniciales, unas condiciones de contorno en las fronteras del dominio. Estas condiciones de contorno se consiguen interpolando los valores de salida de un modelo atmosférico global. La ventaja que tienen frente a los modelos atmosféricos globales es el menor esfuerzo computacional en su ejecución (por estar limitados espacialmente).

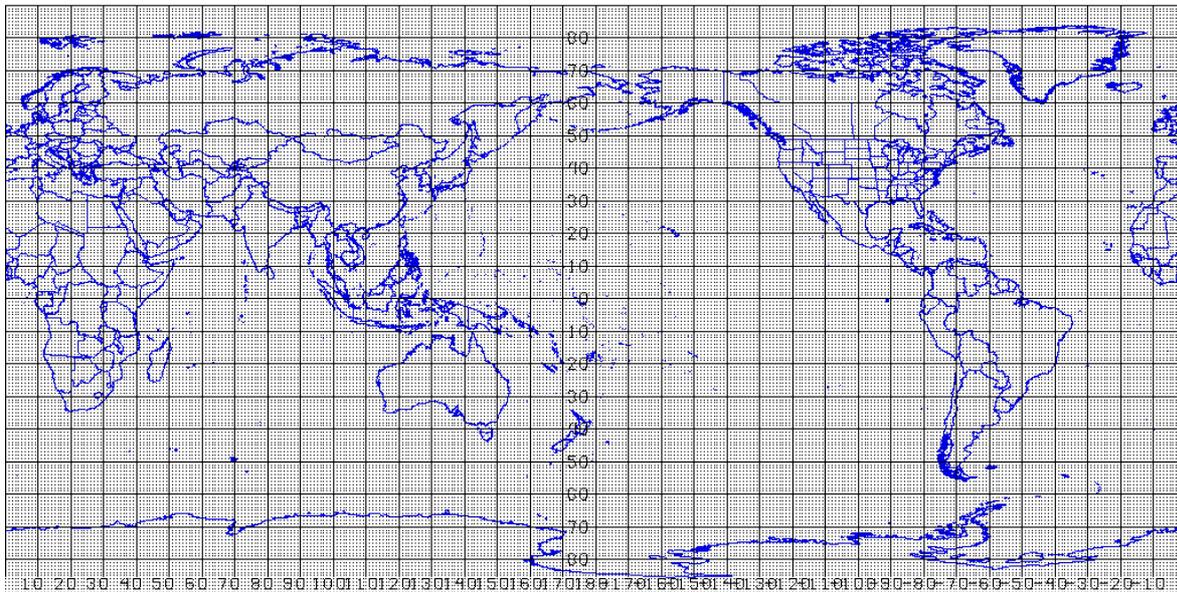


Figura 4.3. Proyección de los puntos de la rejilla de análisis para el modelo atmosférico global GFS 1°.

El modelo atmosférico HIRLAM, citado en el capítulo anterior, corresponde a un modelo atmosférico regional (o global de área limitada). Su resolución horizontal es de 16 km (distancia entre puntos de la rejilla), y su alcance temporal 72 horas. Presenta también características similares a los modelos atmosféricos de mayor resolución como el anidamiento (que se describirá más adelante), permitiéndole resoluciones espaciales de unos 5 km y alcances de 36 horas.

Pero la predicción de la dinámica del viento, en el corto plazo, para una localización geográfica concreta exige la utilización de un modelo atmosférico con mayor resolución espacial que la de los modelos atmosféricos de macro-escala, aunque sean modelos atmosféricos regionales. La solución está en la utilización de modelos atmosféricos meso-escalares.

Los modelos atmosféricos meso-escalares, dada su mayor resolución horizontal, suelen trabajar con coordenadas cartesianas, si bien para la coordenada vertical se utiliza la coordenada σ . Estos modelos atmosféricos no usan la aproximación hidrostática (sin evolución vertical) que utilizan los modelos atmosféricos macro-

escalares, lo que obliga a parametrizar esa posible evolución y definir mucho mejor las características del terreno sobre el que se realizan las predicciones atmosféricas.

Son varios los modelos atmosféricos no hidrostáticos desarrollados por la comunidad científica, alguno de ellos implementados en software de acceso público. Entre estos pueden citarse el modelo atmosférico WRF (Weather Research and Forecast Model), el MM5 (Penn State/NCAR Mesoscale Model, Version 5), COAMPS (US Navy's Coupled Ocean/Atmosphere Mesoscale Prediction System), ARPS (Advanced Regional Prediction System), MASS (Mesoscale Atmospheric Simulation System), MC2 (Mesoscale Compressible Community Model), Meso-NH (Mesoscale no-Hidrostatic Atmospheric Simulation System) y MEMO (Mesoscale Model). La mayoría de los modelos atmosféricos están programados en Fortran, con algunas rutinas en C, y pensados para ser ejecutados en ordenadores con sistema operativo Unix.

Una característica muy interesante que incluyen muchos de los modelos atmosféricos citados, es la posibilidad de anidamiento. Cuando se pretende obtener una gran resolución espacial (lo que se necesita para la predicción de la velocidad del viento), sobre una zona geográfica concreta, el mayor problema viene impuesto por la necesidad de establecer las condiciones de contorno para las fronteras del dominio (zona geográfica estudiada). La baja resolución de los modelos atmosféricos globales hace que estas condiciones de contorno, en dominios muy pequeños, no queden bien definidas. Por ejemplo, la resolución espacial del modelo atmosférico GFS 1° para la zona de España es de unos 100 km; si queremos utilizar este modelo atmosférico para obtener las condiciones de contorno para un dominio de 50x50 km, a lo sumo dispondremos de dos puntos sobre los que interpolar, lo cual difícilmente podrá servir para realizar una simulación verosímil de la evolución atmosférica.

El anidamiento consiste en definir unos dominios menores, con mayor resolución espacial, dentro de otros dominios. Volviendo al ejemplo anterior, el dominio de 50x50 km puede corresponder a un dominio anidado a otro mayor 150x150 km, que a su vez está anidado con otro mayor de 450x450 km, que a su vez está anidado con otro de 1350x1350 km. El dominio mayor toma las condiciones de contorno de un modelo global. Para la latitud de España significará unos 13 puntos distintos, por lo que los puntos del dominio de 1350x1350 km situados en la frontera presentarán valores diferenciados, resultado de la interpolación con 13 puntos, lo que permitirá obtener una simulación más precisa. Si el dominio de 1350x1350 km tiene una rejilla de análisis con una resolución horizontal de 45 km, puede transferir unas buenas condiciones de contorno al siguiente dominio de 450x450 km, y así sucesivamente.

El anidamiento de dominios puede ser unidireccional o bidireccional. El anidamiento es unidireccional cuando la información pasa del dominio de mayor dimensión al de menor dimensión. Por el contrario, el anidamiento es bidireccional cuando la transferencia de información (resultados de las integraciones de las ecuaciones primitivas) es en ambos sentidos.

4.3. EL MODELO ATMOSFÉRICO MM5

El modelo atmosférico MM5 es un modelo atmosférico de meso-escala, desarrollado por investigadores de la universidad estatal de Pennsylvania y el National Center for Atmospheric Research (PSU/NCAR), desde la década de los 80 del siglo pasado. Es un modelo no-hidrostatico, con una definición de coordenadas verticales σ . Dado que estas coordenadas “siguen” la altura del terreno sobre el que se está haciendo la simulación atmosférica, éste tiene que estar perfectamente definido a la resolución espacial del modelo atmosférico (orografía, uso del suelo, vegetación dominante, etc.).

El modelo atmosférico MM5 puede utilizar anidamientos múltiples, con transferencia de información entre dominios de tipo unidireccional o bidireccional, en función de la opción escogida por el usuario. Éste, puede también escoger entre una gran cantidad de opciones en la parametrización del modelo atmosférico y en las simulaciones a realizar [Dudhia05].

El modelo atmosférico está implementado en más de 100000 líneas de código programado en Fortran 77, Fortran 90 y C. El modelo atmosférico está soportado por un conjunto de programas, que pueden ser compilados para funcionar de forma paralela en varias máquinas (un “cluster” de máquinas). Los programas más importantes son:

- Terrain: programa para interpolar horizontalmente en longitud y latitud, la altura del terreno y el uso del suelo (vegetación mayoritaria) sobre los dominios elegidos en la simulación. Este programa crea una nueva rejilla tridimensional que será la utilizada en la simulación posterior de la evolución atmosférica.
- Regrid: es un programa que lee los ficheros de salida de un modelo atmosférico global con los resultados para las variables meteorológicas para cada uno de los niveles de presión (coordenadas σ), e interpola estos

resultados en la rejilla horizontal del dominio definido por Terrain.

- Interpf: este programa establece las condiciones de contorno del dominio a partir de los resultados de algún modelo atmosférico global. Estas condiciones son calculadas por interpolación, seguida por un análisis de coherencia de valores y corrección de los mismos.
- MM5: este programa es el corazón del modelo atmosférico con el mismo nombre. En él se realiza la resolución de las ecuaciones primitivas para todos los puntos de la rejilla definida por Terrain. La integración se realiza por diferencias finitas, a partir de un paso de integración definido por el usuario, quien también define el horizonte final de predicción y el intervalo de tiempo entre predicciones intermedias.

El programa MM5 permite el anidamiento de dominios, si bien con la condición de que los vértices que definen el dominio de mayor resolución correspondan a puntos de la rejilla del dominio “padre” (dominio de menor resolución, pero mayor área abarcada). La relación recomendada entre tamaño de la proyección horizontal del dominio padre y el de mayor resolución es 3:1, haciendo que el paso de integración en el dominio padre sea 3 veces mayor que el del dominio que contiene. Pueden anidarse dominios en varios niveles, tal como se muestra en la Figura 4.4, donde se puede observar cuatro dominios creados con Terrain para la ejecución del modelo MM5 en una zona de la región de La Rioja. El dominio indicado con “D1” en la Figura 4.4, es el dominio padre para el indicado con “D2”, que a su vez lo es para el indicado con “D3”, y éste es, a su vez, dominio padre para el indicado con “D4”. La resolución del último dominio, “D4” es de aproximadamente 1 km (separación horizontal entre puntos de la rejilla de análisis).

Concretamente, para los cuatro dominios representados en la Figura 4.4, el tamaño de las rejillas y resoluciones para cada dominio son los indicados a continuación.

1. Dominio D1: rejilla de 86x80 y resolución aproximada 27 km.
2. Dominio D2: rejilla de 100x100 y resolución aproximada 9 km.
3. Dominio D3: rejilla de 45x50 y resolución aproximada 3 km.
4. Dominio D4: rejilla de 31x40 y resolución aproximada 1 km.

El programa MM5 guarda los resultados, para todos los puntos de la rejilla de

cada uno de los dominios, y para cada uno de los intervalos temporales hasta alcanzar el horizonte final de predicción, en un conjunto de ficheros de texto en formato GRIB con el nombre MMOUT_DOMAINx, donde la x corresponde al número del dominio. Las variables meteorológicas guardadas incluyen las dos componentes del viento (este-oeste y norte-sur) para cada uno de los puntos (para cada punto en el plano y para todos los niveles de presión). Adicionalmente, el programa MM5 ofrece la posibilidad de generar una serie temporal de variables meteorológicas (entre ellas las dos componentes del viento a 10 metros de altura o menor nivel de presión), para cada uno de los pasos de integración hasta alcanzar el horizonte final. El paso de integración para el menor de los dominios (“D4”) es de apenas unos segundos.

Para determinar el valor de predicción de cualquiera de las variables meteorológicas generadas en los ficheros MMOUT_DOMAIN, para una localización concreta que no coincida con ninguno de los puntos de análisis en ninguno de los dominios, puede realizarse por medio de interpolación entre los cuatro puntos más cercanos de la rejilla del dominio de mayor resolución que contenga la localización concreta. La interpolación usada es la llamada bilineal, expresada por la ecuación (4.1), donde $f(x,y)$ corresponde al valor de predicción de la variable deseada para una localización dada por las coordenadas horizontales x e y ; x_i y x_{i+1} corresponden al valor de la coordenada x para cada pareja de puntos de la rejilla más próximos; análogamente y_j e y_{j+1} corresponden al valor de la coordenada y para cada pareja de puntos de la rejilla más próximos.

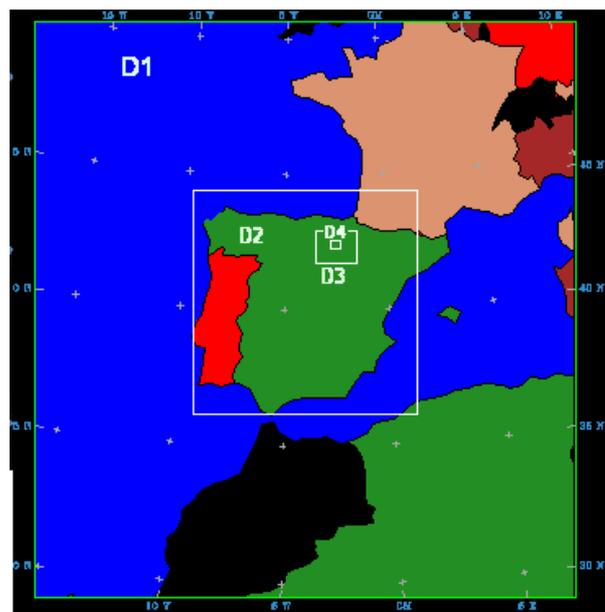


Figura 4.4. Anidamiento de dominios para la ejecución del modelo atmosférico MM5 con una elevada resolución espacial para una zona de La Rioja

$$f(x, y) = (1-t)(1-u)f(x_i, y_j) + t(1-u)f(x_{i+1}, y_j) + t \cdot u f(x_{i+1}, y_{j+1}) + u(1-t) f(x_i, y_{j+1}) \quad (4.1)$$

Los valores de t y u son los expresados por las ecuaciones (4.2) y (4.3), respectivamente; y las condiciones que debe cumplir el punto de coordenadas (x, y) , que son tener como puntos más próximos los de coordenadas (x_i, y_j) , (x_{i+1}, y_j) , (x_i, y_{j+1}) y (x_{i+1}, y_{j+1}) , están expresadas por relaciones de (4.4).

$$t = \frac{x - x_i}{x_{i+1} - x_i} \quad (4.2)$$

$$u = \frac{y - y_j}{y_{j+1} - y_j} \quad (4.3)$$

$$\begin{aligned} x_i < x < x_{i+1} \\ y_j < y < y_{j+1} \end{aligned} \quad (4.4)$$

Por último, resaltar las experiencias en la puesta a punto del modelo atmosférico MM5 en las instalaciones del laboratorio de investigación de Planificación, Operación y Control de Sistemas de Energía Eléctrica, del Departamento de Ingeniería Eléctrica de la Universidad de La Rioja. Los programas informáticos que conforman el modelo atmosférico MM5 fueron instalados y compilados en 2005. Se instaló inicialmente en una máquina Pentium IV, 1.5 GHz, 256 MB de RAM, bajo el sistema operativo Linux Red Hat 9. Se utilizaron compiladores de Fortran 90 de Intel (gratuitos para usos académicos). A finales del año 2006 se migró a un nuevo ordenador con doble procesador Xeon, 3GHz y 4 GB de RAM, con el mismo sistema operativo anterior.

Se utilizan datos del modelo atmosférico global GFS 1°, correspondiente a los datos atmosféricos de las 0 horas GMT, para establecer las condiciones de contorno e interpolar las variables a los puntos de las rejillas de análisis de cada dominio. Los ficheros con los resultados del modelo atmosférico GFS 1°, para horizontes de predicción de 0 a 180 horas, en intervalos de 3 horas, son descargados del servidor ftp del NCEP americano cada mañana. El tamaño total de los ficheros del modelo atmosférico GFS 1° es de 1.56 GB. Para limitar el tamaño de los ficheros se desarrolló un programa informático para la reducción de los resultados del modelo atmosférico global y dejar exclusivamente los de los puntos de la rejilla de este modelo atmosférico

global que abarquen la zona comprendida entre los meridianos 30° Oeste y 20° Este, y los paralelo 20° Norte y 60° Norte. Bajo estas condiciones, toda la información con resultados del modelo atmosférico GFS 1° para un horizonte de predicción de hasta 180 horas ocupa unos 56.6 MB.

En el caso de los dominios de la Figura 4.4, el tiempo de ejecución con el procesador Xeon arriba citado (para obtener las predicciones atmosféricas con un horizonte de hasta 72 horas), es de 12 horas, lo que representa 1 hora de cálculo computacional para cada 6 horas de predicciones atmosféricas. Se puede reducir el tiempo de computación mediante clusters de computadores: si se compila el programa MM5 para ser ejecutado en paralelo en varios ordenadores de una misma red de comunicación, los tiempos de cálculo y procesado se reducirán proporcionalmente al número de ordenadores trabajando en paralelo (supuestas características similares).

4.4. CONCLUSIONES

El desarrollo tecnológico de computadores cada vez más potentes ha permitido la implementación, en forma de programas, de los modelos numéricos de predicción atmosférica (modelos atmosféricos). Estos programas permiten la asimilación masiva de datos meteorológicos para poder determinar las condiciones iniciales de la atmósfera, tarea esencial para poder predecir el futuro estado meteorológico de la misma.

Existen distintos modelos atmosféricos operativos, que se diferencian fundamentalmente en la resolución espacial y temporal. Cuanto mayor es la resolución espacial de un modelo atmosférico, menor es su alcance temporal (plazo de validez de las predicciones). Los modelos atmosféricos clasificados como de macro-escala son los que tienen una resolución espacial menor. Están implementados en los llamados modelos atmosféricos globales o en los grandes modelos atmosféricos regionales.

Para la predicción de la velocidad del viento en una localización geográfica concreta (situación de un parque eólico), los modelos atmosféricos que ofrecen una mayor resolución son los modelos atmosféricos meso-escalares, alguno de ellos implementado en software de libre distribución. El modelo atmosférico MM5 es uno de estos modelos con mayor aceptación en la comunidad científica. Proporciona, como resultado de la simulación, el valor de determinadas variables meteorológicas para todos los puntos de la rejilla tridimensional de análisis, y para todos los intervalos temporales definidos por el usuario. Dentro de las múltiples parametrizaciones que admite, para

poder particularizar la simulación dinámica de la evolución de la atmósfera para cualquier caso, ofrece la de generar series temporales de varias variables meteorológicas, entre ellas el viento.

Actualmente se sigue trabajando en la predicción de variables meteorológicas, utilizando el modelo atmosférico MM5 instalado en los computadores del laboratorio de investigación de Planificación, Operación y Control de Sistemas de Energía Eléctrica, perteneciente al Área de Ingeniería Eléctrica de la Universidad de La Rioja, para aplicaciones en la predicción a corto plazo de la producción eléctrica de parques eólicos.

5. MODELOS PARA LA PREDICCIÓN DE PRODUCCIÓN ELÉCTRICA EN PARQUES EÓLICOS

En este capítulo se presentan las técnicas más usuales en el diseño de modelos para la predicción de futuros valores de series temporales, incidiendo en sus características, ventajas y desventajas. Se ha considerado conveniente incluir en el mismo capítulo las técnicas clásicas más importantes (modelos ARIMA), con otras técnicas meta-heurísticas, históricamente recientes (redes neuronales artificiales y sistemas de inferencia difusa), entroncadas en el conocido mundo del llamado “soft-computing”. Tras una breve descripción de las características de cada una de las técnicas, y de los tipos de modelos que permiten diseñar, se comenta su aplicación a la predicción a corto plazo de la producción eléctrica en parques eólicos.

5.1. SERIES TEMPORALES

La predicción del futuro ha sido uno de los sueños del hombre desde que lo podemos considerar como tal. En la historia de la humanidad podemos encontrar múltiples ejemplos de oráculos, augures, visionarios, etc. que, más bien con poco acierto, vaticinaban acerca del futuro con un gran impacto social. Evidentemente, ese sueño de conocer el futuro arranca de la ventaja que supone evitar tomar decisiones bajo una situación de incertidumbre, arranca de la ventaja de conocer de antemano las consecuencias que nuestras acciones o decisiones van a tener. Pero los tiempos pasados no son una excepción; ese sueño continúa, y más cuando entran en juego cuestiones económicas. De hecho, ha sido en el campo económico donde las técnicas de predicción de valores futuros de una serie temporal han sido más intensamente utilizadas.

Una serie temporal (también conocida como serie histórica, cronológica o de tiempo) es un conjunto de valores, correspondientes a una variable, ordenados según el instante en que fueron tomados, y normalmente tomados a intervalos regulares de tiempo. Las ventas mensuales de vehículos nuevos en España o el consumo diario de agua en la ciudad de Logroño, son dos ejemplos de serie temporal. Los sucesivos valores contenidos en una serie temporal se designan por un número de índice, correspondiente al número de observación (número del instante en que fue tomado el valor). Así, suponiendo una serie temporal X , está vendría expresada como indica la ecuación (5.1).

$$X = \{X_1, X_2, X_3, \dots\} \quad (5.1)$$

La predicción del valor futuro de una serie temporal significa determinar los siguientes valores que tomará la variable que forma la serie temporal a partir del instante presente. Esa determinación del valor futuro de la variable puede realizarse desde dos perspectivas indicadas seguidamente.

1. La predicción de valores futuros se puede realizar por extrapolación de los valores pasados.
2. Se puede establecer un modelo matemático que explique las relaciones entre los valores de la variable en el pasado y utilizar ese modelo para predecir los futuros valores de la variable.

Precisamente nos centraremos en la segunda perspectiva comentada, la de extraer información del pasado. Presentaremos el modelo que mejor explique las relaciones entre los valores pasados de la variable, para ser utilizado posteriormente en la predicción de los valores futuros.

Pero, en muchas ocasiones se dispone de otras informaciones que pueden ser útiles en la determinación de los valores futuros de una variable, de la que disponemos de una serie temporal. La variable que tratamos de predecir puede ser función de otras variables cuyos valores podemos conocer, lo que las hace variables “explicativas” de la variable a predecir. En función de que usemos, o no, esa información con los valores de las posibles variables explicativas, podemos distinguir dos métodos para establecer los valores futuros de la variable a predecir señalados a continuación.

- Métodos de predicción univariante (también conocidos como métodos de series temporales): estos métodos pretenden establecer la relación entre los propios valores pasados de la variable a predecir. Una vez establecida la relación, con el modelo se realizan proyecciones hacia el futuro para obtener predicciones de la variable. Matemáticamente, el valor en un futuro inmediato de la variable a predecir es una función de sus valores pasados, como indica la ecuación (5.2).

$$X_{t+1} = f(X_t, X_{t-1}, \dots, X_{t-d}) \quad (5.2)$$

- Métodos de predicción multivariante (también conocidos como métodos

causales): estos métodos pretenden establecer la relación entre los valores pasados de la variable a predecir y los valores, también pasados, de las posibles variables explicativas. Una vez establecida la relación (determinado el modelo) puede ser utilizada para obtener valores futuros de la variable a predecir. La ecuación (5.3) expresa el valor en un futuro inmediato de la variable a predecir como una función de los valores futuros de otras dos variables (valores futuros, pero conocibles de antemano), u y v .

$$X_{t+1} = f(u_{t+1}, v_{t+1}) \quad (5.3)$$

En el caso más general de métodos multivariantes, el valor futuro de la variable a predecir puede ser función de los valores pasados de la propia variable, del valor futuro de las variables explicativas y de los valores pasados de estas variables, tal y como indica la ecuación (5.4).

$$X_{t+1} = f(X_t, X_{t-1}, \dots, X_{t-d}, u_{t+1}, u_t, u_{t-1}, \dots, u_{t-m}, v_{t+1}, v_t, v_{t-1}, \dots, v_{t-n}) \quad (5.4)$$

Entre las diferentes técnicas utilizadas para establecer el modelo de predicción para la producción eléctrica de parques eólicos, en los trabajos de investigación recogidos en este documento se han utilizado tres:

- la metodología ARIMA, también conocida como Box-Jenkins [Box76],
- las redes neuronales artificiales, y
- los sistemas de inferencia difusa.

5.2. MODELOS ARIMA

La metodología ARIMA es el método más usado para identificar y construir modelos de predicción para series temporales univariantes y estacionarias. Los modelos ARIMA incluyen tres componentes: el auto-regresivo (AR), el integrado (I) y el de movimiento a la media o media móvil (MA). Un modelo ARIMA queda determinado por el orden de cada una de sus tres componentes: el orden del término auto-regresivo, p , el orden del término integrador o diferenciador, d , y el orden del término de movimiento a la media, q . Así, un modelo puede expresarse, y quedar totalmente definido, como ARIMA(p, d, q). En las próximas páginas se definirá brevemente cada uno de esos tres términos del modelo general ARIMA, para describir finalmente el

modelo completo, incluyendo la posible estacionalidad. También se perfilarán las bases de los modelos ARIMA multivariantes.

5.2.1. Modelos auto-regresivos.

Decimos que un modelo es auto-regresivo si el valor de la variable en un instante cualquiera de tiempo puede ser explicado por los valores en los instantes anteriores más un término de error. Así, la ecuación (5.2) responde a lo que podría ser un modelo auto-regresivo, con la única particularidad de que los modelos ARIMA son modelos lineales, por lo que la ecuación (5.2) se transforma en la (5.5), en la que los factores ϕ representan los coeficientes del modelo, p el orden y e_t el término de error para el instante t .

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t \quad (5.5)$$

El modelo representado por la ecuación (5.5) quedaría definido como un modelo AR(p), es decir un modelo auto-regresivo de orden p . La ecuación (5.5) puede escribirse de forma abreviada como indica la ecuación (5.6), donde $\phi_p(B)$ se conoce como el operador polinomial de retardos, cuya equivalencia viene expresada en la ecuación (5.7).

$$\phi_p(B)X_t = \phi_0 + e_t \quad (5.6)$$

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (5.7)$$

El operador retardo, B , aplicado al valor de una variable X_t , determina el valor de dicha variable un periodo antes, como muestra la ecuación (5.8); aplicando p veces B a la variable X_t se obtiene el valor de dicha variable p periodos antes, como muestra la ecuación (5.9).

$$BX_t = X_{t-1} \quad (5.8)$$

$$B^p X_t = X_{t-p} \quad (5.9)$$

Para el modelo auto-regresivo más simple, el AR(1), existe un límite para el

coeficiente ϕ_1 llamado el límite de estacionariedad, de tal forma que se debe cumplir la condición (5.10) para que la serie temporal representada por el modelo AR(1) sea estacionaria.

$$|\phi_1| < 1 \quad (5.10)$$

Por otra parte, el término de error, e_t , de las ecuaciones (5.5) y (5.6), en el caso de estar correctamente modelada la variable, corresponde a lo que se conoce como un “ruido blanco”, esto es, presenta, a lo largo del tiempo, una media nula, su varianza es constante y la covarianza es nula entre errores de instantes de tiempo diferentes.

El valor de predicción de un modelo AR(p) viene dado por la ecuación (5.11), donde \hat{X}_t representa el valor de predicción para la variable X en el periodo t .

$$\hat{X}_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} \quad (5.11)$$

Los coeficientes del modelo pueden calcularse por el método de mínimos cuadrados siempre que el número de observaciones (valores) de la serie temporal sea superior al doble del número de coeficientes a determinar.

5.2.2. Modelos de media móvil.

Un modelo de media móvil es aquel en que el valor de la variable para un instante t es función de un término independiente y de una sucesión ponderada de errores correspondientes a los instantes precedentes. Estos modelos se denotan con las siglas MA, seguidos, como en el caso de los modelos auto-regresivos, del orden entre paréntesis. Así, un modelo con q términos de error, MA(q), responde a la ecuación (5.12), donde e_{t-k} representa el error para la variable X en el instante $t-k$ (valor del error para el instante $t-k$), y μ una constante numérica.

$$X_t = \mu - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_k e_{t-k} - \dots - \theta_q e_{t-q} + e_t \quad (5.12)$$

Utilizando la misma notación utilizada en la ecuación (5.6) con el operador retardo, B , la ecuación (5.12) se transforma en la (5.13), donde el valor del operador polinomial de retardos viene expresado en la ecuación (5.14).

$$X_t = \theta_q(B)e_t + \mu \quad (5.13)$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (5.14)$$

5.2.3. Modelos ARMA.

La combinación de los modelos auto-regresivo y de media móvil produce el modelo conocido como ARMA (auto-regresivo y de media móvil). Si combinamos las ecuaciones (5.5) y (5.12), que definen los modelos AR(p) y MA(q), obtendremos la expresión del modelo ARMA(p, q), tal como expresa la ecuación (5.15).

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} + e_t \quad (5.15)$$

El término μ que aparecía en el modelo MA(q) queda ahora integrado en el coeficiente ϕ_0 . La ecuación (5.15) puede expresarse de una forma abreviada, tal como aparece en la ecuación (5.16) utilizando los operadores retardo. Las expresiones correspondientes a $\phi_p(B)$ y a $\theta_q(B)$ son las expresadas en las ecuaciones (5.7) y (5.14), respectivamente. Muchos autores suelen prescindir en la notación de la ecuación (5.16) del término ϕ_0 , como en [Wei90].

$$\phi_p(B)X_t = \phi_0 + \theta_q(B)e_t \quad (5.16)$$

Los modelos ARMA, como vimos en el Capítulo 3, han sido utilizados en la predicción de la velocidad media horaria del viento en el muy corto plazo.

5.2.4. Modelos ARIMA.

Los modelos ARMA, caso general de un modelo AR o MA, presentan una deficiencia importante y es la imposibilidad de conservar en las predicciones un comportamiento periódico o cuasi-periodico, si la serie temporal lo incluye. El comportamiento periódico significa que hay patrón que se repite en la serie temporal con un periodo que es múltiplo del que existe entre observaciones consecutivas de la serie temporal. Muchos autores se refieren a este comportamiento periódico como estacionalidad.

Una serie temporal que presenta ese comportamiento periódico o una serie temporal que presenta una tendencia determinística y, por tanto, no es estacionaria, puede transformarse en una serie representada por un modelo ARMA si se le aplica el operador diferencia. En el fondo convertimos la serie temporal original X (la de la variable que queremos predecir) en otra, Z , que ya no presentará ese comportamiento periódico o que es ya estacionaria, tal como muestra la ecuación (5.17), donde τ representa el periodo.

$$Z_t = X_t - X_{t-\tau} \quad (5.17)$$

Lo más usual es emplear esta técnica de diferenciación para transformar en estacionarias las series temporales que no lo son, lo que en la práctica se reduce a aplicar el operador diferencia una o dos veces. Con este operador diferencia aplicado al modelo ARMA, se define el modelo ARIMA que incluye los tres términos citados anteriormente, el auto-regresivo, el integrado (correspondiente al operador diferencia ahora comentado) y el de media móvil. Un modelo ARIMA(p, d, q) responde a la ecuación (5.18), donde d representa el número de diferencias aplicadas hasta convertir la serie temporal en estacionaria.

$$\phi_p(B)(1-B)^d X_t = \phi_0 + \theta_q(B)e_t \quad (5.18)$$

En un modelo ARIMA, una vez establecido el número de diferencias necesarias para convertir la serie temporal en estacionaria, d , los parámetros de los términos auto-regresivo y de media móvil se estiman por procedimientos de mínimos cuadrados no lineales.

5.2.5. Modelos SARIMA.

A pesar de que el operador diferencia permitía eliminar el comportamiento periódico de una serie temporal, no era suficiente para poder modelar series temporales que incluyeran también tendencias en la estacionalidad (por ejemplo, el patrón de comportamiento periódico va incrementándose con el tiempo). Para esto, en la ecuación (5.18) se incluyen unos nuevos términos para representar ese comportamiento estacional, siendo el índice de estacionalidad, s , el periodo estacional, que es múltiplo del que existe en observaciones sucesivas de la variable en la serie temporal X . Los nuevos términos introducidos representan a un comportamiento auto-regresivo, con los valores retrasados en el tiempo un múltiplo del índice de estacionalidad, una

diferenciación con dichos valores y una media móvil referida también a dichos valores. El modelo resultante, expresado en la ecuación (5.19) se designa como modelo SARIMA (modelo ARIMA con estacionalidad multiplicativa) o $ARIMA(p, d, q)(P, D, Q)_s$.

$$\phi_p(B)\Phi_P(B^s)(1-B^s)^D(1-B)^d X_t = \phi_0 + \theta_q(B)\Theta_Q(B^s) e_t \quad (5.19)$$

5.2.6. Identificación de modelos ARIMA.

En [Box76], [Wei90], [DeLurgio98] y [Makridakis98] se puede encontrar una detallada descripción del método para identificar el orden de los procesos o componentes que definen el modelo ARIMA: p, d, q, P, D y Q . El índice de estacionalidad vendrá determinado por la naturaleza de la variable de la serie temporal. Por ejemplo, para el caso de la predicción de la velocidad media del viento el índice de estacionalidad, en el caso de que pueda ser considerado, será el correspondiente a 24 horas. La serie temporal de velocidad media horaria del viento no presentará, en general, componentes estacionales, excepto para los casos donde aparezca un fenómeno repetitivo, como es el fenómeno de las brisas, cuya repetición (dirección del viento durante una parte del día en un sentido, y en sentido contrario durante otra parte), día tras día, se puede modelar como una componente estacional.

Básicamente, los pasos que hay que seguir para identificar un modelo ARIMA son los indicados seguidamente.

1. Representar gráficamente la serie temporal. Esto permite evidenciar outliers y un comportamiento no estacionario en la media o en la varianza. Los outliers pueden ser sustituidos por otros valores más ajustados al resto de valores, el comportamiento no estacionario en la varianza puede corregirse por una transformación (logarítmica, Box-Cox, etc.) y el de la media puede corregirse aplicando diferenciación.
2. Examinar las gráficas de la función de auto-correlación (ACF) y de la función de auto-correlación parcial (PACF) para confirmar si hace falta aplicar otra diferenciación para alcanzar la estacionareidad de la serie temporal.
3. Cuando se ha alcanzado la estacionareidad de la serie, observar de nuevo las gráficas de ACF y PACF para determinar si queda algún patrón en la misma: si

hay que considerar estacionalidad (los valores de la ACF y de la PACF para múltiplos del índice de estacionalidad son significativamente diferentes de cero). A partir de las gráficas de ACF y PACF se puede determinar también los valores de los parámetros p y q del modelo ARIMA.

Algunos expertos consideran la selección de la estructura del modelo ARIMA como un arte más que una ciencia [Wei90], evidenciando que se requiere mucha experiencia para poder ajustar bien la estructura más adecuada del modelo. Para aquellos casos en que es posible ajustar varios modelos, cumpliendo las condiciones de estacionariedad de la serie y cuando además las gráficas de ACF y PACF no quedan suficientemente claras como para determinar inequívocamente el orden de los parámetros p y q (órdenes de las componentes auto-regresivas y de media móvil), puede utilizarse un criterio de selección adecuado entre los posibles modelos. Uno de estos criterios es el criterio de información de Akaike (AIC) [Akaike76], que penaliza a aquellos modelos que requieren un mayor número de parámetros. Una de las leyes básicas en la elaboración de modelos de predicción es el principio de parsimonia (si una serie temporal puede ser explicada por dos modelos, el más simple es el mejor). En el caso de usar el error medio cuadrático de los residuos (diferencia entre valores de la serie temporal y el valor obtenido aplicando el modelo), el valor del AIC corresponde a la ecuación (5.20), donde el MSE corresponde al error medio cuadrático de los residuos, n al número de observaciones y m al número de parámetros del modelo ($m=p+q+P+Q$) [DeLurgio98]. Se selecciona aquel modelo que presenta un menor valor del AIC.

$$AIC = n \ln(MSE \cdot n) + 2m \quad (5.20)$$

5.2.7. Modelos ARIMA multivariantes.

Cuando se pretende predecir los nuevos valores de varias series temporales, cuyos valores están relacionados entre sí, puede realizarse mediante modelos ARIMA multivariantes. Los modelos multivariantes tienen expresiones similares a los univariantes, con la excepción de usar vectores y matrices (vectores para las series temporales, matrices para los coeficientes).

Estos modelos no van a ser usados en los modelos de predicción de la producción eléctrica de un parque eólico, ya que en los casos de modelos de predicción de la segunda familia (según la clasificación del Capítulo 3), las otras series temporales

de valores para las variables meteorológicas son obtenidas independientemente mediante un modelo atmosférico.

5.2.8. Modelos con variables exógenas.

Cuando un modelo ARIMA incluye como variable explicativa otra serie temporal distinta a la que queremos modelar, el modelo recibe el nombre de ARIMA con variable exógena o externa, ARIMAX (algunos autores los han denominado modelos de regresión dinámica). Con este tipo de modelos, la variable para la que construimos el modelo es función de sus valores pasados y de los valores presentes y pasados de la variable o variables explicativas.

En los siguientes párrafos se va a comentar el caso más general de relación de una variable (la variable a predecir) con una variable explicativa. Los modelos ARIMA, entendidos como los modelos que pueden tener todas o alguna de las componentes citadas (auto-regresiva, integradora y de media móvil) no son más que casos particulares del modelo general.

Si X es la variable que queremos predecir, y U es la variable explicativa, se puede expresar la relación entre ambas por medio de la ecuación (5.21) [Ljung99], donde aparece también la variable error, e . El valor de los operadores polinomiales de retardo, $A(B)$, $H(B)$, $I(B)$, $J(B)$ y $K(B)$ aparece reflejado en las ecuaciones (5.22).

$$A(B) X_t = \frac{H(B)}{K(B)} U_t + \frac{I(B)}{J(B)} e_t \quad (5.21)$$

$$\begin{cases} A(B) = 1 + a_1 B + a_2 B^2 + \dots + a_{na} B^{na} \\ H(B) = h_1 B + h_2 B^2 + \dots + h_{nh} B^{nh} \\ I(B) = 1 + i_1 B + i_2 B^2 + \dots + i_{ni} B^{ni} \\ J(B) = 1 + j_1 B + j_2 B^2 + \dots + j_{nj} B^{nj} \\ K(B) = 1 + k_1 B + k_2 B^2 + \dots + k_{nk} B^{nk} \end{cases} \quad (5.22)$$

La ecuación (5.21) representa el caso más general de la relación entre las dos variables. Esta relación vendría determinada por los parámetros na , nh , ni , nj y nk , además de todos los coeficientes de los operadores polinomiales. En función de que alguno de los parámetros fuera nulo, se pueden particularizar distintos modelos que

reciben denominaciones como ARX (auto-regresivo con variable exógena), ARMAX (auto-regresivo con media móvil y variable exógena), ARIMAX (auto-regresivo integrado con media móvil y variable exógena) o “output error” (cuando $J(B) = I(B)$).

Para identificar un modelo con estructura de este tipo, se diseñan modelos con diferentes valores para los parámetros, determinándose posteriormente el más adecuado por medio de algún criterio, como el AIC (existen otros criterios que también pueden usarse como el criterio de información bayesiano, BIC).

5.2.9. Ventajas e inconvenientes de los modelos ARIMA.

La principal ventaja que presentan los modelos ARIMA es la posibilidad que ofrecen de proporcionar intervalos de confianza en sus predicciones con una base estadística fiable. Por otra parte, son unos modelos ampliamente difundidos en los ámbitos de predicción económica, por lo que es fácil encontrar herramientas informáticas para su desarrollo.

Como inconvenientes se pueden citar que son modelos lineales, lo que hace difícil que puedan modelar suficientemente bien todos los problemas (el caso de la producción eléctrica en parques eólicos tiene características de no linealidad). Además, la adecuada selección de la estructura del modelo exige una gran experiencia por parte del desarrollador. También se puede citar que son modelos muy sensibles a outliers, lo que obliga a revisar la serie temporal para detectarlos y corregirlos antes de diseñar el modelo de predicción.

5.2.10. Aplicación de los modelos ARIMA para la predicción de la producción eléctrica en parques eólicos en el corto plazo.

En el Capítulo 3 se han mostrado varios ejemplos de artículos publicados en la literatura internacional en los que se usan modelos ARIMA, fundamentalmente para la predicción de la velocidad media del viento en el corto o muy corto plazo [Kamal97], [Schlink98], [Sfetsos00], [Sfetsos02], [Poggi03] y [Torres05]. Todos estos ejemplos corresponden a modelos, de los clasificados en dicho capítulo, como pertenecientes a la primera familia de modelos (sólo usan la serie temporal de la variable). Son por tanto modelos univariantes.

También aparece un ejemplo con un modelo de la segunda familia. En

[Sanchez06] se utiliza la metodología ARIMA para construir varios modelos cuyos valores de predicción son utilizados posteriormente para calcular, por un procedimiento de medias ponderadas con coeficientes variables con el tiempo, el valor final de predicción. Los modelos usados corresponden a auto-regresivos puros y a auto-regresivos con variables exógenas (predicciones de un modelo atmosférico).

Así, los modelos ARIMA constituyen una opción a considerar en el diseño de modelos de predicción de la producción eléctrica en parques eólicos, al menos como un modelo más con el que poder comparar los nuevos modelos propuestos en los trabajos de investigación recogidos en este documento.

5.3. REDES NEURONALES ARTIFICIALES

Los modelos ARIMA, descritos en el anterior apartado, presentan como una limitación el hecho de que sean modelos lineales, lo que los hace relativamente imprecisos en la predicción de variables de naturaleza no lineal. En la literatura especializada se puede encontrar una gran colección de modelos, derivados de los creados bajo la metodología de los modelos ARIMA, pero de carácter no lineal. Así existen modelos auto-regresivos no-lineales (NAR), modelos auto-regresivos no-lineales de media móvil (NARMA), etc. Estos últimos dos tipos de modelos están basados en las series de Volterra, y aunque son modelos no lineales, son modelos lineales en los parámetros, lo que hace que se puedan estimar éstos con los mismos métodos que se emplean en los modelos lineales (por ejemplo por mínimos cuadrados).

Pero incluso los modelos NARMA tienen limitaciones: sólo son precisos representando un tipo específico de no linealidades, por lo que no se pueden aplicar a cualquier tipo de problema de predicción. Las redes neuronales artificiales superan esa limitación ya que son auténticos “aproximadores universales” de funciones incluyendo, evidentemente, las no lineales.

Las redes neuronales artificiales son sistemas inspirados en los sistemas neuronales biológicos (cerebro animal). Las redes neuronales artificiales (en lo sucesivo redes neuronales) constan de un conjunto de elementos simples de procesamiento, llamados neuronas o nodos, conectados entre sí por conexiones que tienen asignado valor numérico modificable llamado peso. Cada neurona puede recibir información de otras neuronas (que actúan para ella como entradas) o de una entrada externa, y a su vez generar información en una salida que puede servir de información de entrada a otras

neuronas. El procesamiento que cada neurona hace con la información que recibe consiste en generar un valor (información) de salida que es función de una combinación lineal de los valores de sus entradas, tal como recoge la ecuación (5.23).

En la ecuación (5.23), y_i representa el valor de la salida de la neurona i , x_j es el valor de la entrada j (valor de la salida de otra neurona, también designada con el índice j); w_{ij} es el peso de la conexión entre las neuronas i y j ; θ_i es el umbral o sesgo (bias) de la neurona; y f_i es la función de activación. Esta función de activación puede ser lineal, pero usualmente es de tipo no lineal (sigmoideal, gaussiana, función de Heaviside, etc.). Esa no linealidad de la función de activación constituye la base para que una red neuronal pueda ser el aproximador universal de funciones no lineales. La estructura de una neurona de una red neuronal es la mostrada en la Figura 5.1.

$$y_i = f_i \left(\sum_{j=1}^n w_{ij} x_j - \theta_i \right) \quad (5.23)$$

Las redes neuronales pueden dividirse en redes unidireccionales (feedforward) y redes recurrentes, en función de cómo estén realizadas las conexiones entre neuronas: en las redes unidireccionales las neuronas están conectadas siguiendo el sentido de entrada hacia salida; las redes recurrentes tienen conexiones entre neuronas en el sentido contrario. Formalmente, una red neuronal unidireccional sería aquella en que numerando correlativamente las neuronas comenzando por las que constituyen la entrada de la red y acabando por las que constituyen las salidas, no hay ninguna neurona que tenga como entrada la salida de otra neurona a la que se le haya asignado un número mayor al suyo. En las redes recurrentes es imposible realizar ese esquema de numeración.

La estructura de una red neuronal viene determinada por la disposición de las neuronas y de sus funciones de activación. Las neuronas se agrupan en capas, considerándose como “ocultas” aquellas que no corresponden a la capa de entradas (las neuronas que actúan como entrada de señales o valores externos) y la capa de salidas (las neuronas que proporcionan su salida al exterior). El conjunto de valores que se presentan a las neuronas de la capa de entradas (un valor por neurona), correspondientes a las entradas de la propia red neuronal, recibe la denominación de valor del vector de entradas.

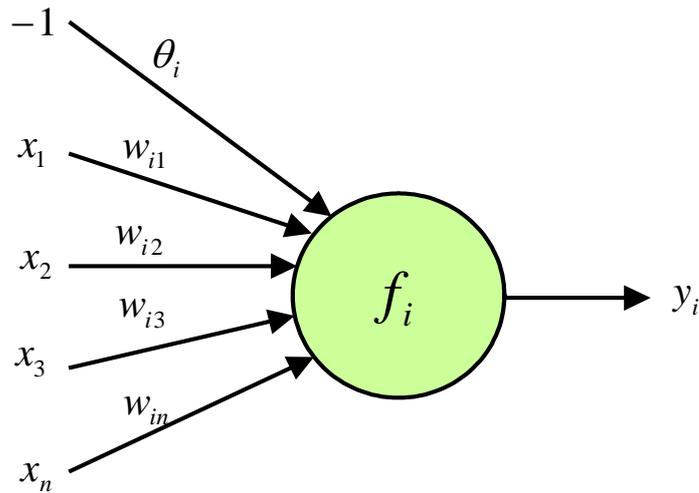


Figura 5.1. Procesamiento de la información en una neurona.

La principal característica de las redes neuronales es su capacidad de aprendizaje. Este aprendizaje consiste en el ajuste de los pesos de las conexiones entre neuronas de una forma iterativa para conseguir un efecto deseado, el de “aprender” las relaciones existentes entre pares de valores de vectores de entradas-salidas que se les suministren, o encontrar un patrón común en datos de entrada para poderlos clasificar de acuerdo a los distintos patrones. Se pueden distinguir tres tipos de aprendizajes en redes neuronales comentados seguidamente.

- **Aprendizaje supervisado:** se suministran a la red neuronal tanto el valor del vector de entradas, como el valor del vector de salidas (valor del vector de salidas esperado correspondiente a cada valor del vector de entradas). Durante el proceso de aprendizaje, la red neuronal ajusta sucesivamente el peso de las conexiones de forma que se minimiza el error entre los valores del vector de salidas de la red neuronal y los valores del vector de salidas esperado. Este tipo de aprendizaje es usual en aproximación de funciones no lineales con redes neuronales.
- **Aprendizaje no supervisado:** la red neuronal no recibe ningún valor esperado del vector de salidas (de ahí lo de aprendizaje no supervisado), pero busca en los valores del vector de entradas alguna propiedad o correlación que los permita clasificar de acuerdo a la misma.
- **Aprendizaje reforzado:** es un caso especial de aprendizaje supervisado en el que el valor del vector de salidas esperado no es suministrado: únicamente se

suministra a la red neuronal la información precisa de si el valor del vector de salidas es igual o no al esperado.

En el desarrollo de los trabajos de investigación de este documento se han utilizado los tipos de redes neuronales más prometedores para la predicción de la producción eléctrica en parques eólicos: redes neuronales descritas en los próximos apartados.

5.3.1. Perceptrones multicapa.

El perceptrón multicapa o red neuronal MLP (del inglés “Multilayer Perceptron”) es un tipo de red neuronal unidireccional en el que las neuronas están organizadas en capas, de forma que una neurona situada en una capa intermedia sólo puede tener como entradas las salidas de la capa precedente y su salida sirve de entrada para neuronas de la siguiente capa. Su representación gráfica es la recogida en la Figura 5.2, en la que las neuronas de color amarillo forman la capa de entradas (cada neurona de la capa de entradas recibe un valor externo), las neuronas de color naranja forman una capa intermedia o capa oculta, y las neuronas rojas forman la capa de salidas (el valor de salida de cada una de las neuronas constituye una de las salidas de la red neuronal).

Una red neuronal MLP puede tener varias capas ocultas, si bien con una sola capa oculta sería capaz de aproximar, con error arbitrariamente pequeño, cualquier función continua (lineal o no) acotada, y con dos capas ocultas cualquier función continua [Cybenko89]. En los trabajos de investigación de este documento se han utilizado redes neuronales MLP de una capa oculta y sólo en los casos de predicción más complejos se han usado redes neuronales MLP con dos capas ocultas.

La red neuronal MLP puede utilizarse en aplicaciones de clasificación (dado un valor del vector de entradas, determinar a que clase o grupo predefinido pertenecen) o de aproximación de funciones (aprovechando su característica de no linealidad de las funciones de activación de las neuronas).

El diseño de una red neuronal MLP implica la selección del número de capas ocultas y del número de neuronas de cada capa. Cuanto mayor sea el número de capas y de neuronas, mayor será la capacidad de la red neuronal MLP de ajustar cualquier

función, pero aumentará el tiempo necesario para entrenarla (para que realice el aprendizaje) y, sobre todo, se corre el riesgo de sobre-entrenar la red. El sobre-entrenamiento se traduce en que la red neuronal ajusta perfectamente los valores que se le han dado en el vector de entradas, proporcionando el valor deseado del vector de salidas o un valor relativamente muy próximo, pero se ha “adaptado” demasiado a esos datos perdiendo su capacidad de generalización en la predicción. Cualquier valor del vector de entradas diferente de los utilizados en el entrenamiento de la red neuronal puede dar resultados pobres.

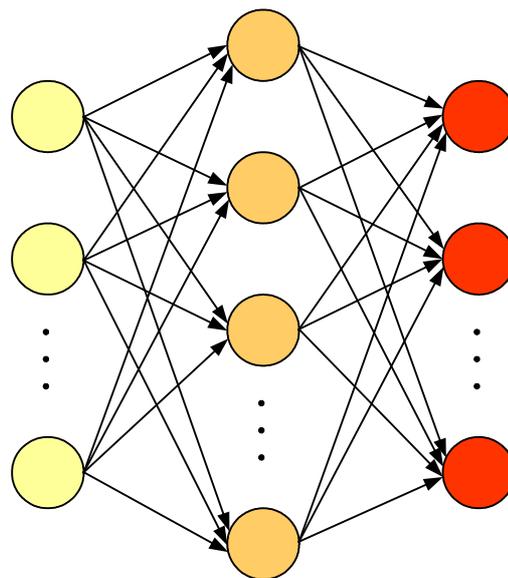


Figura 5.2. Perceptrón multicapa con una capa oculta.

Existen diversas técnicas para el aprendizaje o entrenamiento de las redes neuronales MLP. El más extendido es el de retropropagación del error (backpropagation) [Haykin94], que ha sido adaptado al entrenamiento de más tipos de redes neuronales, aparte de las MLP. En las próximas líneas se comenta brevemente el caso de una red neuronal con varias entradas y varias salidas, y una única capa oculta, si bien la generalización a cualquier número de capas ocultas es inmediata. El conjunto de valores, en un instante dado, de las entradas es lo que llamamos valor del vector de entradas. El conjunto de valores, en un instante dado, de las salidas es lo que llamamos valor del vector de salidas.

La técnica de retropropagación de error consiste, básicamente, en actualizar los pesos de las conexiones entre neuronas proporcionalmente a su contribución al error. Existen dos versiones en la aplicación de esta técnica, la conocida como versión “on-line” (también conocida como “instance-based”) y la versión “off-line” (también

llamada “batch”). En la versión “on-line” de entrenamiento, la actualización de los pesos de la red neuronal se realiza tras la presentación de cada valor del vector de entradas. En la versión “off-line”, la actualización de los pesos de la red neuronal se realiza tras la presentación de todo el conjunto de valores del vector de entradas de los datos de entrenamiento (lo que constituye un “epoch”). En los trabajos de investigación recogidos en este documento se ha utilizado siempre este segundo método.

Si la actualización de los pesos de la red neuronal se realiza tras la presentación de cada valor del vector de entradas (entrenamiento “on-line”), la variación del peso para el valor del vector de entradas n , $\Delta w_{ij}(n)$, de la conexión entre las neuronas i y j (neurona i situada en una capa posterior a la de la neurona j), viene dada por la ecuación (5.24), donde η corresponde al llamado factor de aprendizaje (una constante numérica, también llamada “step”), δ_j es el gradiente del error local en la neurona j , $o_i(n)$ es el valor de la salida de la neurona i para el valor del vector de entradas n , α es un coeficiente llamado momento y $\Delta w_{ij}(n-1)$ representa la variación del peso de la conexión entre las neuronas i y j para el valor del vector de entradas $n-1$.

$$\Delta w_{ij}(n) = \eta \delta_j o_i(n) + \alpha \Delta w_{ij}(n-1) \quad (5.24)$$

El gradiente de error local para una neurona de la capa de salidas es el producto de la derivada de la función de error escogida (normalmente el error cuadrático) y de la función de activación de la neurona. El gradiente de error local para una neurona de una capa oculta es igual a la suma ponderada de los gradientes de error local de las neuronas a las que conectan las conexiones que salen de dicha neurona (el peso de cada conexión actúa de factor de ponderación). El valor del peso de la conexión entre las neuronas i y j para el siguiente valor de vector de entradas, $(n+1)$, viene dado por la ecuación (5.25). La expresión para la actualización de los pesos para neuronas de capas ocultas es muy similar a la ecuación (5.24). Puede consultarse en [Haykin94].

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (5.25)$$

En la ecuación (5.24), aparecen dos coeficientes, η y α , que juegan un papel especialmente importante en el entrenamiento de la red neuronal. Ambos son dos constantes numéricas, con una importancia primordial en el entrenamiento de la red neuronal. Un alto valor del factor de aprendizaje, η , hace que la red neuronal reaccione rápidamente a cambios en las entradas, pero hace que la red neuronal sea inestable si es

demasiado alto. Por el contrario, un factor de aprendizaje bajo hace que el tiempo necesario para el entrenamiento de la red neuronal se incremente.

El coeficiente α o momento describe la porción de cambio de peso que se añade en la siguiente actualización de pesos. Un bajo valor del momento hace que el entrenamiento se vuelva inestable, ocasionando oscilaciones en los valores de los pesos de las conexiones e impidiendo el aprendizaje de la red neuronal. El máximo valor para el momento es la unidad, y los valores usuales los comprendidos entre 0.5 y 0.9.

Para evitar el sobre-ajuste o sobre-entrenamiento (over-fitting) de la red neuronal, existen varias técnicas como la limitación en el número de epochs en el proceso de entrenamiento (“early-stopping”), la adición de ruido blanco sobre los valores de las variables de entrada, o la utilización de la validación cruzada, siendo esta última la técnica más extendida. Con esta última técnica se dividen los datos disponibles en dos grupos: entrenamiento y validación cruzada. Los datos del grupo de entrenamiento se utilizan para el ajuste de los parámetros de la red neuronal utilizando el método de entrenamiento elegido. Los datos del grupo de validación cruzada no se utilizan en la determinación del error que se pretende minimizar con el entrenamiento de la red neuronal, sino que se presentan a la red neuronal tras cada epoch de entrenamiento. El proceso de entrenamiento se interrumpe cuando haya transcurrido un número predeterminado de epochs sin reducción del error con los datos de validación cruzada.

En la predicción de nuevos valores para series temporales, la red neuronal MLP puede sacar partido de su capacidad de aproximar funciones no lineales, permitiendo diseñar fácilmente modelos auto-regresivos no lineales: sólo es necesario utilizar como variables de entrada para la red neuronal los valores pasados de la propia variable que forma la serie temporal.

5.3.2. Red neuronal unidireccional multicapa generalizada.

La red neuronal unidireccional multicapa generalizada, FFBP (de feedforward backpropagation, ya que éste es el método de entrenamiento usualmente utilizado), comparte las características de la red neuronal MLP, ya que está formada por capas de neuronas con conexiones unidireccionales (de entradas a salidas), pero a diferencia de la red neuronal MLP permite que haya conexiones entre neuronas de capas no consecutivas. Así, para el caso de una red neuronal FFBP con una sola capa oculta, las

neuronas de la capa de entradas también se conectan a las neuronas de la capa de salidas (en la red neuronal MLP no sería posible), tal como muestra la Figura 5.3.

Para el caso de una red neuronal FFBP con dos capas ocultas, las neuronas de la capa de entradas se conectan a las dos capas ocultas y a la capa de salidas; las neuronas de la primera capa oculta se conectan a la segunda capa oculta y a la capa de salidas; y las neuronas de la segunda capa oculta se conectan solamente a la capa de salidas.

La ventaja de la red neuronal FFBP sobre la red neuronal MLP es que, para una misma estructura (capas y neuronas), necesita menos tiempo en el entrenamiento al adaptarse más rápidamente los pesos de las conexiones de las neuronas de las primeras capas de la red.

5.3.3. Red neuronal modular.

La red neuronal modular constituye otra variante sobre la red neuronal MLP. Una red neuronal modular está compuesta de varias redes neuronales multicapa en paralelo (sub-redes neuronales), es decir, comparten la capa de entradas y la capa de salidas, pero las capas ocultas son independientes. El valor del vector de salidas de la red neuronal modular se obtiene por combinación lineal de los valores de los vectores de salidas de las sub-redes neuronales.

En el entrenamiento, cada sub-red neuronal puede “especializarse” en ajustar los valores de su vector de salidas para un conjunto de valores del vector de entradas, lo que mejora el funcionamiento de la red neuronal en su conjunto. Por otra parte, para un mismo número total de neuronas, el entrenamiento de una red neuronal modular es más rápido que el de una red neuronal MLP, ya que el número de conexiones entre neuronas es menor para la red neuronal modular (al no existir conexiones entre neuronas de sub-redes neuronales distintas).

Pueden implementarse diferentes redes neuronales modulares haciendo que cada sub-red neuronal sea MLP o FFBP. Evidentemente, no se permiten conexiones de las neuronas de una sub-red neuronal a las de otra.

Una variante, relativamente más compleja, es la descrita en [Weigend96], donde un conjunto de redes expertas (neuronales o no) compiten entre sí, siendo una red no lineal la encargada de seleccionar la red experta que debe activarse en cada periodo de

tiempo (para cada valor del vector de entradas). Esta última red no lineal es la llamada red maestra.

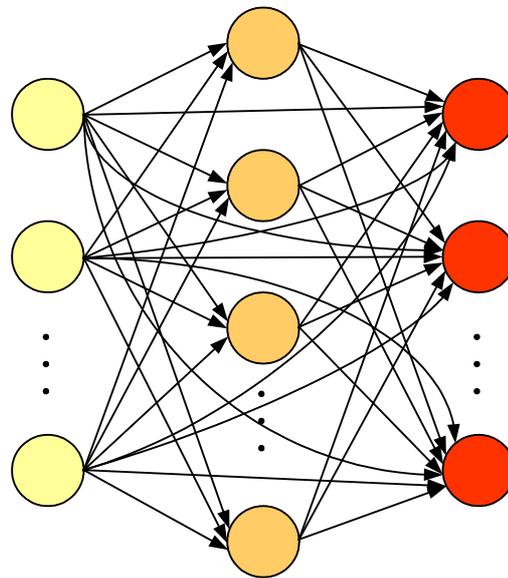


Figura 5.3. Red neuronal unidireccional multicapa generalizada, FFBP, con una capa oculta.

5.3.4. Red neuronal de función de base radial.

Se trata de un tipo de red neuronal que se caracteriza por tener un aprendizaje híbrido, es decir, combina el aprendizaje no supervisado con el aprendizaje supervisado. La red neuronal con función de base radial, RBF (radial basis function), es una red con tres capas de neuronas: capa de entradas, capa oculta y capa de salidas. Además, es una red neuronal unidireccional y sólo se permiten conexiones entre neuronas de capas consecutivas en el orden de entrada a salida, tal como muestra la Figura 5.4. Su diferencia fundamental respecto a la red neuronal MLP está en la función de activación de las neuronas de la capa oculta.

En una red neuronal MLP, las neuronas de la capa oculta realizan una suma ponderada de los valores de las entradas (el factor de ponderación son los pesos de las conexiones) y aplica a esta suma una función de activación, generalmente de tipo sigmoideal. En una red neuronal RBF las neuronas de la capa oculta calculan la distancia euclídea entre el valor del vector de entradas y un valor del vector previamente almacenado en cada neurona, que recibe el nombre de centro, y sobre esa distancia aplica una función con forma gaussiana.

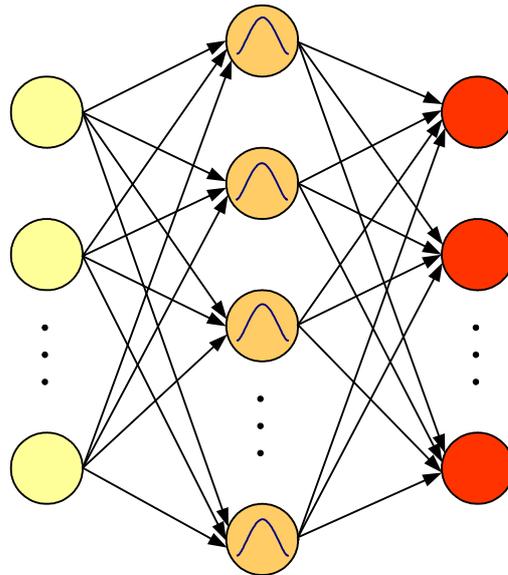


Figura 5.4. Red neuronal de función de base radial.

Para que cada una de las neuronas de la capa oculta puedan calcular esa distancia euclídea, hace falta que almacene previamente el valor del centro. Esto se consigue por un aprendizaje no supervisado aplicando el algoritmo “K-means” [Bishop95], que trata de reflejar en los centros las agrupaciones naturales entre los valores del vector de entradas. Con este método de aprendizaje, y establecido previamente el número de neuronas en la capa oculta, es decir, el número de centros, se selecciona un número igual de valores del vector de entradas para actuar de centros. Cada valor del vector de entradas pertenecerá a un único grupo, correspondiente al del centro más cercano (en términos de distancia euclídea), y el centro de cada grupo es el centroide (valor medio) de los valores del vector de entradas que pertenecen a dicho centro. La varianza de las funciones gaussianas se calcula como una fracción del valor medio de la distancia euclídea a los centros más cercanos.

Puesto que una función gaussiana es una función no lineal, una única capa oculta es suficiente para aproximar cualquier tipo de función [Park89], si bien para los trabajos de investigación de este documento, y con el fin de mejorar la capacidad de aproximación a funciones no lineales, se ha dispuesto de una o dos capas ocultas adicionales. Estas capas ocultas son similares a las de una red neuronal MLP, es decir, con función de activación sigmoideal en cada neurona.

Así, el entrenamiento de una red neuronal RBF consta de dos fases. En una primera fase se realiza el aprendizaje no supervisado (no se le aporta a la red neuronal el valor esperado del vector de salidas) por medio del algoritmo K-means. En la segunda

fase se realiza un aprendizaje supervisado, donde se actualizan los pesos de las neuronas de la capa de salidas (con función de activación lineal) y los de las capas ocultas (con función de activación sigmoïdal), si es que estas capas han sido utilizadas.

La ventaja de las redes neuronales RBF frente a las redes neuronales MLP se centra en su mayor rapidez en el entrenamiento y en la necesidad de un menor número de neuronas para aproximar con igual grado de exactitud una función no lineal.

5.3.5. Redes neuronales de análisis de componentes principales.

En muchos sistemas de procesamiento de información es necesario extraer las características principales e inherentes de un conjunto de entradas de alta dimensión. Esta extracción de información de características principales conduce a una reducción de la dimensión de los datos. Un ejemplo típico puede ser el procesamiento de una imagen: la gran cantidad de datos que significan todos los píxeles de la imagen puede reducirse a un conjunto menor que contenga lo realmente relevante y específico de la imagen.

Una de las técnicas propuestas para la extracción genérica de esas características es el análisis de componentes principales, PCA, técnica estadística que trata de encontrar las componentes principales de una población o conjunto de datos [Jolliffe02].

Un ejemplo simple de reducción de las dimensiones sería el representado en la Figura 5.5. Disponemos de un sistema con dos variables, x_1 y x_2 , y un conjunto de observaciones de ambas variables representadas por puntos en el plano (parte izquierda de la figura). La misma información (mismos puntos), pero con un solo eje está representada en la parte derecha, que contiene una dimensión menos que antes (ahora es suficiente una variable para determinar cada punto; antes necesitábamos dos).

Para un caso general n -dimensional, las componentes principales son las proyecciones sobre las direcciones principales o de mayor varianza. La eliminación de direcciones donde la varianza es muy pequeña para la población, se traduce en una reducción de dimensión en la representación de los datos. Además esa reducción es óptima en el sentido del error medio cuadrático (es posible la reconstrucción de los datos originales con un error MSE mínimo).

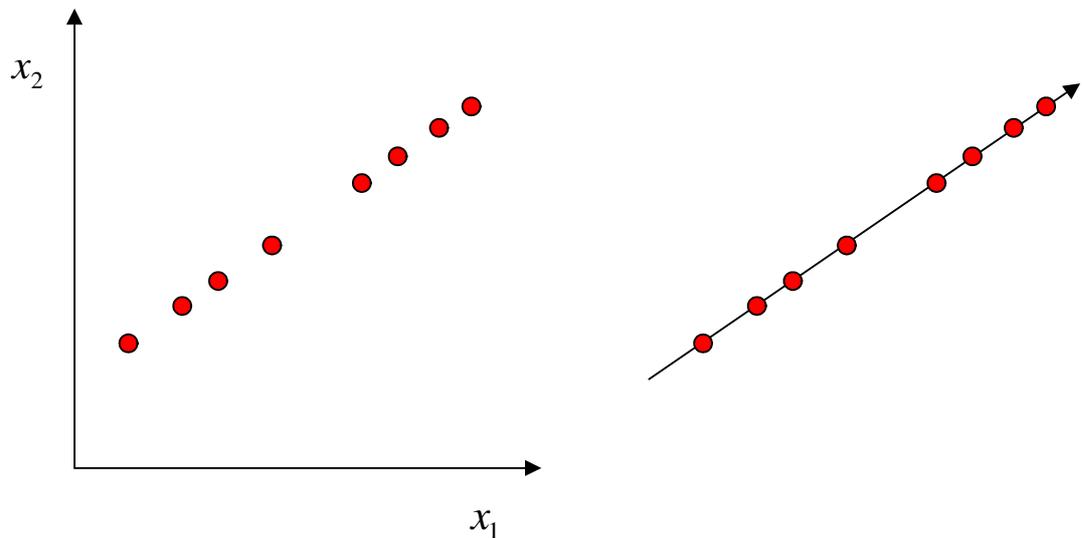


Figura 5.5. Ejemplo de reducción dimensional.

Las redes neuronales con análisis de componentes principales (en lo sucesivo redes neuronales PCA) realizan esa reducción de la dimensión del espacio de entrada (reducción de la dimensión del vector de entradas) mediante la proyección de la información de D neuronas de entrada sobre N neuronas de una capa oculta, tal como muestra la Figura 5.6. La capa de neuronas ocultas, representadas en color naranja, tendrán como vector de salidas las N primeras componentes principales correspondientes a las D variables de entrada. El valor del vector de salidas se calcula por combinación lineal, es decir, sin función de activación.

La reducción del espacio de entrada se puede realizar aplicando la regla de Oja [Oja89], o el método de Sanger [Sanger89]. En ambos casos se trata de un aprendizaje no supervisado en el que se realiza el ajuste de los pesos de las conexiones entre la capa de entradas de la red neuronal y la primera capa oculta, de forma que las salidas de esta primera capa oculta corresponden al valor de las N componentes principales del vector de entradas.

Las salidas de esta primera capa oculta pueden servir como capa de entradas de una red neuronal MLP. El resultado obtenido es una red neuronal MLP que utiliza menos entradas que las del conjunto de la red neuronal PCA, ya que con la aplicación del análisis de componentes principales se ha producido una reducción efectiva en el número de variables de entrada (reducción de dimensiones), con la ventaja de que los valores de estas componentes principales carecen de correlación entre sí.

El entrenamiento de este tipo de redes neuronales incluye una primera fase con

entrenamiento no supervisado, para adecuar los pesos de las conexiones entre las dos primeras capas, de forma que se obtengan a la salida el valor de las N primeras componentes principales correspondientes al vector de entradas. A esta primera fase le sigue un entrenamiento supervisado para ajustar los pesos de la red neuronal MLP que sigue a esa primera capa oculta.

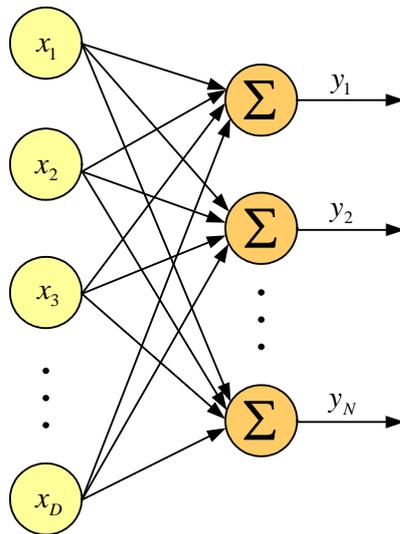


Figura 5.6. Principio de la red neuronal PCA.

La ventaja de las redes neuronales PCA sobre las redes neuronales MLP es que están menos afectadas por el problema conocido como “curso de dimensionalidad” [Härdle90]: el número de valores del vector de entradas distintos necesarios para entrenar adecuadamente una red neuronal crece exponencialmente con el número de dimensiones del vector de entradas (con el número total de entradas de la red neuronal). En una red neuronal MLP, un número limitado de valores del vector de entradas hace que la red neuronal no ajuste bien los pesos de las conexiones entre neuronas. La reducción de dimensiones conseguida con la red neuronal PCA ayuda a evitar este problema.

5.3.6. Redes neuronales recurrentes.

Las redes neuronales recurrentes, como comentábamos en páginas anteriores, son aquellas que tienen conexiones entre neuronas que no siguen la dirección de capa de entradas a capa de salidas. Existen varios tipos de redes neuronales recurrentes, algunos de los cuales han sido utilizados en los trabajos de investigación recogidos en este documento. Estos tipos son los que van a ser brevemente descritos en los próximos

apartados.

5.3.6.1. Redes neuronales de Elman.

Las redes neuronales de Elman [Elman90] son un tipo de redes recurrentes, que el propio Elman describió como “redes recurrentes simples”. Las redes neuronales de Elman parten de la estructura de una red neuronal MLP pero añaden una capa oculta adicional, llamada capa de neuronas de contexto, cuya misión es extraer información temporal de los datos. Esta extracción se realiza copiando el contenido de una capa oculta en la capa de neuronas de contexto. La capa de neuronas de contexto tiene como entradas sus propias salidas y las salidas de las neuronas de la capa oculta. La capa oculta tiene como entradas las neuronas de la capa de entradas y las neuronas de la capa de contexto. Las salidas de las neuronas de la capa oculta son las entradas de la capa de salida. La Figura 5.7. muestra una red neuronal de Elman con una capa oculta: en color amarillo se representan las neuronas de la capa de entrada, en color rojo las neuronas de la capa de salida, en color naranja las neuronas de la capa oculta y en color azul las neuronas de contexto.

El valor de la salida de una neurona de contexto para el valor del vector de entradas n , viene dado por la ecuación (5.26), en la que τ representa la llamada constante de tiempo y $x(n)$ representa el valor de entrada de la neurona cuando se presenta el n -ésimo valor del vector de entradas.

$$y(n) = \sum_{i=0}^{n-1} x(n-i) \tau^i \quad (5.26)$$

Así, los valores previos de las entradas de las neuronas de contexto tienen un peso que va decayendo exponencialmente. La constante de tiempo es un valor comprendido entre 0 y 1. Su valor tiene influencia directa en la “capacidad de memorización” de la red neuronal. Un bajo valor de la constante de tiempo (por ejemplo 0.1) hace que el peso de las últimas entradas se reduzca muy rápidamente en función de su antigüedad. El valor óptimo de la constante de tiempo es $1-(1/D)$, donde D es la llamada profundidad de memoria, que corresponde al número de valores previos del vector de entradas de la red neuronal que ésta habrá “memorizado”.

La ventaja de la red neuronal de Elman es su capacidad de actuar dinámicamente: las salidas de la red neuronal no dependen del valor del vector de

entradas en un instante determinado, sino también de los valores del vector de entradas previos, representando relaciones temporales que una red neuronal MLP no podría recoger. Como desventaja se puede citar que la profundidad de memoria necesaria para un problema concreto no siempre es conocida de antemano. Además la constante de tiempo (que limita la profundidad de memoria) es un valor constante que no se adapta con el tiempo.

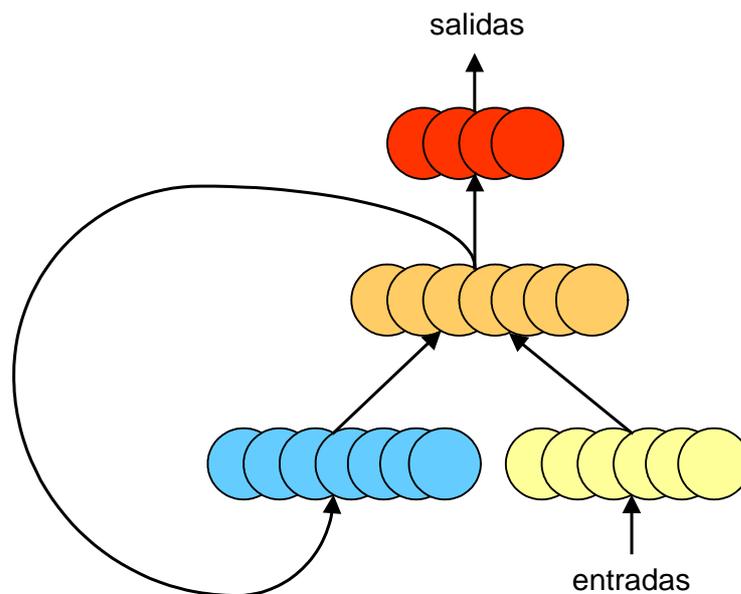


Figura 5.7. Red neuronal de Elman con una capa oculta.

5.3.6.2. Redes neuronales retardadas en el tiempo.

Las redes neuronales retardadas en el tiempo constituyen un tipo de redes neuronales recurrentes en las que se establecen conexiones entre neuronas de una misma capa [Principe00]. Se puede distinguir entre redes neuronales retardadas en el tiempo focalizadas y no focalizadas. Una red neuronal retardada en el tiempo focalizada es aquella en la que las conexiones entre neuronas de la misma capa se producen en la capa de entradas. En las redes no focalizadas las conexiones entre neuronas de la misma capa se producen en la capa de entradas y en las capas ocultas.

La Figura 5.8 representa una red neuronal retardada en el tiempo focalizada. Este tipo de red se conoce como TDNN focalizada (Time Delay Neural Network). En dicha figura está representada una red neuronal TDNN focalizada con tres neuronas “de entrada” en la capa de entradas (de color amarillo), es decir, con tres entradas. Hay una

capa oculta (color naranja) y una capa de salidas (color rojo). Pero lo específico de la red neuronal TDNN focalizada es que la capa de entradas cuenta con varias neuronas más que las entradas de la red neuronal: el número de neuronas en la capa de entradas es un múltiplo de las entradas de la red. En la Figura 5.8 se ha representado una red neuronal con tres entradas, pero la capa de entradas presenta nueve neuronas, tres grupos de tres neuronas. El primer grupo de tres neuronas recibe el valor de las entradas de la red. El segundo grupo de tres neuronas de la capa de entradas almacena el valor previo del vector de entradas de la red neuronal (esas tres neuronas están unidas a través de unidades de retardo, designadas por z^{-1} , a las tres neuronas que reciben el valor del vector de entradas de la red neuronal). El tercer grupo de tres neuronas almacena los valores previos de las neuronas del segundo grupo (por tanto, el valor del vector de entradas de la red neuronal con una antigüedad de dos unidades de tiempo). Así, la red TDNN, y en general las redes neuronales retardadas en el tiempo, cuentan con una estructura que le permite mantener una memoria de la actividad de las neuronas de la red con valores previos del vector de entradas.

Una red neuronal TDNN focalizada, con k unidades de retardo y dos entradas $x_1(t)$ y $x_2(t)$, es equivalente a una red neuronal MLP que tuviese por entradas $x_1(t)$, $x_2(t)$, $x_1(t-1)$, $x_2(t-1)$, $x_1(t-2)$, $x_2(t-2)$, ..., $x_1(t-k)$ y $x_2(t-k)$. La red neuronal TDNN no focalizada presenta unidades de retardo no sólo en la capa de entrada, sino también en las capas ocultas, es decir, tiene memoria de los valores previos de las entradas y de los valores previos en las neuronas de las capas ocultas.

Las redes neuronales retardadas en el tiempo tienen como principal aplicación el procesamiento de información que cambia con el tiempo, como puede ser la predicción de series temporales. Con una única entrada externa, (el valor previo de la variable de la serie temporal), la red neuronal TDNN configura un modelo auto-regresivo no lineal, más simple en el caso de la red neuronal focalizada y más complejo en el caso de la red neuronal no focalizada.

Además de la red neuronal TDNN existen otras configuraciones para las redes neuronales retardadas en el tiempo, según sea la estructura asociada a las neuronas que implementan la memoria de la red neuronal. Una de las configuraciones usa la estructura de memoria Gamma [DeVries92], por lo que se denominan redes neuronales con retardo en el tiempo Gamma. La configuración del retardo entre neuronas de memoria para una red neuronal con retardo en el tiempo Gamma es la mostrada en la Figura 5.9, donde se representa la salida de la neurona de memoria correspondiente a un retardo k (y_0 representaría la salida de la neurona de entrada). El funcionamiento de la

unidad de retardo, z^{-1} , hace que el valor presente en su entrada pase a su salida con la presentación a la red neuronal de un nuevo valor del vector de entradas.

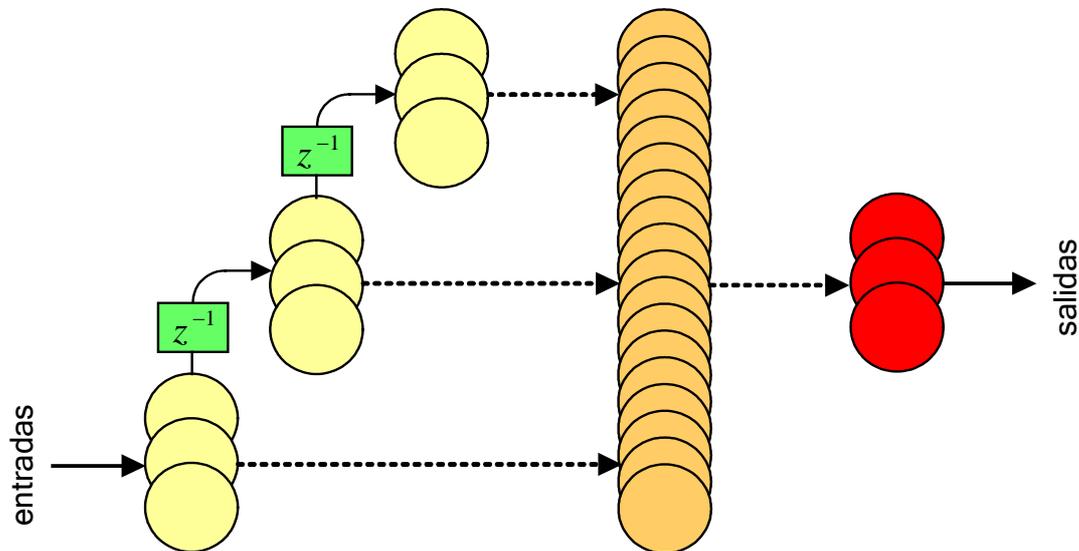


Figura 5.8. Red neuronal TDNN focalizada.

El funcionamiento de la memoria Gamma guarda parecido con la neurona de contexto de la red de Elman. En la memoria Gamma, el valor de entrada se multiplica por un factor, μ , llamado factor de realimentación y se le suma el valor previo de salida de la neurona multiplicado por $(1-\mu)$. El valor de la salida de la neurona correspondiente al retardo k , cuando a la red neuronal se le presenta el n -ésimo valor del vector de entradas, es el expresado en la ecuación (5.27).

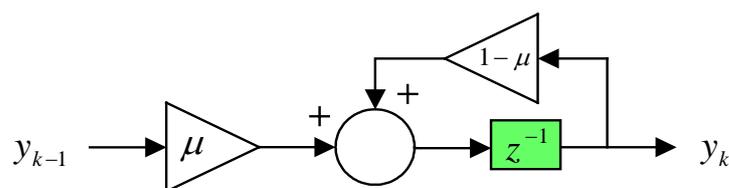


Figura 5.9. Estructura de la unidad de retardo k para una memoria Gamma.

$$y_k(n) = \mu y_{k-1}(n-1) + (1-\mu)y_k(n-1) \quad (5.27)$$

El valor del factor de realimentación también puede ser adaptado en el proceso de entrenamiento de la red neuronal, cambiando así la profundidad de memoria (que vimos para la red neuronal de Elman), sin necesidad de añadir o eliminar unidades de retardo. Los valores del factor de realimentación deben estar comprendidos entre 0 y 2.

Otra variante de memoria para redes neuronales retardadas en el tiempo es la memoria Laguerre [Back96]. La unidad de retardo para una memoria Laguerre es la presentada en la Figura 5.10. El único parámetro de la unidad de retardo sigue siendo el factor de realimentación, μ , con los mismos límites que para la memoria Gamma (entre 0 y 2).

Las memorias Laguerre tienen como ventaja, frente a las memorias Gamma, que las salidas de las distintas neuronas no están correlacionadas entre sí (ortogonaliza el espacio de entrada), lo que las hace especialmente útiles en problemas con necesidad de grandes profundidades de memoria.

El entrenamiento de las redes neuronales retardadas en el tiempo se realiza, fundamentalmente, con la técnica conocida como retropropagación de error a través del tiempo (backpropagation through time, BPTT) [Werbos90]. El entrenamiento es más lento que el de las redes unidireccionales y requiere más recursos informáticos (memoria).

La principal ventaja de las redes recurrentes retardadas en el tiempo es su menor tamaño en comparación con una red neuronal MLP que use la misma información temporal (utilizando entradas adicionales para introducir los valores previos de las entradas), con un menor tiempo de entrenamiento. Esta ventaja se ve acrecentada si además la red puede adaptar en el entrenamiento la profundidad de memoria necesaria para representar la relación temporal entre los valores del vector de entradas (memorias Gamma y Laguerre). El inconveniente principal de las redes neuronales recurrentes con retardo en el tiempo es la posibilidad de alcanzar un mínimo local en el proceso de entrenamiento.

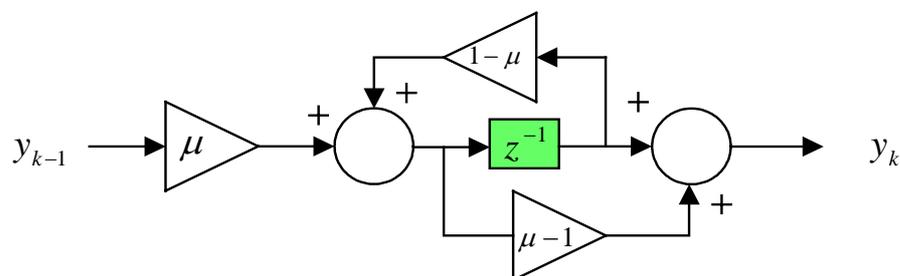


Figura 5.10. Estructura de una unidad de retardo para una memoria Laguerre.

Esa capacidad de extraer información útil de los valores pasados es la que

permite usar este tipo de redes neuronales para implementar modelos no-lineales auto-regresivos con media móvil (modelos NARMA).

5.3.6.3. Redes neuronales recurrentes generales.

Las redes neuronales recurrentes generales son aquellas redes neuronales en las que las neuronas de una o varias capas se realimentan a sí mismas por medio de unos pesos ajustables (la salida previa de la neurona, actúa como una entrada más, con una conexión determinada por un peso). Se puede distinguir entre redes neuronales totalmente recurrentes y redes parcialmente recurrentes. Son redes neuronales totalmente recurrentes aquellas en las que cada una de las neuronas de una capa se realimenta a sí misma. Una red neuronal es parcialmente recurrente cuando sólo algunas neuronas de la capa tienen realimentación.

La ventaja de este tipo de redes frente a las redes neuronales con retardo en el tiempo (las del apartado anterior), consiste en su gran profundidad de memoria (sólo limitada por el peso de las conexiones). Las redes neuronales con retardo en el tiempo tienen una profundidad limitada (igual al número de unidades de retardo en el caso de las TDNN, o fijado por el número de unidades de retardo y el factor de realimentación, en el caso de las redes neuronales retardadas en el tiempo con memorias Gamma o Laguerre). La Figura 5.11 muestra una red neuronal recurrente general. Además de las capas de entradas, salidas y oculta, tiene otra capa adicional donde se realimentan las neuronas.

Como desventaja, estas redes neuronales recurrentes pueden ser inestables durante el proceso de entrenamiento, y su entrenamiento es más lento que el del resto de redes neuronales, en parte debido a la técnica utilizada (por ejemplo, retropropagación del error a través del tiempo).

5.3.7. Mapas auto-organizados.

Los mapas auto-organizados son redes neuronales en cuyo entrenamiento se utiliza aprendizaje no supervisado. Puesto que a la red neuronal no se le suministra, durante el entrenamiento, ningún valor esperado de salida, ella debe procesar los datos de entrada de forma autónoma; de ahí que se afirme que se auto-organiza.

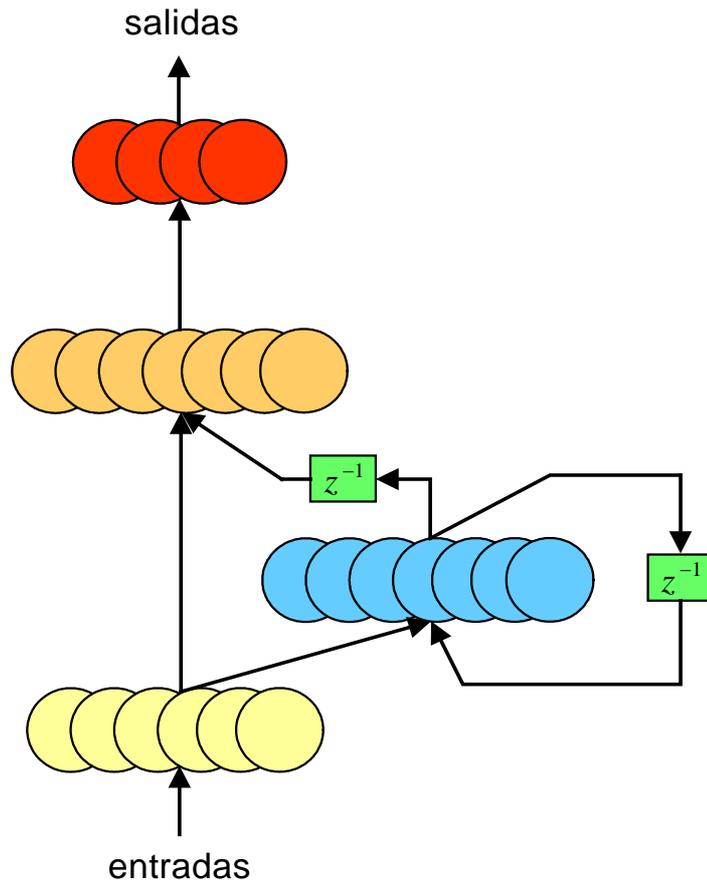


Figura 5.11. Estructura de una red neuronal recurrente general.

Los mapas auto-organizados o SOM (self organizing map) realizan una transformación de un espacio de entrada continuo, y con dimensión arbitraria, a un espacio discreto de una o dos dimensiones (mapa), manteniendo las propiedades topológicas de la entrada. Esto se realiza por reconocimiento, en los valores del vector de entradas, de unas características o patrones que son representadas después en el plano como un mapa de dichas características o patrones.

La Figura 5.12 representa la estructura de una red neuronal SOM. Un total de D neuronas de entrada (en color amarillo) se conectan a las neuronas de la capa de salidas del mapa auto-organizado (en este caso una estructura de 4x2 neuronas). En dicha figura sólo se han representado una parte de las conexiones: todas las neuronas de la capa de entradas se unen a todas las neuronas que conforman la capa de salidas del mapa auto-organizado.

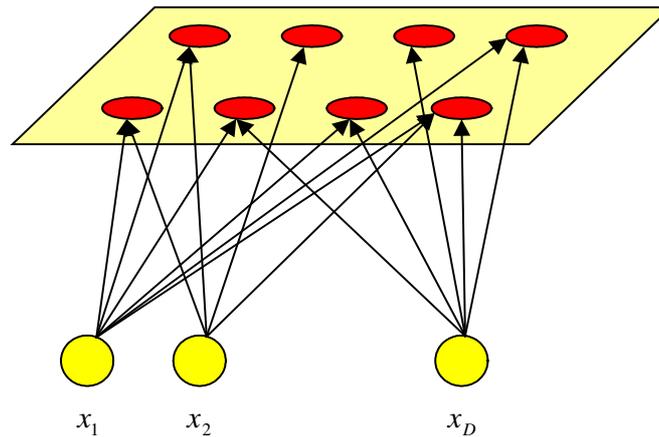


Figura 5.12. Estructura de un mapa auto-organizado.

Los mapas auto-organizados se basan en las llamadas redes neuronales competitivas, donde en función del valor del vector de entradas se activa únicamente una neurona de la capa de salidas (la unidad ganadora). En una red neuronal competitiva las neuronas de la capa de salidas que implementa ese funcionamiento (el competitivo), suelen ser de tipo binario, en el sentido de que sólo tienen dos posibles valores de salida: un 0 cuando están desactivadas y un 1 cuando están activadas. Si el número de entradas de la red neuronal es D , la neurona i de la capa de funcionamiento competitivo (capa de salidas) tiene asociado un vector de pesos sinápticos $(w_{i1}, w_{i2}, \dots, w_{iD})$, con el que va a ponderar los valores que le lleguen de las neuronas de la capa de entradas.

En un mapa auto-organizado el funcionamiento es similar al de una red competitiva, excepto que no sólo se activa una neurona, sino que se modifican también los valores de los pesos de las conexiones (vectores sinápticos) de las neuronas vecinas, aunque en menor medida, según su proximidad a la neurona ganadora, acercándose al patrón de entradas. Ello garantiza que las neuronas vecinas tengan sus vectores sinápticos parecidos, es decir, se preserve la topología del espacio de entradas. Para el entrenamiento de una red neuronal SOM se utiliza el método de aprendizaje no supervisado de Kohonen [Kohonen95].

El objetivo de una red neuronal SOM es clasificar los valores del vector de entradas en grupos de características similares, de tal forma que los valores del vector de entradas del mismo grupo activarán siempre las mismas neuronas de la capa de salidas del mapa auto-organizado. De ahí su principal aplicación como clasificador, o sistema para agrupamiento o clusterización de datos.

5.3.8. Ventajas e inconvenientes de los modelos con redes neuronales.

La principal ventaja que aportan los modelos con redes neuronales se deriva del hecho de que no hace falta conocer la naturaleza del conjunto de datos que se quiere representar, sino que la propia red neuronal, a través del proceso de entrenamiento, recoge las características esenciales del conjunto de datos a representar. Otras ventajas que se pueden citar son las indicadas seguidamente.

- No linealidad: el comportamiento no lineal de las funciones de activación de las neuronas posibilita que las redes neuronales puedan actuar como aproximadores universales de funciones no lineales.
- Son adaptables a la evolución de cualquier tipo de representación de datos sin más que volverlas a entrenar.
- Presentan una menor dependencia del conocimiento de los “expertos” en el desarrollo de modelos. Por ejemplo, el desarrollo de modelos ARIMA exigía un alto grado de experiencia que no resulta necesario para crear modelos con redes neuronales.

Como desventajas se pueden citar las señaladas seguidamente.

- Necesidad de disponer de gran cantidad de datos: el entrenamiento de redes neuronales con muchas entradas exige un número elevado de valores del vector de entradas y de los correspondientes valores del vector de salidas. Hay muchos problemas en los que resulta difícil obtener gran cantidad de datos, necesaria para que la red neuronal sea entrenada adecuadamente.
- Posibilidad de que el proceso de entrenamiento conduzca a algún mínimo local.

5.3.9. Aplicación de los modelos de redes neuronales en la predicción de la producción eléctrica en parques eólicos en el corto plazo.

Los modelos basados en redes neuronales artificiales se han manifestado como candidatos ideales en la aproximación de funciones no lineales. En el caso de las redes neuronales unidireccionales es sencillo construir un modelo auto-regresivo no lineal, sin

más que utilizar los valores pasados de la serie temporal como variables de entrada de la red neuronal. En el caso de redes neuronales recurrentes es posible construir modelos más complejos, capaces de proporcionar una salida distinta no sólo en función del valor del vector de entradas, sino en función del valor de dicho vector en el pasado próximo.

Para el caso concreto de la predicción de la producción eléctrica en parques eólicos, como problema de características no lineales que es, los modelos basados en redes neuronales constituyen una de las opciones, a priori, más interesantes a considerar. Los modelos que se puedan desarrollar van a responder a dos planteamientos diferentes, en función del horizonte de predicción del modelo.

- Para horizontes de predicción en el muy corto plazo (hasta 6 horas) y sin predicciones de un modelo atmosférico, los modelos de predicción que se pueden crear deben contener como entrada los valores pasados de la variable a predecir (potencia eléctrica media generada, por ejemplo), lo que dará lugar a modelos auto-regresivos no lineales (NAR) usando redes neuronales unidireccionales, y modelos auto-regresivos no lineales con media móvil (NARMA) usando redes neuronales recurrentes.
- Para horizontes de predicción en el muy corto plazo (hasta 6 horas) y con predicciones de un modelo atmosférico, los modelos de predicción que se pueden crear deben contener como entrada los valores pasados de la variable a predecir (potencia media eléctrica generada), y las predicciones del modelo atmosférico para el horizonte de predicción e instantes anteriores, lo que dará lugar a modelos auto-regresivos no lineales con variable exógena (NARX) usando redes neuronales unidireccionales, y modelos auto-regresivos no lineales con media móvil con variable exógena (NARMAX) utilizando redes neuronales recurrentes.
- Para horizontes de predicción en el corto plazo (por encima de las 6 horas) y con predicciones de un modelo atmosférico, los modelos de predicción a desarrollar corresponden a una aproximación de funciones no lineales, en la que las variables de entrada al modelo corresponden a las predicciones del modelo atmosférico para el horizonte de predicción y para los instantes previos. Los modelos de predicción a crear pueden corresponder tanto redes neuronales unidireccionales, como a redes recurrentes; éstas últimas podrían ser más adecuadas para la predicción de la producción eléctrica en los instantes siguientes a altos valores de velocidad de viento (cuando puede

producirse la desconexión de los aerogeneradores por razones de seguridad).

Los modelos basados en redes neuronales artificiales han sido utilizados en varios de los trabajos publicados en la literatura internacional. Así, han sido utilizados en [Kariniotakis96], [Alexiadis99], [Sfetsos00], [Maqsood05] y [Flores05] en el caso de modelos pertenecientes a la primera familia (los modelos que usan como entrada los valores pasados de la variable a predecir). En el caso de modelos pertenecientes a la segunda familia de modelos (los que usan como variables de entrada las predicciones de un modelo atmosférico) se han usado en [Ramirez06], [Larson06] y [Sideratos07]. Los modelos más usados han sido las redes neuronales MLP y las redes neuronales RBF.

5.4. SISTEMAS DE INFERENCIA DIFUSA

Los modelos basados en la lógica fuzzy o difusa [Zadeh65], junto con los modelos basados en redes neuronales artificiales, constituyen un conjunto de herramientas de representación o modelización que pertenecen al campo conocido como “soft computing”. La lógica fuzzy es una técnica para la incorporación del conocimiento estructurado humano en algoritmos eficientes [Kecman01].

Muchos de los procesos intelectuales humanos están basados en un razonamiento inductivo, cuyo mejor exponente son los razonamientos que lingüísticamente responden a estructuras lógicas del tipo “SI ... ENTONCES ...”. Podemos, así, argumentar que el conocimiento humano se estructura en reglas del tipo “SI ... ENTONCES ...”, y que esa combinación de reglas lleva a acciones, toma de decisiones, etc. Una regla de este tipo es una función de múltiples variables que relaciona las variables asociadas al “SI” y denominadas como antecedentes o causas, con las variables asociadas al “ENTONCES”, denominadas consecuentes o efectos.

También, una de las características del conocimiento humano es la de caracterizar las situaciones, elementos, propiedades, etc., en unos términos vagos o difusos. Así, para la mente humana, por ejemplo, un rascacielos no es un edificio de más de 100 metros de altura, sino que es un edificio “muy alto”. Pero en esta última idea hay una incertidumbre inherente: ¿cuál es la frontera entre un edificio muy alto y otro que no lo es? Esta incertidumbre puede ser acotada en términos matemáticos mediante lo que denominamos función de pertenencia, que es la expresión matemática del grado (entre 0 y 1) en que se valora la pertenencia de un elemento a un conjunto determinado por un específico término difuso (en este caso “muy alto”).

Un conjunto difuso se representa por una etiqueta, por ejemplo Z , que toma valores dentro del universo de discurso X , o límites entre los que está definido, y lo caracteriza mediante una función de pertenencia μ_Z que tomará valores comprendidos entre 0 y 1. La expresión matemática para un conjunto difuso es la reflejada en la ecuación (5.28).

$$A = \{(x, \mu_Z(x)) \mid x \in X\} \quad (5.28)$$

$$\mu_Z : X \rightarrow [0,1]$$

La función de pertenencia es una función que asocia a cada elemento (del conjunto difuso) un valor dentro del intervalo $[0,1]$ y dentro del universo de discurso. Las funciones de pertenencia más usuales son: triangular, trapezoidal, gaussiana, sigmoideal y campana. La Figura 5.13 muestra tres funciones de pertenencia, para otros tantos conjuntos difusos, definidos para un universo de discurso en el intervalo $[0,20]$. Las tres funciones de pertenencia de dicha figura son del tipo gaussiano, con centros en los valores 0, 10 y 20, y varianza de valor 4.

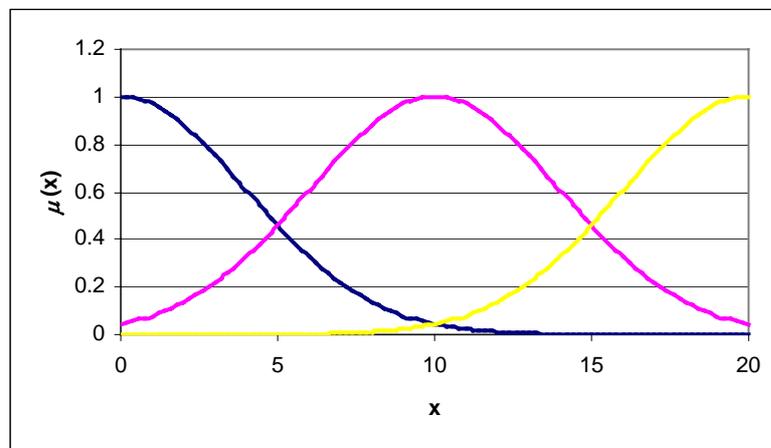


Figura 5.13. Ejemplo de funciones de pertenencia.

5.4.1. Estructura de un sistema de inferencia difusa.

La Figura 5.14 muestra la estructura general de un sistema de inferencia difusa. El vector de entradas está representado por \vec{X} y corresponde al conjunto de valores numéricos de las variables de entrada del sistema. La primera tarea del sistema es convertir esas variables numéricas de entrada en variables que pueda manipular ese sistema de inferencia difusa. Esta labor la realiza el bloque “fuzzyficador” que calcula

los niveles de pertenencia de los valores de entrada a los diferentes conjuntos difusos en los cuales se ha dividido el universo de discurso de las diferentes variables de entrada al sistema.

Como se ha comentado anteriormente, la base del razonamiento difuso son las reglas del tipo “SI ... ENTONCES ...”. El conjunto de reglas (base de reglas) constituye el conocimiento lingüístico que “tiene” el sistema de inferencia difuso y que le permite resolver el problema. Cada una de las reglas responde a una expresión como la reflejada en la ecuación (5.29), donde x_1 y x_2 representan dos de las variables de entrada e y representa la variable de salida.

$$\text{SI } x_1 \text{ es } A \text{ Y } x_2 \text{ es } B \text{ ENTONCES } y \text{ es } C \quad (5.29)$$

En la ecuación (5.29) A , B y C son conjuntos difusos definidos en el universo de discurso de x_1 , x_2 e y , y representados por la etiqueta lingüística correspondiente. A la proposición difusa “ x_1 es A Y x_2 es B ” se le llama antecedente o premisa y a la proposición difusa “ y es C ” se le llama consecuente o conclusión. Una regla expresa un tipo de relación entre A , B y C que se denomina implicación lógica.

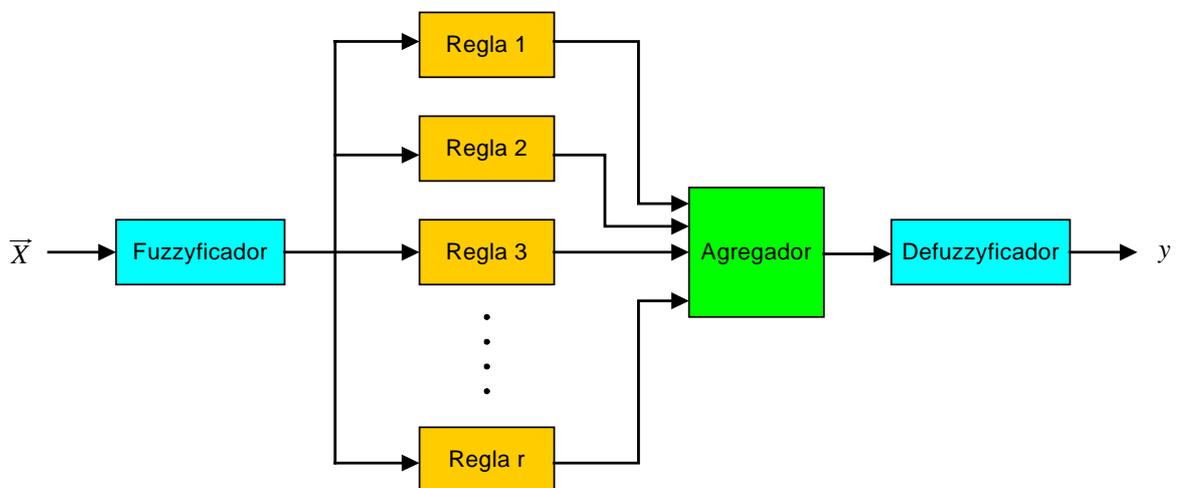


Figura 5.14. Estructura de un sistema de inferencia difusa.

El bloque agregador manipula los resultados de las distintas reglas para producir un resultado, todavía en términos difusos, que el bloque “defuzzificador” se encarga de convertir en una variable numérica.

Existen varios tipos de sistemas de inferencia difusa que difieren,

fundamentalmente, en la parte consecuente de las reglas, el método de agregación y la correspondiente “defuzzyficación” posterior. Entre estos sistemas se pueden citar los de Mandani, Takagi-Sugeno y Tsukamoto [Jang97]. De ellos, el utilizado en los trabajos de investigación reflejados en este documento ha sido el sistema de inferencia difusa de Takagi-Sugeno, cuyas características se revisarán en el próximo apartado.

5.4.2. Sistemas de inferencia difusa de Takagi-Sugeno.

Los sistemas de inferencia difusa de Takagi-Sugeno [Takagi85] son uno de los sistemas de inferencia difusa de aplicación más extendida a día de hoy. Se conoce también como sistema de inferencia difusa TSK (Takagi-Sugeno-Kang). Su principal característica reside en la formulación de la parte consecuente de las reglas. En los sistemas de inferencia difusa TSK estas partes consecuentes son combinaciones lineales (polinomios) de las variables de entrada al sistema.

En función del grado del polinomio de la parte consecuente de las reglas se puede distinguir entre sistemas de orden cero (el polinomio es de grado cero, es decir, una constante numérica), sistemas de primer orden (el polinomio es de primer grado), o sistemas de orden superior (polinomio de segundo grado o superior). En la práctica los sistemas se limitan a orden cero o primer orden, ya que el esfuerzo en la obtención de los polinomios de la parte consecuente con orden igual o superior a 2 no justifica la posible mejora en los resultados. Los sistemas de orden cero son, bajo determinadas restricciones, equivalentes a redes neuronales de función de base radial [Jang93b].

La ecuación (5.30) representa la expresión de la regla i para un sistema de inferencia difusa TSK de primer orden, donde x_1 y x_2 son las entradas del sistema, y_i es la salida de la regla i , y p_{0i} , p_{1i} y p_{2i} son constantes numéricas.

$$\text{SI } x_1 \text{ es } A_i \text{ Y } x_2 \text{ es } B_i \text{ ENTONCES } y_i = p_{0i} + p_{1i} \cdot x_1 + p_{2i} \cdot x_2 \quad (5.30)$$

Dado que la salida de cada regla es numérica, no hace falta bloque “defuzzyficador”. La salida y_s del sistema de inferencia difusa se obtiene como media ponderada de la salida de cada una de las reglas, tal como expresa la ecuación (5.31).

$$y_s = \frac{\sum_{i=1}^r w_i \cdot y_i}{\sum_{i=1}^r w_i} \quad (5.31)$$

En la ecuación (5.31) el número total de reglas del sistema es r , y_i representa la salida para la regla i , y w_i representa el producto (o el valor mínimo) de los valores de pertenencia de acuerdo a la ecuación (5.32). En la ecuación (5.32) $\mu_{A_i}(x_1)$ representa el valor de pertenencia de la variable x_1 al conjunto difuso A_i , y $\mu_{B_i}(x_2)$ representa el valor de pertenencia de la variable x_2 al conjunto difuso B_i .

$$w_i = \mu_{A_i}(x_1) \cdot \mu_{B_i}(x_2) \quad (5.32)$$

El número de reglas para cada sistema depende directamente de la partición del espacio de entrada, es decir, de cómo se hayan escogido los conjuntos difusos con sus familias de pertenencia para cada variable de entrada, y de cómo se hayan compuesto las reglas. Dos son los métodos más utilizados que se describen brevemente en el siguiente apartado.

5.4.3. Partición del espacio de entrada.

Los dos métodos más utilizados en la partición del espacio de entrada son el conocido como partición en rejilla (grid partitioning), que divide el rango de cada variable en un número determinado de intervalos (partes), y el conocido como partición esparcida o por agrupamiento (scatter partitioning), que selecciona los lugares del espacio de entrada con mayor densidad de puntos por medio del agrupamiento o clusterización. Cada valor del vector de entradas al sistema de inferencia difusa representa un punto de un espacio multi-dimensional (con tantas dimensiones como número de entradas).

La partición en rejilla se aplica en sistemas que contengan puntos distribuidos por todo el espacio y no existan zonas con alta densidad de puntos donde es más conveniente utilizar el otro método de partición.

El número máximo de reglas es igual al de combinaciones que se pueden formar con los intervalos en que hemos dividido cada una de las variables de entrada, tomando

un único intervalo por variable. Así, un sistema con tres variables de entrada y en el que se ha partido el espacio de entrada en rejilla con cada variable en 2 partes, dará un total de $2 \cdot 2 \cdot 2 = 8$ reglas. La parte consecuente de cada regla se calcula como la ecuación del hiperplano que mejor se ajusta (regresión por mínimos cuadrados) a los puntos pertenecientes a todos los intervalos de las variables de entrada correspondientes a la combinación escogida.

Para la partición del espacio de entrada por agrupamiento o clusterización se utilizan diferentes métodos, siendo los más utilizados los de K-means, fuzzy C-means y el substractivo [Jang97]. Todos estos métodos son procedimientos iterativos que determinan las posiciones de los centros de los grupos o clusters que mejor agrupan en torno a sí los valores del vector de entradas. Mientras que en el caso de los métodos K-means y fuzzy C-means el número de clusters es establecido de antemano, en el método del agrupamiento substractivo el número de clusters se calcula durante el proceso de agrupamiento.

El método de agrupamiento substractivo [Chiu94] se basa en la selección de los centros de los grupos o clusters en función de la “densidad” de cada punto. Si tomamos cada valor del vector de entradas al sistema de inferencia difusa como un punto de un espacio multi-dimensional, se define la densidad de un punto como una medida de la cercanía de otros puntos. Inicialmente se calcula la densidad de todos los puntos (valores del vector de entradas) y se toma como centro de un grupo al punto con mayor densidad (con mayor número de puntos vecinos cercanos). Seguidamente, la densidad de los puntos situados dentro de un radio establecido del centro del grupo se reduce y el algoritmo busca un nuevo centro (siguiente punto con mayor densidad). Este proceso continúa hasta que la densidad de los puntos restantes es menor a un valor predeterminado. El resultado final es un conjunto de puntos (valores del vector de entradas) correspondientes a los centros de los grupos. En definitiva, el agrupamiento substractivo identifica las regiones del espacio con mayor densidad de puntos de entrada, caracterizándolas por uno de sus puntos: el centro del grupo.

Los centros establecidos con el método de agrupamiento substractivo difieren de los obtenidos con los otros métodos en que corresponden a puntos concretos del espacio de entrada (los centros corresponden a alguno de los datos suministrados). Otra diferencia significativa radica en que el número de centros no es conocido a priori en un proceso de agrupamiento substractivo; depende de los parámetros que lo definen, en especial del llamado radio de vecindad, r_a . Un valor pequeño para el radio de vecindad puede conducir a obtener un excesivo número centros y ocasionar un posterior sobre-

ajuste del modelo. En algunas ocasiones en lugar de utilizar el radio de vecindad, r_a , se utiliza en su lugar el llamado “rango de influencia” cuyo significado es equivalente al de radio de vecindad, pero con la condición de que todas las variables de entrada han sido normalizadas a valores entre 0 y 1 (es decir, el espacio de entrada se ha reducido a un hiper-cubo unitario).

Una vez seleccionados los centros de los clusters usando el agrupamiento substractivo, los siguientes pasos para la construcción de un sistema de inferencia difusa son la de determinación del número de conjuntos difusos por variable de entrada, el ajuste de las correspondientes funciones de pertenencia a cada conjunto difuso, el establecimiento del número de reglas y el cálculo de los coeficientes de las partes consecuentes de las reglas. La solución más utilizada consiste en tomar, para cada variable de entrada, un número de conjuntos difusos igual al número de clusters identificados (determinados con el agrupamiento substractivo), utilizar funciones de pertenencia gaussianas y tomar, como centro de cada función de pertenencia, el valor de la variable considerada en cada uno de los centros de los cluster (por ejemplo, en un sistema con n clusters, para una variable x_1 , los centros de las n funciones de pertenencia gaussianas correspondientes a los n conjuntos difusos definidos para la variable x_1 , tienen como valores los de la variable x_1 para cada uno de los centros de los clusters). La varianza de las funciones de pertenencia gaussianas se toman iguales para todos los conjuntos difusos de una misma variable de acuerdo al valor de la ecuación (5.33), en la que σ_{ij} representa la varianza de la función de pertenencia del conjunto difuso i de la variable j , x_j corresponde a la variable de entrada j , y r_a es el radio de vecindad definido para el agrupamiento substractivo.

$$\sigma_{ij} = r_a \frac{\max(x_j) - \min(x_j)}{\sqrt{8}} \quad (5.33)$$

El cálculo de las funciones de pertenencia es mucho más sencillo utilizando el agrupamiento substractivo que cualquier otro de los dos algoritmos citados (K-means y fuzzy C-means), ya que la forma bien definida de las funciones de pertenencia gaussianas evita la pérdida de información. Además, los centros calculados por el algoritmo substractivo son tomados directamente como centros de las funciones de pertenencia gaussianas de la parte antecedente de las reglas y la varianza es calculada directamente al escoger el valor del radio de vecindad para los grupos, r_a . Por ello, el agrupamiento substractivo es el algoritmo ideal para obtener un modelo difuso inicial de cualquier sistema a modelar. Este modelo inicial puede ser optimizado posteriormente mediante técnicas neuro-difusas [Ferreyra05].

5.4.4. Sistemas neuro-difusos.

La combinación, en una misma estructura, de la representación del conocimiento de un sistema de inferencia difusa con la capacidad de aprendizaje supervisado de una red neuronal, conduce a un sistema neuro-difuso. En [Jang97] estos sistemas son denominados Adaptive Neuro-Fuzzy Inference Systems, ANFIS.

Supongamos un sistema de inferencia difusa TSK de primer orden con dos variables de entrada, X_1 y X_2 ; dos conjuntos definidos para cada una de las variables de entrada (conjuntos A_1 y A_2 para la variable X_1 ; y conjuntos B_1 y B_2 para la variable X_2); y dos reglas. Las familias de pertenencia para los conjuntos definidos para cada una de las variables de entrada están representadas en la Figura 5.15. El valor de salida de cada una de las dos reglas viene determinado por las ecuaciones (5.34), donde los coeficientes θ son constantes numéricas, y x_1 y x_2 son los valores que toman las variables X_1 y X_2 , respectivamente.

$$\begin{aligned}y_1 &= \theta_{01} + \theta_{11} x_1 + \theta_{21} x_2 \\y_2 &= \theta_{02} + \theta_{12} x_1 + \theta_{22} x_2\end{aligned}\tag{5.34}$$

El valor de la salida del sistema viene dado por la ecuación (5.35), donde el valor de los coeficientes w_1 y w_2 corresponde al producto de los valores de pertenencia de x_1 en A_1 y de x_2 en B_1 , y al producto de los valores de pertenencia de x_1 en A_2 y de x_2 en B_2 , respectivamente, tal y como indican las ecuaciones (5.36).

$$y = \frac{w_1 \cdot y_1 + w_2 \cdot y_2}{w_1 + w_2}\tag{5.35}$$

$$\begin{cases}w_1 = \mu_{A_1}(x_1) \cdot \mu_{B_1}(x_2) \\w_2 = \mu_{A_2}(x_1) \cdot \mu_{B_2}(x_2)\end{cases}\tag{5.36}$$

La ecuación (5.36) puede expresarse también como la ecuación (5.37), donde el valor de los nuevos coeficientes viene determinado por la ecuación (5.38).

$$y = \overline{w}_1 \cdot y_1 + \overline{w}_2 \cdot y_2\tag{5.37}$$

$$\bar{w}_i = \frac{w_i}{w_1 + w_2} \quad (i = 1, 2) \quad (5.38)$$

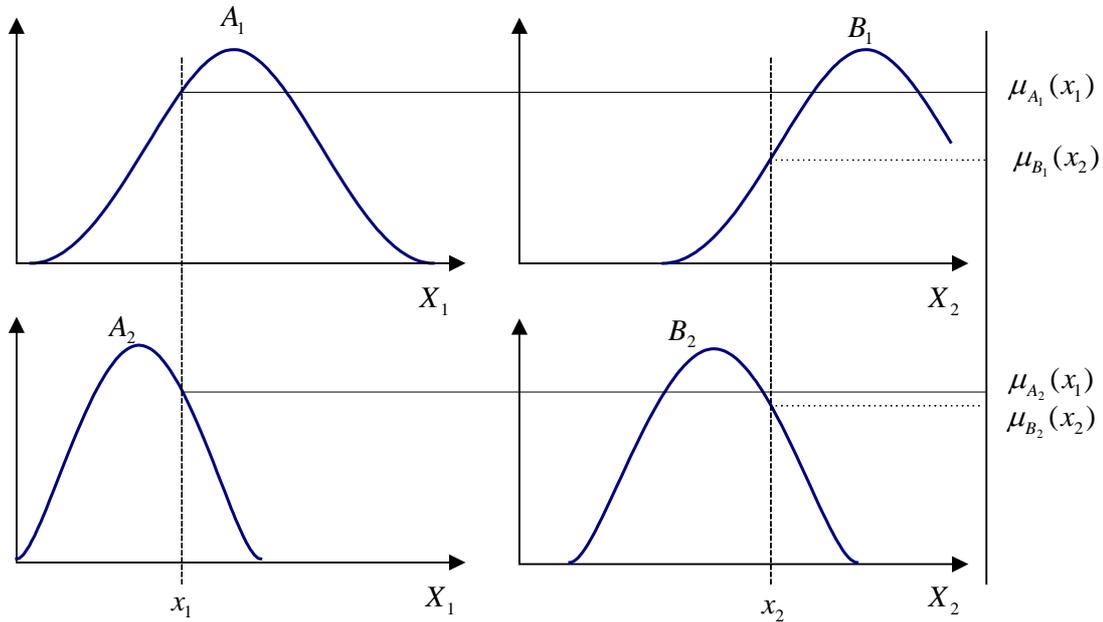


Figura 5.15. Sistema de inferencia difusa con dos entradas y dos reglas.

A la misma expresión se puede llegar con la red representada en la Figura 5.16, en la que en los nodos de la capa de entradas calculan el valor de pertenencia de las variables de entrada a cada uno de los conjuntos difusos definidos para cada entrada; los nodos de las dos siguientes capas calculan los coeficientes de la ecuación (5.38); los nodos siguientes calculan el valor de la salida de cada regla; y el nodo de salida calcula el valor de salida de la red (idéntico al valor de salida del sistema de inferencia difusa).

La red de la Figura 5.16, que no es estrictamente hablando una red neuronal sino la implementación de un sistema de inferencia difusa TSK en forma de red adaptativa, puede entrenarse como una red neuronal bajo entrenamiento supervisado, de forma que se ajusten el valor de sus parámetros, y se minimice el error (diferencia entre valor esperado de la salida y el valor de salida que proporciona la red). Así, esta red será entrenada bajo el equivalente al entrenamiento supervisado de las redes neuronales artificiales.

La técnica de entrenamiento utilizada es una técnica híbrida [Jang97], derivada del método de retropropagación de error de las redes neuronales. El resultado final del proceso de entrenamiento es el ajuste de los coeficientes de los polinomios de la parte consecuente de las reglas, y de los parámetros de las premisas (funciones de pertenencia

para los conjuntos definidos en las entradas). Para evitar el sobre-ajuste de los modelos se utilizan las mismas técnicas que las usadas en el entrenamiento de redes neuronales, en especial la de uso del grupo de validación cruzada, interrumpiéndose el proceso de entrenamiento cuando el error con los datos del grupo de validación cruzada no se ha reducido tras un número predeterminado de epochs.

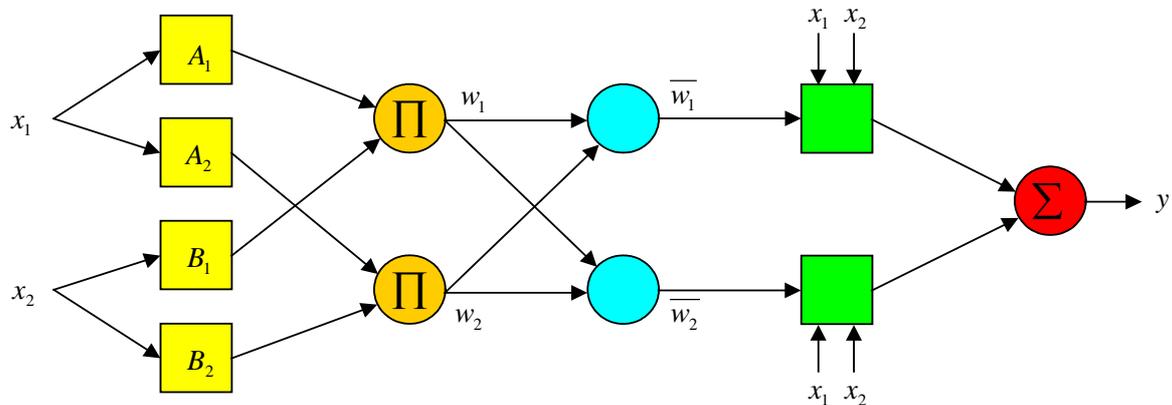


Figura 5.16. Red equivalente al sistema de inferencia difusa de dos entradas y dos reglas.

Sin embargo, una de las deficiencias detectadas (en las pruebas realizadas en los trabajos de investigación) es que el método de entrenamiento es propenso a caer en mínimos locales, siendo el resultado final del sistema entrenado muy dependiente de los datos iniciales del sistema (parámetros de las funciones de pertenencia y de los coeficientes de los polinomios iniciales). En las pruebas realizadas, utilizando la partición del espacio de entrada con el agrupamiento substractivo, se comprobó que los mejores resultados finales se obtenían ajustando el valor del rango de influencia, bien al valor mínimo o bien al valor máximo, que proporcionase un número determinado de clusters. Por ejemplo, si los valores del rango de influencia que proporcionan 2 clusters son los comprendidos entre 0.64 y 0.78, con uno de esos dos valores (y no con otro) se obtiene el sistema con mejores resultados de entre todos los posibles con 2 clusters. En todas las pruebas el resto de los parámetros del agrupamiento substractivo se mantuvieron constantes: 1.25 para el factor de reducción de densidad (factor r_b), 0.5 para el factor de aceptación ($\bar{\varepsilon}$), y 0.15 para el factor de rechazo ($\underline{\varepsilon}$).

5.4.5. Ventajas e inconvenientes de los modelos basados en sistemas de inferencia difusa.

La principal ventaja que reportan los modelos basados en sistemas de inferencia difusa radica en la facilidad que representa para la mente humana entender el propio

modelo. A diferencia de los modelos basados en redes neuronales (que se comportan como una “caja negra”), en el caso de los modelos de inferencia difusa, llegar a comprender el modelo es más sencillo.

Seguidamente indicamos otras ventajas.

- La posibilidad de los sistemas de inferencia difusa de integrar la irremplazable experiencia de expertos humanos. Esta experiencia se puede implementar fácilmente en las reglas del tipo “SI...ENTONCES...”.
- Los sistemas de inferencia difusa pueden integrar términos lingüísticos en el razonamiento (conjuntos difusos para alguna de las variables).
- La posibilidad que ofrece un sistema de inferencia difusa de ser entrenado (modelo neuro-difuso) permite ajustar los parámetros de dicho sistema, y reducir los errores si se utiliza como aproximador de funciones (lineales o no).

Los sistemas de inferencia difusa presentan limitaciones: desde el principio se deben fijar las reglas que establecen el funcionamiento del sistema; y, en ocasiones, un cambio de condiciones puede requerir la creación de un nuevo sistema de inferencia difusa (puede aparecer una imposibilidad en adaptar el sistema al cambio de condiciones).

Los sistemas neuro-difusos permiten obviar las desventajas de los sistemas de inferencia difusa (por ejemplo, no ser capaces de aprender), al poder ajustar, por lo general, los parámetros del sistema de forma adaptativa. En este sentido, los sistemas neuro-difusos combinan las ventajas que ofrecen las redes neuronales con la de los sistemas de inferencia difusa.

5.4.6. Aplicación de modelos basados en sistemas de inferencia difusa en la predicción de la producción eléctrica en parques eólicos en el corto plazo.

Los modelos basados en sistemas de inferencia difusa han mostrado, a lo largo de los últimos años, una aplicabilidad tan universal como la de los modelos basados en redes neuronales artificiales. Comparten con estas la capacidad de aproximación de

funciones no lineales, y les aventajan en la posibilidad de que diseñador pueda adaptar el modelo de predicción debido a su mayor sencillez conceptual.

La predicción de la producción eléctrica en parques eólicos utilizando sistemas de inferencia difusa es similar a la realizada con redes neuronales. En función del horizonte de predicción, se usan (o no) los últimos valores de la variable a predecir (normalmente la potencia eléctrica media generada) como entrada al sistema de predicción. El resto de entradas corresponderá a las predicciones de un modelo atmosférico, en el caso de que estuvieran disponibles. Los modelos serán de Takagi-Sugeno de primer orden, dada la simplicidad y generalización que ofrecen. Además estos modelos pueden adaptar sus parámetros utilizando técnicas de modelos neuro-difusos.

En cuanto al método de partición del espacio de entrada, en principio se pueden usar cualquiera de los dos expuestos, es decir, la partición en rejilla y la partición por agrupamiento o clusterización. Cuando se use la partición en rejilla no se sobrepasará de los tres conjuntos difusos por entrada, para evitar un gran número de reglas, que precisarían a su vez de un gran número de datos de entrada para su construcción. Para la partición del espacio de entrada por agrupamiento, la solución ideal será la utilización del algoritmo de agrupamiento substractivo, de tal forma que el número de clusters y reglas no sea elevado (inferior a 6). Tanto en el caso de partición en rejilla o por agrupamiento, pueden utilizarse técnicas neuro-difusas para la adaptación de los parámetros de los modelos y la reducción de los errores de predicción.

En varios trabajos publicados en la literatura internacional, se han utilizado modelos de predicción de la producción eléctrica en parques eólicos basados en sistemas de inferencia difusa. Recordando las dos familias de modelos presentadas en el Capítulo 3, podemos encontrar modelos de la primera familia (usando como variable de entrada los valores pasados de la variable a predecir) en [Sfetsos00], [Hatziaargyriou01], [Ramirez01], [Ramirez03], [Damousis04], y [Potter06]. En cuanto a modelos de la segunda familia, sólo aparece un modelo basado en sistemas de inferencia difusa [Sideratos07], y únicamente como un módulo más del sistema de predicción propuesto.

5.5. CONCLUSIONES

Los modelos expuestos en este capítulo constituyen los más usados en la predicción de futuros valores para series temporales, y por tanto resultan muy

adecuados para la predicción de la producción eléctrica en parques eólicos, bien como modelos de predicción completos, o bien como elementos principales constituyentes de los sistemas de predicción propuestos en este documento.

En todos los casos, las condiciones que se deben observar antes de desarrollar los modelos deben incluir los aspectos señalados a continuación.

- Detección y corrección de outliers en la base de datos disponible: los outliers pueden empeorar significativamente la calidad de las predicciones obtenidas con cualquiera de los modelos desarrollados. Los modelos más sensibles a los outliers, sobre todo en los datos de entrenamiento o ajuste, son los modelos ARIMA y los modelos basados en redes neuronales. Los modelos basados en sistemas de inferencia difusa con técnicas de agrupamiento substractivo parecen más inmunes al efecto de los outliers.
- División de los datos disponibles para el desarrollo de los modelos en, al menos, dos grupos: en los modelos ARIMA, o en los sistemas de inferencia difusa, se ajustan los parámetros de los modelos de predicción a los datos utilizados en su construcción; en los modelos de redes neuronales, o en los modelos neuro-difusos, los parámetros se adaptan a los datos de entrenamiento. Para la evaluación de los modelos deben usarse otros datos (datos para el testeo) diferentes de los utilizados en el ajuste o entrenamiento de los modelos de predicción.
- En el caso de seleccionar un modelo de una misma familia (modelos ARIMA, por ejemplo), debe utilizarse algún criterio que penalice los modelos con mayor número de parámetros (como el criterio de información de Akaike).
- Para el caso de redes neuronales o sistemas neuro-difusos, los datos que se vayan a utilizar en el proceso de entrenamiento deben dividirse en dos grupos: los datos que van a ser usados en el ajuste de los parámetros, que los llamaremos grupo de datos de entrenamiento, y los datos utilizados como grupo de validación cruzada. Cuando ya no se reduce el error en la salida con los datos del grupo de validación cruzada tras un número determinado de epochs, se termina el proceso de entrenamiento. En todo caso, los parámetros del modelo de predicción corresponderán a los que presentaba el modelo cuando el error con los datos del grupo de validación era mínimo.

6. OPTIMIZACIÓN DE MODELOS DE PREDICCIÓN CON ALGORITMOS GENÉTICOS

En este capítulo se presentan las técnicas (algoritmos genéticos) utilizadas para la optimización de modelos de predicción basados en redes neuronales o sistemas de inferencia difusa. En el desarrollo de ambos tipos de modelos de predicción, muchos de los parámetros que los definen quedan a juicio del diseñador, como por ejemplo el número de neuronas en capas ocultas o los parámetros que definen las funciones de pertenencia de cada conjunto difuso definido para una variable de entrada. La utilización de procesos de optimización basados en el uso de algoritmos genéticos permitirá una selección óptima de valores para los parámetros que definen la estructura de cada uno de los modelos de predicción. También se describe un original método creado para la optimización de los parámetros de sistemas de inferencia difusa basado en la aplicación de un nuevo algoritmo genético.

6.1. ALGORITMOS GENÉTICOS

Los algoritmos genéticos constituyen unas herramientas de optimización que han demostrado ser muy útiles para la búsqueda de soluciones satisfactorias en problemas de diversos campos científicos: ingeniería, economía, medicina, etc. El algoritmo genético, GA (Genetic Algorithm), fue desarrollado por J.H. Holland en la década de los 70 del pasado siglo como una técnica de búsqueda basada en los mecanismos de la selección natural. La publicación de [Goldberg89] atrajo la atención sobre el potencial de esta técnica como eficaz herramienta para la optimización, la adaptación y el aprendizaje, necesarias en la resolución de los más variados problemas.

En el desarrollo de modelos de predicción, los algoritmos genéticos han sido utilizados como técnica de optimización para el diseño automático de dichos modelos. Los algoritmos genéticos pueden ser usados en múltiples procedimientos de diseño y requieren muy poca información específica para poder realizar la optimización. Han sido utilizados en la optimización de redes neuronales [Yao99], en la identificación de modelos ARIMA estacionales [Blake95], o como método de eliminación de reglas menos importantes en sistemas de inferencia difusa [Leng06].

De forma resumida, los procesos de optimización con algoritmos genéticos

comienzan con la creación de una población inicial de individuos, generados de una forma aleatoria. Cada individuo constituye una solución potencial al problema que se esté considerando. El conjunto de individuos va evolucionando de una manera iterativa. Cada uno de los pasos del proceso iterativo recibe el nombre de generación. En cada generación, todos los individuos de la población son evaluados usando un criterio especificado de medida de la bondad de la solución al problema que representan (aptitud o fitness). La población para la siguiente generación se crea, a partir de la población existente, mediante la aplicación de operadores genéticos. El proceso iterativo, con sucesivas generaciones, termina cuando se cumple un criterio específico establecido previamente.

Cada solución debe estar codificada en forma de una cadena numérica o alfanumérica, que recibe el nombre de individuo. La codificación más usada es la numérica, con una representación de los números en formato binario, bien en forma de números enteros o de coma flotante. Así, la cadena estará compuesta por un conjunto de bits. A su vez, esta cadena puede considerarse como un cromosoma compuesto por n genes, tal y como representa la expresión (6.1), donde n representa los diferentes parámetros o características que definen la solución contenida en la cadena. La estructura de la cadena recibe el nombre de genotipo y su contenido el nombre de fenotipo [Sakawa02].

$$s_1 s_2 \cdots s_j \cdots s_n \quad (6.1)$$

En problemas de maximización, la aptitud (criterio de bondad) puede consistir en el valor que proporciona cada solución potencial (codificada en forma de cromosoma) para la variable a maximizar. Las soluciones con mayor aptitud son mejores soluciones al problema. En problemas minimización (por ejemplo minimización del error de predicción), la aptitud puede consistir en la diferencia entre un valor constante y el valor que proporciona cada solución (individuo) para la variable a minimizar. La combinación de aptitud y tres operadores fundamentales proporciona, tras un número razonable de generaciones, la solución óptima o aproximadamente óptima al problema de optimización.

Seguidamente indicamos los tres operadores fundamentales usados en la creación de una nueva generación.

1. Reproducción: la probabilidad de reproducción de cada individuo de una población está ligada a su valor de aptitud. Básicamente la reproducción consiste

en realizar un número de copias de cada individuo relacionado con su valor de aptitud. Un método habitual es usar una ruleta, conocida como ruleta de Holland [Holland92], con un número determinado de casillas. Las diferentes casillas de la ruleta son ocupadas por los individuos cuyo número de repeticiones está ligado a su valor de aptitud. El individuo con mayor valor de aptitud se copiará más veces que el individuo con peor valor. Con esto se crea una población intermedia entre la de la última generación y la siguiente.

2. Cruce: dos individuos seleccionados aleatoriamente de entre los que componen la población intermedia intercambian parte de sus cromosomas con una probabilidad definida y conocida como el índice de cruce o tasa de cruce. El intercambio se realiza seleccionando de forma aleatoria (probabilidad uniforme) un punto o dos puntos de las cadenas, e intercambiando los fragmentos de las mismas desde el punto seleccionado al principio o final de la cadena, o intercambiando los fragmentos comprendidos entre los dos puntos. A partir del cruce se crean dos nuevos individuos para la siguiente generación. Dado que en la población intermedia los individuos con mayor aptitud estaban más veces repetidos que los de menor aptitud, los individuos con mayor aptitud tienen mayor probabilidad de ser seleccionados para la operación de cruce que aquellos con un menor valor de la misma.
3. Mutación: Consiste en la alteración de una o más partes de un individuo seleccionado, con una probabilidad definida y conocida como índice o tasa de mutación. En este caso, los individuos se seleccionan de la población de la última generación, no de la población intermedia de la ruleta.

Este operador proporciona un carácter aleatorio a la búsqueda de soluciones, de forma que se evita que, tras unas pocas generaciones, todas las soluciones sean muy similares entre sí. El operador mutación facilita una vía para “escapar” de posibles óptimos locales (de la aptitud) en el proceso de búsqueda de la solución óptima.

La selección de valores para las tasas de cruce y mutación tiene una gran influencia en las soluciones obtenidas con el algoritmo genético. Un elevado valor de la tasa de cruce se traduce en un proceso dirigido de búsqueda de soluciones, mientras que un alto valor de la tasa de mutación se traduce en un proceso de búsqueda aleatorio. Para no caer en ninguno de los dos extremos, búsqueda dirigida propensa a encerrarse en óptimos locales y búsqueda aleatoria que no converge a soluciones satisfactorias, se

utilizan tasas de cruce inferiores al 90%, y tasas de mutación muy pequeñas (sobre un 1%).

Con todo, los algoritmos genéticos constituyen una eficaz técnica de optimización, que se diferencia de las técnicas clásicas en cuatro aspectos fundamentales [Goldberg89].

1. Los algoritmos genéticos trabajan con un código del conjunto de parámetros del problema a optimizar (no con el conjunto mismo), el cual debe estar codificado en cadenas finitas sobre un determinado alfabeto.
2. Trabajan con una población de soluciones (no con una única solución). Manteniendo una población adecuada, se reduce la probabilidad de caer en óptimos locales de la función de aptitud (mínimos o máximos locales del problema de optimización).
3. Emplean una función de aptitud y no precisan derivadas de funciones ni aproximaciones. De este modo ganan eficiencia y generalidad.
4. Se valen de métodos de transición estocásticos, no deterministas. La búsqueda de la mejor solución está controlada por unos operadores que actúan de forma aleatoria.

6.2. OPTIMIZACIÓN DE MODELOS DE PREDICCIÓN BASADOS EN REDES NEURONALES

El diseño idóneo de una red neuronal puede plantearse, en cierta forma, como un problema de optimización. Una de las primeras decisiones que hay que tomar al crear una red neuronal es fijar su estructura. Por ejemplo, al crear una red neuronal MLP hay que decidir el número de capas ocultas, el tipo de funciones de activación de las neuronas, el número de neuronas en cada capa, los parámetros del algoritmo de entrenamiento escogido, etc. Una buena parte de estos parámetros pueden ser optimizados con un algoritmo genético.

En general, y una vez escogido el tipo de la red neuronal (MLP, FFBP, RBF, etc.), un algoritmo genético adecuado permite determinar el valor óptimo de varios parámetros, algunos de los cuales son indicados seguidamente.

- Entradas de la red neuronal. Los mejores modelos de predicción son aquellos que usan aquellas variables de entrada, de entre las disponibles, que realmente aportan información relevante para la predicción.
- Número de capas ocultas: en principio podrían construirse redes neuronales con un número de capas ocultas ilimitado. Sin embargo, para aproximación de funciones no lineales, una o dos capas ocultas son suficientes en el caso de las redes neuronales MLP.
- Función de activación de las neuronas: en el caso más complejo podríamos seleccionar una función de activación diferente por neurona, escogiendo entre sigmoideal, lineal, tangente hiperbólica, etc. En los casos más simples se escoge la función de activación común para todas las neuronas de una capa.
- Número de neuronas en cada capa: el número de neuronas en las capas ocultas tiene relación directa con la capacidad de aproximación no lineal de la red neuronal. Cuanto mayor sea el número de neuronas, mayor es esa capacidad, pero también es mayor la probabilidad de sobre-entrenar la red neuronal. La selección del valor óptimo del número de neuronas en cada capa se puede realizar mediante un algoritmo genético adecuado.
- Valores de inicialización de pesos: para entrenar una red neuronal hay que partir de unos valores iniciales de los pesos de las conexiones entre neuronas. La asignación de valores a estos pesos se hace de forma aleatoria, pero su valor inicial influye en el comportamiento final de la red neuronal. Con unos valores iniciales u otros, el comportamiento de la red neuronal, una vez entrenada, puede ser completamente diferente. Un adecuado algoritmo genético permite la selección idónea del valor medio de los pesos iniciales y de su varianza. Posteriormente se asignan de forma aleatoria valores iniciales a los pesos, siguiendo una distribución estadística con aquel valor medio y aquella varianza.
- Parámetros del algoritmo de entrenamiento utilizado: en la ecuación (5.24) aparecían dos parámetros de crucial importancia en el proceso de entrenamiento de la red neuronal, en el caso de utilizar la técnica de retropropagación de error. Estos parámetros son el factor de aprendizaje, o step, y el momento. Cada capa de neuronas puede tener sus propios valores, pero su valor óptimo puede ser determinado con la ayuda del algoritmo genético. En el caso de entrenamiento

no supervisado, otros parámetros que lo definen pueden tener el mismo tratamiento.

- Número de unidades de retardo: en las redes neuronales recurrentes con retardo en el tiempo, el número de unidades de retardo también es susceptible de ser escogido por medio de un algoritmo genético.
- Número de componentes principales: en las redes neuronales PCA el número de componentes principales, que se van a considerar en las entradas, tiene influencia en el posterior funcionamiento de la red neuronal. Su valor óptimo también puede ser seleccionado por medio de un algoritmo genético.
- Número de neuronas de función de base radial: en una red neuronal RBF las neuronas de la capa siguiente a la capa de entradas tienen una función de activación de tipo gaussiano, cuyo valor de centro, y de varianza, es determinado por aprendizaje no supervisado. Sin embargo, queda a elección del diseñador el número de neuronas de ese tipo. La elección óptima de su valor puede lograrse con un algoritmo genético adecuado.

La codificación de toda la información de estos parámetros correspondientes a una red neuronal puede ser fácilmente realizada en forma de una cadena de bits, representando cada grupo de bits (que definen un parámetro) un gen, y el conjunto de todos ellos el cromosoma o cadena completa.

Los parámetros cuyos valores son números enteros (número de neuronas en una capa, número de neuronas de retardo, etc.) se pueden representar directamente mediante código binario, sin más que utilizar un número de bits que garantice la posibilidad de alcanzar un límite máximo predefinido. Por ejemplo, para limitar el número de neuronas en cada capa a un número entre 64 y 127 deberían usarse 7 bits.

La selección de la función de activación de las neuronas se puede codificar también mediante código binario como un número entero. Por ejemplo, se puede asignar un valor 0 a la función sigmoideal, un 1 a la lineal, un 2 a la tangente hiperbólica, etc.

Aquellos otros parámetros numéricos que no corresponden a números enteros se pueden representar en el formato de coma flotante. Este formato es una codificación

binaria, con una sub-cadena de longitud fija (32 bits para el caso más simple contemplado en la norma ANSI/IEEE Std 754-1985).

En todos los casos en que la codificación se ha realizado por un número binario, antes de aceptar un nuevo individuo para la siguiente generación se debe comprobar la consistencia de cada uno de los números, evitando la incorporación de nuevos individuos (nuevas redes neuronales) cuyos parámetros excedan los límites admisibles (por ejemplo, un valor de momento negativo o mayor que la unidad).

La selección de entradas se puede codificar mediante código binario, con un número de bits igual al de posibles variables de entrada, interpretándose su selección con el valor 1 y su no selección con el valor 0. Así, si se dispone de 8 posibles variables de entrada, el número binario 10100100 representaría que las variables realmente seleccionadas como entradas son la primera, tercera y sexta de las 8 disponibles.

Una vez configurada cada red neuronal, definida por un cromosoma completo, debe ser entrenada, de acuerdo a los posibles valores de los parámetros que definen el entrenamiento. Parámetros como el número máximo de epochs, criterio de interrupción de entrenamiento o criterio de evaluación deben haber sido definidos desde el primer momento por el diseñador. El criterio de evaluación corresponde a la función objetivo que marca la aptitud de cada uno de los individuos. El criterio puede ser el error cuadrático medio con los datos de entrenamiento, con los datos de validación, o una media ponderada de los dos. Evidentemente, en este caso lo que busca la optimización es minimizar el error.

En los trabajos de investigación recogidos en este documento, se ha utilizado una herramienta [Principe00] que, mediante un algoritmo genético, permite optimizar la mayor parte de los parámetros antes indicados. El diseñador sólo ha de fijar de antemano el número de capas ocultas y el tipo de función de activación. El criterio de evaluación seleccionado ha sido el error cuadrático medio con los datos del grupo de validación.

6.3. OPTIMIZACIÓN DE MODELOS DE PREDICCIÓN BASADOS EN SISTEMAS DE INFERENCIA DIFUSA

En las pruebas realizadas en los trabajos de investigación recogidos en este documento, los mejores resultados con modelos de predicción basados en sistemas de

inferencia difusa se han obtenido con aquellos que utilizaban agrupamiento substractivo para determinar el número conjuntos difusos definidos para cada entrada, y para el ajuste de las respectivas funciones de pertenencia. Los modelos de predicción basados en sistemas de inferencia difusa, con partición del espacio de entrada por medio de rejilla, han mostrado en todos los casos un funcionamiento más pobre (mayor error de predicción) que sus homólogos contruidos sobre la base del agrupamiento substractivo.

La posibilidad de ajuste de los parámetros del propio sistema de inferencia difusa por medio del entrenamiento como sistema neuro-difuso [Jang97] hace que el diseñador apenas tenga que decidir la estructura que va a dar al sistema: sólo tiene que establecer los parámetros del algoritmo substractivo, que son los que van a imponer el número de reglas del sistema de inferencia difusa. El posterior entrenamiento, de tipo híbrido, permite ajustar los valores de los coeficientes de la parte consecuente, así como la posición de los centros y varianzas de las funciones de pertenencia de tipo gaussiano. Aun así, este método tiende a caer en mínimos locales.

La Figura 6.1 representa los errores RMSE para un modelo de predicción construido con un modelo de inferencia difusa utilizando el agrupamiento substractivo. La variable a predecir es la potencia eléctrica media horaria generada en un parque eólico para la próxima hora, $p(t+1)$. Las variables de entrada corresponden a los valores de esa misma variable en el momento actual y una hora antes, $p(t)$ y $p(t-1)$. Los datos utilizados en la construcción del modelo han sido 840 datos como grupo de entrenamiento, 280 datos como grupo de validación y otros 280 datos como grupo de testeo, lo que corresponde a un total de casi dos meses de datos.

Para la elaboración de la Figura 6.1 se han cambiado sucesivamente los parámetros del algoritmo substractivo, concretamente el rango de influencia de 0.5 a 1, manteniendo constantes el resto de parámetros. La función utilizada fue la “subclust” [Jang95]. En la citada figura se representan los errores RMSE para cada uno de los tres grupos de datos, entrenamiento, validación y testeo. El error indicado como ponderado corresponde a la media ponderada entre los errores del grupo de entrenamiento y el grupo de validación. La media ponderada está calculada en función al porcentaje relativo de ambos grupos con respecto a la suma de los dos (75% para el grupo de entrenamiento y 25% para validación).

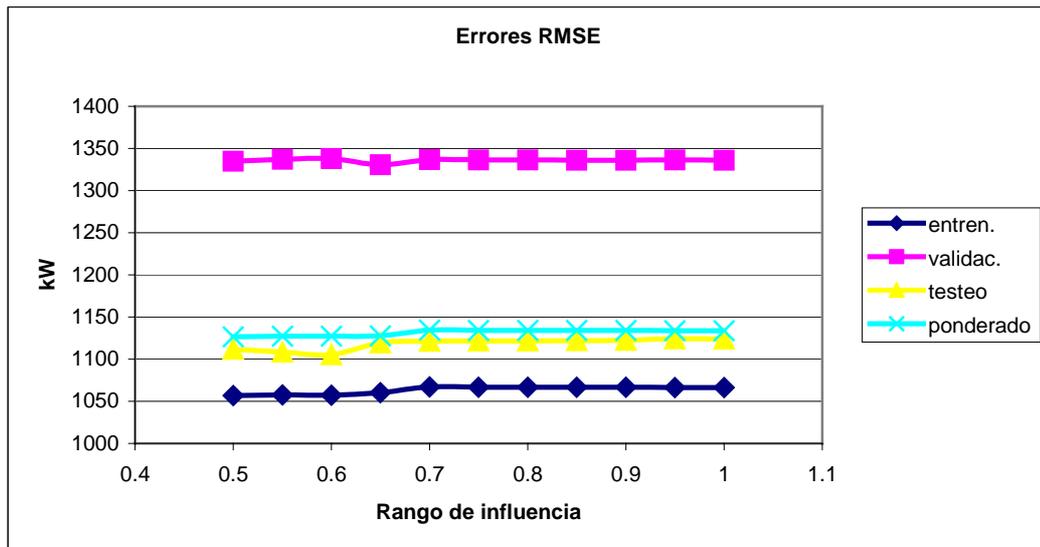


Figura 6.1. Errores RMSE en la predicción de potencia eléctrica media horaria.

Los dos primeros valores de la Figura 6.1 (rango de influencia de 0.5 y 0.55) corresponden a sistemas de inferencia difusa con 4 reglas y otros tantos conjuntos difusos definidos para ambas variables de entrada. Los siguientes dos puntos (rangos de influencia de 0.6 y 0.65) corresponden a sistemas con 3 reglas y otros tantos conjuntos difusos. Finalmente, los últimos 7 puntos (rangos de influencia de 0.7 a 1) corresponden a sistemas con 2 reglas y otros tantos conjuntos difusos por entrada. La Figura 6.2 representa los valores del criterio de información de Akaike (AIC), calculados por medio de la ecuación (5.20) y utilizando los datos de los grupos de entrenamiento y validación, en función del rango de influencia del agrupamiento substractivo empleado.

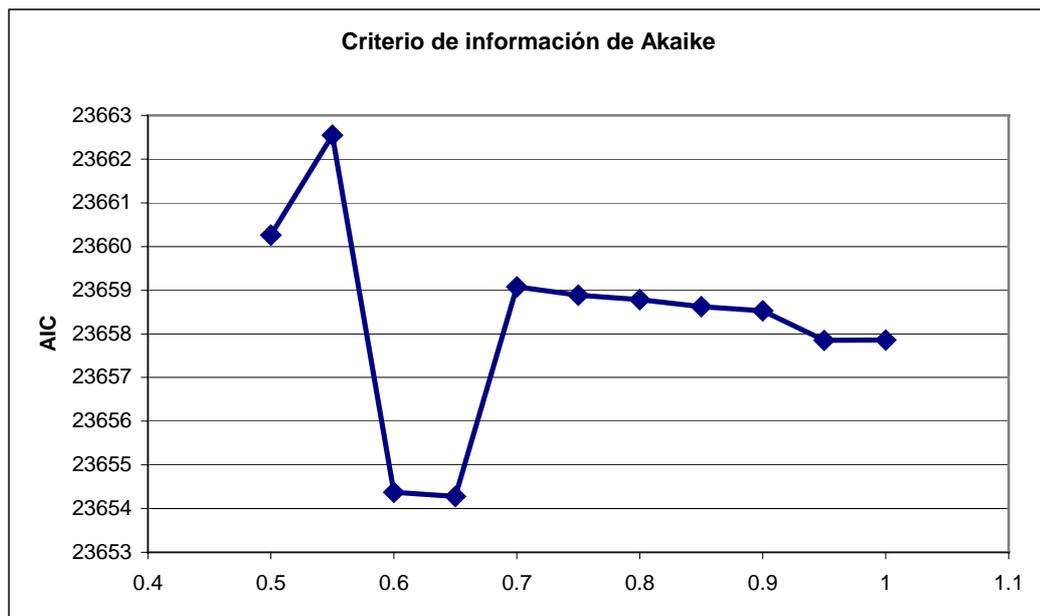


Figura 6.2. Valor del AIC en la predicción de potencia eléctrica media horaria.

El menor valor del AIC corresponde a un rango de influencia en el agrupamiento substractivo de 0.65, por lo que debería ser el sistema de inferencia creado con ese valor del rango de influencia el que debiéramos usar como modelo de predicción (corresponde a un sistema con 3 reglas). Sin embargo, el sistema creado con un rango de influencia de 0.6 también tiene 3 reglas, también ha sido entrenado, y no proporciona los mismos valores de error de predicción para ninguno de los tres grupos en que hemos dividido los datos (ver Figura 6.1). Es decir, en el entrenamiento del modelo neuro-difuso, el entrenamiento se ha interrumpido en un mínimo local.

Para evitar estos mínimos locales, se ha propuesto la utilización de algoritmos genéticos en la optimización de sistemas de inferencia difusa, en lugar de utilizar la técnica de entrenamiento neuro-difusa. Así, en [Damousis04] se utiliza un algoritmo genético en la optimización de un sistema de inferencia difusa TSK de primer orden para la predicción de la velocidad media del viento, para un parque eólico, en el muy corto plazo (minutos), utilizando como datos de entrada los valores medidos en estaciones meteorológicas cercanas. La partición propuesta para el espacio de entrada es en rejilla y las funciones de pertenencia son gaussianas. Los parámetros del sistema de inferencia difusa que forman parte del cromosoma propuesto son:

1. los centros de las funciones de pertenencia gaussiana para cada uno de los conjuntos difusos en que se divide cada variable;
2. la varianza de dichas funciones de pertenencia; y
3. los coeficientes de las partes consecuentes de las reglas (tantos como el número de reglas multiplicado por el número de entradas más uno).

Todos los parámetros son codificados de forma binaria, ocupando cada uno de ellos 8 bits, sin que los autores indiquen su significado. La función objetivo es el error RMSE en la predicción de la velocidad media del viento para el horizonte de predicción con los datos del grupo de entrenamiento.

6.3.1. Sistema de optimización propuesto.

Para la optimización de los sistemas de inferencia difusa, en el caso de no optar por su entrenamiento como sistemas neuro-difusos, se ha desarrollado un sistema de

optimización original, basado en algoritmo genético, cuyas características se indican a continuación.

- Las variables de entrada y salida deben estar escaladas a valores comprendidos entre 0.1 y 0.9.
- La codificación de los parámetros del sistema de inferencia difusa no es con números binarios, sino con números enteros, concretamente hasta 16 dígitos por parámetro.
- Inicialmente usa funciones de pertenencia gaussianas (podría extenderse para funciones de pertenencia triangulares, trapezoidales, sigmoideas, etc.).
- Cada cromosoma está compuesto como indica la Figura 6.3, en la que c_{ij} representa el centro de la función de pertenencia gaussiana correspondiente al conjunto difuso j de la entrada i (el que interviene en la regla j); σ_{ij} representa la varianza de la función de pertenencia gaussiana del conjunto difuso j de la entrada i ; θ_{0j} representa el coeficiente independiente del polinomio de primer orden de la regla j ; θ_{ij} representa el coeficiente que multiplica al valor de la entrada i para la regla j ; y m es el número de entradas y n el número de reglas.

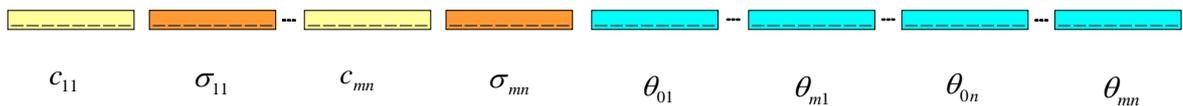


Figura 6.3. Estructura propuesta para el cromosoma.

Así, en la estructura del cromosoma representado en la Figura 6.3, los centros de las funciones de pertenencia están representados en amarillo, las varianzas de las mismas en naranja y los coeficientes de los polinomios en azul.

- El significado de cada dígito, en el gen representativo de los centros de las funciones de pertenencia, es el representado en los ejemplos de la Figura 6.4. Para el caso del gen que representa el centro de la función de pertenencia gaussiana, los 16 dígitos corresponden todos a parte decimal (las entradas están escaladas entre 0.1 y 0.9). El gen que representa la varianza de las funciones de pertenencia gaussianas contiene también 16 dígitos, el primero corresponde a la parte entera y los otros 15 a la parte decimal. El gen que representa a los

coeficientes de los polinomios de primer grado tiene 18 dígitos: el primero representa la parte entera de la mantisa, los quince siguientes la parte decimal de la mantisa, el 17° dígito representa tanto el signo de la mantisa como del exponente y el último dígito representa el exponente decimal.

El significado para el dígito que representa el signo de la mantisa y del exponente para los coeficientes es el indicado en la Tabla 6.1.

- El operador de cruce intercambia un número aleatorio de dígitos entre un cromosoma y otro (intercambia los comprendidos entre dos posiciones elegidas al azar).

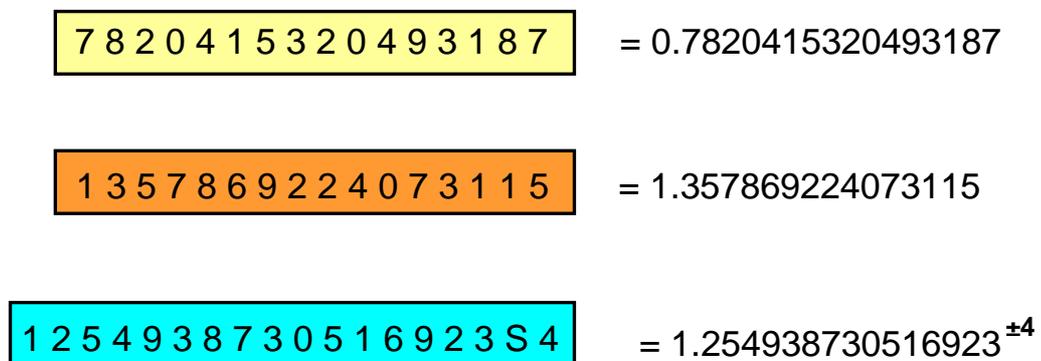


Figura 6.4. Interpretación del valor de cada gen.

- El operador mutación sustituye el valor de un dígito seleccionado por cualquier otro escogido al azar.

| Séptimo dígito | Significado |
|----------------|--------------------------------------|
| 0, 1 ó 2 | Mantisa positiva, exponente positivo |
| 3, 4 ó 5 | Mantisa negativa, exponente positivo |
| 6 ó 7 | Mantisa positiva, exponente negativo |
| 8 ó 9 | Mantisa negativa, exponente negativo |

Tabla 6.1. Significado del séptimo dígito del gen de los coeficientes.

- Se aplica elitismo, de forma que de una generación a otra se conserva un porcentaje preestablecido (hasta un 4%) de los mejores individuos de la generación anterior. Con esto se consigue que la aptitud del mejor individuo no empeore entre una generación y su consecutiva.

- La función objetivo es el error RMSE medio entre los valores de salida del sistema de inferencia difusa TSK, definido por los parámetros contenidos en cada cromosoma, y los valores reales. Se trata, por tanto, de un problema de minimización de la función objetivo. La aptitud de cada individuo se calcula como la diferencia entre un valor constante y preestablecido y el error RMSE medio del sistema de inferencia difusa.
- Para la reproducción se utiliza una ruleta de Holland, en la que los individuos se repiten en función de su posición en el ranking establecido según el valor de aptitud en la última generación, de acuerdo a la ecuación (6.2).

$$\text{número de repeticiones} = \text{individuos por generación} + 1 - \text{posición en el ranking} \quad (6.2)$$

- Se utiliza un nuevo operador, el de rotación, con una finalidad similar al operador de mutación (promover un proceso más aleatorio de búsqueda de soluciones). Seleccionado un individuo y un número entero aleatorio, n , inferior a la longitud del cromosoma, intercambia la posición de los dos fragmentos en que se divide el cromosoma al partir por la posición (dígito) n , lo que equivale a rotar todos los dígitos del cromosoma hacia la izquierda n veces (el primer dígito pasa a ser el último en cada rotación).
- Las tasas de cruce y de mutación son seleccionables por el usuario. Los valores por defecto utilizados corresponden a un 4% de individuos escogidos por elitismo, un 84% por cruce y un 10% por mutación. Para este último caso, la probabilidad de mutar para cada dígito del cromosoma es, por defecto, del 2%.
- El número total de genes (parámetros) contenidos en un cromosoma viene dado por la ecuación (6.3), en la que m representa el número de entradas y n el número de reglas.

$$\text{número de genes} = 2 \cdot m \cdot n + n(m + 1) \quad (6.3)$$

- Para obtener individuos que representen a una buena solución desde la primera generación, el primer individuo en ser creado puede corresponder al de parámetros obtenidos por la aplicación del agrupamiento substractivo a los datos de entrenamiento o ajuste, de tal forma que el número de clusters obtenido sea igual al de reglas. Los centros de las funciones de pertenencia corresponden a

los centros de los clusters, su varianza se calcula de acuerdo a la ecuación (5.33) y los coeficientes de los polinomios se obtienen aplicando regresión lineal, con el conjunto de puntos (valores del vector de entradas) pertenecientes a cada cluster (cada valor del vector de entradas pertenece al cluster más cercano, en términos de distancia euclídea).

Una vez obtenido el primer individuo, el resto de individuos de la primera generación se puede obtener por mutación del cromosoma de este primer individuo.

El tipo de codificación escogido evita tener que realizar comprobaciones de la consistencia de cada uno de los parámetros, ya que cualquier combinación de dígitos en cualquier gen contiene un valor válido para cada parámetro.

6.3.2. Resultados del sistema de optimización propuesto.

Tras crear este original sistema de optimización, se realizaron diversas pruebas en la optimización de un sistema de inferencia difusa usado para la predicción de un nuevo valor para una serie temporal. Los resultados obtenidos con los sistemas de inferencia difusa, optimizados mediante el sistema de optimización propuesto, eran equiparables a los obtenidos con un sistema neuro-difuso utilizando el mismo número de entradas y reglas.

A modo ilustrativo, a continuación presentamos una de las experiencias realizadas, mediante un sistema de inferencia difusa, sobre la predicción de la potencia media horaria generada en un parque eólico. El fragmento de serie escogido para la realización de la experiencia correspondió a 500 datos de la “Serie II”, que fueron previamente normalizados a valores comprendidos entre 0.1 y 0.9. El objetivo perseguido con el desarrollo de un modelo de predicción, fue el de la predicción de la potencia eléctrica media horaria generada en el parque eólico para la hora siguiente, tomando como entradas los últimos tres valores conocidos de la propia variable, es decir, se trató de desarrollar un modelo auto-regresivo no lineal. El número de reglas se limitó a dos.

Para comparar los resultados de predicción obtenidos con el sistema de inferencia difusa, cuyos parámetros fueron optimizados con el original sistema propuesto, se desarrollaron otros dos modelos: un modelo consistió en una red neuronal

MLP con una capa oculta cuyos parámetros también se optimizaron con un algoritmo genético; y el otro modelo consistió en un sistema neuro-difuso (ANFIS) con el mismo número de reglas (fue construido utilizando el algoritmo de agrupamiento substractivo y la función “anfisedit” [Jang95]).

Los datos disponibles fueron divididos en tres grupos: entrenamiento (60% de los datos), validación (20% de los datos) y testeo (20% restante). Para el entrenamiento de la red neuronal MLP se utilizaron los datos de los grupos de entrenamiento y validación, al igual que para el entrenamiento del sistema neuro-difuso. En ambos casos, los valores de los parámetros optimizados correspondían a aquellos que proporcionaban el mínimo error cuadrático medio con los datos del grupo de validación. Para la optimización de la red neuronal MLP se estableció un total de 50 generaciones con 50 individuos por generación; el número máximo de epochs se estableció en 1500.

Para el desarrollo del sistema de inferencia difusa, optimizado con el sistema propuesto, se utilizaron conjuntamente los datos de los grupos de entrenamiento y validación, y se estableció un total de 500 generaciones con 100 individuos por generación. Todos los individuos de la primera generación se crearon aleatoriamente. Se usó elitismo para el 4% de la población; el 84% de los individuos de la nueva generación se creó por cruce; el 10% por mutación (un 2% de probabilidad de cambio de cada dígito del cromosoma correspondiente); y el 2% por rotación. Los datos utilizados en la evaluación de la aptitud de cada individuo fueron el conjunto de los de entrenamiento y validación.

Los resultados obtenidos, usando los datos del grupo de testeo con los tres modelos probados, son los recogidos en la Tabla 6.2, en la que se muestra que el indicador RMSE con los datos del grupo de testeo es menor para el sistema de inferencia difusa optimizado con el sistema propuesto (FIS genético). En la Tabla 6.2 también se refleja el tiempo empleado en el desarrollo de los modelos con un ordenador Pentium D a 3 GHz.

| | RMSE Entrenamiento | RMSE Validación | RMSE Testeo | Tiempo |
|--------------|--------------------|-----------------|-------------|------------|
| Red MLP | 0.093790 | 0.077507 | 0.093228 | 15 minutos |
| ANFIS | 0.093433 | 0.089050 | 0.094495 | 8 minutos |
| FIS genético | 0.092481 | | 0.087146 | 11 minutos |

Tabla 6.2. Resultados obtenidos en la prueba del sistema de optimización propuesto.

A pesar de que resultados obtenidos con el sistema de optimización propuesto

son muy esperanzadores, su uso para grandes grupos de datos precisa de tiempos de optimización significativos. Esto se debe a que la función objetivo utilizada es el error RMSE medio con los datos de los grupos de entrenamiento y de validación. El sistema de optimización calcula el valor de este indicador tomando todos esos datos, para cada individuo de cada generación, lo que hace que, si el número de datos de los grupos de entrenamiento y validación es elevado, se precisen de tiempos de tiempos de cálculo considerables. En el momento actual se sigue trabajando en la mejora del sistema de optimización propuesto.

6.4. CONCLUSIONES

Los métodos de optimización, basados en algoritmos genéticos, aplicados a la selección de parámetros de modelos de predicción, constituyen una interesante herramienta para el desarrollo de dichos modelos.

En los trabajos de investigación recogidos en este documento se han utilizado métodos de optimización para la mejora de los modelos basados en redes neuronales artificiales, permitiendo la selección de los parámetros de la mejor red neuronal que minimizan los errores de predicción para los datos del grupo de validación.

Para la optimización de parámetros de los sistemas de inferencia difusa, se ha creado un sistema de optimización original, cuyos resultados son comparables a los obtenidos usando las redes neuronales citadas anteriormente y a los obtenidos utilizando sistemas neuro-difusos. Actualmente se sigue trabajando en la mejora del nuevo sistema de optimización creado.

7. NUEVOS MODELOS AVANZADOS PARA LA PREDICCIÓN A CORTO PLAZO DE LA PRODUCCIÓN ELÉCTRICA EN PARQUES EÓLICOS

En este capítulo se presentan nuevos modelos avanzados de predicción a corto plazo de producción eléctrica en parques eólicos. Se indican en primer lugar los datos disponibles para la realización de los trabajos de investigación, con los que se han creado los nuevos modelos de predicción. Posteriormente se indican la preparación previa y selección realizada con los datos. A continuación se describen dichos modelos originales, se definen sus características fundamentales y las técnicas utilizadas.

Los modelos son presentados haciendo referencia a la diferente cantidad y tipo de datos disponibles usados en su desarrollo. Los primeros modelos están basados en series temporales y, por tanto, con horizontes de predicción de unas horas. Los últimos modelos de predicción de la producción eléctrica están basados en el uso de valores de predicción de variables meteorológicas, obtenidas de modelos atmosféricos, extendiéndose el horizonte de predicción hasta los tres días.

7.1. CARACTERÍSTICAS DE LOS PARQUES EÓLICOS. DATOS PARA LA PREDICCIÓN

La obtención de datos ha sido una labor relevante en los trabajos de investigación. Así, fueron necesarias diversas gestiones con responsables de entidades como Iberdrola S.A., Red Eléctrica de España S.A., Energía Hidroeléctrica de Navarra S.A., Eólicas de La Rioja S.A., además de otras gestiones fuera de nuestras fronteras, para conseguir datos para la predicción.

Los datos referentes a un parque eólico, con un interés significativo para el desarrollo de los modelos de predicción de la producción eléctrica en el mismo, son los indicados seguidamente.

1. Serie temporal de la potencia eléctrica media generada en el parque eólico con una resolución mínima de una hora.
2. Registros de los sistemas SCADA del parque con la información sobre funcionamiento de cada aerogenerador conteniendo, al menos, la potencia

eléctrica media generada con la misma resolución mínima que la serie temporal de la potencia eléctrica media generada en todo el parque. Estos datos permiten determinar los aerogeneradores que están fuera de servicio o bajo unas condiciones de funcionamiento anómalas (por avería, actuación de los sistemas de seguridad, parada programada de mantenimiento, etc.) y no contribuyen en algunos periodos de tiempo a la generación de potencia eléctrica del parque.

3. Registros de las estaciones meteorológicas existentes en el parque eólico. Las variables de interés son los valores medios de velocidad y dirección del viento, la temperatura, la presión atmosférica y la humedad. La resolución mínima debe ser también de una hora.
4. Descripción de la configuración del parque eólico: situación, número y características de los aerogeneradores y localización concreta de cada aerogenerador. También es necesaria la orografía del terreno sobre el que se sitúa el parque eólico. Estos datos permiten el desarrollo de modelos “físicos”, como los descritos en [Landberg99], [Beyer99] o [Tambke05].
5. Predicciones de los valores numéricos de variables meteorológicas obtenidas por medio de un modelo atmosférico (HIRLAM, MM5, ALADDIN, SIKRON, etc.). Lo ideal es disponer de estos valores para los cuatro puntos de la rejilla tridimensional de análisis correspondiente a los niveles de presión atmosférica más cercanos a la superficie del terreno, para poder interpolarlos en la posición asociada al parque. Las variables de interés son los valores medios de velocidad y dirección del viento, la densidad del aire, la temperatura y la presión atmosférica a nivel de superficie del terreno. Igualmente, la resolución mínima debe ser también de una hora.

Sin embargo, en diversas ocasiones la información de datos es insuficiente y/o incompleta, por lo que se hace necesario el desarrollo de modelos avanzados idóneos para situaciones de incertidumbre en la información de partida. En los próximos párrafos se indica la relación de datos obtenidos para la realización de los trabajos de investigación, así como una breve descripción de los mismos y de sus características. Cuando se omite la localización geográfica concreta del parque eólico es porque no fue facilitada por el suministrador de los datos o porque existe compromiso de confidencialidad que impide su concreción.

- Serie I: serie temporal de los valores medios horarios de la velocidad y dirección

del viento tomados en una estación meteorológica de La Rioja. La serie cubre un periodo de tres años.

- Serie II: serie temporal de la potencia eléctrica media horaria de dos parques eólicos del norte de España (provincias de Burgos y Navarra). Las series temporales sólo incluyen el dato de la hora y la potencia eléctrica media horaria generada en cada parque para dicha hora. Los datos corresponden a un total de 18 meses.
- Serie III: serie temporal de datos de un parque eólico en el norte de la península ibérica correspondiente a un periodo de 14 meses. El parque consta de 10 aerogeneradores de 1.8 MW. Para los últimos 68 días se dispone además de las predicciones atmosféricas medias horarias de velocidad y dirección del viento, para un punto cercano al parque eólico, obtenidas con el modelo atmosférico MM5. Los datos del parque eólico incluyen los registros del equipo SCADA (valores medio, mínimo y máximo de la potencia eléctrica generada, velocidad de giro, velocidad y dirección del viento en cada aerogenerador). Las predicciones atmosféricas del modelo MM5 cubren un horizonte de predicción de 0 a 24 horas (los datos corresponden a las predicciones atmosféricas hechas a las 0 horas de cada día y cubren hasta las 24 horas).
- Serie IV: serie temporal de datos de un parque eólico en la provincia de Navarra, compuesto de 50 aerogeneradores, correspondiente a un año completo. La serie temporal incluye los registros medios horarios del sistema SCADA y las previsiones de valores medios de velocidad y dirección del viento, para cada seis horas y para un punto del parque, obtenidas del modelo de previsión ECMWF (modelo atmosférico regional). La serie no está completa, observándose deficiencias en los registros del sistema SCADA (fundamentalmente faltan registros de la potencia media horaria generada en cada aerogenerador).
- Serie V: serie temporal de datos de un parque eólico situado en el norte de la península ibérica, con un total de 12 aerogeneradores y una potencia instalada de 21.6 MW, para un año. Se incluyen los datos de predicciones de valores medios de velocidad y dirección del viento, temperatura y presión atmosférica obtenidas con el modelo MM5 con horizontes de predicción de 0.5 a 72 horas en intervalos de 0.5 horas. La serie recoge los valores de potencia eléctrica media entregada en la subestación cada media hora, así como los registros SCADA de cada aerogenerador. También se dispone de la curva de potencia del modelo de aerogenerador instalado. La serie no está completa, faltando datos intermedios.

Los resultados de los trabajos de investigación desarrollados están condicionados por las características de los datos disponibles, arriba señalados, pudiéndose asegurar que los resultados obtenidos con los nuevos modelos avanzados son mejores cuanto mayor es la cantidad y calidad de los datos utilizados.

7.2. SELECCIÓN Y NORMALIZACIÓN DE VARIABLES PARA LOS MODELOS DE PREDICCIÓN

7.2.1. Selección de variables de entrada para el modelo.

La selección de las variables de entrada para un sistema de predicción de la producción eléctrica en un parque eólico constituye una de las fases más laboriosas del proceso de diseño del mismo. Resulta evidente que la incorporación de una variable explicativa con una elevada correlación con la variable a predecir puede mejorar significativamente el funcionamiento (mejora significativa de exactitud) del sistema de predicción. Por otra parte, la inclusión como variable de entrada en el sistema de predicción de una variable con muy baja correlación con la variable a predecir únicamente ralentizará el sistema de predicción, ya que éste debe procesar esa información suministrada por la variable de entrada aunque su contribución al valor de la salida (variable a predecir) sea muy pequeña. Así, aunque en la fase de diseño pueda parecer conveniente considerar todas las variables de entrada posibles, esta decisión no suele ser acertada.

Además el hecho de considerar más variables explicativas que las estrictamente imprescindibles provoca la necesidad de disponer una cantidad mayor de muestras en lo que se conoce como “curso de dimensionalidad” [Härdle90], lo que significa que para poder modelizar adecuadamente la predicción de la producción eléctrica, el número de datos disponibles debe crecer exponencialmente con el número de variables de entrada. Así, resulta importante reducir el número de entradas a las que realmente tienen una influencia notable en el valor de la salida del modelo de predicción.

Sin un conocimiento apriorístico de la influencia de las posibles variables de entrada podríamos intentar construir tantos modelos como fueran posibles, considerando en cada uno de ellos diferentes combinaciones de las posibles variables de entrada. De esta forma podríamos determinar empíricamente el modelo con la mejor combinación de entradas, siendo éste el que presentando un error de predicción menor, también ha requerido menor tiempo de ajuste o entrenamiento. Pero esto resulta una

labor especialmente laboriosa: por ejemplo, supongamos que el número de posibles variables explicativas es de 4; podemos construir un total de 15 modelos diferentes considerando todas las posibles combinaciones de las entradas (uno con las cuatro entradas, 4 modelos considerando tres entradas, 6 modelos considerando 2 entradas y 4 modelos considerando una de las entradas). El número de combinaciones crece también exponencialmente con el número de posibles variables de entrada; así para 8 posibles entradas el número de modelos diferentes (sólo en cuanto a las entradas consideradas) es de 255, y para 12 entradas de 4095. Si n es el número de posibles entradas, entonces $2^n - 1$ representa el total de combinaciones que se pueden realizar con esas entradas y, por tanto, de diferentes sistemas que se podrían construir.

El número de posibles variables explicativas para ser utilizadas en un problema de predicción a corto plazo puede ser muy grande. Así, para un problema con algunas similitudes al de la predicción de la producción eléctrica en un parque eólico, como puede ser el de la previsión a corto plazo de la demanda de energía eléctrica (“short-term load forecasting” en la literatura especializada), el número de variables explicativas puede ser de hasta 500, como en el caso descrito en [Mohammed95]. La reducción del número de entradas acorta considerablemente el tiempo necesario para ajustar o entrenar los sistemas de predicción. Un ejemplo comparativo de la reducción de tiempo obtenida, aunque en el campo de la clasificación con redes neuronales, puede encontrarse en [Ceylan07].

El enfoque clásico para la selección de variables, utilizado fundamentalmente en la construcción de modelos de regresión lineal múltiple, consiste en crear una relación ordenada de entradas en función de los resultados obtenidos con los modelos desarrollados, considerando las entradas una a una (sólo se considera una entrada por modelo). La variable de entrada que haya dado lugar al modelo con menor error se retiene como variable de entrada, formándose con ella y otra variable más todos los modelos posibles con dos entradas. La variable correspondiente al modelo con menor error (evidentemente otra distinta a la seleccionada anteriormente) conforma la segunda variable seleccionada. El proceso continúa construyendo todos los modelos posibles con tres entradas (dos de ellas son las ya seleccionadas) y escogiendo la tercera variable de acuerdo al criterio seguido con las dos primeras. El proceso de selección finaliza cuando no se obtiene una mejora preestablecida en los resultados de los nuevos modelos como consecuencia de añadir una nueva variable de entrada.

El proceso de selección ordenada, comentado en el párrafo anterior, también puede realizarse en sentido inverso, esto es, comenzando con un modelo constituido con todas las posibles variables de entrada, para construir después todos los posibles

modelos con una entrada menos. La primera variable en ser eliminada como posible variable explicativa es aquella que no está incluida en el modelo que origine menor error. El proceso continúa eliminando una a una las posibles variables hasta que la diferencia de errores entre un paso y otro consecutivo es menor que un valor prefijado de antemano. Los dos procesos de selección ordenada pueden conjugarse a la vez, permitiendo la selección de variables de entrada de una forma más rápida, aunque con un gran número de variables posibles resulta todavía lento.

Otra alternativa a la selección de entradas se basa en el llamado estudio de sensibilidad. Se construye un modelo utilizando como entradas todas las disponibles. Una vez ajustado o entrenado el modelo se prueba su funcionamiento haciendo que el valor de todas las entradas coincida con su valor medio respectivo, excepto para una de las entradas en las que su valor varía ligeramente en torno a su valor medio. Las variables de entrada que con su variación en torno a su valor medio producen mayores variaciones en la salida, es decir, presentan mayor sensibilidad, son aquellas que tienen un mayor peso relativo en el valor de la misma. Aquellas otras entradas cuya variación respecto a su valor medio apenas tienen influencia en la variación del valor de salida, denotando una baja sensibilidad, podrían eliminarse como posibles variables de entrada.

Son varias las técnicas utilizadas con objeto de lograr una selección más rápida de variables de entrada para sistemas lineales, como el análisis de correlaciones y auto-correlaciones entre las variables o la reducción del número de variables de entrada (espacio de entrada) por medio del análisis de componentes principales [Jolliffe02] u otra técnica de transformación del espacio de entrada que permita la selección de variables como en [Drezga98]. Pero en general estos métodos, que están concebidos para sistemas lineales, son relativamente poco adecuados para sistemas no lineales, lo que ha dado pie al desarrollo de nuevos métodos de selección de variables de entrada basadas en técnicas meta-heurísticas próximas a los métodos de inteligencia artificial: redes neuronales [Linkens99], lógica fuzzy [Hadjili02] y [Sindelar04], técnicas evolutivas para el diseño de redes neuronales artificiales [Yao99], o para el diseño de sistemas de inferencia difusa [Xiong02].

7.2.2. Normalización de variables de entrada.

Las variables de entrada a un modelo de predicción presentan rangos y características muy distintas, por lo que se precisa de alguna técnica de normalización de los valores de estas variables, de tal forma que el modelo de predicción pueda dar un tratamiento interno similar a todas las variables de entrada. El proceso de normalización

se aplica usualmente en modelos de predicción basados en aprendizaje, como es el caso de redes neuronales artificiales.

El proceso de normalización de una variable consiste en “escalar” el rango de valores de dicha variable de forma que esté comprendido dentro de los límites aceptables por el sistema. Esta existencia de límites aceptables viene impuesta por la necesidad de generalización, de poder aplicar las mismas herramientas a casos de diferente naturaleza. El conjunto de algoritmos, rutinas o procedimientos implementados en un modelo de predicción precisa que las variables de entrada presenten unos valores acotados entre unos límites aceptables.

Existen varias técnicas de normalización y de uso muy común en el desarrollo de redes neuronales artificiales. La más sencilla consiste en “escalar” todas las variables de entrada entre los valores 0 y 1 (especialmente utilizada en el caso de trabajar con funciones sigmoideas en la red neuronal). Por ejemplo, el valor normalizado de la variable e para la muestra j de un total de N muestras de la variable, supuestas todas positivas, viene dado por la ecuación (7.1).

$$e_{jnor} = \frac{e_j}{\max(e_k)_{k=1\dots N}} \quad (7.1)$$

En la ecuación (7.1) el denominador representa el máximo valor de la variable e para el total de las N muestras.

Una mejor normalización que la expresada en la ecuación (7.1) consistirá en ajustar al valor 0 el mínimo valor de la variable a normalizar y ajustar a 1 el máximo valor de la variable a normalizar, tal y como se recoge en la ecuación (7.2), donde ahora el valor de la variable e no tiene por que ser siempre positivo.

$$e_{jnor} = \frac{e_j - \min(e_k)_{k=1\dots N}}{\max(e_k)_{k=1\dots N} - \min(e_k)_{k=1\dots N}} \quad (7.2)$$

Otra alternativa de normalización es la de “escalar” los valores entre -1 y 1 , que es muy común en el caso de utilizar redes neuronales en las que la función de activación sea la tangente hiperbólica (\tanh , también conocida como sigmoidal bipolar). En este caso, el valor normalizado de la variable e para la muestra j de un total de N muestras de la variable, viene dado por la ecuación (7.3).

$$e_{jnor} = 2 \frac{e_j - \min(e_k)_{k=1\dots N}}{\max(e_k)_{k=1\dots N} - \min(e_k)_{k=1\dots N}} - 1 \quad (7.3)$$

Sin embargo, en muchas aplicaciones no podemos asegurar que el rango del grupo de datos de entrenamiento o ajuste realmente incluye todos los valores posibles de las variables; en tal caso lo común es normalizar entre -0.9 y 0.9 , ofreciendo así un margen, tanto por encima del máximo de los valores del grupo de datos de entrenamiento como por debajo del mínimo (la función de activación ofrece valores entre -1 y 1). La normalización entre -0.9 y 0.9 viene dada por la ecuación (7.4).

$$e_{jnor} = 1.8 \frac{e_j - \min(e_k)_{k=1\dots N}}{\max(e_k)_{k=1\dots N} - \min(e_k)_{k=1\dots N}} - 0.9 \quad (7.4)$$

7.2.3. Técnicas de selección y normalización utilizadas.

Se ha utilizado una técnica de selección de variables de entrada, basada en algoritmos genéticos, para los modelos de predicción creados sobre redes neuronales artificiales. Señalamos seguidamente algunas razones que han conducido a elegir esta técnica.

- El desarrollo de modelos de predicción basados en redes neuronales artificiales requiere que se especifiquen desde el primer momento parámetros como el número de entradas, el número de neuronas por capa, los valores del factor de aprendizaje y del momento para las neuronas de cada capa, el número de unidades de retardo (para una red neuronal recurrente retardada en el tiempo), el número de componentes principales (para una red neuronal PCA), etc. Un adecuado sistema de optimización automática, basado en el uso de algoritmos genéticos, permite determinar una buena parte de las decisiones idóneas sobre la estructura óptima de una red neuronal.
- El procedimiento de selección de variables es automatizable: las operaciones de cruce y mutación permiten dicha selección.
- El procedimiento puede combinarse con técnicas de reducción del espacio de entradas: para el caso de modelos de predicción de producción eléctrica con muchas entradas, la selección de entradas puede ser previa a la aplicación de técnicas de reducción.

De esta forma, la selección de las variables de entrada se realiza de forma automática, dejando que sea la evolución de las soluciones (modelos de predicción) la que conduzca a una selección óptima de éstas.

Una vez optimizado el modelo de predicción, puede realizarse un estudio de sensibilidad, lo que aún puede permitir eliminar alguna variable de entrada más de las consideradas (alguna variable con baja sensibilidad).

En cuanto al rango de las variables utilizadas para la construcción de los modelos de predicción, la Tabla 7.1 recoge, a modo de ejemplo, los rangos correspondientes a los datos utilizados en el desarrollo de los modelos para mayor horizonte de predicción (Serie V). En dicha tabla, P_{inst} corresponde a la potencia total instalada en el parque eólico; la potencia eléctrica media es la medida en la subestación de conexión con la red eléctrica y los valores de las variables meteorológicas corresponden a las predicciones del modelo atmosférico.

| Variable | Valor mínimo | Valor máximo |
|--------------------------|-------------------------|------------------------|
| Potencia eléctrica media | $-0.005 \cdot P_{inst}$ | $1.012 \cdot P_{inst}$ |
| Velocidad viento (m/s) | 0 | 27.2 |
| Dirección viento (°) | 0 | 360 |
| Temperatura (°C) | -7.3 | 34.6 |
| Presión atmosférica (Pa) | 86479 | 90030 |
| Hora del día | 0 | 23 |

Tabla 7.1. Valores máximo y mínimo de las variables.

A pesar de que los datos disponibles están acotados entre los límites indicados en la Tabla 7.1, podría darse la circunstancia de que con otro conjunto de datos de la misma fuente (mismo parque eólico) se sobrepasaran esos límites. Por ejemplo, podría darse un valor de velocidad de viento por encima de los valores disponibles actualmente; lo mismo podría suceder con otras variables como la potencia eléctrica media, la temperatura o la presión atmosférica.

Para asegurar un cierto margen en la normalización de las entradas, se ha optado por una de las normalizaciones indicadas a continuación.

- Normalizar los valores de todas las entradas entre -0.9 y $+0.9$. Esta normalización ha sido aplicada para todos los datos utilizados en el desarrollo de modelos de predicción basados en redes neuronales artificiales.
- Normalizar los valores de todas las entradas entre $+0.1$ y $+0.9$. Esta normalización ha sido aplicada para todos los datos utilizados en el desarrollo de modelos de predicción basados en sistemas de inferencia difusa.

7.3. NUEVOS MODELOS DE PREDICCIÓN DE LA PRODUCCIÓN ELÉCTRICA BASADOS EN SERIES TEMPORALES

Los primeros modelos de predicción de la producción eléctrica en un parque eólico, desarrollados en los trabajos de investigación de este documento, corresponden a modelos de la primera familia (modelos basados en series temporales) de acuerdo a la clasificación de modelos, extraída del análisis de las publicaciones internacionales sobre el tema, realizada en el Capítulo 3.

Como consecuencia de los trabajos de investigación, se han desarrollado tres nuevos modelos basados en series temporales de la potencia eléctrica media horaria. La variable utilizada en el desarrollo de los modelos ha sido exclusivamente una serie temporal. En un caso se ha usado la serie temporal de velocidad del viento y en los otros dos casos la serie temporal de potencia eléctrica media horaria generada en el parque eólico considerado. La variable de salida ha sido la potencia eléctrica media horaria o energía generada para las próximas horas. En las siguientes páginas se describen los detalles y aplicaciones de cada uno de los modelos desarrollados.

7.3.1. Modelo de predicción basado en descomposición de Fourier.

En [Atiya99] se describe un sistema de predicción de series temporales basado en redes neuronales, donde se proponen diferentes estructuras de redes neuronales para la predicción del caudal de agua en el río Nilo, con un horizonte de predicción de varios días. Una de las estructuras propuestas presenta como entradas los coeficientes obtenidos con la aplicación de la transformada discreta de Fourier, DFT, sobre los últimos valores conocidos de la serie temporal de nivel del agua del río. Para cada uno de los coeficientes se entrena una red neuronal que realiza la predicción del valor futuro de dicho coeficiente. La transformación inversa (transformada inversa discreta de Fourier, IDFT) con los valores de predicción de todos los coeficientes proporciona el valor del nivel del agua del río para el horizonte de predicción propuesto.

Utilizando la transformada DFT, hemos desarrollado un nuevo modelo de predicción de la velocidad media del viento con horizonte de 1 hora descrito en [Ramirez01] y ampliado en [Ramirez04a]. El análisis de los valores obtenidos de la función de auto-correlación con los datos de la velocidad media horaria del viento arrojan una clara estacionalidad, como se desprende de la variación sinusoidal recogida en la Figura 7.1. Obviamente el valor de la estacionalidad de la serie temporal de

velocidad media del viento corresponde a un valor de 24 horas.

Cualquier función periódica puede ser descompuesta en una serie de funciones sinusoidales conocida como serie de Fourier. En el caso de funciones periódicas o no periódicas pero muestreadas a intervalos regulares de tiempo, es posible realizar dicha descomposición por medio de la transformada discreta de Fourier (DFT), que permite calcular los coeficientes del desarrollo en serie de las funciones sinusoidales.

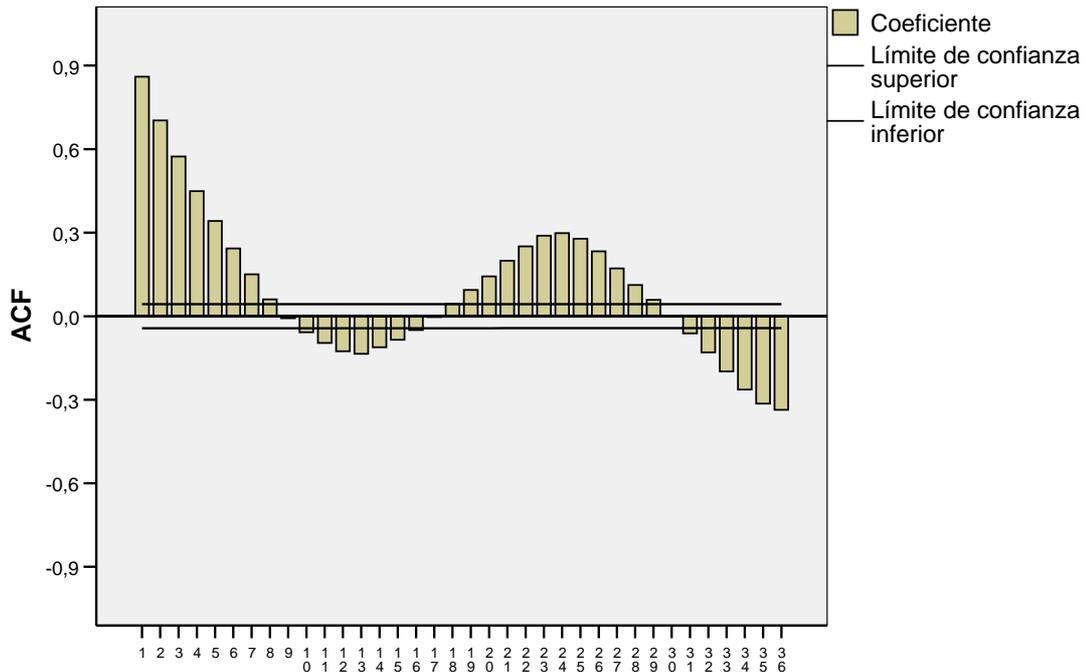


Figura 7.1. Coeficientes de auto-correlación para la serie temporal de velocidad media horaria del viento del conjunto de datos de la Serie I.

Si a los 24 últimos valores conocidos de la serie de velocidad media horaria del viento se le aplica la transformada DFT, se obtienen otros tantos coeficientes, que en general son números complejos, $a_k + jb_k$, con k entre 0 y 23. Estos coeficientes pueden calcularse por medio de las ecuaciones (7.5) y (7.6), donde $x(n)$ representa al término n de los 24 últimos valores conocidos de la serie temporal de la velocidad media horaria del viento.

$$a_k = \sum_{n=1}^{24} x(n) \cos\left(\frac{2\pi(n-1)k}{24}\right) \quad (7.5)$$

$$b_k = -\sum_{n=1}^{24} x(n) \operatorname{sen}\left(\frac{2\pi(n-1)k}{24}\right) \quad (7.6)$$

De los 24 coeficientes complejos, resultados de la aplicación de la transformada DFT, dos no tienen más que parte real. Concretamente son nulos los coeficientes b_0 y

b_{12} . Por otra parte, casi la mitad de los coeficientes son el complejo conjugado de la otra mitad, tal y como expresa la ecuación (7.7).

$$a_k + jb_k = a_{24-k} - jb_{24-k} \quad k = 1, 2, \dots, 11 \quad (7.7)$$

El número total de coeficientes (incluyendo como tales las partes real e imaginaria de los 24 complejos originales) a determinar queda así reducido a 24: 13 coeficientes correspondientes a las partes reales y otros 11 coeficientes correspondientes a las partes imaginarias. Aún es posible reducir un coeficiente más, concretamente el coeficiente a_6 . En la ecuación (7.8) se recoge el valor del coeficiente a_6 para el instante siguiente al de realizar la predicción, es decir, para $t+1$. El valor de este coeficiente, $a_6(t+1)$, coincide con el valor del coeficiente b_6 en el instante de predicción, $b_6(t)$, cuyo valor aparece en la ecuación (7.9). Ambas ecuaciones, (7.8) y (7.9) arrojan el mismo resultado.

$$a_6(t+1) = -\sum_{n=1}^{24} x(t+1-24+n) \cos\left(\frac{\pi(n-1)}{2}\right) \quad (7.8)$$

$$b_6(t) = -\sum_{n=1}^{24} x(t-24+n) \operatorname{sen}\left(\frac{\pi(n-1)}{2}\right) \quad (7.9)$$

Con esto, el número total de coeficientes reales que permiten recuperar perfectamente la serie de 24 valores original, $x(n)$, se reduce a 23, según se reflejan en la Tabla 7.2.

| | | | | | |
|----------|----------|-------------|----------|----------|-------------|
| $a_0(t)$ | $a_4(t)$ | $a_9(t)$ | $b_1(t)$ | $b_5(t)$ | $b_9(t)$ |
| $a_1(t)$ | $a_5(t)$ | $a_{10}(t)$ | $b_2(t)$ | $b_6(t)$ | $b_{10}(t)$ |
| $a_2(t)$ | $a_7(t)$ | $a_{11}(t)$ | $b_3(t)$ | $b_7(t)$ | $b_{11}(t)$ |
| $a_3(t)$ | $a_8(t)$ | $a_{12}(t)$ | $b_4(t)$ | $b_8(t)$ | |

Tabla 7.2. Coeficientes independientes.

Los restantes 25 coeficientes necesarios para recomponer los 24 valores de la serie temporal original, la de los últimos 24 valores conocidos de la serie temporal de la velocidad media del viento, pueden calcularse a partir de los coeficientes independientes de acuerdo a los valores de la Tabla 7.3. Además debemos tener en cuenta que los coeficientes $b_0(t)$ y $b_{12}(t)$ son nulos en todo instante.

Así, la idea de construir un sistema de predicción del siguiente valor de una serie temporal (puede ser tanto la velocidad media horaria del viento como la potencia eléctrica media horaria de un parque eólico), se basa en la predicción de los 23 coeficientes independientes para, a partir de sus valores y aplicando la inversa de la

transformada discreta de Fourier (IDFT), determinar el valor de predicción de la variable original, tal y como muestra la Figura 7.2.

| | | | |
|-------------------------|----------------------|--------------------------|-----------------------|
| $a_6(t) = b_6(t-1)$ | $a_{18}(t) = a_6(t)$ | $b_{13}(t) = -b_{11}(t)$ | $b_{19}(t) = -b_5(t)$ |
| $a_{13}(t) = a_{11}(t)$ | $a_{19}(t) = a_5(t)$ | $b_{14}(t) = -b_{10}(t)$ | $b_{20}(t) = -b_4(t)$ |
| $a_{14}(t) = a_{10}(t)$ | $a_{20}(t) = a_4(t)$ | $b_{15}(t) = -b_9(t)$ | $b_{21}(t) = -b_3(t)$ |
| $a_{15}(t) = a_9(t)$ | $a_{21}(t) = a_3(t)$ | $b_{16}(t) = -b_8(t)$ | $b_{22}(t) = -b_2(t)$ |
| $a_{16}(t) = a_8(t)$ | $a_{22}(t) = a_2(t)$ | $b_{17}(t) = -b_7(t)$ | $b_{23}(t) = -b_1(t)$ |
| $a_{17}(t) = a_7(t)$ | $a_{23}(t) = a_1(t)$ | $b_{18}(t) = -b_6(t)$ | |

Tabla 7.3. Coeficientes dependientes.

Para el módulo de predicción pueden probarse distintas tecnologías que incluyen modelos ARIMA, redes neuronales artificiales, sistemas de inferencia difusa, etc. En el Capítulo 8 del presente documento se comentarán los resultados obtenidos con este sistema de predicción, utilizando uno de los conjuntos de datos disponibles (Serie I).

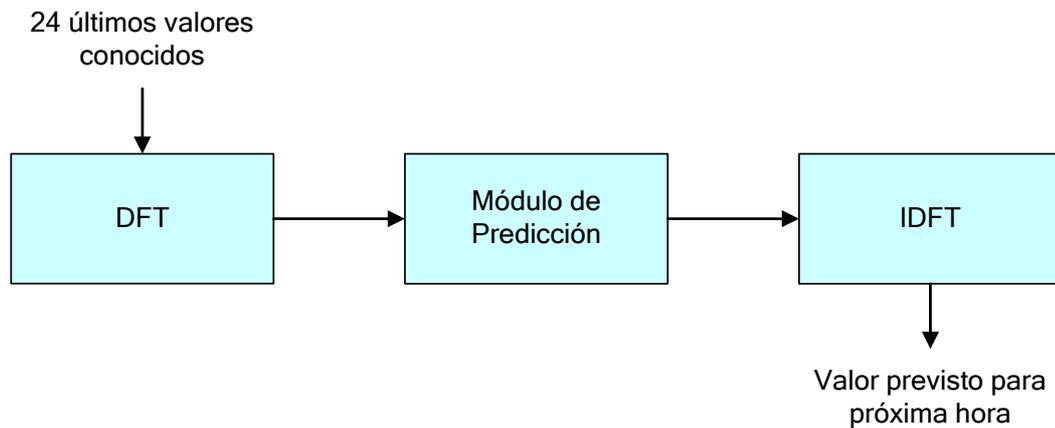


Figura 7.2. Modelo de predicción propuesto para series temporales con una componente estacional importante.

La obtención de la potencia eléctrica media generada a partir del valor de predicción de la velocidad del viento puede realizarse, en una primera aproximación, por medio de la curva de potencia de cada aerogenerador instalado en el parque eólico y sumando los valores de potencia obtenidos para cada uno de ellos.

La aplicación de este modelo, dentro del contexto de predicción a corto plazo de la producción eléctrica en parques eólicos, se refiere a aquellos casos en que no se dispone de más información que la velocidad media del viento medida en alguna estación meteorológica del propio parque. Otra situación en la que este modelo resulta interesante es para realizar las predicciones tras los primeros días de puesta en marcha de un parque eólico, cuando no se dispone de una serie temporal de potencia eléctrica media generada, pero si de dispone de una serie temporal de velocidad media del viento para ese mismo lugar (puede ser la obtenida de las estaciones o “postes de medida” que

se instalan en los lugares candidatos a la instalación del parque para evaluar la calidad de los recursos eólicos).

7.3.2. Modelo de predicción basado series temporales fuzzy.

Las series temporales fuzzy (o series temporales difusas) representan un campo relativamente nuevo de sistemas de predicción propuesto en los últimos 15 años [Song93], [Song97], [Hwang98]. Presentan como particularidad la capacidad de incluir datos históricos en las series como valores lingüísticos.

Sea U el universo del discurso, $U = \{u_1, u_2, \dots, u_n\}$. La definición para universo del discurso consiste en la familia de todos los términos usados en un discurso o problema concreto. Estos términos pueden ser lingüísticos o semánticos. Un conjunto difuso (o fuzzy) A ([Zadeh65]) de U queda definido mediante la ecuación (7.10).

$$A = \mu_A(u_1)/u_1 + \mu_A(u_2)/u_2 + \dots + \mu_A(u_m)/u_m \quad (7.10)$$

En la ecuación (7.10) μ_A es la función de pertenencia de A ; $\mu_A(u_i)$ especifica el grado de pertenencia de u_i en A ; el símbolo “+” corresponde al operador de unión y el símbolo “/” corresponde a un separador del grado de pertenencia del elemento.

Supongamos que $X(t)$ es un subconjunto de números reales y que constituye el universo del discurso; y que los conjuntos fuzzy $\mu_i(t)$ están definidos en $X(t)$. Supongamos que $F(t)$ es un grupo de $\mu_i(t)$. Entonces, $F(t)$ es una serie temporal fuzzy, FTS, de $X(t)$. La serie temporal fuzzy $F(t)$ es una función del tiempo, puesto que su valor cambia con él, y $\mu_i(t)$ son los valores lingüísticos de $F(t)$. Si el valor de $F(t)$ depende del valor de $F(t-1)$, se puede expresar esta relación mediante la ecuación (7.11).

$$F(t) = F(t-1) \circ R(t, t-1) \quad (7.11)$$

En la ecuación (7.11) $R(t, t-1)$ es la relación fuzzy entre $F(t)$ y $F(t-1)$, y el símbolo “ \circ ” corresponde al operador de composición min-max (valor mínimo o valor máximo, aunque sería posible aplicar cualquier otro). Si $R(t, t-1) = R(t-1, t-2)$ para cualquier valor de t , $F(t)$ es una serie temporal fuzzy invariante con el tiempo y su valor no depende de t . Si, por el contrario, $R(t, t-1)$ depende de t , $F(t)$ es una serie temporal fuzzy dependiente del tiempo. Obsérvese que el concepto que encierra la ecuación (7.11) es similar al de suponer que el valor actual de una variable depende del valor de

dicha variable en el pasado, propio de modelos auto-regresivos.

Para configurar un sistema de predicción utilizando series temporales fuzzy lo único necesario es convertir la serie temporal original en serie temporal fuzzy, lo que constituye un proceso de “fuzzyficación”. Una vez convertida en fuzzy la serie temporal original, es posible manipular con técnicas difusas (cálculos fuzzy) los valores de esta serie para obtener una predicción del valor de la misma (de la serie temporal fuzzy) para el horizonte de predicción considerado. Únicamente será necesario convertir el nuevo valor de la serie temporal fuzzy en un valor numérico por medio de una “defuzzyficación”. Si en los cálculos difusos sólo interviene una serie temporal fuzzy entonces decimos que se trata de un modelo de serie temporal fuzzy con un sólo factor. Cuando son n las series temporales fuzzy que intervienen en el modelo, decimos que se trata de un modelo de serie temporal fuzzy de n factores.

En el Capítulo 8 se presentarán los resultados obtenidos con dos modelos de predicción basados en series temporales fuzzy en la predicción de la energía total generada en un parque eólico para las siguientes 24 horas. Se mostrarán los resultados obtenidos mediante un modelo con una FTS de un solo factor y los obtenidos mediante un segundo modelo con una FTS de dos factores, en la que la segunda FTS corresponde a una información meteorológica expresada en términos lingüísticos.

En los siguientes dos apartados se comenta brevemente las bases de los modelos de predicción basados en FTS de uno y dos factores.

7.3.2.1. Series temporales fuzzy de un factor.

La Figura 7.3 muestra esquemáticamente un modelo de predicción para una serie temporal usando series temporales fuzzy. Este modelo (Figura 7.3) corresponde a una serie temporal fuzzy de un sólo factor. Los valores numéricos de la serie temporal son convertidos en términos fuzzy mediante el proceso de “fuzzyficación”, obteniéndose así la serie temporal fuzzy. Tras manipulaciones adecuadas, con técnicas difusas, se obtiene un nuevo valor para la serie temporal fuzzy correspondiente al horizonte de predicción considerado. La “defuzzyficación” permite obtener el valor de predicción, $\hat{f}(t + 1)$, de la serie temporal original.

El proceso de “fuzzyficación” propuesto para este modelo es similar al propuesto en [Hwang98] y [Chen00], si bien en ambos casos para problemas de predicción diferentes al planteado en el presente documento. La serie temporal original se transforma en otra aplicando la primera diferencia, es decir, el valor de la posición n

de la nueva serie corresponde a la diferencia entre los valores de posiciones $n+1$ y n de la serie original. El rango de valores de esta segunda serie temporal es dividido en k intervalos iguales, siendo identificado cada uno de ellos con un término lingüístico (cada término corresponde a un conjunto difuso) y puede expresarse de forma similar a la recogida en la ecuación (7.10). Así, para los k conjuntos difusos se tienen las ecuaciones (7.12).

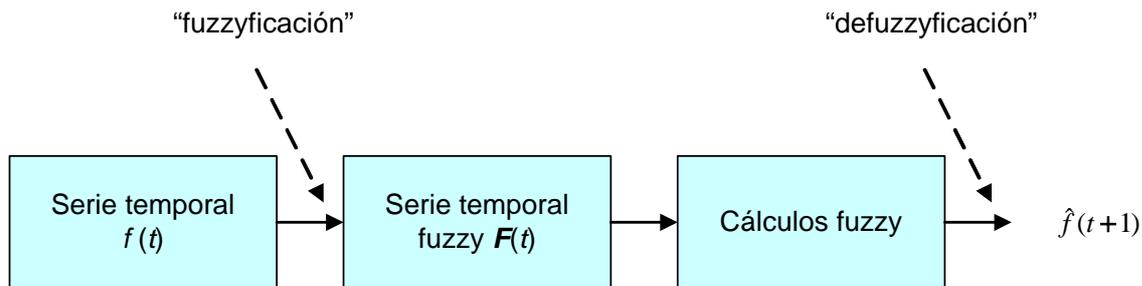


Figura 7.3. Modelo de predicción de una serie temporal fuzzy de un solo factor.

$$\begin{aligned}
 A_1 &= \mu_{A_1}(u_1)/u_1 + \mu_{A_1}(u_2)/u_2 + \dots + \mu_{A_1}(u_k)/u_k \\
 A_2 &= \mu_{A_2}(u_1)/u_1 + \mu_{A_2}(u_2)/u_2 + \dots + \mu_{A_2}(u_k)/u_k \\
 &\vdots \\
 A_k &= \mu_{A_k}(u_1)/u_1 + \mu_{A_k}(u_2)/u_2 + \dots + \mu_{A_k}(u_k)/u_k
 \end{aligned}
 \tag{7.12}$$

En las ecuaciones (7.12) los términos lingüísticos varían desde “un muy gran decremento” para el conjunto difuso A_1 , hasta “un muy gran incremento” para el último, A_k . Así, cualquier valor numérico de la serie temporal, cuyo valor está comprendido en el intervalo u_i y corresponde al término lingüístico A_i tiene un valor difuso, $F(t)$, que se puede expresar mediante la ecuación (7.13).

$$F(t) = [\mu_{A_i}(u_1), \mu_{A_i}(u_2), \dots, \mu_{A_i}(u_k)]
 \tag{7.13}$$

En los modelos auto-regresivos de series temporales, cualquier elemento de la serie puede expresarse en función de los valores de varios elementos anteriores de la misma. Para la mayoría de los casos prácticos, el peso de los elementos más recientes es mayor que el de los más lejanos en el tiempo, salvo que la serie temporal tenga una importante componente estacional. Algo similar ocurre con las series temporales fuzzy, donde el peso del elemento más reciente de la serie tiene mayor peso que los más antiguos. Ese último elemento de la FTS, el de mayor peso en el valor del nuevo elemento, es el utilizado para la formación del llamado “vector de criterio”, $C(t)$. Cada una de las componentes de $C(t)$ representan el grado de pertenencia a cada uno de los k

conjuntos difusos, tal como muestra la ecuación (7.14).

$$\mathbf{C}(t) = \mathbf{F}(t-1) = [C_1, C_2, \dots, C_m] \quad (7.14)$$

Si en un modelo clásico de auto-regresión el valor de predicción para el instante t depende de los valores de la serie en los instantes $t-1, t-2, \dots, t-w$, entonces para una serie temporal fuzzy podemos establecer una relación similar. Para ello definimos la matriz de operación $\mathbf{O}^w(t)$, creada con los últimos $w-1$ elementos de la serie, tal como expresa la ecuación (7.15). La fila j de la matriz de operación representa el conjunto difuso A que corresponde al intervalo que incluye el valor de $f(t-j-1)$.

$$\mathbf{O}^w(t) = \begin{bmatrix} F(t-2) \\ F(t-3) \\ \vdots \\ F(t-w) \end{bmatrix} = \begin{bmatrix} O_{11} & O_{12} & \dots & O_{1k} \\ O_{21} & O_{22} & \dots & O_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ O_{(w-1)1} & O_{(w-1)2} & \dots & O_{(w-1)k} \end{bmatrix} \quad (7.15)$$

Definimos la matriz de relación, $\mathbf{R}(t)$, de la ecuación (7.16) como la obtenida multiplicando columna a columna el vector de criterio y la matriz de operación, tal y como expresa la ecuación (7.17). El valor de cada elemento en la matriz de relación es el expresado en la ecuación (7.18).

$$\mathbf{R}(t) = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1k} \\ R_{21} & R_{22} & \dots & R_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ R_{(w-1)1} & R_{(w-1)2} & \dots & R_{(w-1)k} \end{bmatrix} \quad (7.16)$$

$$\mathbf{R}(t) = \mathbf{O}^w(t) \otimes \mathbf{C}(t) \quad (7.17)$$

$$R_{ij} = O_{ij} \times C_j \quad (7.18)$$

El valor de predicción en la FTS se calcula por medio de la ecuación (7.19). Evidentemente este valor es de igual naturaleza que el resto de los elementos de la FTS. Para obtener el valor de predicción de la serie temporal hay que “defuzzyficar” este último valor, lo cual puede realizarse sin más que hacerlo coincidir con el intervalo que representa. La obtención del valor de predicción en la serie temporal original no implicará más que sumar el valor “defuzzyficado” al último valor conocido de la serie.

$$F(t) = \left[\begin{array}{l} \text{Max}(R_{11}, R_{21}, \dots, R_{(w-1)1}) \\ \text{Max}(R_{12}, R_{22}, \dots, R_{(w-1)2}) \dots \\ \text{Max}(R_{1k}, R_{2k}, \dots, R_{(w-1)k}) \end{array} \right] \quad (7.19)$$

7.3.2.2. Series temporales fuzzy de dos factores.

Cuando podemos identificar otras variables explicativas para los valores de la serie temporal que queremos predecir, es posible utilizar las series temporales fuzzy de estas nuevas variables en el proceso de cálculo. En la Figura 7.4 se muestra esquemáticamente la estructura de un sistema de predicción de una serie temporal, $f(t)$, usando una variable explicativa adicional, $g(t)$. Ambas series temporales, $f(t)$ y $g(t)$, son convertidas en sus respectivas series temporales fuzzy, $F(t)$ y $G(t)$, mediante el proceso de “fuzzyficación”. La serie temporal fuzzy correspondiente a la variable a predecir es la llamada FTS del factor principal; la otra FTS, la correspondiente a la variable explicativa, es la llamada FTS del factor secundario. Tras el cálculo fuzzy entre las dos FTS, se obtiene un nuevo valor para la FTS del factor principal. La “defuzzyficación” de este valor proporciona la información necesaria para la predicción de la variable correspondiente al horizonte considerado.

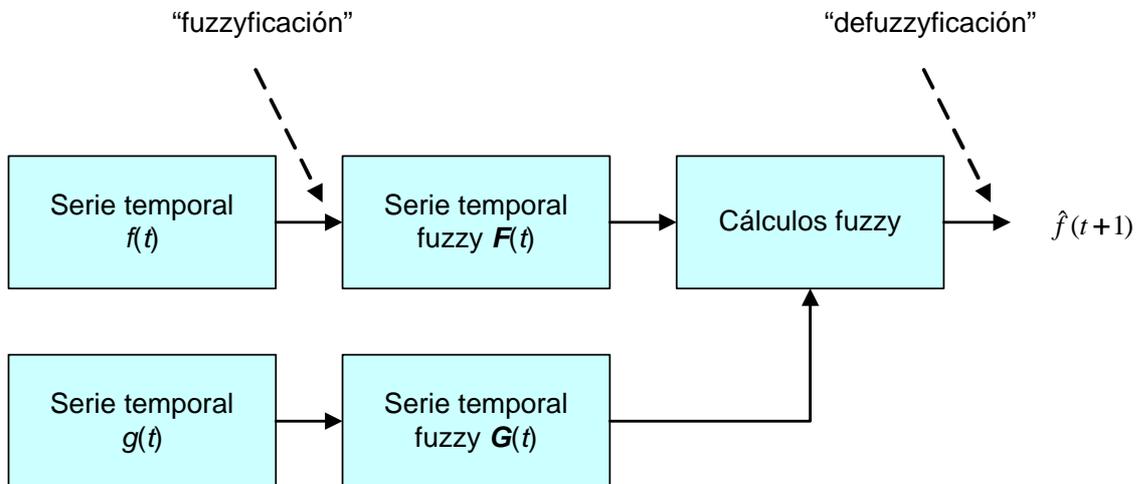


Figura 7.4. Modelo de predicción de una serie temporal fuzzy de dos factores.

Cuando se desarrolla un modelo de predicción FTS con dos factores, el procedimiento seguido es similar al de una FTS de un solo factor, con una única diferencia en el cálculo de la matriz de relación. En este caso (FTS con dos factores), los últimos valores conocidos de la segunda FTS actúan de manera similar al vector de criterio. La matriz de relación se calcula por medio de la ecuación (7.20) y cada uno de sus elementos por medio de la ecuación (7.21).

$$\mathbf{R}(t) = \mathbf{O}^w(t) \otimes \mathbf{C}(t) \otimes \mathbf{G}(t) \quad (7.20)$$

$$R_{ij} = O_{ij} \times C_j \times G_j \quad (7.21)$$

Los modelos de predicción basados en series temporales fuzzy son interesantes para aquellos casos en que se pretenda calcular la energía total producida en un periodo a corto plazo (día siguiente, tres días siguientes, etc.) y se pueda disponer de predicciones de tipo lingüístico para la velocidad de viento, por ejemplo, información tan difusa como “calma”, “viento flojo”, “viento fuerte”, etc.

7.3.3. Modelo de predicción basado en regresión lineal.

En un sistema de inferencia difusa TSK el valor de salida se obtiene como media ponderada del valor obtenido con cada una de las reglas. El valor proporcionado por cada una de las reglas es el obtenido mediante una combinación lineal de los valores de las entradas del sistema (ecuaciones de hiperplanos). Tomando como inicio este planteamiento se ha diseñado un nuevo modelo de predicción que no hace uso de técnicas difusas, sino de técnicas de regresión lineal.

La base del nuevo sistema propuesto consiste en el agrupamiento de los valores del vector de entradas del grupo de entrenamiento con características similares por medio de la aplicación del agrupamiento substractivo [Chiu94]. Una vez localizados los centros de los clusters y establecida la pertenencia de cada uno de los valores del vector de entradas del grupo de entrenamiento a un cluster u otro (el más cercano en términos de distancia euclídea), se calcula para cada uno de los grupos un modelo de regresión lineal múltiple. Cada uno de estos modelos de regresión lineal corresponde a la ecuación de un hiperplano que tiene por variables cada una de las entradas del sistema. El valor de la salida de todo el sistema se calcula por medio de una media ponderada de las salidas de todos los modelos de regresión lineal múltiple. Los coeficientes de ponderación son las inversas de las distancias euclídeas a los centros de cada cluster.

La estructura del sistema de predicción propuesto es la reflejada en la Figura 7.5, en la que el agrupamiento substractivo ha detectado n clusters dentro del conjunto de valores del vector de entradas del grupo de entrenamiento. Se obtienen así n grupos de datos, uno para cada uno de los clusters, perteneciendo cada valor del vector de entradas sólo al cluster cuyo centro es el más cercano. Para cada uno de estos grupos se construye un modelo de regresión lineal múltiple. La salida de todo el sistema se obtiene por medio de la media ponderada de las salidas los n modelos de regresión

lineal múltiple.

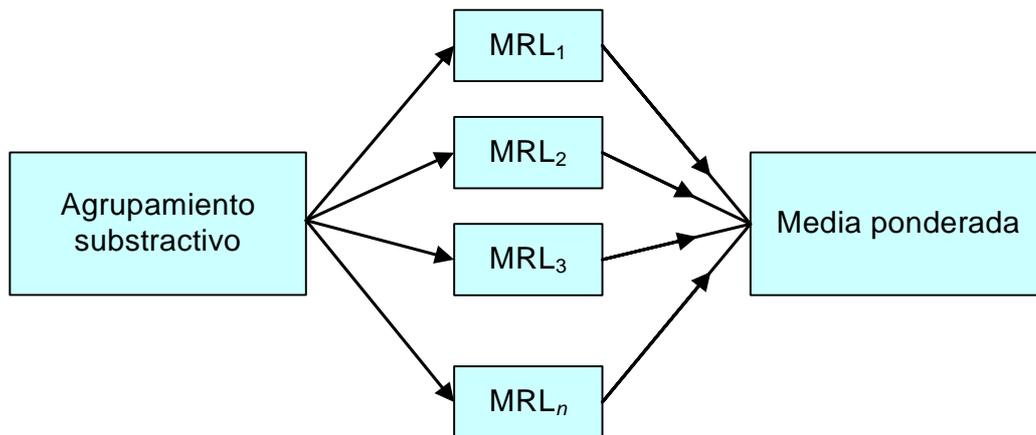


Figura 7.5. Modelo de predicción basado en modelos de regresión lineal.

Como se indicó anteriormente, los coeficientes de ponderación se calculan como la inversa de la distancia de cada valor del vector de entradas al centro de cada cluster, tal como se recoge en la ecuación (7.22). En dicha ecuación, s_j representa la salida para el valor del vector de entradas j , s_{jk} representa la salida del modelo de regresión lineal k (el formado para los datos del grupo k , es decir, pertenecientes al cluster k); d_{jk} es la distancia euclídea del valor del vector de entradas j al centro del cluster k y c es el número total de clusters.

$$s_j = \frac{\sum_{k=1}^c \frac{1}{d_{jk}} s_{jk}}{\sum_{k=1}^c \frac{1}{d_{jk}}} \quad (7.22)$$

La distancia euclídea se puede calcular por medio de la ecuación (7.23) en la que d_{jk} es la distancia euclídea del vector de entradas j al centro del cluster k ; p_{ij} corresponde al valor de la entrada i del valor del vector de entradas j y c_{ik} corresponde al valor de la entrada i para el centro del cluster k .

$$d_{jk} = \left[(p_{1j} - c_{1k})^2 + (p_{2j} - c_{2k})^2 + \dots + (p_{nj} - c_{nk})^2 \right]^{\frac{1}{2}} \quad (7.23)$$

Evidentemente es posible construir múltiples modelos distintos para cada caso con un conjunto de datos concretos para la predicción. La diferencia de un modelo a otro puede estar en el número de entradas y el número de clusters. Para la selección del modelo que proporcione la mejor generalización, se ha optado por construir diferentes modelos de este tipo variando el número de entradas entre 2 y 6 y el número de clusters

también entre 2 y 6. Todos los modelos son construidos considerando el mismo conjunto de datos (datos de entrenamiento). Tomando otro grupo de datos diferente al de entrenamiento, el grupo de datos de validación, se calcula el error de predicción obtenido para este grupo de datos. Se escoge el sistema que proporciona un menor valor del criterio de información de Akaike [Akaike76] en la predicción del conjunto de datos de los grupos de entrenamiento y validación.

Una variante probada, que ha demostrado experimentalmente su eficacia en la reducción de los errores de predicción, ha sido la de modificar el grupo de datos que forma cada cluster antes de calcular el modelo de regresión lineal múltiple para cada cluster. La variante ha sido repetir, en el grupo, cada valor del vector de entradas de forma inversamente proporcional a su distancia al centro del cluster: el más alejado aparece una vez, los más cercanos aparecen varias veces (hasta centenares de veces más). La idea que subyace bajo esta variante es la de dar mayor peso, en el cálculo de los sistemas de regresión lineal múltiple, a aquellos valores del vector de entradas que se encuentran más cercanos al centro de cada cluster y menor peso a aquellos valores del vector de entradas más alejados del centro del cluster. El número de veces que ha de repetirse cada valor del vector de entradas perteneciente al cluster k viene dado por la ecuación (7.24), en la que n_{jk} representa el número de veces que hay que repetir el valor del vector de entradas j ; d_{jk} es la distancia euclídea del valor del vector de entradas del grupo k más alejado del centro del cluster; d_{jk} representa la distancia del valor del vector de entradas j al centro del cluster k ; e (int) representa la parte entera del valor entre corchetes.

$$n_{jk} = (\text{int}) \left[2 \frac{d_{fk}}{d_{jk}} \right] \quad (7.24)$$

Este modelo propuesto es de aplicación para aquellos casos en que se quiera realizar la predicción de la potencia eléctrica media horaria, con un horizonte de predicción limitado a unas horas, y no se disponga de más información que la propia potencia eléctrica generada en las últimas horas.

7.4. NUEVOS MODELOS DE PREDICCIÓN DE LA PRODUCCIÓN ELÉCTRICA BASADOS EN PREDICCIONES DE MODELOS ATMOSFÉRICOS

En la revisión bibliográfica realizada en el Capítulo 3 se puso de manifiesto que los modelos de predicción de la producción eléctrica en parques eólicos con mejores

resultados en el corto plazo, eran aquellos que utilizaban como entradas los valores de predicción para ciertas variables meteorológicas (fundamentalmente la velocidad y dirección media del viento en la superficie del terreno, obtenidas de un modelo atmosférico). En este apartado se presentan dos nuevos modelos de predicción de este tipo, uno específico para su uso cuando la serie temporal de predicciones del modelo atmosférico es limitada en cuanto a su longitud, y un segundo modelo para su uso cuando la serie temporal de predicciones del modelo atmosférico no presenta tal limitación.

7.4.1. Nuevo modelo de predicción con información limitada del modelo atmosférico.

Para aquellos casos en que se disponga de una serie completa de registros del sistema SCADA de un parque eólico, pero no se disponga de una serie de similar duración (en periodo de tiempo) de predicciones de modelos atmosféricos, es posible desarrollar nuevos modelos de predicción que combinen ambos tipos de series. Este es el caso presentado en [Ramirez06], en el que se desarrollan varios modelos de predicción de la potencia eléctrica media horaria generada en un parque eólico del que se disponían de 14 meses con los registros del sistema SCADA, así como los datos de velocidad y dirección media horaria del viento obtenidos del modelo MM5 pero sólo para 68 días.

Este caso es común para parques eólicos en los que se comienzan a desarrollar modelos de predicción: los datos del modelo atmosférico empiezan a estar disponibles a partir de una fecha posterior a la de entrada en servicio del parque, por lo que la serie correspondiente a los registros del equipo SCADA del parque eólico cubre un periodo mucho mayor que el cubierto por el modelo atmosférico.

Los nuevos modelos desarrollados específicamente para este caso son todos neuronales, incluyendo los tipos de redes neuronales siguientes: perceptrones multicapa (MLPs), redes unidireccionales multicapa generalizadas (FFBP), redes neuronales de función de base radial (RBF), redes neuronales con reducción del espacio de entrada con análisis de componentes principales (PCA), redes neuronales híbridas compuestas de un módulo con aprendizaje no supervisado y otro módulo con aprendizaje supervisado, redes neuronales modulares, redes neuronales de Elman y redes neuronales con retardo en el tiempo (TDNN). Los resultados obtenidos con cada una de las redes neuronales en función de los horizontes de predicción (de 1 a 24 horas) se presentan en el Capítulo 8. En los siguientes párrafos se comentan las novedades más interesantes

introducidas en estos modelos.

La disponibilidad completa de los registros SCADA del parque eólico facilita la obtención de una nueva serie temporal, la del factor de capacidad, FC . Dicho factor FC para un instante t se define como el cociente entre la potencia eléctrica generada en ese instante, $P(t)$, y la capacidad de generación disponible también en ese instante, $C_d(t)$, tal y como refleja la ecuación (7.25).

$$FC(t) = \frac{P(t)}{C_d(t)} \quad (7.25)$$

La obtención de la serie temporal del factor FC resulta sencilla: sólo es necesario disponer de los registros de potencia eléctrica media generada en cada aerogenerador en intervalos regulares de tiempo (1 hora o menos). Si uno o varios aerogeneradores presentan un valor de potencia generada nulo, mientras que la potencia media generada por el resto de aerogeneradores es significativa, podemos considerar que los aerogeneradores con potencia nula estaban desconectados por avería o mantenimiento, reduciéndose la capacidad de generación disponible (será exclusivamente la suma de las potencias instaladas de los aerogeneradores que se encuentran funcionando).

Al disponer de esta nueva serie de factor de capacidad, se estimó que la predicción de nuevos valores de esta serie resulta, para los gestores de un parque eólico, mucho más interesante que la de potencia eléctrica media generada. A partir del valor de predicción del factor de capacidad se puede calcular el valor de predicción de la potencia eléctrica media sin más que multiplicar por la potencia instalada de los aerogeneradores que van a estar en funcionamiento (de antemano es posible saber los aerogeneradores que se encontrarán parados por labores de mantenimiento).

Así, la primera novedad que presentan los modelos de predicción basados en redes neuronales desarrollados es la predecir el factor de capacidad del parque eólico, en lugar de la potencia eléctrica media. La segunda novedad radica en la selección de las entradas a los modelos. Recordemos que los datos disponibles de predicción del modelo atmosférico son mucho más limitados, en periodo temporal cubierto (sólo 68 días), que los datos de la propia serie de factor de capacidad del parque. Esto imposibilita el poder utilizar como entradas de los modelos de predicción las variables obtenidas del modelo atmosférico (velocidad y dirección media horaria del viento), ya que no son suficientes para entrenar y validar un modelo de este tipo. Sin embargo, al analizar la relación existente entre las predicciones del modelo atmosférico y los valores registrados en el sistema SCADA, se observó cierta discrepancia entre los valores de una serie y otra,

pero esa diferencia de valores era significativamente menor si en lugar de comparar los valores de las dos series para un mismo instante de tiempo, comparáramos la variación relativa de ambas con respecto a un instante que tomásemos de referencia. En la Tabla 7.4 se recogen unos valores reales para uno de los 68 días para los que se dispone de datos, que pueden aclarar esa menor diferencia de valores entre las dos series.

Así, mientras que para los datos de la Tabla 7.4 se observa una diferencia relativa importante entre los valores de velocidad media horaria del viento real (medida en el parque eólico) y prevista (obtenida con el modelo atmosférico MM5), esa diferencia es significativamente menor para el cociente entre dos valores de la propia serie. El dato “cociente” de la Tabla 7.4 representa el cociente de los valores de velocidad media horaria a las 17 horas y a las 0 horas, tanto reales (1.105) como previstos (1.109).

| | Velocidad real (m/s) | Velocidad prevista (m/s) |
|--------------------|----------------------|--------------------------|
| Valor para las 17h | 8.4 | 6.1 |
| Valor para las 0h | 7.6 | 5.5 |
| Cociente | 1.105 | 1.109 |

Tabla 7.4. Comparación entre valores reales y previstos.

La conclusión que se extrae del hecho anterior, del que los datos de la Tabla 7.4 son un pequeño ejemplo ilustrativo, es que las predicciones del modelo atmosférico MM5 no se ajustan suficientemente a los valores reales futuros, pero si reflejan bien las variaciones temporales de un instante a otro. Por ello se consideró, como una posible entrada de los modelos de predicción neuronales, el cociente entre la velocidad para el instante correspondiente al horizonte de predicción y la velocidad para el instante actual. Esto permite entrenar las redes con los valores reales, sin que esto afecte significativamente a la calidad de las predicciones realizadas cuando se use como cociente el de los valores de predicción del modelo atmosférico.

7.4.2. Nuevo modelo de predicción con información completa del modelo atmosférico.

Cuando se dispone de un conjunto completo de datos que incluye tanto los registros SCADA de un parque eólico, como las predicciones a corto plazo de un modelo atmosférico, es posible desarrollar unos modelos más completos que los presentados en las páginas anteriores.

La disponibilidad de predicciones de velocidad y dirección media del viento reduce significativamente el error de predicción, en especial para horizontes de

predicción por encima de unas 6 horas, tal y como queda reflejado en la literatura internacional que comentamos en el Capítulo 3. Pero las predicciones realizadas usando como variables de entrada las predicciones de los modelos atmosféricos, con horizontes de predicción de 3 ó 4 horas, son incluso inferiores a las obtenidas con el modelo persistente.

Para mejorar el comportamiento, con horizontes de predicción pequeños (por debajo de las tres horas) de los sistemas de predicción de la producción eléctrica en un parque eólico, se ha incluido un nuevo filtro Kalman [Kalman60] cuya misión es la de mejorar el valor de la predicción de la velocidad media del viento, $\hat{v}(t+k)$, proporcionada por el modelo atmosférico para el horizonte k de predicción, siendo t el instante en que se realiza dicha predicción.

Así, se ideó inicialmente un modelo de predicción que respondía al esquema de la Figura 7.6. Sólo el valor de predicción de la velocidad media del viento del modelo atmosférico, y únicamente para las primeras horas, es procesado por el novedoso filtro Kalman (sólo para mejora en las predicciones a muy corto plazo). El resto de previsiones del modelo atmosférico (dirección media del viento, temperatura y presión atmosférica), sin “filtrar”, forman parte de las entradas de una red neuronal o sistema de inferencia difusa. Estas entradas se completan con los últimos valores conocidos de la potencia eléctrica producida en el parque eólico: concretamente, en la Figura 7.6, $p(t)$ representa el último valor conocido de la potencia eléctrica media entregada en la subestación de conexión a la red eléctrica; y $p(t-m)$ representa la potencia eléctrica media entregada en la subestación m periodos (de media hora) antes del instante actual. Además en dicha Figura 7.6 $\hat{p}(t+k)$ representa el valor de predicción, para la potencia eléctrica media entregada en la subestación, para el instante correspondiente al horizonte de predicción.

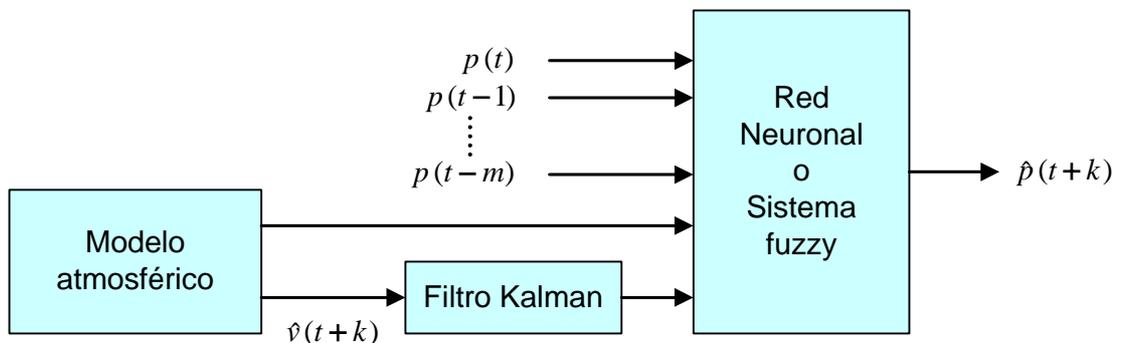


Figura 7.6. Modelo de predicción con información completa de modelo atmosférico.

A partir del modelo inicialmente propuesto, mostrado en la Figura 7.6, se ha

desarrollado un sistema global de predicción de la producción eléctrica en parques eólicos recogido en la Figura 7.7. Básicamente este sistema global consiste en un conjunto de modelos de predicción, relativos a cada rango de valores del horizonte de predicción considerado (hay modelos de predicción específicos para el muy corto plazo, modelos especializados en realizar las predicciones correspondientes al día siguiente, etc.), seleccionándose, en función de la hora en la que se hacen las predicciones (junto con el horizonte de predicción), los modelos idóneos de forma que el error de predicción sea el mínimo posible.

En la Figura 7.7, se representan r modelos específicos en realizar predicciones en el muy corto plazo. De ellos se seleccionan tres modelos de forma que el valor de predicción es la combinación, por medio de un sistema de inferencia difusa, de los valores de predicción aportados por los tres modelos seleccionados. Existen hasta s sistemas de inferencia difusa, cada uno específico para un horizonte de predicción.

Además en la Figura 7.7 aparecen otros tres modelos A, B, y C especializados para las predicciones en el corto plazo (concretamente en el periodo de 48-72 horas, en el periodo de 24-48 horas y otro periodo que llega desde el muy corto plazo hasta las 24 horas).

El valor de predicción de todo el sistema global se obtiene bien por selección de entre los valores de predicción obtenidos en los sistemas de inferencia difusa (para el muy corto plazo), o bien por selección de los valores de los modelos especializados para el corto plazo. La selección se realiza en función de la hora en la que se realiza la predicción y del horizonte de predicción.

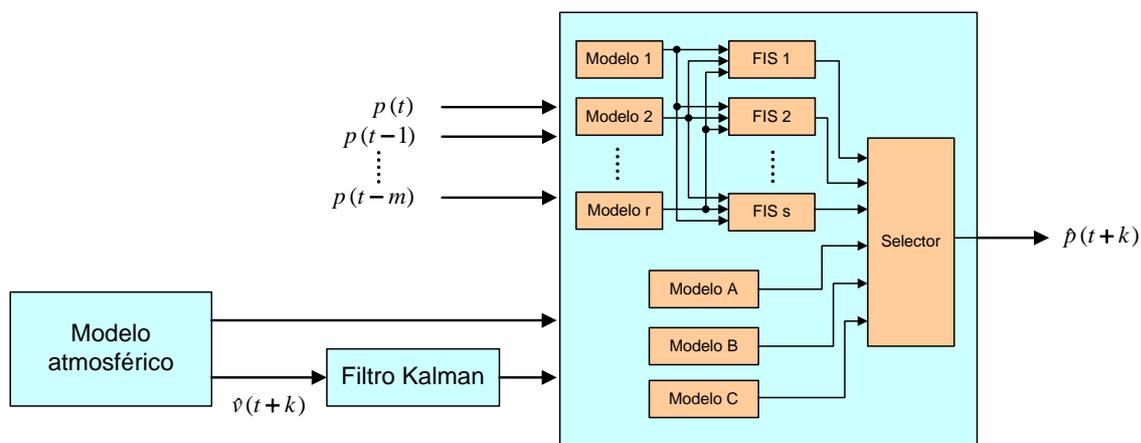


Figura 7.7. Estructura del sistema global de predicción propuesto.

En el próximo capítulo se recogen comentarios adicionales sobre el diseño de

este sistema de predicción global y de los resultados logrados con dicho sistema, en la predicción de la producción eléctrica en parques eólicos con un horizonte de predicción de hasta 72 horas. También se ilustra un ejemplo de los beneficios económicos que puede proporcionar, al gestor de un parque eólico, el uso del sistema global de predicción propuesto.

7.5. CONCLUSIONES

A lo largo de este capítulo se han presentado los diferentes modelos creados como consecuencia de los trabajos de investigación desarrollados. Los tres primeros modelos corresponden a la familia de series temporales, tal y como clasificábamos los modelos de predicción en el Capítulo 3. Los dos últimos modelos corresponden, parcialmente (cuarto modelo) o totalmente (quinto modelo), a la familia de modelos que utilizan predicciones de valores de variables meteorológicas obtenidas mediante modelos atmosféricos. Todos los modelos han sido diseñados idóneamente para el tipo de datos existentes para su creación. Los datos que propiciaron el desarrollo de cada uno de los modelos son los indicados en los párrafos siguientes.

- *Modelo de predicción basado en descomposición de Fourier.*

Únicamente se disponía del conjunto de información designado como “Serie I”, correspondiente a una serie temporal de velocidad y dirección media horaria del viento en una estación meteorológica. Aún con tan escasa información, se pudo desarrollar un modelo original con horizonte de predicción de una hora. Este modelo podría ser de utilidad en la predicción de la potencia eléctrica media horaria generada en un parque eólico cuando la información histórica disponible sólo consiste en la serie temporal del viento (fundamentalmente velocidad media).

- *Modelo de predicción basado en series temporales fuzzy.*

Se disponía del conjunto de información designado como “Serie II”. Con la serie temporal disponible de potencia eléctrica media horaria en un parque eólico, y ubicación concreta desconocida (sólo se conocía la provincia en la que se encuentra), y con la información lingüística de previsión del tiempo (obtenida, por ejemplo, de diarios de prensa u otro medio de difusión), se desarrolló un modelo de predicción de la energía total generada en el parque eólico para el día

siguiente.

- *Modelo de predicción basado en regresión lineal.*

Es un modelo desarrollado también a partir del conjunto de información designado como “Serie II”. En este caso el modelo proporciona la potencia eléctrica media horaria generada en un parque eólico con un horizonte de predicción de hasta 6 horas utilizando como variables de entrada los últimos valores conocidos de la misma.

- *Modelo de predicción con información limitada de modelo atmosférico.*

Se disponía de la información designada como “Serie III”. Los datos de predicciones del modelo atmosférico (concretamente predicciones para 68 días de la velocidad y dirección media horaria del viento, obtenidas con el modelo atmosférico MM5) eran insuficientes para poder desarrollar un modelo que utilizase esas predicciones como entradas. Aprovechando la disponibilidad de una serie temporal de mayor longitud con los registros del sistema SCADA del parque eólico, se optimizó un modelo de predicción basado en redes neuronales entrenado, exclusivamente, con valores históricos del propio parque.

- *Sistema avanzado global de predicción con información completa de un modelo atmosférico.*

La disponibilidad de un conjunto completo de información, “Serie V”, propició el desarrollo de un sofisticado sistema global de predicción de la potencia eléctrica media formado por un amplio conjunto de modelos de predicción, con una resolución de hasta media hora y un horizonte de predicción de hasta 72 horas. El sistema utiliza además un filtro Kalman para la mejora de la predicción de la velocidad media del viento, proporcionada por el modelo atmosférico, para horizontes de predicción pequeños.

8. RESULTADOS COMPUTACIONALES

En este capítulo se presentan los resultados logrados en la predicción de la producción eléctrica en el corto plazo para parques eólicos, utilizando los nuevos modelos creados y que han sido descritos en el capítulo anterior. También se presentan los resultados obtenidos, utilizando los mismos datos, con otros modelos de predicción usados en trabajos publicados en la literatura internacional, para poder comparar la bondad de las predicciones logradas con los nuevos modelos desarrollados.

También en este capítulo se indica el beneficio bruto que se puede obtener, por parte del gestor de un parque eólico, al usar las predicciones de potencia eléctrica media suministradas por el nuevo sistema de predicción, tanto para la venta de la energía eléctrica al mercado eléctrico, como por la venta de la misma a tarifa regulada.

En todos los casos presentados, los modelos originales desarrollados reúnen unas características significativamente diferentes a las de los modelos publicados en la literatura internacional y logran un menor error de predicción.

8.1. PREDICCIÓN DE LA VELOCIDAD MEDIA DEL VIENTO A MUY CORTO PLAZO

En el Capítulo 7 se presentó un modelo de predicción basado en la descomposición de Fourier de la serie temporal de velocidad media del viento, para casos en los que los datos disponibles corresponden a series históricas de velocidad y dirección media del viento. Este modelo está diseñado para la obtención de la velocidad media del viento con un horizonte de predicción de 1 hora, cuyo interés se centra en labores de mantenimiento a muy corto plazo de la instalación eólica.

8.1.1. Datos disponibles.

En el desarrollo de este modelo los únicos datos necesarios son la serie temporal de velocidad media horaria del viento, medida en el emplazamiento para el cual se va a realizar la predicción. Para los resultados presentados en este documento, esa serie temporal corresponde a los datos denominados como Serie I en el Capítulo 7, es decir,

la serie temporal de tres años con la velocidad y dirección media horaria del viento. La Figura 8.1 muestra la velocidad media horaria (en el eje vertical) a lo largo de las dos primeras semanas (en el eje horizontal) de julio de 1990 (primer año de los disponibles en la serie temporal), donde se puede apreciar el carácter aparentemente caótico de la serie.

Los datos disponibles fueron inicialmente divididos en tres grupos para la generación de los modelos de predicción. Los datos de la serie temporal de velocidad media horaria del viento para el primer año (1990) conformaron el grupo de datos de entrenamiento de los modelos. Los datos del segundo año (1991) conformaron el grupo de los datos de validación y los del tercer año (1992) el grupo de datos de testeo. No se observaron diferencias significativas en los valores medios y varianzas de los datos de cada uno de los tres grupos, tal y como refleja la Tabla 8.1.

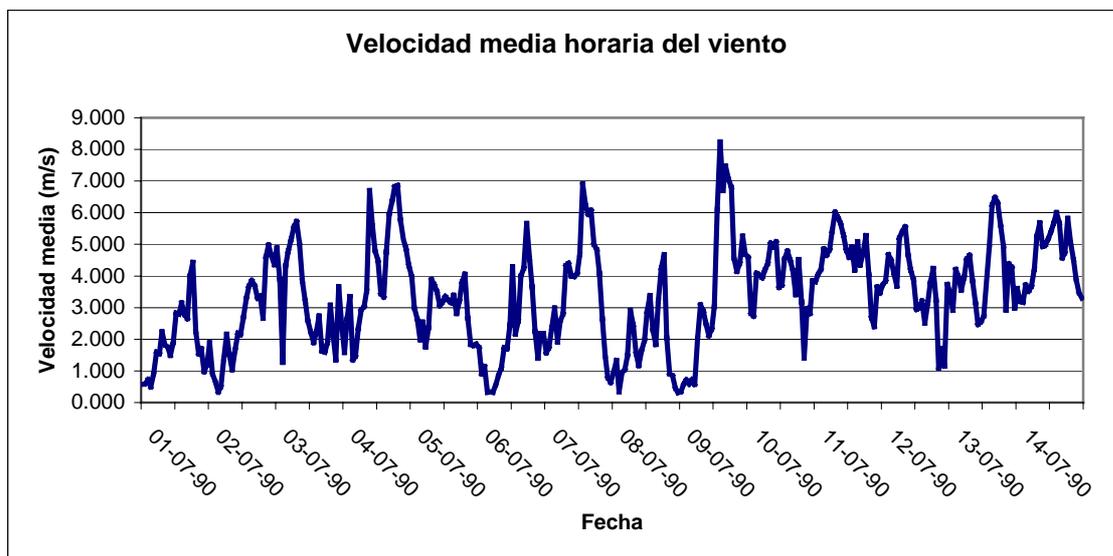


Figura 8.1. Velocidad media horaria del viento para las dos primeras semanas de julio de 1990.

| Grupo | Valor medio (m/s) | Desviación estándar (m/s) |
|---------------|-------------------|---------------------------|
| Entrenamiento | 2.803 | 1.640 |
| Validación | 2.864 | 1.642 |
| Testeo | 2.678 | 1.596 |

Tabla 8.1. Valores medios y desviaciones estándar relativos a los datos de velocidad del viento.

El objetivo inicial es dar un valor de predicción para la velocidad media del viento con un horizonte de predicción de 1 hora (muy corto plazo). Posteriormente, con este valor se puede realizar la predicción de la potencia eléctrica media horaria del parque eólico.

8.1.2. Modelos de predicción analizados.

Para la obtención del valor de predicción (velocidad media horaria del viento) se consideraron distintos tipos de modelos hasta un total de siete. Estos modelos de predicción fueron los indicados seguidamente.

1. *Modelo persistente.*

Es el más simple de los modelos de predicción ya que el valor de predicción coincide con el valor del último dato conocido de la variable a predecir. Es decir, el valor de predicción de la velocidad media horaria del viento para $t+1$, $\hat{v}(t+1)$, es igual al valor de la velocidad media horaria del viento en el instante (hora) t , $v(t)$, tal como expresa la ecuación (8.1). Habitualmente utilizaremos t como indicador de la hora a la que se refiere la variable analizada.

$$\hat{v}(t+1) = v(t) \quad (8.1)$$

El modelo persistente está considerado en la literatura científico-técnica internacional [Alexiadis99], [Sfetsos00], [Damousis04], [Torres05], como el modelo a batir en la predicción a muy corto plazo.

2. *Modelo del valor medio.*

Es el modelo equivalente al persistente pero para largo plazo. En el muy corto plazo, este modelo podría ser utilizado aun desconociendo el último valor de la serie temporal (en este caso $v(t)$). El valor de predicción con este modelo corresponde a la media de los valores conocidos de la serie temporal comprendidos en una “ventana” temporal. Esta ventana temporal puede ser fija o móvil. En el caso de una ventana temporal móvil, el valor medio viene dado por la ecuación (8.2), en la que N es la longitud de la ventana temporal utilizada para el cálculo del valor medio, y $v(t)$ la velocidad media horaria en el instante t .

$$\hat{v}(t+1) = \frac{1}{N} \sum_{j=1}^N v(t-j) \quad (8.2)$$

3. Modelo “nueva referencia”.

Este modelo ha sido propuesto en [Nielsen00] como alternativo al modelo persistente en las predicciones a muy corto y corto plazo, sobre todo a efectos de comparación de la bondad de las predicciones con nuevos modelos (de ahí el denominarlo como modelo de referencia). Este modelo, que también puede ser utilizado para realizar predicciones para horizontes de predicción cortos y muy cortos, proporciona un valor próximo al obtenido por el modelo persistente, mientras que para horizontes mayores (varias horas) su valor de predicción se va aproximando progresivamente al valor medio de la serie temporal.

El valor de predicción con este modelo de “nueva referencia” puede calcularse por medio de la ecuación (8.3), en la que μ representa el valor medio de la serie temporal (velocidad media horaria) y a_1 es un coeficiente de ponderación para el último término conocido de la serie temporal, $v(t)$.

$$\hat{v}(t+1) = a_1 v(t) + (1 - a_1) \mu \quad (8.3)$$

El valor del coeficiente a_1 corresponde al coeficiente de correlación entre $v(t)$ y $v(t+1)$, que puede calcularse por medio de la ecuación (8.4), en la que $\tilde{v}(t)$ representa la diferencia entre la velocidad media horaria en el instante t y el valor medio de dicha variable (ecuación (8.5)).

$$a_1 = \frac{\frac{1}{N} \sum_{t=1}^{N-1} \tilde{v}(t) \tilde{v}(t+1)}{\frac{1}{N} \sum_{t=1}^{N-1} \tilde{v}^2(t)} \quad (8.4)$$

$$\tilde{v}(t) = v(t) - \mu \quad (8.5)$$

4. Modelo de Holt-Winters.

Este modelo de predicción, también conocido como el modelo suavizado exponencial de tres parámetros de Winters [DeLurgio98], es uno de los modelos clásicos más utilizados en el campo de la predicción económica. Puede ser fácilmente automatizado, requiere un mínimo esfuerzo computacional y ofrece buenos resultados de predicción [Makridakis00]. El

modelo de Holt-Winters logra una mejora sobre el modelo de Holt de dos parámetros mediante el ajuste de la estacionalidad por medio de un tercer parámetro. El modelo de Holt-Winters puede ser aditivo o multiplicativo, en función del comportamiento como factores o como sumandos de dos de los tres términos que componen el modelo. Estos tres términos son el valor suavizado de la serie temporal, el valor suavizado de la tendencia y el valor suavizado de la estacionalidad, expresados en las ecuaciones (8.6), (8.7) y (8.8), respectivamente.

$$S(t) = \alpha \frac{v(t)}{I(t-L)} + (1-\alpha)[S(t-1) + b(t-1)] \quad (8.6)$$

$$b(t) = \beta[S(t) - S(t-1)] + (1-\beta)b(t-1) \quad (8.7)$$

$$I(t) = \gamma \frac{v(t)}{S(t)} + (1-\gamma)I(t-L) \quad (8.8)$$

En las ecuaciones (8.6), (8.7) y (8.8), $S(t)$ representa el valor suavizado para la variable a predecir (velocidad media horaria) para el instante t ; α es la constante de suavizado para la velocidad media horaria; $v(t)$ representa la velocidad media horaria para el instante t ; $I(t)$ representa el índice suavizado de estacionalidad para el instante t ; L representa la longitud del ciclo estacional (en este caso 24, las horas de un día); $b(t)$ representa el valor suavizado de la tendencia de la serie para el instante t ; β es la constante de suavizado para la tendencia; y γ es la constante de suavizado para el índice de estacionalidad.

Así, son tres los parámetros a determinar para definir el modelo, las constantes α , β y γ , de ahí que se le denomine al modelo de Holt-Winters como modelo suavizado de tres parámetros. Si consideramos un comportamiento aditivo para el término de tendencia, el modelo puede ser aditivo (según recoge la ecuación (8.9)) o multiplicativo (según recoge la ecuación (8.10)), en función del comportamiento del término de índice suavizado de estacionalidad.

$$\hat{v}(t+1) = S(t) + b(t) + I(t-L+1) \quad (8.9)$$

$$\hat{v}(t+1) = [S(t) + b(t)]I(t-L+1) \quad (8.10)$$

En este caso se han considerado ambos modelos, el aditivo y multiplicativo para realizar la predicción de la velocidad media horaria.

5. *Modelo ARIMA.*

Bajo la metodología de Box-Jenkins [Box76], también conocida como ARIMA, se han desarrollado una serie de modelos de predicción de la velocidad media horaria del viento. Este tipo de modelos ha sido utilizado en diversos trabajos publicados como modelo de predicción de la velocidad media horaria en el muy corto plazo [Sfetsos02], [Poggi03], [Torres05]. Los resultados obtenidos con aquellos modelos ARIMA que ofrecían un menor valor de error se recogen en el apartado 8.1.3 de este documento.

6. *Modelo basado en una red neuronal artificial.*

Los modelos de predicción basados en redes neuronales artificiales también han sido utilizados desde hace años en la predicción de la velocidad media horaria del viento en el muy corto plazo [Alexiadis99], [Maqsood05], [Flores05]. En este caso se ha optado por desarrollar cinco tipos de modelos de redes neuronales: perceptrones multicapa (MLPs), red neuronal multicapa generalizada (FFBP), red neuronal de función de base radial (RBF), red neuronal con retardo en el tiempo no focalizada (TDNN) y red neuronal recurrente.

Para cada uno de los tres tipos de redes se ha utilizado un algoritmo genético en la optimización de los parámetros de la red neuronal. Concretamente, mediante el proceso de optimización (con algoritmo genético) se seleccionaron las entradas que se iban a utilizar en la red neuronal de un total de 24 posibles (los 24 últimos valores de velocidad media horaria del viento). En el proceso de optimización también se seleccionaron el número de neuronas en capas ocultas, y los valores de factor de aprendizaje y momento de cada capa para el algoritmo de retropropagación de error. En el caso de la red neuronal con retardo en el tiempo también se optimizó el número de retardos en las entradas (número de valores previos de cada entrada que se toman como entrada adicional) y en la capa oculta. El número

de epochs se limitó en todos los casos a 2000, el de individuos por generación a 40 y el número de generaciones a 50.

7. *Modelo basado en un sistema de inferencia difusa.*

Los sistemas de inferencia difusa, con y sin adaptabilidad como modelo neuro-difuso [Jang93], han sido también seleccionados como modelos a considerar. En este caso se diseñaron un total de 45 modelos del tipo TSK de primer orden, considerando diferentes combinaciones de entradas y reglas. Las entradas seleccionadas son valores pasados de la velocidad media horaria, concretamente $v(t)$, $v(t-1)$, $v(t-2)$, $v(t-3)$, $v(t-23)$, $v(t-24)$ y $v(t-25)$; de esta forma, se diseñaron los sistemas de inferencia difusa con las entradas indicadas en la Tabla 8.2.

| Sistemas | Entradas |
|----------|--|
| 1 | $v(t)$, $v(t-1)$, $v(t-2)$ y $v(t-3)$ |
| 2 | $v(t)$, $v(t-1)$, $v(t-2)$ y $v(t-23)$ |
| 3 | $v(t)$, $v(t-1)$, $v(t-22)$ y $v(t-23)$ |
| 4 | $v(t)$, $v(t-1)$, $v(t-23)$ y $v(t-24)$ |
| 5 | $v(t)$, $v(t-22)$, $v(t-23)$ y $v(t-24)$ |
| 6 | $v(t)$, $v(t-1)$ y $v(t-2)$ |
| 7 | $v(t)$, $v(t-1)$ y $v(t-23)$ |
| 8 | $v(t)$ y $v(t-1)$ |
| 9 | $v(t)$ y $v(t-23)$ |

Tabla 8.2. Modelos de inferencia difusa diseñados y entradas consideradas en cada caso.

Para cada uno de los nueve sistemas, se dividió el espacio de entradas en forma de rejilla (“grid partitioning”) considerando 2 y 3 funciones de pertenencia por entrada, o se dividió el espacio de entrada por medio del agrupamiento substractivo (“subtractive clustering”) formando 2, 3 ó 4 clusters. Es decir, para cada una de los 9 sistemas seleccionados, se probaron 5 opciones diferentes, lo que hace que el total de sistemas de inferencia difusa probados sea de 45 como ya se indicó anteriormente.

8. *Modelo de predicción basado en descomposición de Fourier.*

Las características básicas de este modelo fueron presentadas en el apartado 7.3.1 de este documento. Para el módulo de predicción, que conforma una de las tres partes del modelo, se ensayaron computacionalmente las siguientes

opciones: una única red neuronal de tipo MLP con 23 entradas (los 23 coeficientes independientes para el instante t) y 23 salidas (los 23 coeficientes independientes para el instante $t+1$); 23 redes neuronales del tipo MLP, cada una con una única salida (una neuronal red por coeficiente, siendo la salida el valor del coeficiente correspondiente para $t+1$); 23 sistemas de inferencia difusa con valor de salida cada uno de los 23 coeficientes independientes para $t+1$.

8.1.3. Desarrollo de los modelos de predicción.

El modelo persistente no ofreció ninguna dificultad en su desarrollo. El valor de predicción de la velocidad media horaria para el instante $t+1$ no es más que el valor de dicha variables para el instante t . Para la evaluación del modelo se utilizó exclusivamente el grupo de datos de testeo (datos del último año).

En el modelo del valor medio, dicho valor medio se considera como valor de predicción para el instante $t+1$, y se obtuvo calculando el valor medio de la velocidad media horaria del viento para los datos de los grupos de entrenamiento y validación. Para la evaluación del modelo se consideró que el valor de predicción, para los datos del grupo de testeo, era el valor medio calculado anteriormente.

En el modelo “nueva referencia” se calcularon los valores del coeficiente a_1 y del valor medio μ (es el valor calculado para el modelo del valor medio) con los datos de los grupos de entrenamiento y validación. Ambos valores se utilizaron para la predicción de la velocidad media horaria, con los datos del grupo de testeo, utilizando la ecuación (8.3).

Para el desarrollo de las dos variantes del modelo de Holt-Winters, la multiplicativa y la aditiva, se utilizó el programa informático SAS/ETS de SAS Institute Inc., bajo el entorno del paquete estadístico SAS diseñado para la elaboración de modelos de predicción econométricos. Se utilizaron los datos de los grupos de entrenamiento y validación para el ajuste de los parámetros del modelo. Con los valores de los parámetros α , β y γ ya determinados, se evaluó el modelo para los datos del grupo de testeo.

Con los datos de los grupos de entrenamiento y validación, y siguiendo la metodología Box-Jenkins, se desarrollaron varios modelos ARIMA considerando y no

considerando la estacionalidad. La selección de un único modelo ARIMA se realizó aplicando el criterio de información de Akaike.

En el desarrollo de los modelos de redes neuronales artificiales todos los modelos fueron entrenados con los datos de entrenamiento. Para evitar el sobreentrenamiento se utilizó el grupo de validación: tras cada epoch se presentaba a la red neuronal los valores del vector de entradas del grupo de validación; el entrenamiento se interrumpía si el error de predicción con los valores del vector de entradas del grupo de validación no se reducía tras 100 epochs consecutivos; en todo caso, al final del proceso de entrenamiento, la red neuronal correspondía a aquella que ofrecía el menor error RMSE con el grupo de valores del vector de entradas de validación (los 100 epochs adicionales que podían empeorar el error con los datos del grupo de validación no se consideran a efectos de modificación de los parámetros de la red neuronal).

Como se ha comentado anteriormente, todos los modelos de redes neuronales fueron optimizados por medio de un algoritmo genético permitiendo la selección de entradas (de entre los últimos 24 valores de la velocidad media horaria del viento) y parámetros de la propia red. Una vez optimizada la red neuronal se determinó el error con los datos del grupo de testeo (datos del último año).

Los modelos de inferencia difusa (FIS) fueron desarrollados con la herramienta informática “fuzzy toolbox” de matlab y siguiendo una metodología similar a la de las redes neuronales. El diseño de los modelos se realizó utilizando los datos de entrenamiento. La herramienta utilizada, a partir de las especificaciones de partida, realizó el cálculo de los parámetros del modelo según se indica a continuación.

- Para modelos con partición del espacio de entrada en rejilla todas las funciones de pertenencia fueron gaussianas. La herramienta calculó el centro y dispersión de cada una de las funciones de pertenencia de las entradas para cubrir uniformemente todo el rango de valores de las variables de entrada. Se consideró el mismo número de funciones de pertenencia por entrada para todas las entradas. El número de reglas fue igual al número de funciones de pertenencia por entrada elevado al número de entradas.

El número de funciones de pertenencia consideradas varió entre 2 y 3, y el número de entradas entre 2 y 4, por lo tanto, el número mínimo de reglas fue 4 y el máximo fue de 81.

- Para modelos con partición del espacio de entrada por agrupamiento substractivo, también todas las funciones de pertenencia fueron gaussianas. La herramienta calculó el centro y dispersión de cada una de las funciones de pertenencia de las entradas y la parte consecuente de cada una de las reglas. El número de reglas coincidió con el número de clusters obtenido.

Para el diseño de los distintos modelos, bajo esta partición del espacio de entrada por agrupamiento substractivo, únicamente se varió el parámetro del rango de influencia del mismo. Se buscaron los valores de rango de influencia que marcaban el límite entre obtener 2, 3 ó 4 clusters. Esto se hizo así al comprobar experimentalmente que se obtenían sistemas con menor error: un sistema con tres clusters obtenido justo con el valor de rango de influencia muy próximo a aquél que hace que el número de clusters sea 2, presentaba un error menor que cualquier otro sistema con tres clusters.

Una vez construido el modelo, se realizó un ajuste de los parámetros por medio de la técnica adaptativa neuro-difusa descrita en [Jang93]. El principio de esta técnica es ajustar los parámetros del sistema de inferencia difusa de la misma forma que se hace en el entrenamiento de una red neuronal. Una vez más se utilizaron los datos del grupo de validación para fijar los parámetros del modelo, esto es, el sistema fija los parámetros que producen el menor error con los datos del grupo de validación.

Por último se desarrolló el modelo de predicción basado en descomposición de Fourier. Según lo descrito en el apartado 7.3.1. de este documento el sistema consta de tres módulos: el primero es el módulo que aplica la transformada discreta de Fourier, DFT, a los 24 últimos datos conocidos de la velocidad media horaria del viento para obtener los 23 coeficientes independientes; el segundo es un módulo que realiza la predicción de los 23 coeficientes independientes para la hora siguiente; el tercer módulo aplica la transformada inversa discreta de Fourier sobre los coeficientes para obtener el valor de la velocidad media del viento para la siguiente hora.

Para el módulo que realiza la predicción se desarrollaron tres soluciones distintas, según se describe seguidamente.

1. Una única red neuronal del tipo perceptrón multicapa, MLP, con 23 entradas, 23 salidas y dos capas ocultas. Las 23 entradas correspondían a los 23 coeficientes independientes del resultado obtenido tras la aplicación de la DFT a la serie temporal de velocidad media horaria del viento (24 últimos valores). Para la

selección de entradas, selección del número de neuronas en las dos capas ocultas de esta red y en la selección de los parámetros del algoritmo de retropropagación de error, se utilizó una optimización con algoritmo genético con 40 individuos y 50 generaciones.

2. Una red neuronal del tipo perceptrón multicapa por coeficiente a predecir. Cada una de las redes podía utilizar como entradas los 23 coeficientes independientes y proporcionaba como salida el valor de predicción de uno de ellos. Cada una de las 23 redes fue optimizada de la misma manera que la red neuronal de 23 entradas y 23 salidas.
3. Un sistema de inferencia difusa por coeficiente a predecir. Cada uno de los sistemas utilizó como entradas los 23 coeficientes independientes y utilizó partición del espacio de entrada por medio del agrupamiento substractivo con 2 ó 3 clusters. Estos sistemas de inferencia difusa fueron desarrollados de igual manera que los sistemas inferencia difusa con agrupamiento substractivo expuestos en los párrafos anteriores de este apartado.

8.1.4. Resultados obtenidos.

El análisis del error RMSE en la predicción de la velocidad media horaria del viento obtenido con el modelo persistente, con los datos del grupo de testeo y con horizonte de predicción de una hora, fue de 0.753 m/s.

El valor medio de la serie temporal de velocidad del viento correspondiente a los datos de los grupos de entrenamiento y validación fue de 2.8334 m/s. La aplicación de este valor de predicción (modelo del valor medio) sobre los datos del grupo de testeo arrojó un error RMSE del 1.774 m/s, considerablemente mayor que el error con el modelo persistente.

Los dos parámetros del modelo “nueva referencia” fueron determinados con los datos de los grupos de entrenamiento y validación. El valor del coeficiente de correlación fue 0.86299 y el valor medio 2.8334 m/s. El error RMSE en la predicción de la velocidad media horaria del viento para el grupo de testeo con un horizonte de predicción de una hora fue 0.737 m/s, ligeramente más bajo que el error del modelo persistente.

Los tres parámetros del modelo Holt-Winters en sus variantes de estacionalidad aditiva y estacionalidad multiplicativa se determinaron con los datos de los grupos de entrenamiento y validación. La aplicación de ambos modelos en la predicción de la velocidad media horaria del viento con los datos del grupo de testeo, arrojó unos errores RMSE del 0.744 m/s para el modelo Holt-Winters de estacionalidad aditiva, y del 0.749 m/s para el modelo de estacionalidad multiplicativa.

Los modelos desarrollados bajo la metodología ARIMA son los recogidos en la Tabla 8.3. Para la selección de un único modelo representativo de este grupo de modelos se utilizó el criterio de información de Akaike, es decir, se seleccionó aquel modelo que presentaba un menor valor para este criterio tras ajustarlo con los datos del grupo de entrenamiento y validación. Este criterio asegura la selección del modelo más “parsimonioso”. La Tabla 8.3 también recoge los valores del criterio de información de Akaike, AIC, para todos los modelos desarrollados dentro de la familia ARIMA.

| Modelo | AIC |
|----------------------|---------|
| AR(1) | -1651.7 |
| AR(2) | -1741.3 |
| AR(3) | -1748.1 |
| AR(4) | -1746.9 |
| ARMA(1,1) | -1749.0 |
| ARMA(2,1) | -1749.3 |
| ARMA(2,2) | -1747.5 |
| ARIMA(1,1,1) | -1387.3 |
| ARIMA(2,1,2) | -1745.7 |
| ARIMA(1,0,0)(1,0,0)s | -1770.8 |
| ARIMA(2,0,0)(1,0,0)s | -1831.7 |
| ARIMA(3,0,0)(1,0,0)s | -1839.7 |
| ARIMA(4,0,0)(1,0,0)s | -1838.1 |
| ARIMA(1,0,1)(1,0,0)s | -1838.2 |
| ARIMA(2,0,1)(1,0,0)s | -1840.5 |
| ARIMA(2,0,2)(1,0,0)s | -1838.5 |
| ARIMA(1,1,1)(1,0,0)s | -1465.6 |
| ARIMA(2,1,2)(1,0,0)s | -1834.8 |

Tabla 8.3. Modelos ARIMA diseñados y valor del criterio de información de Akaike para cada uno de ellos con los datos de entrenamiento y validación.

De acuerdo al valor del criterio de información de Akaike el modelo seleccionado de la familia ARIMA es el ARIMA(2,0,1)(1,0,0)s (también era el modelo que ofrecía un menor error RMSE de entre todos los diseñados, tanto para los datos de

los grupos de entrenamiento y validación como para los del grupo de testeo). El error RMSE en la predicción de la velocidad media horaria del viento para los datos del grupo de testeo fue de 0.70032 m/s.

La red neuronal que presentaba un menor error RMSE con los datos del grupo de testeo es el perceptrón multicapa (MLP), aunque con muy poca diferencia con respecto a la red neuronal multicapa generalizada (FFBP). Las características y valor del error RMSE con el grupo de testeo de las diferentes redes neuronales optimizadas se refleja en la Tabla 8.4.

| Red neuronal | RMSE | Entradas | Características |
|--------------|--------|---|---|
| MLP | 0.6990 | $v(t)$, $v(t-2)$, $v(t-4)$ y $v(t-22)$ | 15 neuronas en capa oculta |
| FFBP | 0.6997 | $v(t)$, $v(t-1)$ y $v(t-22)$ | 16 neuronas en capa oculta |
| RBF | 0.7161 | $v(t)$, $v(t-21)$, $v(t-22)$ y $v(t-23)$ | 77 neuronas en capa oculta 23 centros de gaussianas |
| TDNN | 0.7015 | $v(t)$ y $v(t-21)$ | 13 unidades de retardo (taps) 27 neuronas en capa oculta |
| Recurrente | 0.7268 | $v(t)$, $v(t-1)$, $v(t-2)$, $v(t-3)$, $v(t-4)$, $v(t-21)$ y $v(t-23)$ | 38 neuronas en capa oculta 2 unidades de retardo en entrada y 1 en capa oculta |

Tabla 8.4. Características de las redes neuronales y error RMSE en la predicción de la velocidad media horaria del viento con horizonte de predicción 1 hora.

Los resultados obtenidos con los diferentes modelos de inferencia difusa desarrollados están recogidos en la Tabla 8.5. En esta tabla se encuentran los resultados, para cada una de las nueve combinaciones de entradas estudiadas, de los dos modelos con partición del espacio de entrada en rejilla con 2 y 3 funciones de pertenencia (mfs) por entrada, y de los tres modelos con partición del espacio de entradas por agrupamiento substractivo con 2, 3 y 4 clusters. También figura en la citada tabla el número de epochs que fue posible entrenar cada uno de los sistemas [Jang93]. El entrenamiento (ajuste de los parámetros de los modelos) se realizó con los datos del grupo de entrenamiento, interrumpiendo el proceso en el momento en que los errores de predicción con los datos del grupo de validación comenzaban a incrementarse. Todos los modelos representados son, en realidad, sistemas híbridos neuro-difusos, al combinar las características de aprendizaje de las redes neuronales supervisadas con la de los sistemas de inferencia difusa. La selección de un único modelo neuro-difuso se realizó determinando el modelo que proporcionaba un menor error medio (RMSE medio en la Tabla 8.5) con los datos de los grupos de entrenamiento y validación.

| Entradas | mfs | reglas | epochs | RMSE Ent. | RMSE Val. | RMSE medio | RMSE testeo |
|-----------|-----|--------|--------|-----------|-----------|------------|-------------|
| $v(t)$ | 2 | 16 | 642 | 0.8183 | 0.7620 | 0.7902 | 0.7053 |
| | 3 | 81 | 93 | 0.8085 | 0.7671 | 0.7878 | 0.7087 |
| | 2 | 2 | 400 | 0.8235 | 0.7713 | 0.7974 | 0.7093 |
| | 3 | 3 | 152 | 0.8191 | 0.7595 | 0.7893 | 0.7035 |
| $v(t-1)$ | 4 | 4 | 160 | 0.8174 | 0.7597 | 0.7886 | 0.7075 |
| | 2 | 16 | 27 | 0.8000 | 0.7479 | 0.7740 | 0.7063 |
| | 3 | 81 | 72 | 0.7449 | 0.8108 | 0.7779 | 0.9127 |
| | 2 | 2 | 207 | 0.8144 | 0.7551 | 0.7847 | 0.7061 |
| $v(t-2)$ | 3 | 3 | 321 | 0.8108 | 0.7462 | 0.7785 | 0.7054 |
| | 4 | 4 | 197 | 0.8086 | 0.7404 | 0.7745 | 0.7056 |
| | 2 | 16 | 4 | 0.7952 | 0.7583 | 0.7768 | 0.7065 |
| | 3 | 81 | 220 | 0.7587 | 0.7982 | 0.7785 | 0.8115 |
| $v(t-22)$ | 2 | 2 | 220 | 0.8109 | 0.7440 | 0.7775 | 0.7004 |
| | 3 | 3 | 270 | 0.8078 | 0.7393 | 0.7736 | 0.7017 |
| | 4 | 4 | 250 | 0.8064 | 0.7375 | 0.7720 | 0.7035 |
| | 2 | 16 | 112 | 0.7934 | 0.7552 | 0.7743 | 0.7017 |
| $v(t-23)$ | 3 | 81 | 35 | 0.7423 | 0.8204 | 0.7814 | 1.0525 |
| | 2 | 2 | 92 | 0.8120 | 0.7479 | 0.7800 | 0.7033 |
| | 3 | 3 | 95 | 0.8091 | 0.7484 | 0.7788 | 0.7033 |
| | 4 | 4 | 17 | 0.8119 | 0.7482 | 0.7801 | 0.7026 |
| $v(t-24)$ | 2 | 16 | 50 | 0.8012 | 0.7461 | 0.7736 | 0.7098 |
| | 3 | 81 | 250 | 0.7599 | 0.8102 | 0.7851 | 0.8188 |
| | 2 | 2 | 300 | 0.8150 | 0.7423 | 0.7787 | 0.7035 |
| | 3 | 3 | 2 | 0.8205 | 0.7471 | 0.7838 | 0.7038 |
| $v(t)$ | 4 | 4 | 90 | 0.8144 | 0.7430 | 0.7787 | 0.7058 |
| | 2 | 8 | 540 | 0.8213 | 0.7684 | 0.7949 | 0.7083 |
| | 3 | 27 | 300 | 0.8172 | 0.7680 | 0.7926 | 0.7069 |
| | 2 | 2 | 100 | 0.8234 | 0.7705 | 0.7970 | 0.7088 |
| $v(t-1)$ | 3 | 3 | 320 | 0.8178 | 0.7566 | 0.7872 | 0.7051 |
| | 4 | 4 | 215 | 0.8150 | 0.7550 | 0.7850 | 0.7055 |
| | 2 | 8 | 300 | 0.8099 | 0.7461 | 0.7780 | 0.7025 |
| | 3 | 27 | 250 | 0.7903 | 0.7637 | 0.7770 | 0.7662 |
| $v(t-2)$ | 2 | 2 | 340 | 0.8160 | 0.7497 | 0.7829 | 0.7058 |
| | 3 | 3 | 350 | 0.8139 | 0.7417 | 0.7778 | 0.7069 |
| | 4 | 4 | 255 | 0.8086 | 0.7426 | 0.7756 | 0.7060 |
| | 2 | 4 | 200 | 0.8195 | 0.7604 | 0.7900 | 0.7055 |
| $v(t-23)$ | 3 | 9 | 170 | 0.8131 | 0.7501 | 0.7816 | 0.7092 |
| | 2 | 2 | 90 | 0.8253 | 0.7674 | 0.7964 | 0.7089 |
| | 3 | 3 | 220 | 0.8234 | 0.7595 | 0.7915 | 0.7101 |
| | 4 | 4 | 400 | 0.8157 | 0.7493 | 0.7825 | 0.7065 |
| $v(t)$ | 2 | 4 | 160 | 0.8281 | 0.7534 | 0.7907 | 0.7090 |
| | 3 | 9 | 100 | 0.8266 | 0.7807 | 0.8037 | 0.7562 |
| | 2 | 2 | 85 | 0.8284 | 0.7537 | 0.7911 | 0.7088 |
| | 3 | 3 | 10 | 0.8278 | 0.7552 | 0.7915 | 0.7097 |
| $v(t-23)$ | 4 | 4 | 4 | 0.8279 | 0.7563 | 0.7921 | 0.7089 |

Tabla 8.5. Resultados, con los modelos de inferencia difusa, del error RMSE de predicción de la velocidad media horaria del viento.

El sistema neuro-difuso con menor valor del error RMSE medio (de los grupos de entrenamiento y validación) es un sistema que utiliza de entradas los valores de $v(t)$, $v(t-1)$, $v(t-22)$ y $v(t-23)$. Está obtenido por agrupamiento substractivo con los datos de los grupos de entrenamiento y validación, presenta 4 clusters y otras 4 reglas, se entrena hasta un total de 250 epochs y presenta un error RMSE en la predicción de la velocidad media horaria del viento de 0.7035 m/s con los datos del grupo de testeo.

En cuanto al nuevo sistema propuesto con descomposición de Fourier y predicción de los nuevos coeficientes para obtener el valor de predicción de la velocidad media horaria, los resultados obtenidos para el error RMSE de la velocidad media horaria con el grupo de testeo fue de 0.712 m/s utilizando una única red neuronal con 23 salidas (los 23 coeficientes independientes). Si en lugar de utilizar una única red neuronal para predecir los 23 coeficientes independientes se utilizan 23 redes, una por coeficiente, el error se reduce a 0.707 m/s. Todas las redes, la de 23 salidas y las 23 redes de una única salida correspondían al tipo MLP con dos capas ocultas, y se optimizaron en un proceso controlado por un algoritmo genético, como en los casos anteriores. Las entradas de todas las redes eran iguales, los 23 coeficientes independientes (ver Tabla 7.2) más el coeficiente a_6 para el instante t .

Si en lugar de utilizar 23 redes neuronales se emplean 23 sistemas de inferencia difusa con las mismas entradas que las consideradas para las redes neuronales con una sola salida, el error RMSE con los datos del grupo de testeo se reduce ligeramente con respecto al caso neuronal, quedando en 0.706 m/s. Para cada coeficiente se repitió la metodología comentada en las anteriores páginas, desarrollando 5 sistemas diferentes (dos con partición del espacio de entrada en rejilla y otros tres con partición del espacio de entrada con agrupamiento substractivo con 2, 3 y 4 clusters). En todos los sistemas de inferencia se realizó un entrenamiento como sistema neuro-difuso, si bien sólo en algunos casos se logró una reducción del error. La selección de los sistemas se realizó por medio del menor error RMSE medio con los datos de los grupos de entrenamiento y validación. Así se seleccionaron los 23 sistemas de inferencia difusa (en 9 de ellos había sido posible reducir el error por medio del entrenamiento).

Como resumen, los resultados obtenidos con los diversos modelos utilizados en la predicción de la velocidad media horaria del viento con horizonte de predicción de una hora están reflejados en la Tabla 8.6. En dicha tabla, aparece el error de predicción para los datos del grupo de testeo, así como la mejora aportada por cada modelo tomando como referencia el modelo persistente.

8.1.4.1. Mejora de los modelos de predicción a muy corto plazo.

Tras el análisis de los resultados obtenidos con los distintos modelos se consideraron diversas soluciones para conseguir reducir los errores de predicción. La solución que ofreció mejores resultados en las primeras pruebas era la de incrementar el número de entradas de los modelos. Concretamente, no se ha utilizado ninguna entrada para indicar, por ejemplo, la hora de predicción. Esta entrada es común en otros sistemas de predicción relacionados con sistemas de energía eléctrica, como es el caso de predicción de la demanda.

| Modelo | RMSE (m/s) | mejora |
|-------------------------------|------------|---------|
| Persistente | 0.753 | --- |
| Valor medio | 1.774 | -135.6% |
| Nueva referencia | 0.737 | 2.12% |
| Holt-Winters (aditivo) | 0.744 | 1.20% |
| Holt-Winters (multiplicativo) | 0.749 | 0.53% |
| ARIMA | 0.700 | 7.04% |
| Red neuronal (MLP) | 0.699 | 7.17% |
| FIS | 0.703 | 6.64% |
| Fourier + 23 MLPs | 0.707 | 6.11% |
| Fourier + 23 FIS | 0.706 | 6.24% |

Tabla 8.6. Error RMSE en la predicción de la velocidad media horaria del viento con los diferentes modelos y mejora con respecto al modelo persistente.

La hora de predicción (u hora del horizonte de predicción) corresponde a un valor entre 0 y 23, si bien no puede usarse de esta forma en los modelos de predicción porque éstos no suelen interpretar correctamente el “salto” entre la hora 23 y la hora 0 (el salto es de sólo una hora, no de 23). Por ello en lugar de usar este valor de la hora directamente, se suele usar en forma de dos entradas distintas, una correspondiente al seno de la porción de día (considerando al día como un ciclo de 2π radianes) que representa la hora, x_1 , y la otra correspondiente al coseno de ese mismo valor de porción de día, x_2 . Por ejemplo, esta solución está adoptada en [Ramirez06]. Las ecuaciones (8.11) expresan estos dos valores, x_1 y x_2 , siendo h la hora del día (de 0 a 23).

$$x_1 = \sin\left(h \frac{2\pi}{24}\right) \quad x_2 = \cos\left(h \frac{2\pi}{24}\right) \quad (8.11)$$

Las primeras pruebas de modelos de predicción con las nuevas variables de entrada lograron algunas ligeras mejoras respecto a las pruebas anteriores. Así, el

modelo de red neuronal que mejores resultados había dado en las pruebas anteriores, el MLP, reduce en una milésimas el error RMSE de predicción con los datos del grupo de testeo. El error con esas dos entradas adicionales, x_1 y x_2 , es de 0.693 m/s.

Otra solución estudiada fue la utilizar como entrada no la hora, sino el valor medio de la velocidad del viento para esa hora (en lo que se denominó como “hora modificada”). Este valor medio es obtenido de los datos de los grupos de entrenamiento y validación. Con esta nueva entrada se entrenaron los dos mejores modelos de redes neuronales que en la experiencia anterior, utilizando exclusivamente los 24 últimos valores de la serie temporal de velocidad media del viento, habían conseguido los mejores resultados de predicción. Estos modelos son una red neuronal MLP y una red neuronal FFBP. Siguiendo los mismos pasos realizados anteriormente (optimización con algoritmo genético, selección de entradas, etc.), se obtuvieron dos modelos optimizados, uno para cada tipo de red neuronal. Los errores RMSE de predicción de la velocidad media horaria con los datos del grupo de testeo con estas redes fueron de 0.695 m/s para la red neuronal MLP y de 0.686 m/s para la red neuronal FFBP.

La utilización de esta entrada adicional de “hora modificada” con el sistema de predicción basado en la descomposición de Fourier con 23 sistemas de inferencia difusa, permitió reducir el error de predicción de la velocidad media horaria del viento en mayor medida que en el resto de sistemas. Los nuevos 23 sistemas de inferencia difusa se desarrollaron siguiendo la metodología descrita anteriormente. El error de predicción con los datos del grupo de testeo se redujo a 0.683 m/s, el valor más reducido de todos los modelos estudiados. La Tabla 8.7 muestra los resultados obtenidos con los tres modelos que incorporaron esta “hora modificada”.

| Modelo | RMSE (m/s) | mejora |
|------------------------------|------------|--------|
| Red neuronal con hora (MLP) | 0.695 | 7.76% |
| Red neuronal con hora (FFBP) | 0.686 | 8.90% |
| Fourier + 23 FIS con hora | 0.683 | 9.29% |

Tabla 8.7. Error RMSE en la predicción de la velocidad media horaria del viento en el grupo de testeo con los modelos que tienen como entrada la “hora modificada”.

Así, la estructura del modelo finalmente propuesto para la predicción de la velocidad media horaria del viento, utilizando como única variable de entrada la propia serie temporal, es el representado en la Figura 8.2. En dicha figura el valor de “hora modificada” es obtenido de la propia serie temporal de velocidad media horaria. El modelo responde a tres módulos interconectados. El primer módulo obtiene los

coeficientes de Fourier (realizando la transformada discreta de Fourier) de los 24 últimos valores conocidos de la velocidad media horaria del viento. El segundo módulo está compuesto por 23 sistemas de inferencia difusa que realizan la predicción de los 23 coeficientes independientes para la próxima hora. El tercer módulo recompone el valor de la velocidad media horaria del viento a partir de los valores de predicción de los coeficientes obtenidos en el segundo módulo.

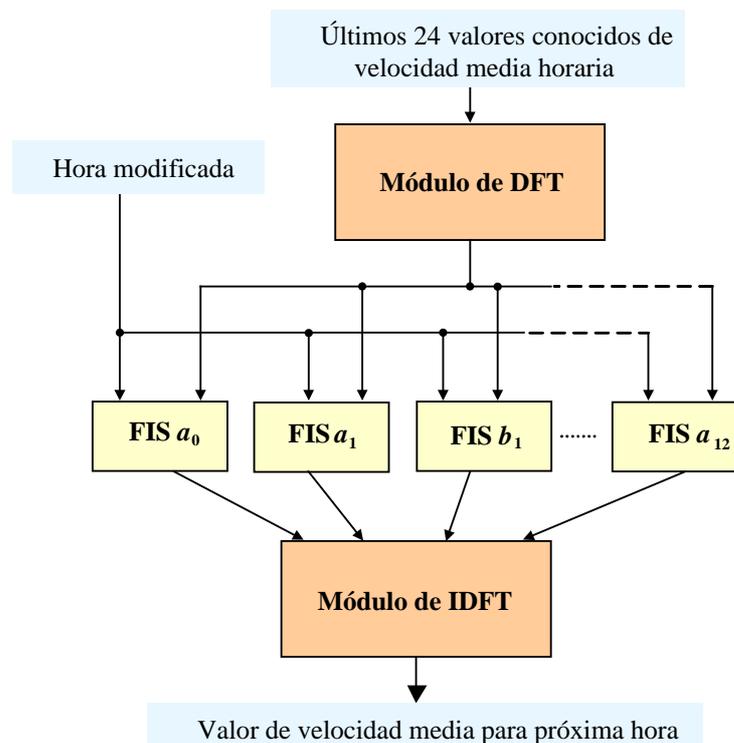


Figura 8.2. Modelo propuesto para la predicción de la velocidad media horaria el viento.

La Figura 8.3 muestra un ejemplo de las predicciones obtenidas con el modelo propuesto. En concreto, en dicha figura están representados los valores real y de predicción, con horizonte de predicción de una hora, de la velocidad media horaria del viento para las dos últimas semanas de junio de 1992, pertenecientes al grupo de testeo.

8.1.4.2. Predicción de la potencia eléctrica media generada.

Una vez obtenido el valor de predicción para la velocidad media horaria del viento con el modelo propuesto, puede calcularse la potencia eléctrica media horaria generada en un parque eólico a partir de dicho valor de predicción. Suponiendo que todos los aerogeneradores fueran del mismo modelo, se puede calcular la potencia

eléctrica media generada en el horizonte de predicción de una hora utilizando la curva de potencia del aerogenerador o una curva de potencia empírica del parque, resultado de medidas sobre el propio parque, en la que se tiene en cuenta la posible dirección del viento. Esta es la solución adoptada en [Joensen97], [Hatziargyriou01] y [Moehrlen02].

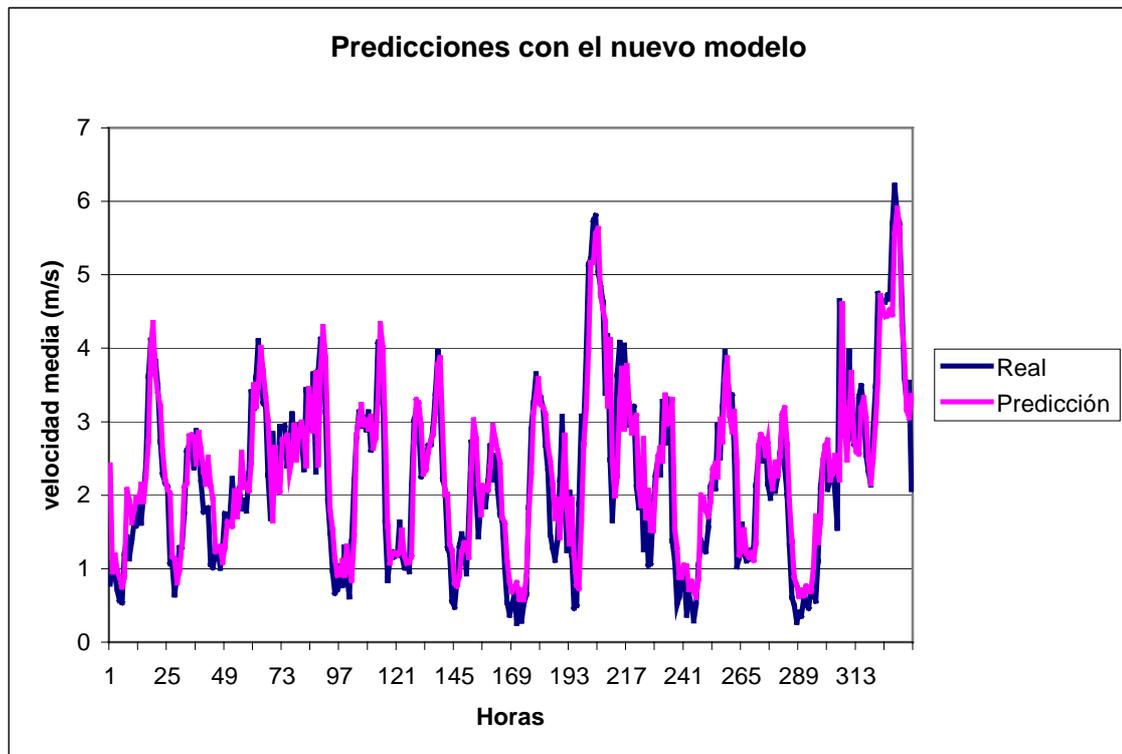


Figura 8.3. Valor real y de predicción de la velocidad media horaria del viento, utilizando el modelo de predicción propuesto.

Para este caso concreto, con un aerogenerador aislado (parque eólico con un único aerogenerador) de potencia instalada de 750 kW, el error RMSE en la predicción de la potencia eléctrica media generada para la próxima hora, calculada con las predicciones obtenidas con el nuevo modelo, es de 32.38 kW para todo el grupo de testeo, lo que representa un 4.32% de la potencia eléctrica instalada. En la Figura 8.4 se presentan los valores de potencia eléctrica media generada real y su valor de predicción, con el modelo propuesto, para las dos últimas semanas del mes de junio de 1992 (pertenecientes al grupo de testeo).

8.2. PREDICCIÓN DE LA ENERGÍA ELÉCTRICA GENERADA EN UN PARQUE EÓLICO A CORTO PLAZO

El segundo problema de predicción planteado es el de la predicción de la energía

eléctrica total generada en un parque eólico en las 24 horas siguientes, contando con una información escasa o bien expresada con incertidumbre mediante términos lingüísticos. En este caso la predicción resulta útil para la programación de labores de mantenimiento en un parque eólico. El modelo de predicción seleccionado está basado en el uso de series temporales fuzzy o difusas, FTS (Fuzzy Time Series).

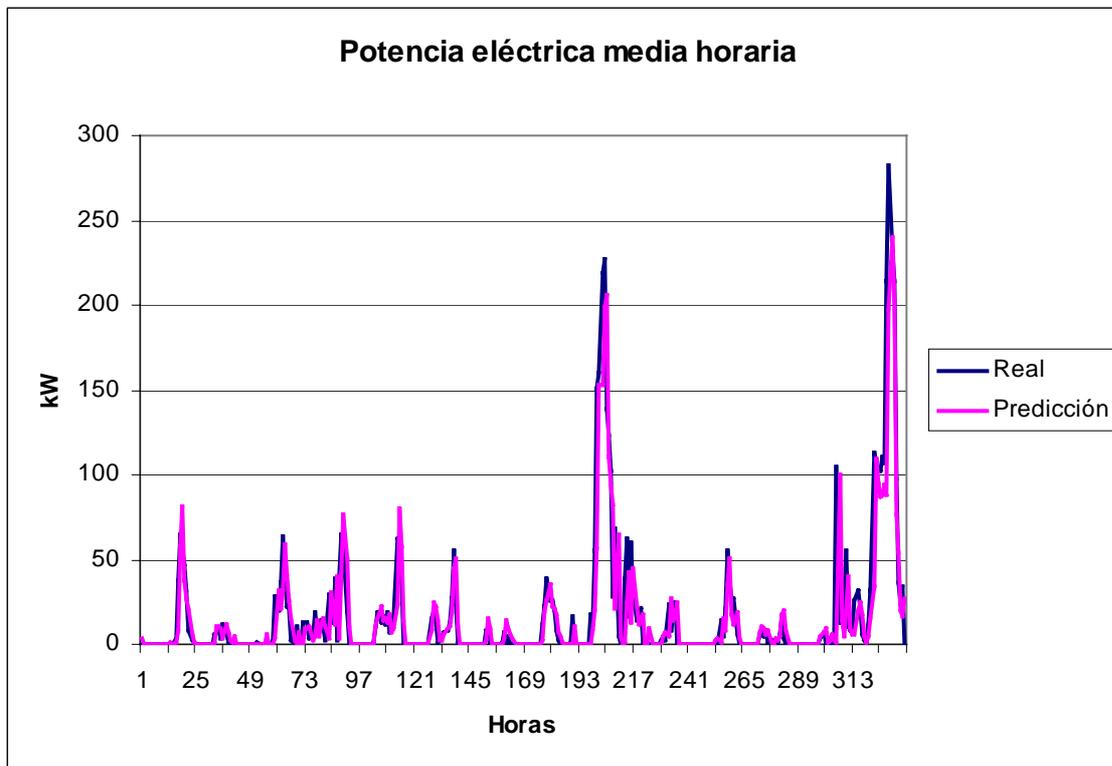


Figura 8.4. Valor real y de predicción de la potencia eléctrica media horaria del viento, utilizando el nuevo modelo de predicción.

8.2.1. Datos disponibles.

Para el desarrollo del sistema de predicción se utilizaron el conjunto de datos designado en el Capítulo 7 como “Serie II”. Se tomaron los datos de nueve meses con la energía eléctrica diaria generada en un parque eólico situado en la provincia de Burgos. Se usaron los datos de 6 meses (181 datos) para el ajuste o entrenamiento de los modelos y los datos de los restantes 3 meses (92 datos) para el testeo y comparación de resultados entre modelos. La Figura 8.5 representa la energía eléctrica diaria generada en este parque para los dos primeros meses de la serie.

8.2.2. Modelos de predicción analizados.

Los modelos analizados para el diseño del sistema de predicción fueron los indicados seguidamente.

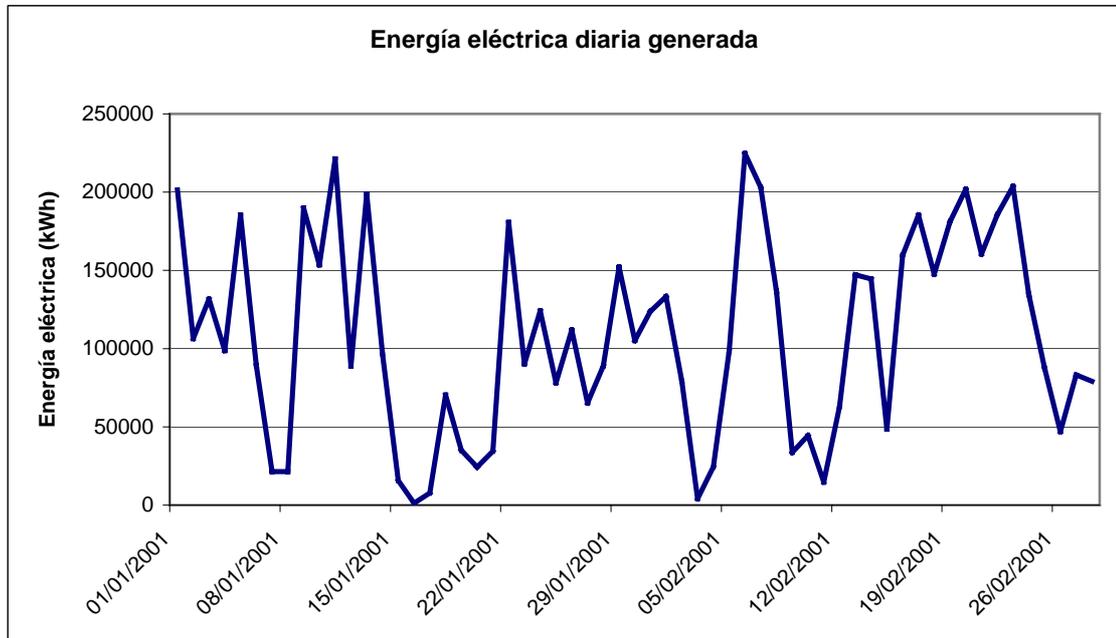


Figura 8.5. Energía eléctrica diaria generada en el parque eólico.

1. *Modelo persistente:* la energía eléctrica generada en las próximas 24 horas es igual a la generada en las últimas 24 horas.
2. *Modelo auto-regresivo:* para cada nueva predicción, se desarrollaron seis modelos auto-regresivos desde el de orden 1, el AR(1), hasta el de orden 6, el AR(6). El modelo auto-regresivo seleccionado para realizar la predicción fue aquel que proporcionó un menor valor del criterio de información de Akaike (AIC) [Akaike76] con el grupo de datos de ajuste (entrenamiento), más los datos correspondientes a los días ya transcurridos del grupo de testeo. El uso de este criterio asegura la selección del modelo de acuerdo al principio de parsimonia [DeLurgio98].
3. *Modelo de red neuronal:* se utilizó una red neuronal del tipo MLP que permitía usar como entradas los últimos 6 valores conocidos de la energía eléctrica diaria generada en el parque eólico. La salida de la red neuronal fue la energía eléctrica generada para el día siguiente (24 horas siguientes a la de la realización de la predicción). Se utilizó un algoritmo genético en la

optimización de la estructura de la red neuronal, de forma que el proceso de optimización realizaba la selección de entradas entre las 6 posibles, seleccionaba el número de neuronas en la capa oculta y seleccionaba los valores del factor de aprendizaje y momento del algoritmo de retropropagación de error.

Para evitar un sobre-entrenamiento de la red neuronal, el grupo de entrenamiento se dividió en dos mitades tras ordenarlo aleatoriamente. La primera mitad se utilizó para realizar el entrenamiento de la red; la segunda mitad se empleó como grupo de validación. El entrenamiento de la red neuronal se interrumpió tras 100 epochs sin reducción de error con el grupo de validación.

4. *Modelo de serie temporal difusa de un factor*: se desarrolló el modelo descrito en el apartado 7.3.2.1 de este documento. Únicamente se hizo uso de la serie temporal de energía eléctrica diaria producida para obtener la serie temporal fuzzy de un factor.
5. *Modelo de serie temporal difusa de dos factores*: se desarrolló el modelo descrito en el apartado 7.3.2.2 de este documento. La primera serie temporal fuzzy fue igual a la usada por el modelo anterior. La segunda serie temporal fuzzy fue una serie de términos lingüísticos que expresaban (con incertidumbre) los valores de predicción de velocidad del viento para el día siguiente en la provincia en la que se encuentra el parque eólico (Burgos).

8.2.3. Desarrollo de los modelos de predicción.

Únicamente el desarrollo de los modelos de series temporales fuzzy requiere una metodología no comentada hasta ahora. En los siguientes párrafos se describe paso a paso la metodología seguida en el desarrollo del modelo de predicción de serie temporal difusa de un sólo factor.

1. *Creación de la serie temporal de diferencias entre energías eléctricas generadas diarias.*

La serie temporal original con la energía eléctrica generada diariamente en el parque eólico no es la adecuada para su transformación en serie temporal fuzzy. Por ello,

obtuvimos una nueva serie temporal que es la diferencia primera de la serie temporal original. Esta diferencia primera denota que el elemento j de la nueva serie temporal corresponde a la diferencia entre los elementos $j+1$ y j de la serie temporal original.

2. *Definición del universo del discurso.*

Tras analizar la nueva serie temporal obtenida en el primer paso, se obtuvieron los valores máximo y mínimo para establecer los límites del universo del discurso. Sin embargo, para asegurar que cualquier valor futuro (y no sólo los correspondientes al grupo de datos de entrenamiento) iba a estar comprendido entre esos límites, se añadió una cantidad al valor máximo y se le restó al valor mínimo.

En el caso aquí desarrollado, los valores límites de la nueva serie temporal eran -80.5 MWh para el valor mínimo y 54.6 MWh para el valor máximo. Estos límites se expresaban de forma abreviada como [-80.5, 54.6] MWh. La cantidad a restar o sumar a cada límite se estimó en torno al 7% de la diferencia entre ambos (límite máximo menos límite mínimo), es decir, una cantidad en torno a los 9.5 MWh. Así el universo del discurso quedó establecido en [-90, 65] MWh.

3. *Determinación de los valores óptimos del número de términos lingüísticos (asociados a la incertidumbre de la predicción) y de la ventana temporal utilizada en el modelo.*

En los modelos de predicción de series temporales fuzzy se consideran exclusivamente los w últimos valores conocidos de la propia serie temporal, es decir, w representa la longitud de la ventana temporal utilizada en la determinación de los datos que configuran el modelo de predicción.

El universo del discurso debe ser partido en k intervalos iguales. A cada uno de estos intervalos se le hace corresponder un término lingüístico. Por ejemplo, si fueran dos los intervalos, los términos lingüísticos podrían ser “decremento” e “incremento”; para cinco intervalos los términos podrían ser “gran decremento”, “decremento”, “igualdad”, “incremento” y “gran incremento”. En todo caso el número de intervalos, k , tiene que ser calculado.

Para la determinación del valor óptimo de ambos parámetros, w y k , se diseñaron todos los sistemas de predicción de serie temporal fuzzy de un factor con valores

entre 2 y 40. El sistema finalmente seleccionado sería aquél que presentase un menor error con los datos del grupo de entrenamiento (los utilizados en ajustar los modelos). Los sistemas se diseñaron en las etapas descritas a continuación.

- A) El universo del discurso se dividió en k intervalos iguales. El primer intervalo fue denominado como u_1 , y el último u_k .

Siguiendo el caso numérico real, al partir el universo de discurso, por ejemplo, en 5 intervalos iguales ($k = 5$), estos fueron:

$$u_1 = [-90, -59] \quad u_2 = [-59, -28] \quad u_3 = [-28, 3] \quad u_4 = [3, 34] \quad u_5 = [34, 65]$$

- B) Se definieron los conjuntos difusos sobre el universo del discurso, asignando un posible valor lingüístico a cada uno de los intervalos. Cada uno de estos conjuntos difusos quedaron definidos como expresa la ecuación (8.12).

$$\begin{aligned} A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + \dots + 0/u_k \\ A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + \dots + 0/u_k \\ &\quad \vdots \\ A_{k-1} &= 0/u_1 + \dots + 0.5/u_{k-2} + 1/u_{k-1} + 0.5/u_k \\ A_k &= 0/u_1 + \dots + 0/u_{k-2} + 0.5/u_{k-1} + 1/u_k \end{aligned} \tag{8.12}$$

- C) Los valores de la nueva serie temporal obtenida en el paso 1 fueron “fuzzyficados” asignándoles el término lingüístico y el conjunto difuso que les correspondía en función del intervalo en el que se encontraba su valor numérico. Se obtuvo así la serie temporal fuzzy, FTS.
- D) Para cada valor de la FTS se calculó el vector de criterio, $\mathbf{C}(t)$ y la matriz de operación $\mathbf{O}^w(t)$, tal y como se reflejaba en el apartado 7.3.2.1 de este documento. La matriz de relación, $\mathbf{R}(t)$, se calculó de acuerdo a la ecuación (7.17), y a partir de sus elementos se calculó el valor de predicción para la serie temporal difusa como expresa la ecuación (7.19).
- E) El nuevo valor de la FTS calculado en la etapa anterior fue “defuzzyficado”. Para este proceso se utilizaron tres reglas:

Regla 1: Si el nuevo valor difuso presentaba un solo máximo para u_i , con

un valor de pertenencia entre 0.25 y 1, entonces el valor de predicción “defuzzyficado” correspondía al valor central del intervalo u_i .

Regla 2: Si el nuevo valor difuso presentaba dos máximos para u_i y u_{i+1} , con un valor de pertenencia de 0.5 en ambos casos, entonces el valor de predicción “defuzzyficado” correspondía a la media de los valores centrales de ambos intervalos.

Regla 3: Si el nuevo valor difuso presentaba todos los valores de pertenencia nulos, entonces el valor de predicción “defuzzyficado” fue 0.

- F)** Se calculó el valor de predicción para la energía eléctrica generada en el parque eólico en las próximas 24 horas por medio de la ecuación (8.13), donde $e_d(t)$ corresponde al último valor de la energía eléctrica diaria generada en el parque eólico (la generada en las últimas 24 horas), y $\hat{e}_d(t+1)$ corresponde al valor de predicción para la energía eléctrica que se generará en las próximas 24 horas.

$$\hat{e}_d(t+1) = e_d(t) + \text{valor de predicción "deffuzzyficado"} \quad (8.13)$$

- G)** Una vez calculados los valores de predicción para los datos del grupo de entrenamiento, se calculó el error RMSE producido en la predicción utilizando los datos de entrenamiento. Se seleccionan los valores de k y w correspondientes al sistema de predicción que ofrezca un menor valor de error.

4. Realización de predicciones.

Una vez seleccionados los valores de longitud de la ventana temporal, w , y del número de términos lingüísticos, k , pudo utilizarse el modelo para realizar las predicciones. En este caso los valores óptimos de ambos parámetros fueron 4 términos lingüísticos ($k = 4$) y una ventana temporal de 22 días ($w = 22$). Los términos lingüísticos fueron “gran decremento”, (A_1), “decremento”, (A_2), “incremento”, (A_3), y “gran incremento”, (A_4), que quedan definidos por las ecuaciones (8.14)

$$\begin{aligned} A_1 &= [1 \ 0.5 \ 0 \ 0] & A_2 &= [0.5 \ 1 \ 0.5 \ 0] \\ A_3 &= [0 \ 0.5 \ 1 \ 0.5] & A_4 &= [0 \ 0 \ 0.5 \ 1] \end{aligned} \quad (8.14)$$

En la construcción de la serie temporal difusa de dos factores se utilizó la información lingüística de predicción general del viento facilitada por el Instituto Nacional de Meteorología para el día siguiente en la provincia de Burgos (se puede obtener la información a partir de medios de comunicación regionales). Los términos utilizados para la predicción del viento fueron los de “flojo”, “moderado”, “fuerte” y “muy fuerte”. Con esta información, el valor difuso para la velocidad media del viento prevista para el día siguiente en la zona del parque eólico podía expresarse como indica la ecuación (8.15).

$$S(t) = [S_1 \ S_2 \ S_3 \ S_4] \quad (8.15)$$

En la ecuación (8.15) aparecen cuatro elementos, que corresponden al valor del número de términos lingüísticos considerados. El valor de cada uno de los elementos viene dado por la Tabla 8.8.

| Término | S₁ | S₂ | S₃ | S₄ |
|----------------|----------------------|----------------------|----------------------|----------------------|
| Flojo | 1 | 1 | 1 | 0 |
| Moderado | 1 | 1 | 1 | 1 |
| Fuerte | 1 | 1 | 1 | 1 |
| Muy fuerte | 0 | 1 | 1 | 1 |

Tabla 8.8. Funciones de pertenencia para la FTS de segundo factor.

El significado real de cada uno de los elementos de la Tabla 8.8 es sencillo de inferir. Si el valor medio del viento para el día siguiente está previsto que sea “flojo” es de suponer que no vaya a existir un importante incremento en la producción de energía eléctrica en el parque eólico con respecto a la situación de hoy. Por ello, el elemento S_4 es nulo. Análogamente, si el valor medio del viento previsto para el día siguiente es “muy fuerte” es lógico esperar que no haya una gran disminución en la energía eléctrica producida en el parque eólico con respecto a la energía eléctrica generada hoy (por lo que S_1 es nulo). Cuando la velocidad media prevista del viento es “moderada” o “fuerte” se permiten las cuatro situaciones que van de “gran decremento” a “gran incremento”.

El comportamiento de la FTS de segundo factor es igual al del vector de criterio

(expresión difusa de la situación a día de hoy), tal y como se expresa en las ecuaciones (7.20) y (7.21). Todos los pasos son similares a los seguidos para la predicción con una FTS de un solo factor, a excepción de una pequeña variante en la etapa D), en concreto en el cálculo de la matriz de relación $R(t)$, que no se calcula de acuerdo a la ecuación (7.17), sino de acuerdo a la ecuación (7.20), siendo ahora $S(t)=G(t)$.

8.2.4. Resultados obtenidos.

Tras el desarrollo de todos los modelos analizados para la predicción de la energía eléctrica generada en un parque eólico con un horizonte de predicción de 24 horas, se evaluó el error de predicción con cada uno de ellos utilizando los datos del grupo de testeo, consistente en los datos de tres meses no utilizados en el desarrollo de los modelos. Para el caso de la serie temporal fuzzy de dos factores se recogió también la información lingüística con la predicción de la velocidad media del viento en la región donde se encuentra situado el parque eólico.

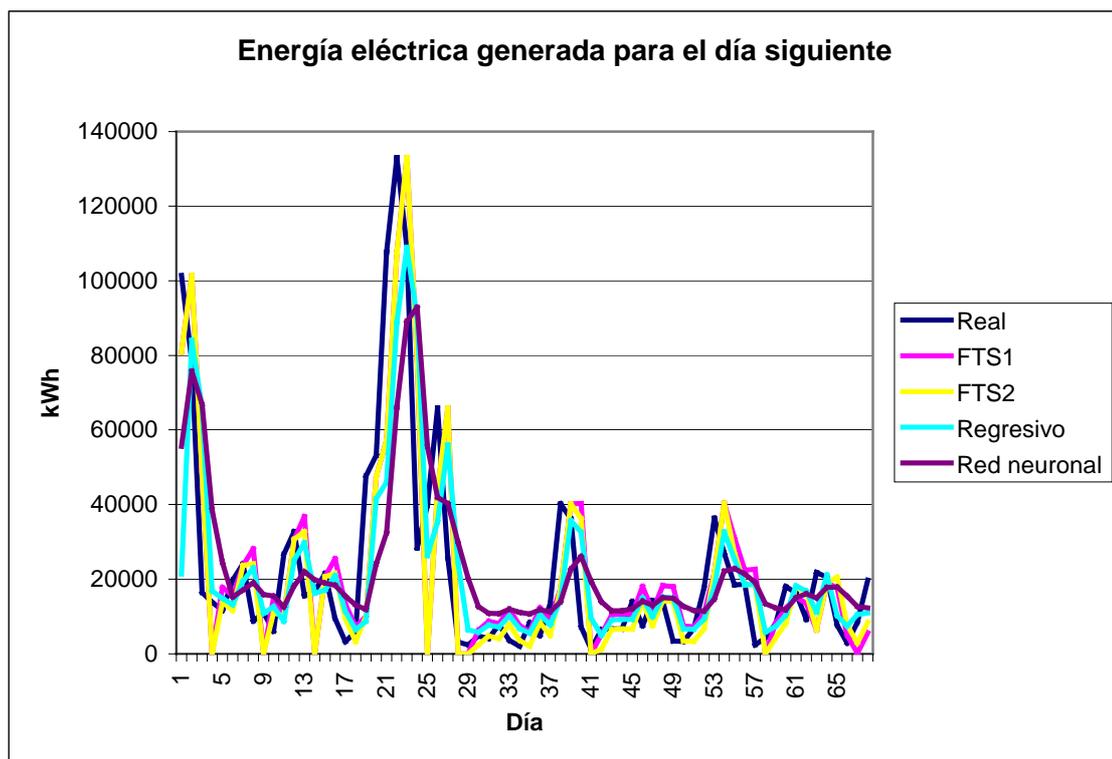


Figura 8.6. Energía eléctrica real generada en el parque eólico y predicciones de la energía eléctrica con cuatro de los modelos de predicción.

La Figura 8.6 muestra las predicciones obtenidas con cuatro de los modelos considerados: auto-regresivo, red neuronal MLP, serie temporal difusa de un factor

(FTS1) y serie temporal difusa de dos factores (FTS2). En la citada figura se representa la energía eléctrica generada en el parque eólico en los últimos 68 días del grupo de testeo y las predicciones para la misma con los distintos modelos de predicción. La energía eléctrica diaria máxima generada en ese periodo fue de 132922 kWh.

La Tabla 8.9 refleja los errores RMSE obtenidos en la predicción con cada uno de los modelos propuestos. En dicha tabla los modelos de series temporales fuzzy vienen reflejados como FTS1 (serie temporal difusa de un factor) y FTS2 (serie temporal difusa de dos factores).

| Modelo | RMSE (MWh) | mejora |
|----------------|-------------------|---------------|
| Persistente | 19.29 | --- |
| Auto-regresivo | 19.67 | -1.97% |
| Red neuronal | 18.04 | 6.48% |
| FTS1 | 16.97 | 12.03% |
| FTS2 | 16.53 | 14.31% |

Tabla 8.9. Error RMSE en la predicción de la energía eléctrica diaria generada en el parque eólico.

Los errores de predicción con los dos modelos de predicción basados en series temporales difusas fueron inferiores a los de los otros tres sistemas considerados. La mejora conseguida con el modelo de predicción basado en la serie temporal difusa de dos factores (FTS2) fue del 14.31% con respecto al modelo persistente.

8.3. PREDICCIÓN DE LA POTENCIA MEDIA HORARIA GENERADA EN UN PARQUE EÓLICO A CORTO PLAZO

Un tercer problema de predicción planteado dentro de los trabajos de investigación fue el de la predicción de la potencia media horaria generada, a muy corto o corto plazo, en un parque eólico. En este caso la única información disponible fue la de la propia serie temporal de potencia media generada en el parque eólico, sin que existiera información adicional de predicciones elaboradas con modelos atmosféricos.

En el Capítulo 3 se indicó que los horizontes de predicción de los modelos que denominábamos como pertenecientes a la primera familia de modelos (los basados exclusivamente en la serie temporal de velocidad media del viento o de potencia

eléctrica media generada en el parque eólico) eran de unas pocas horas. La razón apuntada entonces al porqué de ese limitado horizonte es la del incremento del error de predicción para horizontes cada vez más elevados, lo que les hace competitivamente inferiores a los modelos de predicción que usan predicciones previas de variables meteorológicas obtenidas de un modelo atmosférico. Aun con lo limitado de la información disponible, los modelos de predicción de la primera familia de modelos son superiores a los de la segunda familia con horizontes de predicción pequeños.

El horizonte de predicción que marca la frontera entre la superioridad de un modelo de la primera familia y un modelo de la segunda no es un valor de referencia igual para cualquier parque eólico. El valor de ese horizonte “frontera” puede cambiar de un parque eólico a otro, e incluso para un mismo parque puede cambiar de un periodo a otro; por ejemplo, puede no ser el mismo en invierno que en verano, como consecuencia de cambios en las funciones que rigen las predicciones de los modelos atmosféricos (en un caso debe contemplar situaciones de hielo o nieve que en el otro caso no debe considerar). En los trabajos publicados, en los que se muestran los resultados de la predicción de la potencia media generada en un parque eólico con modelos de las dos familias, ese horizonte frontera es de 4 horas en [Landberg99], de 5 horas en [Ramírez06] o de 6 horas en [Landberg01].

Con el objetivo de complementar las predicciones, en el muy corto plazo y en las primeras horas del corto plazo, de los modelos de predicción de la segunda familia, se ha desarrollado un nuevo modelo de predicción de la potencia media horaria generada en un parque eólico que utiliza solamente los valores pasados de esa variable (modelo de la primera familia). El horizonte máximo de predicción se ha fijado en 6 horas.

8.3.1. Datos disponibles.

Para el desarrollo del sistema de predicción se utilizó parte de los datos designado en el Capítulo 7 como “Serie II”. Se tomaron los datos de 48 días de un parque eólico situado en Navarra. Los datos fueron divididos en tres grupos: entrenamiento, validación y testeo. Los datos de potencia eléctrica media horaria de 25 días fueron designados como datos del grupo de entrenamiento. Los datos de poco más de 6 días fueron designados como datos del grupo de validación. Los datos de los últimos 16 días fueron designados como datos del grupo de testeo. La Tabla 8.10 recoge valores estadísticos de los datos de cada uno de los tres grupos.

La Figura 8.7 representa la potencia eléctrica media horaria para el grupo de datos de entrenamiento (610 horas). Puede observarse la ausencia de patrones temporales y el rango de valores de la serie: potencia eléctrica media máxima de 8900 kW y mínima de 0 kW.

| Grupo | Entrenamiento | Validación | Testeo |
|---------------------|---------------|------------|----------|
| Horas | 1-610 | 611-762 | 763-1147 |
| Número horas | 610 | 152 | 385 |
| Media (kW) | 3452 | 2720 | 2325 |
| Mínima (kW) | 0 | 0 | 0 |
| Máxima (kW) | 8900 | 7400 | 8800 |
| Desviación estándar | 2974 | 2158 | 2573 |

Tabla 8.10. Valores estadísticos de los tres grupos de datos.

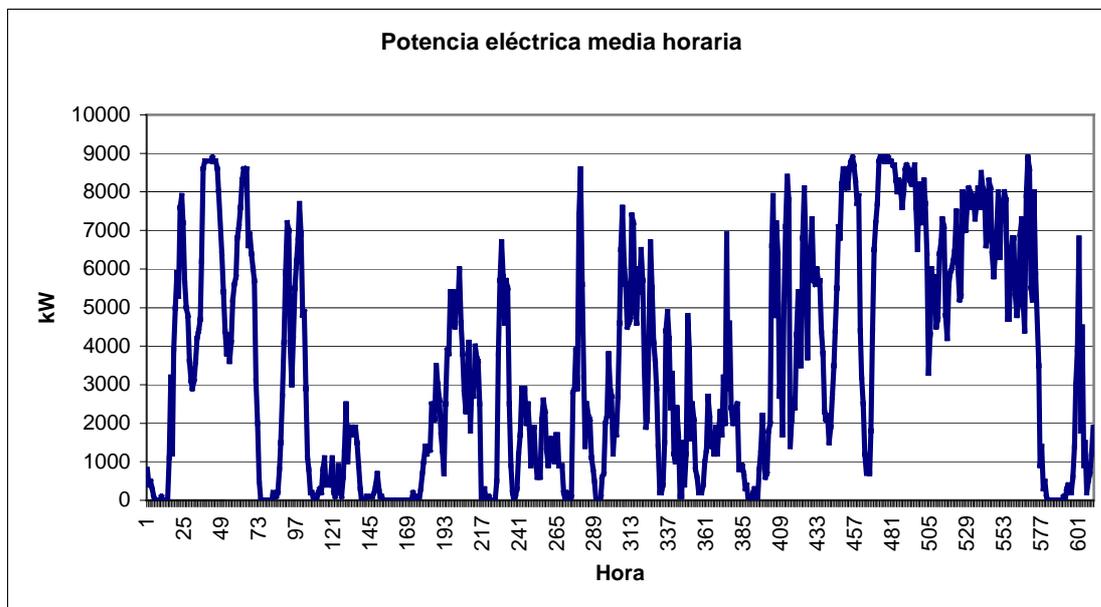


Figura 8.7. Potencia eléctrica media horaria para los datos del grupo de entrenamiento.

8.3.2. Modelos de predicción analizados.

Los modelos analizados para el diseño del sistema de predicción fueron los siguientes:

1. *Modelo persistente.*

La potencia eléctrica media horaria generada para el horizonte de predicción,

$p(t+k)$, es igual a la de la hora actual, $p(t)$.

2. Modelo “nueva referencia”.

Este modelo, propuesto en [Nielsen00] como alternativo al modelo persistente en las predicciones a muy corto y corto plazo, ya ha sido descrito anteriormente en el apartado 8.1.2. En este caso fue usado para realizar la predicción de la potencia eléctrica media horaria.

Los datos de los grupos de entrenamiento y validación fueron usados para determinar los valores de los coeficientes del modelo. En este caso, dado que el horizonte de predicción era variable (1, 2, ..., 6), el valor de predicción con este modelo podía calcularse por medio de la ecuación (8.16). En dicha ecuación k representa el valor del horizonte de predicción en horas; $\hat{p}(t+k)$ es el valor de predicción (potencia eléctrica media horaria) para la hora $t+k$; a_k es el coeficiente de correlación entre $p(t)$ y $p(t+k)$; y μ el valor medio de la serie temporal de potencia eléctrica media horaria.

$$\hat{p}(t+k) = a_k p(t) + (1-a_k)\mu \quad (8.16)$$

El valor del coeficiente de correlación a_k viene dado por la ecuación (8.17), en la que N representa el número total de datos (serie temporal de potencia eléctrica media horaria con los datos de entrenamiento y validación), el valor $\tilde{p}(t)$ representa la diferencia entre la potencia eléctrica media horaria para el instante t y el valor medio de la serie temporal de dicha variable, μ , como expresa la ecuación (8.18).

$$a_k = \frac{\frac{1}{N} \sum_{t=1}^{N-k} \tilde{p}(t) \tilde{p}(t+k)}{\frac{1}{N} \sum_{t=1}^{N-k} \tilde{p}^2(t)} \quad (8.17)$$

$$\tilde{p}(t) = p(t) - \mu \quad (8.18)$$

3. Modelo auto-regresivo.

Se desarrollaron modelos auto-regresivos de ordenes entre 1 y 6, utilizando los

datos de entrenamiento y validación para el ajuste de los modelos. La selección del modelo representativo de esta familia se hizo escogiendo aquél que presentaba un menor valor del criterio de información de Akaike (AIC) con los datos utilizados para su ajuste.

4. *Modelo basado en red neuronal.*

Se desarrollaron modelos basados en redes neuronales de los tipos perceptrón multicapa (MLP), red neuronal multicapa generalizada (FFBP), red neuronal de Elman, red neuronal de función de base radial (RBF), red neuronal con análisis de componentes principales (red neuronal PCA) y red neuronal recurrente.

Para todas las redes neuronales las entradas consideradas fueron los últimos 6 valores de la potencia eléctrica media horaria, de $p(t)$ a $p(t-5)$. Todas las redes neuronales desarrolladas fueron optimizadas por medio de un algoritmo genético que permitía la selección de las entradas (se forzó a que los dos últimos valores de la potencia media horaria, $p(t)$ y $p(t-1)$, estuvieran siempre en los modelos; así en la optimización fue posible escoger adicionalmente cualquiera de las otras cuatro posibles entradas), el número de neuronas en capa oculta, el número de centros (en las redes neuronales RBF), el número de componentes principales (en las redes neuronales PCA) y los parámetros del factor de aprendizaje y momento del algoritmo de retropropagación de error.

Los datos del grupo de validación fueron utilizados para interrumpir el proceso de entrenamiento, si el error no disminuía con los datos de este grupo tras 100 epochs. El límite de epochs de entrenamiento fue de 1500. En caso de interrumpirse el entrenamiento antes de este límite (por incremento del error con el grupo de validación), los parámetros de la red resultante fueron aquellos que ofrecían el menor error con los datos del grupo de validación.

5. *Modelo de inferencia difusa.*

Se desarrollaron diferentes sistemas de inferencia difusa con 6, 5, 4, 3 ó 2 entradas. Las entradas consideradas correspondían siempre a los últimos valores conocidos de la potencia eléctrica media horaria. El espacio de entrada se dividió en rejilla (con 2 y 3 funciones de pertenencia por variable) y por agrupamiento substractivo (con 2, 3 ó 4 clusters). Con esto, el total de sistemas de inferencia difusa desarrollados ascienden a 25.

Una vez construido cada modelo, se intentó realizar un ajuste de los parámetros por medio de la técnica adaptativa neuro-difusa descrita en [Jang93]. Para evitar el sobre-entrenamiento se utilizaron los datos del grupo de validación para interrumpir el proceso de entrenamiento cuando el error con los datos de este grupo comenzaba a incrementarse. El resultado final del proceso fue la construcción de los 25 modelos de inferencia difusa.

6. *Modelo de predicción basado en regresión lineal múltiple.*

Se trata del modelo presentado en el apartado 7.3.3 de este documento. Básicamente este modelo consta de varios módulos de regresión lineal múltiple, obteniéndose el valor de predicción por medio de una media ponderada de los valores aportados por cada módulo.

Inicialmente este modelo no iba a ser contemplado; de hecho la idea que lo generó fue fruto del análisis de los resultados obtenidos con los otros modelos. Se observó un mejor comportamiento en la predicción de los sistemas de inferencia difusa que realizaban un agrupamiento de los valores del vector de entradas por medio del agrupamiento substractivo. De ahí surgió la idea de aplicar la misma técnica de agrupamiento para crear varios grupos de valores del vector de entradas con características similares, seguido del desarrollo de un módulo de regresión lineal múltiple para los valores de cada uno de los grupos.

En un principio se consideró que la salida de todo el sistema de predicción fuera la del modelo de regresión lineal múltiple que correspondiese al cluster al que pertenecía el valor del vector de entradas considerado, pero se observó que los errores se reducían realizando una media ponderada con los valores de salida de todos los modelos de regresión lineal múltiple. En una segunda etapa se comprobó que los resultados mejoraban más si se les daba mayor peso a aquellos valores del vector de entradas más cercanos al centro de cada cluster. Por ello, antes de obtener los coeficientes del modelo de regresión lineal múltiple con los valores del vector de entradas pertenecientes a cada uno de los clusters, se repitieron en el conjunto de valores cada uno de ellos de forma inversamente proporcional a su distancia al centro del cluster (el valor del vector de entradas que correspondía al centro del cluster no se consideró). De esta forma dimos mucho mayor peso a aquellos valores del vector de entradas más cercanos al centro del cluster y redujimos el peso de los más alejados.

Para la selección del mejor modelo de predicción basado en regresión lineal múltiple se desarrollaron un total de 25 modelos. Se desarrollaron todos los modelos con un número de entradas entre 2 y 6 (últimos valores conocidos de la potencia eléctrica media horaria) y con un número de clusters también entre 2 y 6. Todos los modelos fueron construidos con los datos del grupo de entrenamiento. Los datos de validación fueron utilizados para evaluar los 25 modelos. Se seleccionó finalmente el modelo que presentó un menor valor del criterio de información de Akaike (AIC) en la predicción del conjunto de los datos de los grupos de entrenamiento y de validación.

8.3.3. Resultados obtenidos.

Horizonte de predicción de 1 hora

Una vez desarrollados los modelos se seleccionó sólo uno de cada de tipo de modelo analizado. En el caso del modelo persistente el modelo era único y no permitía, por tanto, selección. Lo mismo ocurrió con el modelo “nueva referencia”, en el que se utilizaron los grupos de datos de entrenamiento y validación para el cálculo de los coeficientes de correlación y valores medios, necesarios en el modelo.

En la evaluación de los modelos auto-regresivos, el seleccionado fue el modelo de cuarto orden, AR(4), por ser el que proporcionaba un menor valor del AIC y no presentaba correlación serie en los residuos. Este modelo utilizaba como variables explicativas los últimos 4 valores conocidos de la potencia eléctrica media horaria, de $p(t)$ a $p(t-3)$.

Para los modelos de redes neuronales, el método de optimización por medio de un algoritmo genético proporcionó una única red neuronal de cada uno de los tipos considerados. Cada uno de estos modelos obtenidos, uno por tipo de red neuronal, fue evaluado. Las características de estos modelos están recogidas en la Tabla 8.11.

En los sistemas de inferencia difusa con partición del espacio de entrada con rejilla, el mejor modelo fue el que usaba como entradas los dos últimos valores conocidos de la potencia horaria, $p(t)$ y $p(t-1)$, y tenía dos funciones de pertenencia por entrada y un total de 4 reglas.

Para los sistemas de inferencia difusa con partición del espacio de entrada con agrupamiento substractivo, el mejor sistema usaba también los dos últimos valores conocidos de la potencia eléctrica media horaria generada en el parque, $p(t)$ y $p(t-1)$, tenía dos clusters y, por tanto, dos familias de pertenencia por variable de entrada y un total de dos reglas.

| Tipo de red | Características | Entradas |
|-------------|--|--|
| MLP | 18 neuronas en 1 capa oculta | $p(t), p(t-1), p(t-2), p(t-3)$ |
| FFBP | 8 neuronas en 1 capa oculta | $p(t), p(t-1), p(t-3), p(t-5)$ |
| ELMAN | 9 neuronas en 1 capa oculta | $p(t), p(t-1), p(t-2), p(t-3), p(t-5)$ |
| RBF | 10 clusters, 8 neuronas en 1 capa oculta | $p(t), p(t-1)$ |
| PCA | 3 componentes, 8 neuronas en 1 capa oculta | $p(t), p(t-1), p(t-3)$ |
| Recurrente | 8 neuronas en 1 capa oculta | $p(t), p(t-1)$ |

Tabla 8.11. Características de las redes neuronales optimizadas.

En cuanto al nuevo modelo propuesto de sistema de predicción basado en regresión lineal múltiple, el valor menor del AIC con los datos de entrenamiento y validación se obtuvo con un sistema que utilizaba como entradas los dos últimos valores conocidos de la potencia eléctrica media horaria generada en el parque, $p(t)$ y $p(t-1)$, y presentaba tres clusters. Bajo estas condiciones, la expresión de la ecuación (7.22) que representa el valor de la salida para ese tipo de sistemas, quedó reducida a la ecuación (8.19).

$$s_j = \frac{\frac{1}{d_{j1}} s_{j1} + \frac{1}{d_{j2}} s_{j2} + \frac{1}{d_{j3}} s_{j3}}{\frac{1}{d_{j1}} + \frac{1}{d_{j2}} + \frac{1}{d_{j3}}} \quad (8.19)$$

En la ecuación (8.19) el término d_{jk} representa la distancia euclídea del valor del vector de entradas j al cluster k ; s_{jk} es la salida del modelo de regresión lineal múltiple correspondiente al cluster k ; y s_j es la salida de predicción del sistema para el valor del vector de entradas j . Puesto que la potencia generada por el parque eólico no puede ser negativa, en los casos en que el valor resultante de la ecuación (8.19) fue negativo, se consideró que el valor de la salida s_j era 0.

Los resultados obtenidos con cada uno de los sistemas finalmente seleccionados, en la predicción de la potencia media horaria generada en el parque eólico, para la siguiente hora y con los datos del grupo de testeo, son los recogidos en la Tabla 8.12.

Como se puede apreciar en los datos de la Tabla 8.12 el sistema propuesto basado en regresión lineal múltiple es superior al resto de sistemas de predicción estudiados, presentando una mejora con respecto al modelo persistente del 4.69%. En la Figura 8.8 se puede observar la potencia eléctrica media horaria real y las predicciones obtenidas con los dos de los mejores sistemas (el propuesto y el modelo de inferencia difusa con agrupamiento substractivo para un horizonte de predicción de 1 hora), y para los primeros 7 días del grupo de testeo.

| Modelo | RMSE (kW) | Mejora |
|------------------|-----------|--------|
| Persistente | 966.24 | --- |
| Nueva referencia | 953.38 | 1.33% |
| AR(4) | 938.58 | 2.86% |
| Red MLP | 956.21 | 1.04% |
| Red FFBP | 1002.08 | -3.71% |
| Red de Elman | 1041.14 | -7.75% |
| Red RBF | 964.63 | 0.17% |
| Red PCA | 1010.05 | -4.53% |
| Red recurrente | 1033.01 | -6.91% |
| FIS rejilla | 946.15 | 2.08% |
| FIS substractivo | 941.94 | 2.51% |
| Modelo propuesto | 920.96 | 4.69% |

Tabla 8.12. Errores RMSE en la predicción de la potencia eléctrica media horaria.

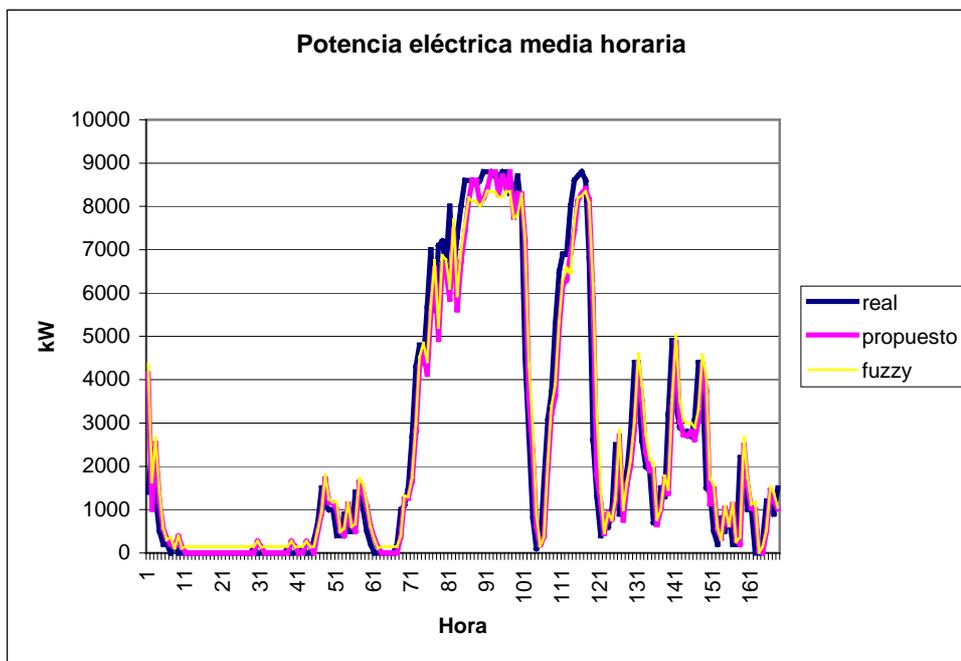


Figura 8.8. Potencia eléctrica media horaria real y predicciones con el modelo propuesto y el mejor modelo de inferencia difusa con horizonte de predicción de 1 hora.

Horizonte de predicción hasta 6 horas

Una vez analizados los errores en la predicción de la potencia eléctrica media horaria de los diferentes modelos con un horizonte de 1 hora, se extendió el horizonte de predicción hasta las 6 horas, sin más que utilizar los valores de predicción obtenidos para horizontes inferiores como valores de entrada en los modelos para horizontes mayores. Veámoslo con un ejemplo, la predicción de la potencia eléctrica media horaria generada para $t+3$ (horizonte de predicción de 3 horas) se puede obtener utilizando como entradas los valores de predicción obtenidos por el propio sistema para $t+1$ y $t+2$, tal como muestra la Figura 8.9.

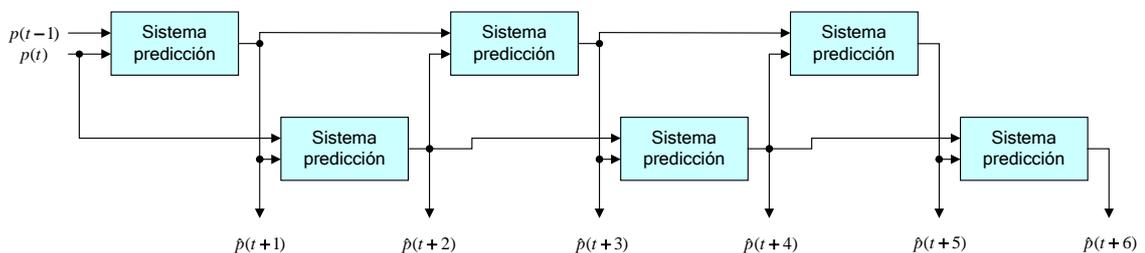


Figura 8.9. Extensión del modelo de predicción hasta un horizonte de predicción de 6 horas.

De cada uno de los tipos de modelos, se seleccionó el de mejor comportamiento en la predicción con el horizonte de una hora y se utilizó iterativamente para completar los horizontes de predicción de 2, 3, 4, 5 y 6 horas. Así, de los modelos de redes neuronales sólo se utilizó la red neuronal MLP y de los sistemas de inferencia difusa el que presentaba dos entradas y utilizaba agrupamiento substractivo.

Los resultados, con la mejora con respecto al modelo persistente obtenida en la predicción de la potencia eléctrica media horaria con horizontes de predicción de hasta 6 horas, aparecen en la Tabla 8.13.

La Figura 8.10 representa los valores reales de la potencia eléctrica media horaria y el valor de predicción de la potencia utilizando el sistema propuesto, con un horizonte de predicción de 6 horas, para todo el grupo de testeo. Aunque el sistema propuesto es el que ofrece mejores resultados, en relación con el error de predicción de entre todos los estudiados, la diferencia visible entre ambos valores (real y de predicción) justifica la búsqueda de nuevos modelos de predicción diferentes de estos (que sólo utilizan la serie temporal de potencia eléctrica media horaria) y que puedan mejorar la calidad de las predicciones cuando el horizonte de predicción se incrementa

por encima de las 4 ó 5 horas.

El modelo propuesto, a pesar de su sencillez (en el sentido de que no precisa de entrenamiento como pueden ser el caso de las redes neuronales artificiales, lo cual hace que se pueda construir en un tiempo mucho menor), se muestra superior al resto de modelos en la predicción de la potencia eléctrica media horaria para todos los horizontes de predicción considerados hasta un máximo de 6 horas.

| Horizonte (horas) | MODELO | | | | |
|-------------------|--------|---------|-------------|------------------|-------------------|
| | AR(4) | Red MLP | Sistema FIS | Nueva referencia | Sistema propuesto |
| 1 | 2.86% | 1.04% | 2.51% | 1.33% | 4.69% |
| 2 | 3.57% | 4.12% | 4.05% | 3.57% | 5.43% |
| 3 | 5.31% | 5.57% | 6.07% | 5.69% | 7.43% |
| 4 | 7.67% | 6.30% | 7.83% | 7.70% | 9.73% |
| 5 | 9.47% | 7.39% | 9.40% | 9.70% | 11.89% |
| 6 | 11.06% | 8.27% | 11.28% | 11.72% | 14.11% |

Tabla 8.13. Mejora porcentual de las predicciones respecto a las del modelo persistente.

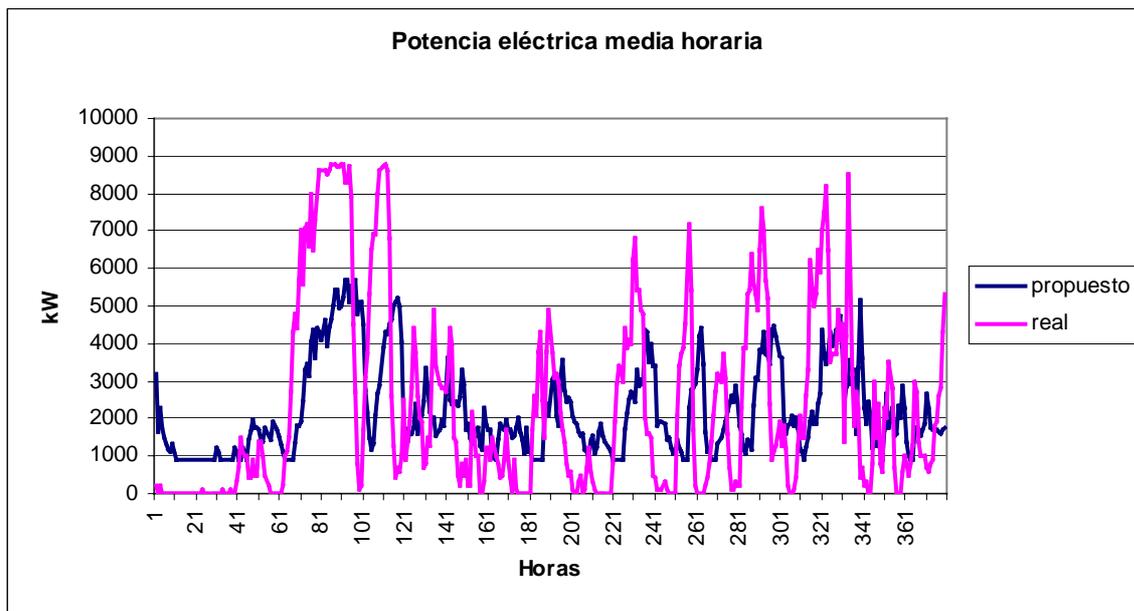


Figura 8.10. Potencia eléctrica media horaria real y predicciones utilizando el modelo propuesto con horizonte de predicción de 6 horas.

Predicción de la energía generada en 6 horas

Disponiendo ya de la herramienta para realizar la predicción de la potencia eléctrica media horaria con un horizonte de predicción de hasta 6 horas, la predicción de

la energía eléctrica total generada para ese horizonte de predicción se obtiene sumando las predicciones de potencia media generada para las 6 horas de horizonte de predicción. Realizando esta suma para los seis horizontes de predicción utilizando los cinco modelos de la Tabla 8.13 más el persistente, y comparando los resultados con los valores reales de producción de energía eléctrica en el parque eólico para cada 6 horas, se obtiene la Tabla 8.14 que recoge los errores RMSE en la predicción de la energía eléctrica generada en el parque eólico para las 6 horas siguientes.

Como se puede apreciar en la Tabla 8.14 la mejora obtenida con el modelo propuesto con respecto al modelo persistente es del 28.63%, casi el doble que la obtenida por el segundo mejor modelo.

| Modelo | Error RMSE (kWh) | Mejora |
|-------------------|------------------|--------|
| Persistente | 12526 | --- |
| AR(4) | 10694 | 14.63% |
| Red MLP | 11782 | 5.94% |
| Sistema FIS | 10893 | 13.04% |
| Nueva referencia | 11243 | 10.24% |
| Sistema propuesto | 8940 | 28.63% |

Tabla 8.14. Errores RMSE en la predicción de la energía generada en 6 horas en el parque eólico.

La Figura 8.11 representa la energía generada en el parque eólico en periodos de 6 horas (real) y los valores de predicción de la misma obtenidos con el modelo propuesto (propuesto) y el modelo auto-regresivo de cuarto orden (AR(4)). Estos dos modelos son los que proporcionan un menor error de predicción de los analizados. Los datos representados corresponden a los últimos 7 días del grupo de testeo. El valor máximo de la energía eléctrica generada en 6 horas en el periodo representado es de 41.6 MWh.

8.4. PREDICCIÓN A CORTO PLAZO DE LA POTENCIA MEDIA HORARIA GENERADA EN UN PARQUE EÓLICO CON INFORMACIÓN LIMITADA DE UN MODELO ATMOSFÉRICO

El cuarto problema de predicción planteado dentro de los trabajos de investigación es el de la predicción de la potencia media horaria generada, a corto plazo en un parque eólico cuando están disponibles las predicciones de un modelo atmosférico. La diferencia de este problema de predicción con el que se analizará en el

apartado 8.5 es que en este caso la información disponible del modelo atmosférico es limitada en el tiempo: los datos de predicciones del modelo atmosférico comienzan a estar disponibles mucho después de la puesta en funcionamiento del parque eólico, por lo que no se dispone de registros históricos de los mismos en cantidad suficiente como para desarrollar modelos de predicción que precisen de aprendizaje o entrenamiento.

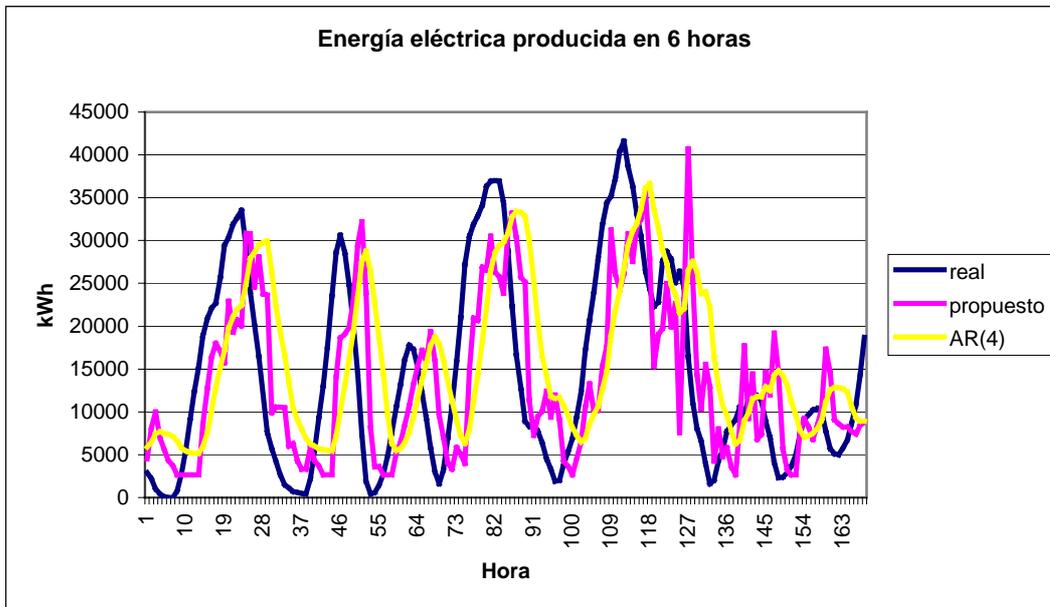


Figura 8.11. Energía eléctrica real producida en intervalos de 6 horas y valores de predicción utilizando el nuevo modelo propuesto y el modelo de “nueva referencia”.

8.4.1. Datos disponibles.

Para el desarrollo del sistema de predicción se utilizaron los datos designados en el Capítulo 7 como “Serie III”. Se trata de la serie temporal de datos de un parque eólico para un periodo de 14 meses, concretamente del 1 de marzo de 2004 al 30 de abril de 2005. Los datos incluyen la potencia eléctrica media horaria, los registros SCADA sobre base horaria de cada aerogenerador (el parque eólico consta de 10 aerogeneradores iguales de 1.8 MW), y las predicciones de velocidad y dirección media horaria del viento para los últimos 68 días. Estas predicciones han sido obtenidas con el modelo atmosférico MM5 y corresponden a un punto geográficamente muy cercano al parque eólico. Las predicciones atmosféricas del modelo MM5 cubren un horizonte de predicción de 24 horas. Los datos corresponden a las predicciones atmosféricas correspondientes a la asimilación de datos de las 0 horas GMT y cubren los horizontes de predicción de las 0 hasta las 23 horas.

La disponibilidad de los registros SCADA de cada aerogenerador aporta una ventaja adicional: podemos determinar, analizando estos registros, cuantos aerogeneradores estaban realmente funcionando en cada momento. En condiciones normales, alguno de los aerogeneradores que componen un parque eólico puede estar fuera de servicio por causas como avería, parada por labores de mantenimiento o parada como consecuencia de una alarma de su equipo de control (puede parar durante unos minutos). No está contemplado como parada el hecho de que el aerogenerador no genere energía como consecuencia de la baja velocidad del viento (baja energía cinética del viento). Cuando la potencia media horaria generada por uno de los aerogeneradores es nula y no lo es para el resto, y además la potencia media de los otros aerogeneradores supera un umbral mínimo (puede ser que la velocidad del viento sea baja, por lo que uno de los aerogeneradores está parado y el resto está generando una potencia muy baja), podemos considerar que ese aerogenerador está fuera de servicio. En las series temporales de potencia media horaria de cada aerogenerador se puede determinar si uno de ellos ha estado funcionando de forma anómala (por ejemplo, parado unos minutos por una alarma). Para ello basta analizar si la potencia de uno de los aerogeneradores representa un outlier en comparación con la potencia generada por el resto de aerogeneradores. La detección de outliers se realiza mediante técnicas estadísticas como el test Q de Dixon. Una estimación de la fracción de hora en la que el aerogenerador ha estado funcionando correctamente, corresponde al cociente entre la potencia media horaria generada por ese aerogenerador dividido por el mismo valor pero referido a la media del resto de aerogeneradores.

Una vez analizados los registros SCADA de los aerogeneradores es posible obtener la serie temporal de capacidad de generación disponible, $C_d(t)$, que representa la máxima capacidad de generación (producto de la potencia nominal del aerogenerador por el número de ellos en correcto funcionamiento, siendo todos los aerogeneradores iguales) en cada intervalo horario. A partir de las series temporales de potencia media horaria generada en el parque, $P(t)$, y la de capacidad de generación disponible, $C_d(t)$, es posible determinar la serie temporal del factor de capacidad, $FC(t)$, tal como expresa la ecuación (7.25).

El objetivo es la predicción del factor de capacidad del parque eólico con un horizonte de predicción de 24 horas, que es justamente el horizonte de las predicciones del modelo atmosférico. A partir de los valores de predicción del factor de capacidad resulta sencillo, para un usuario del sistema de predicción, determinar la potencia eléctrica media generada: sólo es necesario multiplicar el valor previsto del factor de capacidad por la potencia máxima del aerogenerador y por el número de

aerogeneradores que estarán en correcto funcionamiento. Obviamente, siempre puede haber averías imprevistas, pero el gestor del parque puede saber de antemano si alguno de los aerogeneradores estará parado por labores programadas de mantenimiento o por averías todavía no reparadas.

Los datos disponibles fueron divididos en tres grupos: el primer grupo con los datos correspondientes a los primeros 269 días (un 63% del total) fue designado como el grupo de entrenamiento; el segundo con los datos de los siguientes 89 días (un 20.9% del total) fue designado como el grupo de validación; y el tercer grupo con los datos de los últimos 68 días (el 15.9% del total) fue designado como el grupo de testeo. El tamaño del grupo de testeo vino limitado por el periodo de tiempo disponible con predicciones del modelo atmosférico.

8.4.2. Modelos de predicción analizados.

En este caso la mayor parte de los modelos analizados para la predicción del factor de capacidad en el parque eólico estaban basados en redes neuronales, excepto el modelo persistente, un modelo ARIMA y el modelo “nueva referencia”, que se desarrollaron a efectos de referencia comparativa con los otros modelos. Así, los modelos desarrollados fueron los indicados seguidamente.

- A) *Modelo persistente*: el factor de capacidad para el horizonte de predicción, $FC(t+k)$, era igual al de la hora actual, $FC(t)$.
- B) *Modelo ARIMA*: con la metodología Box-Jenkins [Box76] se determinó el mejor modelo ARIMA. Para el ajuste del modelo se utilizaron los datos de los grupo de entrenamiento y de validación (el primer año de datos).
- C) *Modelo “nueva referencia”*: el modelo propuesto en [Nielsen00] fue utilizado aquí en la predicción de la potencia eléctrica media horaria con horizonte de predicción de hasta 24 horas.
- D) *Red neuronal perceptrón multicapa (MLP)*: se desarrollaron redes neuronales MLP con una capa oculta.
- E) *Red neuronal multicapa generalizada (FFBP)*: se desarrollaron redes neuronales FFBP también con una capa oculta.

- F) *Red neuronal modular*: compuesta por dos redes multicapa en paralelo y con una salida obtenida por combinación de las salidas de ambas redes en paralelo.

La Figura 8.12 representa la estructura de la red modular escogida. Cada una de las dos redes en paralelo tiene dos capas ocultas. En ambas redes cada capa está conectada a la anterior y a la posterior. Adicionalmente, las primeras capas ocultas de ambas redes están unidas a la capa de salidas, al igual que la capa de entradas. Esta configuración permite la especialización de ambas redes con un grupo particular de datos del conjunto de entrenamiento, permitiendo, en muchos casos, mejores resultados que los obtenidos con las redes neuronales MLP. La conexión entre capas no significa que todas las neuronas de cada capa estén unidas entre sí.

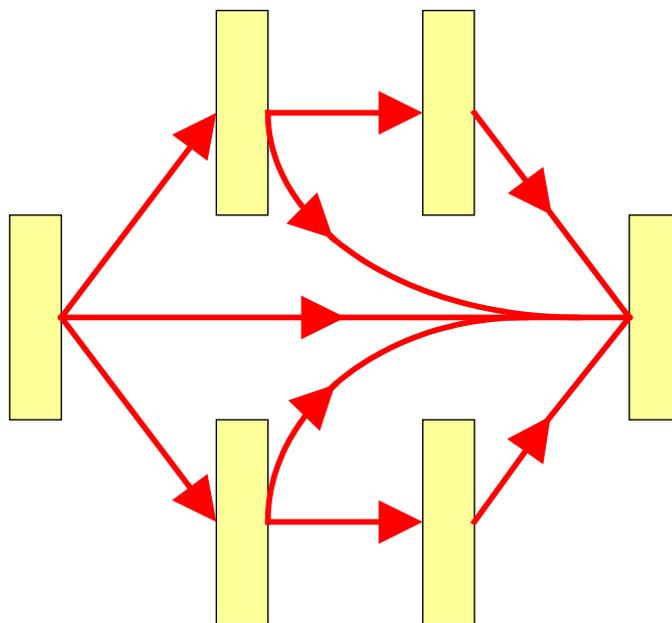


Figura 8.12. Estructura elegida para la red neuronal modular.

- G) *Red neuronal de Elman*: las redes neuronales de Elman [Elman90] son un caso particular de redes recurrentes (el autor las definió como redes recurrentes simples). Responden a la estructura de una red MLP en la que la información de una capa oculta se copia en otra capa de neuronas llamadas neuronas de contexto.

La Figura 8.13 representa la estructura de la red de Elman utilizada, en la que sólo hay una capa oculta. La capa de neuronas de contexto es la situada inferiormente y almacena una copia de la capa oculta, con la particularidad de

“olvidar” está información con un decaimiento exponencial. Así, las salidas de las neuronas de contexto no sólo dependen del valor de las entradas a las mismas en cada momento, sino de los valores previos de esas entradas. El resultado es una media ponderada de los valores de entrada a esas neuronas de contexto, siendo el peso mayor para los valores más recientes y menor cuanto más antiguos son esos valores. El valor escogido para la constante de tiempo de las neuronas de contexto fue 0.8.

- H)** *Red neuronal de función de base radial (RBF)*: las redes neuronales RBF utilizadas en este caso fueron un tipo de redes neuronales híbridas que combinaban aprendizaje supervisado y aprendizaje no supervisado, tal como se definieron en el Capítulo 5 de este documento. Las redes constaban de dos módulos, un primer módulo con una red neuronal RBF pura y un segundo módulo con una red neuronal MLP de una capa oculta. La capa de neuronas de función de base radial de la red neuronal RBF pura actuaba como la capa de entradas para la red neuronal MLP. El entrenamiento de estas redes neuronales exigió entrenar inicialmente el primer módulo con aprendizaje no supervisado (máximo de 500 epochs) y, una vez entrenado, entrenar el segundo módulo con aprendizaje supervisado.

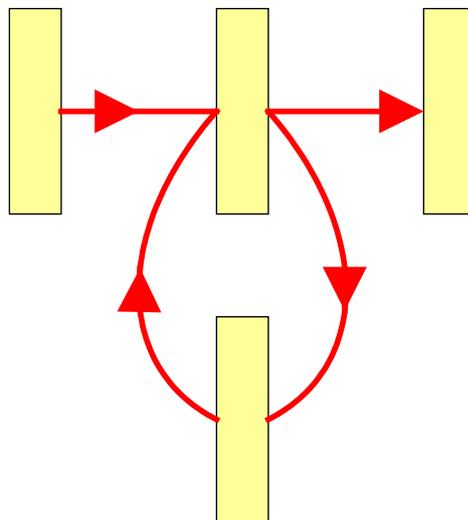


Figura 8.13. Estructura elegida para la red neuronal de Elman de 1 capa oculta.

- I)** *Red neuronal con análisis de componentes principales (PCA)*: la red neuronal PCA utilizada constaba de una capa oculta para el aprendizaje supervisado. El número de componentes principales fue escogido por el proceso de optimización. Para el aprendizaje no supervisado el número de epochs máximo

fue de 500.

- J)** *Red neuronal híbrida con módulo de aprendizaje no supervisado y módulo de aprendizaje supervisado*: se consideraron dos modelos de redes neuronales híbridas, ambas con una estructura similar en cuanto a la disposición de capas de neuronas. Esta estructura combina un mapa auto-organizado o red de Kohonen con una red neuronal MLP. El mapa auto-organizado actúa como la capa de entradas de la red neuronal MLP.

En el primero de los modelos desarrollados el mapa auto-organizado se estructuró en una línea con un total de 11 neuronas, y se denominó como “red híbrida 1”. El segundo modelo consideraba 25 neuronas en el mapa auto-organizado dispuestas en una matriz de 5x5, y se denominó como “red híbrida 2”. Se escogieron estos valores para el tamaño del mapa auto-organizado teniendo en cuenta que un número elevado de neuronas podía ralentizar el proceso de aprendizaje.

- K)** *Red neuronal retardada en el tiempo (Laguerre)*: se utilizó este tipo de red neuronal retardada en el tiempo ([Back96]) ya que reduce el número de neuronas necesarias para almacenar la información de los valores previos de las entradas de la red, disminuyéndose el tiempo de entrenamiento con respecto a una red neuronal TDNN normal. Esta es una red neuronal utilizada en aplicaciones de filtrado de señales.

Todos los modelos de redes neuronales desarrollados presentaban una serie de características comunes. La primera de estas características es que se desarrolló una red neuronal para cada horizonte de predicción, esto es, se desarrollaron un total de 24 redes neuronales de cada tipo, una por cada hora de las 24 horas que conformaba el horizonte máximo de predicción. La segunda característica era que todos los modelos fueron optimizados en un proceso controlado por un algoritmo genético, de igual manera que se ha hecho en el resto de los modelos de redes neuronales desarrollados en los trabajos de investigación recogidos en este documento. La tercera característica común era que se utilizó una función de activación lineal para todas las neuronas de la capa de salidas, con la idea de permitir valores de salida no acotados por las propias funciones de activación del resto de neuronas (la tangente hiperbólica). La cuarta era que todas las redes neuronales usaron las mismas 11 entradas (al menos podían usarlas si las seleccionaba el proceso de optimización). Estas entradas fueron las indicadas a continuación.

- Dos entradas, x_1 y x_2 , para indicar la hora del día correspondiente al horizonte de predicción. El valor de estas dos entradas viene dado por la ecuación (8.20), donde h representa la hora del día (0 a 23).

$$x_1 = \text{sen}\left(h \frac{2\pi}{24}\right) \quad x_2 = \text{cos}\left(h \frac{2\pi}{24}\right) \quad (8.20)$$

- El último valor conocido de la velocidad media horaria del viento, x_3 , recogida del sistema SCADA asociado a la estación meteorológica del parque eólico.
- Dos entradas, x_4 y x_5 , para indicar los últimos valores conocidos de la dirección media horaria del viento. Si d es esa dirección expresada en grados, el valor de las entradas x_4 y x_5 puede calcularse por medio de la ecuación (8.21).

$$x_4 = \text{sen}\left(d \frac{2\pi}{360}\right) \quad x_5 = \text{cos}\left(d \frac{2\pi}{360}\right) \quad (8.21)$$

- Dos entradas, x_6 y x_7 , para indicar los valores de la dirección media horaria del viento para el horizonte de predicción. Estas entradas son similares a las dos anteriores, excepto que ahora la dirección d es la correspondiente para el horizonte de predicción. En el caso de los datos de los grupos de entrenamiento y validación, d es la dirección media horaria del viento para la hora que corresponde al horizonte de predicción; en el caso de los datos de testeo d corresponde a la dirección media horaria del viento prevista para el horizonte de predicción.
- El valor x_8 que representa la relación entre el valor de la velocidad media horaria del viento para el horizonte de predicción y el valor de la velocidad media horaria del viento para la hora actual (momento de realización de la predicción). Si $v(t)$ es la velocidad media horaria del viento para la hora t , esta entrada, x_8 , viene dada por la ecuación (8.22). En el caso de los datos del grupo de testeo x_8 representa la relación (cociente) entre el valor de la velocidad media horaria prevista para el horizonte de predicción y el valor de la velocidad media horaria prevista para la hora actual (valores previstos con el modelo atmosférico).

$$x_8 = \frac{v(t+k)}{v(t)} \quad (8.22)$$

- Tres entradas más, x_9 , x_{10} y x_{11} para los tres últimos valores conocidos del factor de capacidad del parque eólico, tal como muestran las ecuaciones (8.23).

$$\begin{cases} x_9 = FC(t-2) \\ x_{10} = FC(t-1) \\ x_{11} = FC(t) \end{cases} \quad (8.23)$$

8.4.3. Resultados obtenidos.

En el desarrollo del modelo ARIMA, siguiendo la metodología de Box-Jenkins, se determinó en primer lugar que la serie temporal de factor de capacidad era estacionaria. El análisis de los valores de la función de auto-correlación y de la función de auto-correlación parcial sugerían que la serie temporal correspondía a un modelo auto-regresivo puro de orden 1, AR(1). Así el modelo auto-regresivo respondía a la ecuación (8.24), donde los coeficientes b_0 y b_1 fueron estimados usando los datos de los grupos de entrenamiento y validación.

$$\hat{FC}(t+1) = b_0 + b_1 \cdot FC(t) \quad (8.24)$$

Los resultados obtenidos para los coeficientes b_0 y b_1 fueron 0.019909675 y 0.92977611, respectivamente. Para ampliar el horizonte de predicción hasta las 24 horas propuestas se puede emplear la ecuación (8.25), donde k representa el horizonte de predicción ($k = 1, 2, \dots, 24$).

$$\hat{FC}(t+k) = \frac{b_0 \cdot (1-b_1^k)}{1-b_1} + b_1^k \cdot FC(t) \quad (8.25)$$

La Tabla 8.15 muestra los errores RMSE en la predicción del factor de capacidad, FC , del parque eólico para los horizontes de predicción propuestos. En la misma se puede observar que para horizontes de 1 ó 2 horas es mejor el modelo persistente que el modelo auto-regresivo (logra menor error).

En el desarrollo de las redes neuronales se utilizó, como ya se ha hecho en los

trabajos presentados anteriormente, un proceso de optimización basado en un algoritmo genético para optimizar la estructura de dichas redes neuronales. El número de individuos por generación se estableció en 40, al igual que el número de generaciones. La selección de individuos para la siguiente generación se realizó por ruleta ordenada por rango. La probabilidad de cruce fue del 90% y la de mutación del 1%. La función de activación para todas las neuronas fue la tangente hiperbólica. Todos los valores de las variables de entrada fueron previamente normalizados entre los límites de -0.9 y 0.9. El sobre-entrenamiento fue evitado utilizando el grupo de validación: tras cada epoch en el proceso de entrenamiento se determinaba el error medio cuadrático con los datos del grupo de validación; el entrenamiento se interrumpía si el error con el grupo de validación no se había reducido tras 100 epochs y en tal caso se recuperaban los valores de los parámetros (pesos de las conexiones) de la red neuronal que proporcionaban el menor error con los datos de ese grupo. El límite máximo de epochs se estableció en 3000.

| Horizonte | Persistente | AR(1) | Nueva ref. | MLP | FFBP | Modular |
|-----------|-------------|------------|------------|------------|------------|------------|
| 1 | 0.10707883 | 0.10686916 | 0.10559850 | 0.18194846 | 0.18193602 | 0.17026251 |
| 2 | 0.14234080 | 0.13920305 | 0.13871249 | 0.19273889 | 0.19230546 | 0.18823123 |
| 3 | 0.16701084 | 0.16190801 | 0.16124310 | 0.19166814 | 0.19935030 | 0.16008136 |
| 4 | 0.18650475 | 0.17935422 | 0.17864213 | 0.19754125 | 0.20590778 | 0.19557024 |
| 5 | 0.20496332 | 0.19506732 | 0.19451391 | 0.20538657 | 0.20140475 | 0.19949461 |
| 6 | 0.22086475 | 0.20808761 | 0.20770245 | 0.20393468 | 0.20283015 | 0.20010106 |
| 7 | 0.23775027 | 0.22112827 | 0.22099672 | 0.20581962 | 0.20549802 | 0.20673736 |
| 8 | 0.25322618 | 0.23257832 | 0.23264992 | 0.20923083 | 0.21143738 | 0.21147965 |
| 9 | 0.26834086 | 0.24326676 | 0.24349504 | 0.21316957 | 0.21680792 | 0.20867463 |
| 10 | 0.27948686 | 0.25096318 | 0.25123529 | 0.21072670 | 0.21239460 | 0.21459901 |
| 11 | 0.29026748 | 0.25812406 | 0.25837950 | 0.20836734 | 0.21185910 | 0.21179794 |
| 12 | 0.29999008 | 0.26438227 | 0.26459845 | 0.21227405 | 0.21137934 | 0.21302530 |
| 13 | 0.30569911 | 0.26805012 | 0.26823091 | 0.21436649 | 0.21239747 | 0.21299838 |
| 14 | 0.31079362 | 0.27126943 | 0.27138572 | 0.21560208 | 0.21947939 | 0.21377944 |
| 15 | 0.31674118 | 0.27488384 | 0.27487731 | 0.21707344 | 0.22020258 | 0.21736867 |
| 16 | 0.32118273 | 0.27757758 | 0.27744561 | 0.21068648 | 0.21832005 | 0.21161273 |
| 17 | 0.32473266 | 0.27972839 | 0.27947528 | 0.21343458 | 0.21682964 | 0.21872896 |
| 18 | 0.32868766 | 0.28204409 | 0.28165826 | 0.20875882 | 0.20961969 | 0.21533698 |
| 19 | 0.33165841 | 0.28381219 | 0.28328160 | 0.21312092 | 0.21282947 | 0.21254626 |
| 20 | 0.33451206 | 0.28548499 | 0.28481503 | 0.21303731 | 0.21271528 | 0.21610304 |
| 21 | 0.33702684 | 0.28697029 | 0.28616007 | 0.21289858 | 0.21245869 | 0.21444088 |
| 22 | 0.33914196 | 0.28824166 | 0.28728902 | 0.22088985 | 0.21671399 | 0.21161155 |
| 23 | 0.34087882 | 0.28930825 | 0.28820540 | 0.22394940 | 0.21409002 | 0.21081498 |
| 24 | 0.34237335 | 0.29020940 | 0.28896157 | 0.21301606 | 0.22055911 | 0.22177668 |

Tabla 8.15. Errores RMSE en la predicción del factor de capacidad del parque eólico (1ª parte de la tabla) para los diferentes modelos estudiados.

A las redes neuronales de tipo híbrido, entendido éste como la combinación en la misma red neuronal de un módulo con aprendizaje no supervisado y otro módulo de aprendizaje supervisado (red neuronal híbrida 1 y red neuronal híbrida 2), se les limitó la fase de aprendizaje no supervisado a 500 epochs (la del aprendizaje supervisado se mantuvo en 3000 epochs).

Los resultados obtenidos con las distintas redes neuronales se encuentran recogidos en las Tablas 8.15 y 8.16. Estos resultados son los valores de error RMSE en la predicción del factor de capacidad del parque eólico con horizontes de predicción entre 1 y 24 horas. También se incluyen los resultados obtenidos mediante los modelos persistente, auto-regresivo y “nueva referencia”. En dichas Tablas 8.15 y 8.16, se han indicado con color rojo los valores de aquellos sistemas que representan el mínimo valor del error RMSE para el horizonte de predicción correspondiente.

| Horizonte | Elman | PCA | RBF | Híbrida 1 | Híbrida 2 | Laguerre |
|-----------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 1 | 0.16439762 | 0.17999968 | 0.11672340 | 0.15227932 | 0.09994628 | 0.10125509 |
| 2 | 0.14160499 | 0.18531709 | 0.14791208 | 0.15615986 | 0.13445667 | 0.13494480 |
| 3 | 0.19054512 | 0.18408819 | 0.19829544 | 0.18079899 | 0.15438550 | 0.15878878 |
| 4 | 0.19502282 | 0.20381530 | 0.19718057 | 0.17363457 | 0.21347678 | 0.17581342 |
| 5 | 0.20655117 | 0.20374595 | 0.20808760 | 0.19801627 | 0.19292709 | 0.18949036 |
| 6 | 0.20503584 | 0.20266257 | 0.20600837 | 0.19220828 | 0.20552979 | 0.20877604 |
| 7 | 0.20007046 | 0.20750003 | 0.18599254 | 0.20956987 | 0.21374442 | 0.21864103 |
| 8 | 0.20861133 | 0.20759944 | 0.22971582 | 0.19929930 | 0.23366873 | 0.23192569 |
| 9 | 0.20527009 | 0.21241070 | 0.24154768 | 0.22387601 | 0.23727447 | 0.24137546 |
| 10 | 0.21387785 | 0.20864827 | 0.25336866 | 0.21174819 | 0.21591433 | 0.26115435 |
| 11 | 0.20476052 | 0.21223094 | 0.25198706 | 0.21765726 | 0.25176536 | 0.25649314 |
| 12 | 0.21379529 | 0.20946555 | 0.25792525 | 0.21955915 | 0.20353934 | 0.26508295 |
| 13 | 0.21934972 | 0.21215128 | 0.26366905 | 0.21173405 | 0.25207880 | 0.27293061 |
| 14 | 0.20686206 | 0.21211366 | 0.27369989 | 0.21097237 | 0.21447812 | 0.26953354 |
| 15 | 0.22092201 | 0.21413009 | 0.27225389 | 0.21051345 | 0.24049370 | 0.28418919 |
| 16 | 0.21369791 | 0.21447530 | 0.26666388 | 0.21752992 | 0.22103080 | 0.27572120 |
| 17 | 0.20847382 | 0.20929557 | 0.28654467 | 0.20596185 | 0.25507403 | 0.27620374 |
| 18 | 0.21100854 | 0.21435867 | 0.27729851 | 0.21075059 | 0.28664145 | 0.27946133 |
| 19 | 0.21251800 | 0.21305554 | 0.24910156 | 0.21464078 | 0.22802129 | 0.28610735 |
| 20 | 0.21938787 | 0.21755682 | 0.28108772 | 0.23058650 | 0.23572600 | 0.28534878 |
| 21 | 0.19870141 | 0.21608247 | 0.29829838 | 0.22117414 | 0.25848778 | 0.28808219 |
| 22 | 0.23754957 | 0.22091481 | 0.29131910 | 0.22621012 | 0.25236464 | 0.28996193 |
| 23 | 0.21485271 | 0.21057094 | 0.27484348 | 0.25582558 | 0.24186764 | 0.29079469 |
| 24 | 0.21209261 | 0.21404090 | 0.30318847 | 0.21559146 | 0.22664868 | 0.28942022 |

Tabla 8.16. Errores RMSE en la predicción del factor de capacidad del parque eólico (2ª parte de la tabla) para los diferentes modelos estudiados.

En el caso de tener que optar por un único modelo de predicción para todos los

horizontes de predicción, este sería aquél que ofreciese el menor error RMSE medio considerando los 24 horizontes de predicción posibles. La Tabla 8.17 representa el error RMSE medio para todos los modelos estudiados en este trabajo de investigación. La red neuronal de Elman es la que proporciona un menor valor del error RMSE medio, siendo su mejora con respecto a las predicciones del modelo persistente del 25.28% de media para todos los horizontes de predicción (de 1 a 24 horas). Como se observa en la Tabla 8.17, otras redes neuronales, que consiguen resultados cuya bondad es próxima a la de la red neuronal de Elman, son la red neuronal modular, la red neuronal híbrida 1 y la red neuronal PCA.

| Modelo | Error RMSE medio | mejora |
|------------------|-------------------------|---------------|
| Persistente | 0.27463556 | --- |
| AR(1) | 0.24327135 | 11.42% |
| Nueva referencia | 0.24289805 | 11.56% |
| MLP | 0.20873500 | 24.00% |
| FFBP | 0.20997192 | 23.55% |
| Modular | 0.20654889 | 24.79% |
| Elman | 0.20520664 | 25.28% |
| PCA | 0.20775957 | 24.35% |
| RBF | 0.24302971 | 11.51% |
| Híbrida 1 | 0.20692908 | 24.65% |
| Híbrida 2 | 0.21956424 | 20.05% |
| Laguerre | 0.24297900 | 11.53% |

Tabla 8.17. Errores RMSE medios en la predicción del factor de capacidad del parque eólico considerando los 24 horizontes de predicción posibles.

En las Figuras 8.14 y 8.15 están representados los errores RMSE en la predicción del factor de capacidad del parque eólico para los 24 horizontes de predicción con los 12 modelos de predicción analizados (corresponden a la representación gráfica de las tablas 8.15 y 8.16, respectivamente). En dichas figuras se puede apreciar claramente el menor error de predicción obtenido con los modelos basados en redes neuronales (excepto las redes neuronales RBF y de Laguerre) para horizontes de predicción relativamente elevados. También se observa claramente en las figuras que para horizontes de predicción relativamente pequeños (por debajo de las 6 horas), los modelos de predicción basados en la propia serie temporal de factor de capacidad del parque eólico (modelos de la primera familia según la clasificación establecida en el Capítulo 3) presentan menores errores de predicción que los modelos basados en valores de predicción de variables meteorológicas obtenidas de modelos atmosféricos (modelos de la segunda familia).

Es de resaltar que los modelos basados en redes neuronales no fueron entrenados con las predicciones de los modelos atmosféricos, sino con valores reales obtenidos de las series temporales almacenadas en los sistemas SCADA del parque eólico (velocidades y direcciones medias horarias del viento medidas en el propio parque eólico), utilizándose los valores de predicción del modelo atmosférico sólo en los datos del grupo de testeo.

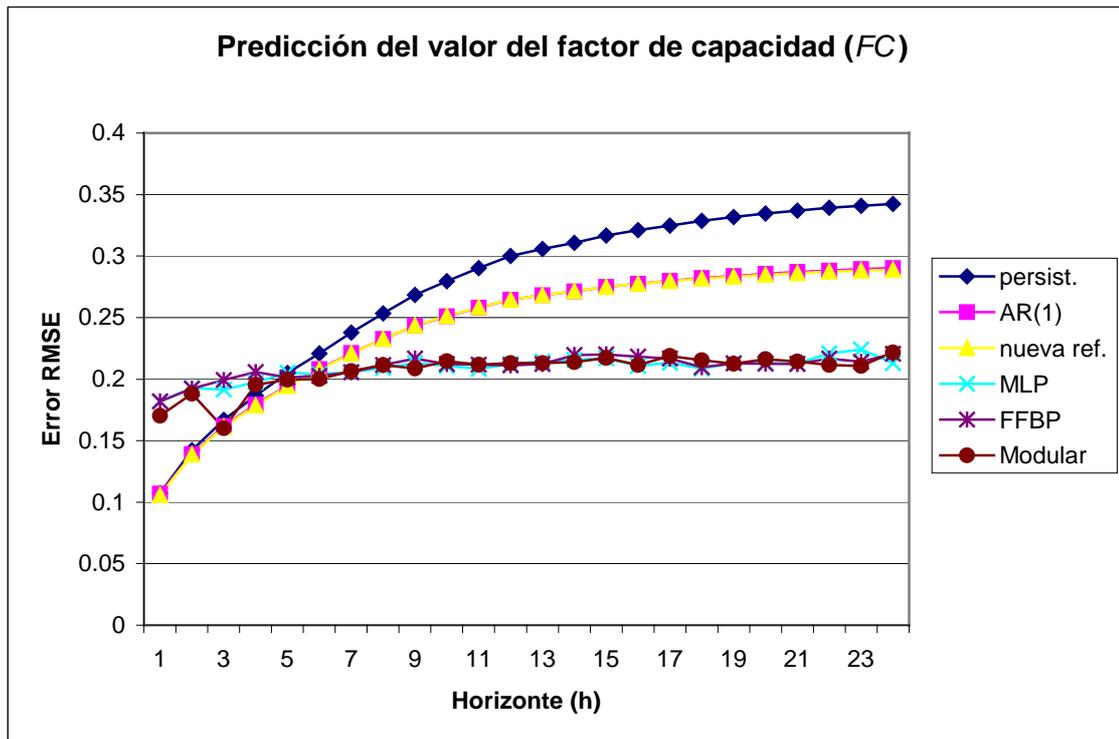


Figura 8.14. Errores RMSE en la predicción del factor de capacidad del parque eólico (1ª parte de los resultados) considerando los 24 horizontes de predicción posibles.

8.5. PREDICCIÓN A CORTO PLAZO DE LA POTENCIA ELÉCTRICA GENERADA EN UN PARQUE EÓLICO CON INFORMACIÓN COMPLETA DE UN MODELO ATMOSFÉRICO

El quinto problema de predicción planteado, dentro de los trabajos de investigación, es el de la predicción de la potencia media horaria generada a corto plazo en un parque eólico cuando están disponibles las predicciones completas de un modelo atmosférico. En el apartado 8.4 de este documento se presentó el mismo problema de predicción, pero con una información limitada en el tiempo del modelo atmosférico. En este caso, las predicciones del modelo atmosférico cubren la totalidad de un año, lo que permite utilizarlas en el desarrollo de modelos de predicción que requieran proceso de

aprendizaje, como el caso de los modelos basados en redes neuronales.

En el apartado 8.4 se comprobó que los modelos de predicción de la primera familia (según la clasificación de modelos de predicción de la producción eléctrica en un parque eólico realizada en el Capítulo 3 de este documento) ofrecían un menor error de predicción que los modelos de la segunda familia cuando el horizonte de predicción estaba por debajo de un límite de horas (6 horas para el caso analizado en el apartado 8.4). En este apartado se va a exponer la experiencia en el desarrollo de un modelo de predicción de la segunda familia, pero con una mejora significativa de los errores de predicción para horizontes en el muy corto plazo.

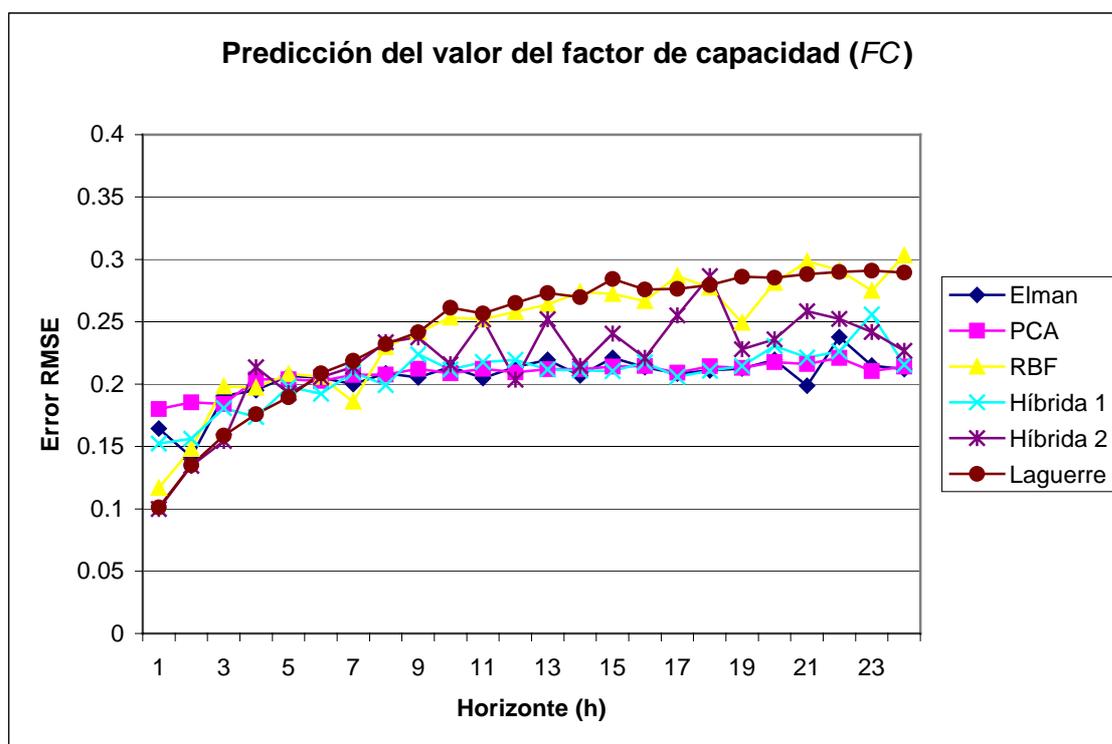


Figura 8.15. Errores RMSE en la predicción del factor de capacidad del parque eólico (2ª parte de los resultados) considerando los 24 horizontes de predicción posibles.

8.5.1. Datos disponibles.

Para el desarrollo del sistema de predicción se utilizaron el conjunto de datos designado en el Capítulo 7 como “Serie V”. Estos datos incluyen las predicciones de valores medios de velocidad y dirección del viento, temperatura y presión atmosférica a 10 metros de la superficie del terreno, obtenidas con el modelo MM5 para horizontes de predicción de 0.5 a 72 horas en intervalos de 0.5 horas. También están disponibles los

valores de potencia eléctrica media entregada en la subestación cada media hora.

La “Serie V” abarca un periodo de un año, si bien no está completa. Faltan datos de días intermedios, siendo el intervalo de tiempo mayor en el que no se presentan vacíos importantes (vacíos no mayores a 3 horas) en las series temporales, el de 149 días (exactamente 7156 registros medio-horarios). Para completar los valores ausentes en el intervalo escogido (un total de 32 “huecos” con duraciones entre media y 3 horas) se utilizó la interpolación lineal entre los datos conocidos en ambos extremos de cada hueco. Se optó por limitar los datos a esos 149 días al objeto de poder desarrollar modelos de predicción basados en valores previos de las variables, como modelos auto-regresivos o modelos basados en redes neuronales recurrentes, y no limitarse a una mera aproximación de funciones no lineales.

Una característica especial de estos datos es la disponibilidad de la serie temporal de la potencia entregada en la subestación que conecta el parque eólico a la red eléctrica. Esta serie, en palabras de los gestores del parque, es mucho más fiable que la serie temporal de potencia eléctrica media generada en el parque (obtenible a partir de la potencia generada por cada aerogenerador) en cuanto a la veracidad de sus registros. Curiosamente esta serie temporal (la de potencia media entregada en la subestación) presenta datos con valores negativos, consecuencia del consumo de las instalaciones del propio parque cuando la generación en él es nula o insuficiente para cubrir su propia demanda. La Figura 8.16 representa la potencia media entregada en la subestación para la primera semana del intervalo de 149 días considerado (semana del 12 al 18 de marzo de 2005).

Para el desarrollo de los modelos de predicción, los datos de los 149 días (7156 registros de media hora de duración) se dividieron en tres grupos, siguiendo el patrón establecido con los cuatro casos de predicción expuestos en las anteriores páginas de este capítulo. Los grupos son los de entrenamiento, con la mayor parte de los datos, y los de validación y testeo, de dimensiones más reducidas. La Tabla 8.18 muestra las características estadísticas de cada uno de estos grupos referidas a la potencia media entregada en la subestación (unidades kW). El “espaciamiento” temporal entre datos es de media hora, esto es, el intervalo entre dato y dato de las series temporales es 30 minutos.

Las predicciones del modelo atmosférico disponibles en el conjunto de datos se refieren a predicciones realizadas con el modelo atmosférico MM5. Estas predicciones corresponden a los valores de las variables meteorológicas de velocidad y dirección

media del viento, temperatura y presión atmosférica, todos para un punto situado en el centro del parque eólico. Se dispone de predicciones realizadas todos los días del año, cubriendo, cada una ellas, tres días completos (desde las 0 horas del primer día, hasta las 24:00 del tercer día) con valores de predicción para cada media hora. Las predicciones están realizadas con los datos obtenidos de otro modelo atmosférico de escala global con la asimilación de datos atmosféricos correspondiente a las 0 horas GMT de cada día.



Figura 8.16. Potencia media generada en el parque eólico y entregada en la subestación durante la semana del 12 al 18 de marzo de 2005.

| | Entrenamiento | Validación | Testeo |
|--------------------------|----------------------|-------------------|---------------|
| Media (kW) | 5706.64794 | 3080.96964 | 4177.11387 |
| Error típico (kW) | 91.9692113 | 112.791172 | 126.690498 |
| Mediana (kW) | 3219.66663 | 1330.83331 | 2506.41675 |
| Moda (kW) | -65 | -61.333302 | 0 |
| Desviación estándar (kW) | 6069.96794 | 4220.25921 | 4740.32439 |
| Rango (kW) | 21686.834 | 21225.6667 | 21329.0007 |
| Mínimo (kW) | -73.5 | -68.166702 | -77.666702 |
| Máximo (kW) | 21613.334 | 21157.5 | 21251.334 |
| Número de medias horas | 4356 | 1400 | 1400 |

Tabla 8.18. Características de la potencia eléctrica media horaria en la subestación de los tres grupos de datos para desarrollar los modelos de predicción.

La disponibilidad, dentro de este conjunto de datos fiables, de predicciones de un modelo atmosférico con un horizonte de predicción de 72 horas, posibilita el desarrollo de modelos de predicción con ese mismo horizonte de predicción. Así, el

objetivo de los modelos desarrollados pasó a ser la predicción de la potencia eléctrica media entregada por el parque eólico en la subestación, con un horizonte máximo de predicción de hasta 72 horas, y con predicciones para intervalos de 30 minutos.

8.5.2. Modelos de predicción analizados.

Los modelos analizados para el desarrollo del sistema de predicción incluían el clásico modelo persistente, un modelo de la familia ARIMA, modelos basados en redes neuronales artificiales y modelos basados en sistemas de inferencia difusa. Los modelos basados en redes neuronales artificiales eran similares a los desarrollados para el anterior problema de predicción expuesto en el apartado 8.4 de este documento. Los modelos analizados fueron los indicados seguidamente.

1. *Modelo persistente*: la potencia eléctrica media entregada en la subestación para el horizonte de predicción k , $p(t+k)$, era igual a la de la hora actual, $p(t)$.
2. *Modelo ARIMA*: con la metodología Box-Jenkins [Box76] se determinó el modelo ARIMA que mejores predicciones ofrecía. Para el ajuste del modelo se utilizaron los datos de los grupos de entrenamiento y de validación.
3. *Modelo “nueva referencia”*: una vez más se utilizó este modelo como modelo de predicción. Los dos parámetros del modelo (coeficiente de correlación y media) se determinaron con los datos de los grupos de entrenamiento y validación.
4. *Red neuronal perceptrón multicapa (MLP)*: se escogió una red neuronal MLP con dos capas ocultas. La estructura de esta red neuronal es la reflejada en la Figura 8.17, en la que las capas a la izquierda y derecha de la figura corresponde a la capa de entradas y a la capa de salidas, respectivamente. Las capas ocultas son las dos capas intermedias. Cada capa sólo está unida (conexiones entre neuronas) a la capa precedente y a la capa que le sigue (si existen).
5. *Red neuronal multicapa generalizada (FFBP)*: se escogió una red neuronal FFBP también con dos capas ocultas. La estructura de esta red es la reflejada en la Figura 8.18, en la que puede apreciar que todas las capas están unidas directamente entre sí.

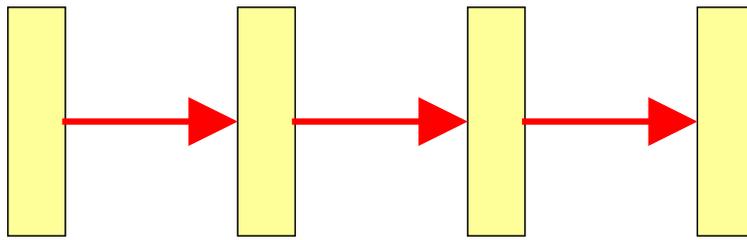


Figura 8.17. Estructura de la red neuronal MLP de dos capas ocultas.

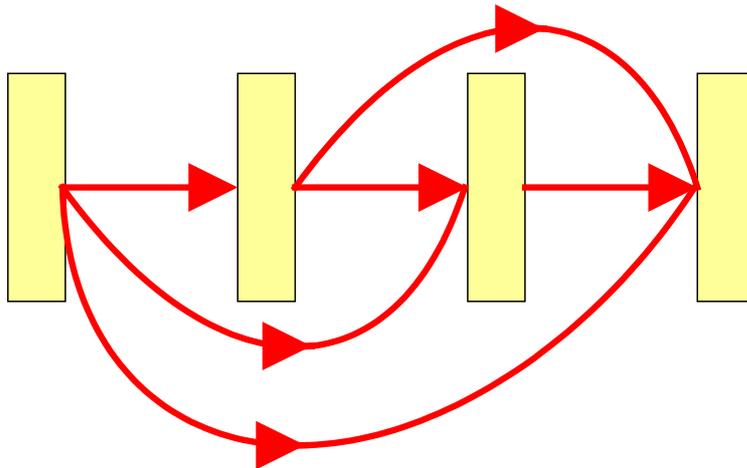


Figura 8.18. Estructura de la red neuronal FFBP de dos capas ocultas.

6. *Red neuronal modular:* se escogió una red neuronal modular con una estructura igual a la reflejada en la Figura 8.12, es decir, con dos pares de capas ocultas.
7. *Red neuronal de Elman:* se escogió una red neuronal de Elman con dos capas ocultas. La capa de neuronas de contexto copia el contenido de la segunda capa oculta y las salidas de esta capa de neuronas de contexto actúan como entradas adicionales de la primera capa oculta. La Figura 8.19 muestra la estructura de la red neuronal de Elman con las dos capas ocultas. Las capas a la izquierda y derecha son las capas de entradas y salidas, respectivamente; la inferior es la capa de neuronas de contexto.
8. *Red neuronal de función de base radial (RBF):* se escogió una red neuronal híbrida compuesta por una red neuronal RBF seguida por una red neuronal MLP con dos capas ocultas. La de neuronas de función de base radial actúa como la capa de entradas de la red neuronal MLP. Ambas redes neuronales se entrenaron separadamente: la red neuronal RBF se entrenó con

aprendizaje no supervisado con un máximo de 800 epochs, con la posibilidad de que fuera menor el número de epochs del entrenamiento no supervisado si el cambio en todos los pesos de las conexiones de las neuronas de la capa de neuronas de función de base radial era menor a 0.0001. La red neuronal MLP se entrenó, posteriormente, con aprendizaje supervisado con un máximo de 3000 epochs.

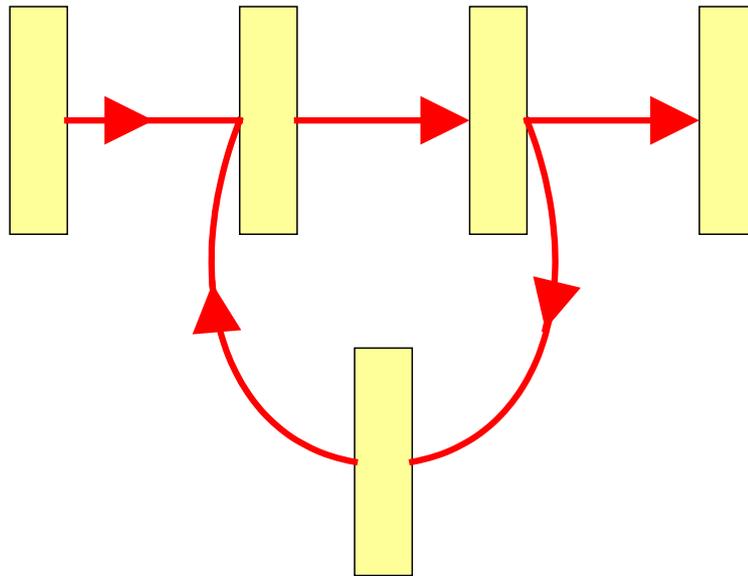


Figura 8.19. Estructura de la red neuronal de Elman con dos capas ocultas.

9. *Red neuronal con análisis de componentes principales (PCA):* la red neuronal PCA escogida constaba de dos capas ocultas para el aprendizaje supervisado. El número de componentes principales era escogido por el proceso de optimización. Para el aprendizaje no supervisado el número de epochs máximo fue de 800, existiendo la posibilidad de ser interrumpido antes con el mismo criterio expuesto en el párrafo anterior.
10. *Red neuronal híbrida con módulo de aprendizaje no supervisado y módulo de aprendizaje supervisado:* tras analizar los resultados de la experiencia anterior, expuesta en el apartado 8.4 de este documento, en la que el comportamiento de la red híbrida de mapa auto-organizado y red neuronal MLP fue mejor cuando el mapa auto-organizado tenía una estructura lineal, se optó por desarrollar aquí solamente redes neuronales con esa estructura. En concreto, el mapa auto-organizado se limitó a una capa de 20 neuronas en línea. Este mapa auto-organizado servía de capa de entradas a una red neuronal MLP con dos capas ocultas. Para el aprendizaje no supervisado el

número de epochs máximo fue de 800, con la posibilidad de interrupción anticipada de este aprendizaje si el cambio en todos los pesos de las conexiones del mapa auto-organizado era menor a 0.0001.

11. *Red neuronal retardada en el tiempo*: se desarrollaron 6 tipos de redes neuronales retardadas en el tiempo. Concretamente se desarrollaron redes neuronales retardadas en el tiempo del tipo TDNN, del tipo Gamma y del tipo Laguerre. A su vez cada una de ellas podía ser focalizada (sólo unidades de retardo en la capa de entradas) y no focalizada (unidades de retardo en todas las capas, excepto la de salidas). Todas las redes neuronales tenían dos capas ocultas y el proceso de optimización podía escoger el número de unidades de retardo de cada red.
12. *Sistema de inferencia difusa*: se desarrollaron tres tipos de sistemas de inferencia difusa (FIS), un sistema con partición del espacio de entrada en rejilla con 2 funciones de pertenencia por entrada, otro sistema similar pero con tres funciones de entrada por entrada, y un tercer sistema con partición del espacio de entrada por agrupamiento sustractivo.

Para el desarrollo del sistema de inferencia difusa con partición del espacio de entrada por agrupamiento substractivo se utilizó el proceso de optimización basado en algoritmo genético descrito en el Capítulo 6. Las funciones de pertenencia utilizadas fueron gaussianas. Los centros de los clusters (y, por tanto, de las familias de pertenencia), la varianza y los coeficientes de los hiperplanos que definían la parte consecuente de las reglas del sistema fueron optimizados en el proceso. La optimización se realizó con 100 generaciones de 100 individuos cada generación. Se aplicó elitismo (2% de cada generación) y el resto de individuos de cada generación se seleccionó por cruce (el 80%) y por mutación (el 20%, con una probabilidad de mutación de cada dígito del 2%). Se utilizaron los datos de los grupos de entrenamiento y validación en el desarrollo de los modelos.

Todos los modelos de redes neuronales desarrollados presentaban una serie de características comunes indicadas a continuación.

- La primera de estas características es que se desarrollaron tres redes neuronales de cada tipo para cubrir todo el horizonte de predicción: la primera con un horizonte de predicción de 0.5 a 24 horas; la segunda red

neuronal con un horizonte de predicción de 24.5 a 48 horas; y la tercera red neuronal con un horizonte de predicción de 48.5 a 72 horas.

La razón que justifica esta decisión es la cantidad de datos disponibles: Si desarrollásemos una red neuronal diferente por horizonte de predicción, los datos disponibles para el entrenamiento resultan ser muy limitados (concretamente 91 datos para el grupo de entrenamiento, 29 datos para el grupo de validación y otros 29 datos para el grupo de testeo), incrementándose la probabilidad de sobre-entrenamiento. El agrupar los horizontes de predicción en tres redes neuronales hace que el tamaño de los grupos de datos se incremente, reduciéndose la probabilidad de sobre-entrenamiento.

- La segunda característica común era que todas las redes neuronales desarrolladas contaban con dos capas ocultas. Se tomó esta decisión al considerar que el problema de predicción era más complejo que el expuesto en el apartado 8.4, en el que se desarrollaba un modelo de red neuronal con una sola capa oculta (excepto para la red neuronal modular) por horizonte de predicción. En este caso cada red neuronal presentaba más de un horizonte de predicción.
- La tercera característica común es que todos los modelos de predicción fueron optimizados en un proceso controlado por un algoritmo genético, de igual manera que se ha hecho en el resto de los modelos basados en redes neuronales recogidos en este documento. El número de individuos por generación se estableció en 50, al igual que el número de generaciones. La selección de individuos para la siguiente generación se realizó por ruleta ordenada por rango. La probabilidad de cruce se estableció en un 90% y la de mutación en el 1%. La función de activación para todas las neuronas fue la tangente hiperbólica. Todos los valores de las entradas fueron previamente normalizados entre los límites de -0.9 y 0.9. El sobre-entrenamiento fue evitado utilizando el grupo de validación. La no disminución del error con los datos del grupo de validación tras 100 epochs en el proceso de entrenamiento lo interrumpía, dejando la red neuronal con los valores de los parámetros (pesos de las conexiones) que proporcionaban el menor error con los datos del grupo de validación. El límite máximo de epochs se estableció en 3000.

8.5.3. Resultados preliminares obtenidos.

Tras el proceso de ajuste o entrenamiento de los modelos, y la aplicación de los mismos en la predicción de la potencia eléctrica media entregada en la subestación a intervalos de media hora, se obtuvieron los resultados que se van a pormenorizar en los próximos párrafos.

El modelo persistente se aplicó directamente sobre los datos del grupo de testeo, considerando los diferentes horizontes de predicción (de 0.5 a 72 horas).

El modelo ARIMA se diseñó a partir de los datos de la potencia media entregada en la subestación, correspondientes a los grupos de entrenamiento y validación. Se diseñaron 6 modelos, incluyendo modelos auto-regresivos puros (con ordenes de 1 a 3) y modelos auto-regresivos de media móvil (el de mayor orden fue el modelo ARMA(2, 2)) sin considerar la estacionalidad; y además, se construyeron otros 6 modelos similares considerando la estacionalidad (con periodo de estacionalidad de 48 muestras, es decir, 24 horas). Una vez diseñados y evaluados los 12 modelos, se escogió aquél que proporcionaba un valor menor del criterio de información de Akaike (AIC) con los datos utilizados en el ajuste del modelo (coincidía en ser también el modelo con menor error RMSE, y las gráficas de auto-correlación y auto-correlación parcial también sugerían que el modelo de predicción debería tener su misma estructura). Este modelo fue el ARMA(1,1). La ecuación (8.26) proporciona el valor de predicción, siendo $p(t)$ la potencia entregada en la subestación en el instante t , $e(t)$ el error en la predicción de la potencia eléctrica entregada para el instante t y donde los valores de los coeficientes ϕ_1 y θ_1 son 0.9669 y 0.2399, respectivamente.

$$\hat{p}(t + 0.5) = 166.575 + \phi_1 p(t) - \theta_1 e(t) \quad (8.26)$$

La ecuación (8.26) permite la predicción de la potencia entregada para la siguiente media hora (con horizonte de predicción de 0.5 horas). Para la predicción con horizontes mayores, sólo hay que aplicar de forma iterativa la ecuación anterior, obteniendo sucesivamente las predicciones para horizontes “intermedios” hasta alcanzar el horizonte deseado. El término de error en la predicción de la ecuación (8.26) para estos horizontes “intermedios” se considera nulo.

Los parámetros del modelo “nueva referencia” fueron calculados con los datos

de los grupos de entrenamiento y validación para cada uno de los horizontes de predicción, y utilizados posteriormente en la predicción de la potencia eléctrica media entregada en la subestación para los datos del grupo de testeo.

En el desarrollo de los modelos basados en redes neuronales artificiales se permitió que el proceso de optimización (con algoritmo genético) pudiera elegir las variables de entrada de las redes neuronales, con la condición de que siempre figurase como una de las entradas seleccionadas, la predicción de la velocidad del viento (referida al horizonte de predicción) obtenida del modelo atmosférico. Para el diseño de cada red neuronal (individuo de una generación para el proceso de optimización), era posible escoger entre diversas variables de entrada. Dichas variables podían ser las indicadas a continuación.

- Predicciones de la velocidad media del viento del modelo atmosférico para los instantes 30, 60, 90, 120, 150 y 180 minutos anteriores al del horizonte de predicción.
- Coseno del ángulo que representa la predicción de la dirección media del viento del modelo atmosférico para los instantes 0, 30 y 60 minutos anteriores al del horizonte de predicción (evidentemente la predicción 0 minutos anterior a la del horizonte de predicción no es más que la predicción para ese horizonte de predicción).
- Seno del ángulo que representa la predicción de la dirección media del viento del modelo atmosférico para los instantes 0, 30 y 60 minutos anteriores al del horizonte de predicción.
- Predicción de la temperatura del modelo atmosférico para los instantes 0, 30 y 60 minutos anteriores al del horizonte de predicción.
- Predicciones de la presión atmosférica del modelo atmosférico para los instantes 0, 30 y 60 minutos anteriores al del horizonte de predicción.
- Antigüedad de la predicción del modelo atmosférico: un número que representaba el número de periodos de 30 minutos que tenían de antigüedad las predicciones del modelo atmosférico. Por ejemplo, si el horizonte de predicción corresponde a una predicción para las 7:00 GMT, el valor de la variable antigüedad es 14, ya que el modelo atmosférico realizó sus predicciones a partir

del estado de la atmósfera a las 0:00 horas GMT.

Para las redes neuronales con retardo en el tiempo (o retardadas en el tiempo) y redes neuronales recurrentes se eliminaron todas las posibles entradas correspondientes a predicciones para instantes previos al del horizonte de predicción, ya que estas redes almacenan por sí mismas los valores previos de las variables de entrada o de los valores de las neuronas de capas intermedias correspondientes a esos valores de las variables de entrada.

Como ya se ha comentado anteriormente, se desarrollaron modelos distintos según el horizonte de predicción. El primer grupo de modelos de predicción fue desarrollado considerando horizontes de predicción entre 0.5 y 24 horas. El segundo grupo de modelos fue desarrollado considerando horizontes de predicción entre 24.5 y 48 horas. El tercer grupo de modelos fue desarrollado considerando horizontes de predicción entre 48.5 y 72 horas. Por esta razón se van a presentar los resultados para cada uno de los tres grupos separadamente.

8.5.2.1. Resultados para horizontes de predicción de 0.5 a 24 horas.

La Tabla 8.19 representa los errores RMSE en la predicción de la potencia eléctrica media entregada en la subestación y correspondiente a la generada en el parque eólico, con horizontes de predicción entre 0.5 y 24 horas, para los datos del grupo de testeo. Los errores están expresados en kW y en porcentaje con respecto a la potencia instalada en el parque. También aparece en la Tabla 8.19 la mejora porcentual de la predicción de cada modelo con respecto a la predicción con el modelo persistente.

Según se puede observar en la Tabla 8.19, el modelo de predicción que exhibe un mejor comportamiento en la predicción es la red neuronal de Elman, al igual que ocurría en la experiencia descrita en el apartado 8.4 de este documento. El comportamiento de la red neuronal de Elman con los datos de los grupos de entrenamiento y validación (los usados en el entrenamiento del modelo) también era superior a la mayoría de los otros modelos estudiados. Concretamente, la red neuronal de Elman era la segunda red con menor error RMSE con los datos de entrenamiento, sólo superada y por escaso margen por la red neuronal con retardo en el tiempo Gamma focalizada.

Las características de la red de Elman que proporciona los errores de predicción

más bajos son las indicadas seguidamente.

- Entradas utilizadas: valores de predicción de la velocidad media del viento, y del seno y del coseno de la dirección media del viento, suministradas por el modelo atmosférico para el horizonte de predicción. Al tratarse de una red neuronal recurrente no se le suministraron como entradas posibles los valores de esas variables para instantes anteriores al horizonte de predicción.
- Seis neuronas en la primera capa oculta y 11 neuronas en la segunda capa oculta. La capa de neuronas de contexto copia la segunda capa oculta. Las salidas de las neuronas de la capa de contexto actúan como entradas adicionales de las neuronas de la primera capa oculta.
- La constante de tiempo en la capa de neuronas de contexto es de 0.8.

| Modelo | Error RMSE (kW) | Error RMSE (%) | mejora |
|------------------------|-----------------|----------------|--------|
| Persistente | 4825.416 | 22.34 | --- |
| ARMA(1,1) | 4442.834 | 20.57 | 7.93% |
| Nueva referencia | 4264.820 | 20.40 | 11.62% |
| MLP | 3615.150 | 16.74 | 25.08% |
| FFBP | 3548.231 | 16.43 | 26.47% |
| Modular | 3478.412 | 16.10 | 27.91% |
| TDNN focalizada | 3682.085 | 17.05 | 23.69% |
| TDNN no focalizada | 3881.480 | 17.97 | 19.56% |
| Gamma focalizada | 3943.071 | 18.25 | 18.29% |
| Gamma no focalizada | 3733.120 | 17.28 | 22.64% |
| Laguerre focalizada | 3540.601 | 16.39 | 26.63% |
| Laguerre no focalizada | 3837.105 | 17.76 | 20.48% |
| Elman | 3441.660 | 15.93 | 28.68% |
| Hibrida (SOM lineal) | 3666.590 | 16.97 | 24.02% |
| PCA | 3502.229 | 16.21 | 27.42% |
| RBF | 3485.931 | 16.14 | 27.76% |
| Recurrente TDNN | 3748.951 | 17.36 | 22.31% |
| FIS rejilla 2 mfs | 3554.909 | 16.46 | 26.33% |
| FIS rejilla 3 mfs | 3621.660 | 16.77 | 24.95% |
| FIS agrupamiento | 3570.568 | 16.53 | 26.00% |

Tabla 8.19. Errores RMSE en la predicción de la potencia eléctrica media entregada en la subestación y mejora con respecto al modelo persistente para horizontes de predicción de 0.5 a 24 horas.

En la Figura 8.20 están representados los valores reales de la potencia eléctrica media entregada en la subestación y el valor de predicción obtenido con el modelo de la red neuronal de Elman, para los últimos 10 días del grupo de testeo. Las predicciones del modelo corresponden a horizontes de predicción entre las 0.5 horas y las 24 horas:

las predicciones para las 00:30 horas de cada día tienen un horizonte de predicción de 0.5 horas, mientras que las predicciones para las 00:00 horas de cada día tienen un horizonte de predicción de 24 horas.

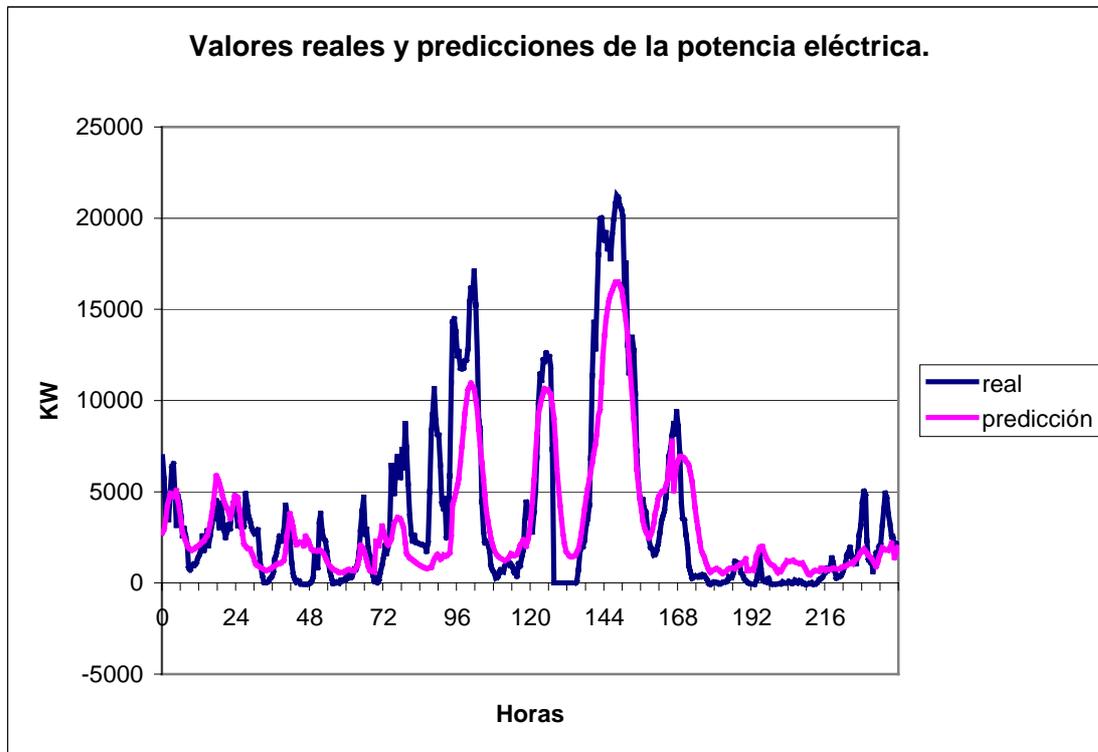


Figura 8.20. Valores reales de potencia eléctrica media entregada en la subestación y predicción obtenida con la red neuronal de Elman para horizontes de predicción entre 0.5 y 24 horas.

8.5.2.2. Resultados para horizontes de predicción de 24.5 a 48 horas.

La Tabla 8.20 representa los errores RMSE en la predicción de la potencia eléctrica media entregada en la subestación y correspondiente a la generada en el parque eólico, con horizontes de predicción entre 24.5 y 48 horas, para los datos del grupo de testeo. Los errores están expresados en kW y en porcentaje con respecto a la potencia instalada en el parque. También aparece en la Tabla 8.20 la mejora porcentual de la predicción de cada modelo con respecto a la predicción obtenida con el modelo persistente.

Una vez más la red neuronal de Elman se perfila como el mejor modelo de predicción de entre todos los analizados. Esta red neuronal de Elman es ligeramente diferente de la red que dio mejores resultados en la predicción con horizontes de 0.5 a 24 horas. Como consecuencia del proceso de optimización, alguno de los parámetros ha

cambiado ligeramente. Las características de esta red neuronal de Elman que da los mejores resultados en la predicción con horizontes de 24.5 a 48 horas son las indicadas a continuación.

| Modelo | Error RMSE (kW) | Error RMSE (%) | mejora |
|------------------------|-----------------|----------------|---------------|
| Persistente | 7102.07 | 32.88 | --- |
| ARMA(1,1) | 4863.25 | 22.52 | 31.53% |
| Nueva referencia | 4976.98 | 23.04 | 29.92% |
| MLP | 3621.67 | 16.77 | 49.01% |
| FFBP | 3706.21 | 17.16 | 47.82% |
| Modular | 3861.36 | 17.88 | 45.63% |
| TDNN focalizada | 3737.84 | 17.30 | 47.37% |
| TDNN no focalizada | 4049.69 | 18.75 | 42.98% |
| Gamma focalizada | 3915.12 | 18.13 | 44.87% |
| Gamma no focalizada | 4055.98 | 18.78 | 42.89% |
| Laguerre focalizada | 3773.43 | 17.47 | 46.87% |
| Laguerre no focalizada | 3805.67 | 17.62 | 46.41% |
| Elman | 3566.53 | 16.51 | 49.78% |
| Hibrida (SOM lineal) | 3878.10 | 17.95 | 45.39% |
| PCA | 3851.35 | 17.83 | 45.77% |
| RBF | 3872.02 | 17.93 | 45.48% |
| Recurrente TDNN | 3882.34 | 17.97 | 45.34% |
| FIS rejilla 2 mfs | 3927.08 | 18.18 | 44.71% |
| FIS rejilla 3 mfs | 4330.57 | 20.05 | 39.02% |
| FIS agrupamiento | 4006.26 | 18.55 | 43.59% |

Tabla 8.20. Errores RMSE en la predicción de la potencia eléctrica media entregada en la subestación y mejora con respecto al modelo persistente para horizontes de predicción de 24.5 a 48 horas.

- Entradas utilizadas: las seis variables de entrada posibles, es decir, valores de predicción de la velocidad media del viento, y del seno y del coseno de la dirección media del viento, temperatura, presión (suministradas todas por el modelo atmosférico para el horizonte de predicción) y antigüedad de las predicciones del modelo atmosférico (un número entre 49 y 98). Al tratarse de una red neuronal recurrente no se le suministraron como entradas posibles los valores de esas variables para instantes anteriores al horizonte de predicción. De todas formas, el análisis de sensibilidad de las entradas reveló que las tres últimas entradas son prácticamente despreciables frente a las tres primeras (valores de predicción de velocidad media del viento, y seno y coseno de su dirección media).
- 56 neuronas en la primera capa oculta, 6 neuronas en la segunda capa oculta y 6 neuronas en la capa de neuronas de contexto (copia de la segunda capa oculta).

- La constante de tiempo en la capa de neuronas de contexto es de 0.8.

En la Figura 8.21 están representados los valores reales de la potencia eléctrica media entregada en la subestación y el valor de predicción obtenido con el modelo de la red neuronal de Elman, para los últimos 10 días del grupo de testeo. Las predicciones del modelo corresponden a horizontes de predicción entre las 24.5 horas y las 48 horas: las predicciones para las 00:30 horas de cada día tienen un horizonte de predicción de 24.5 horas, mientras que las predicciones para las 00:00 horas de cada día tienen un horizonte de predicción de 48 horas.

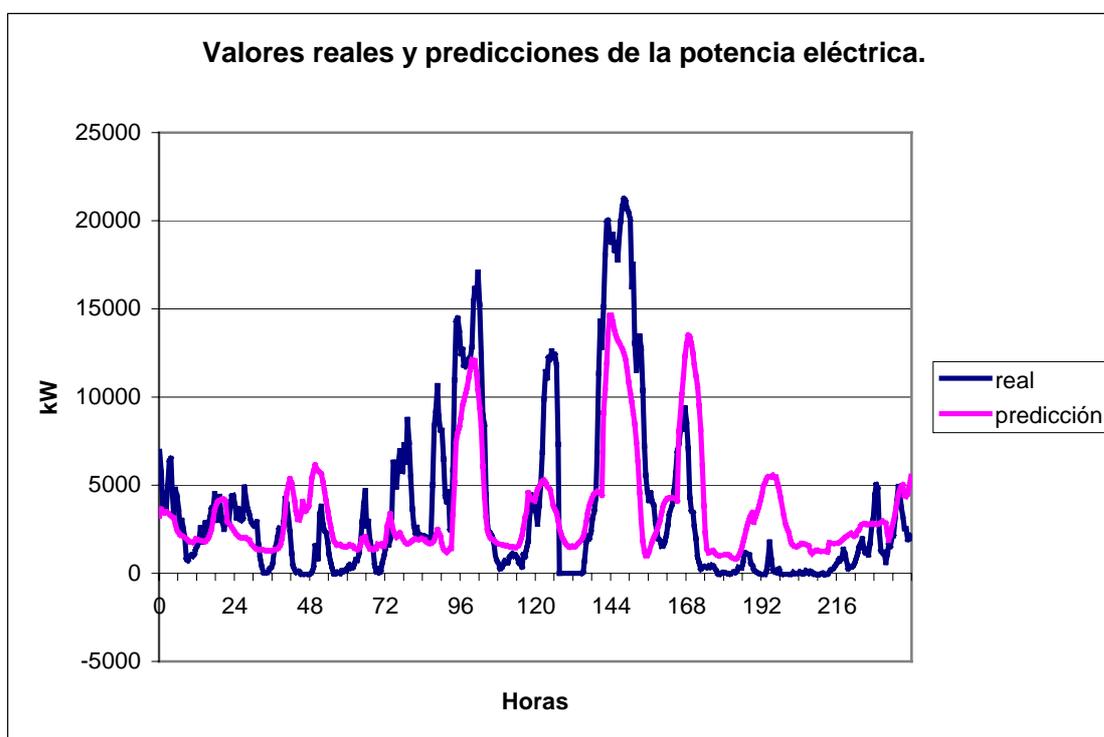


Figura 8.21. Valores reales de potencia eléctrica media entregada en la subestación y predicción obtenida con la red neuronal de Elman para horizontes de predicción entre 24.5 y 48 horas.

8.5.2.3. Resultados para horizontes de predicción de 48.5 a 72 horas.

La Tabla 8.21 representa los errores RMSE en la predicción de la potencia eléctrica media entregada en la subestación y correspondiente a la generada en el parque eólico, con horizontes de predicción entre 48.5 y 72 horas, para los datos del grupo de testeo. Los errores están expresados en kW y en porcentaje con respecto a la potencia instalada en el parque. También está representada en la Tabla 8.21 la mejora de la predicción del modelo con respecto a la predicción obtenida con el modelo persistente.

La red neuronal que proporciona un menor de predicción con los datos del grupo de testeo es una red neuronal modular con las características indicadas seguidamente.

- Entradas: los valores de predicción de la velocidad media del viento para el horizonte de predicción y los valores para los 30, 60, 120, 150 y 180 minutos anteriores; el valor de predicción de la temperatura para el horizonte de predicción y para 30 y 60 minutos antes; la presión atmosférica 30 minutos antes del horizonte de predicción; y la antigüedad. La entrada con mayor sensibilidad es la de predicción de la velocidad media el viento, seguida por el valor de predicción de la temperatura, ambas para el horizonte de predicción.
- Número de neuronas en capas ocultas: En la red superior, 41 neuronas en la primera capa y 3 en la segunda capa. En la red inferior, 38 neuronas en la primera capa y 4 en la segunda capa.

| Modelo | Error RMSE (kW) | Error RMSE (%) | mejora |
|------------------------|-----------------|----------------|--------|
| Persistente | 7231.13 | 33.48 | --- |
| ARMA(1,1) | 4921.15 | 22.78 | 31.94% |
| Nueva referencia | 4977.50 | 23.04 | 31.17% |
| MLP | 4424.30 | 20.48 | 38.82% |
| FFBP | 4376.66 | 20.26 | 39.47% |
| Modular | 4311.63 | 19.96 | 40.37% |
| TDNN focalizada | 4647.99 | 21.52 | 35.72% |
| TDNN no focalizada | 4562.91 | 21.12 | 36.90% |
| Gamma focalizada | 4524.91 | 20.95 | 37.42% |
| Gamma no focalizada | 4830.98 | 22.37 | 33.19% |
| Laguerre focalizada | 4505.28 | 20.86 | 37.70% |
| Laguerre no focalizada | 4326.18 | 20.03 | 40.17% |
| Elman | 4566.53 | 21.14 | 36.85% |
| Hibrida (SOM lineal) | 4465.45 | 20.67 | 38.25% |
| PCA | 4368.22 | 20.22 | 39.59% |
| RBF | 4538.92 | 21.01 | 37.23% |
| Recurrente TDNN | 4485.47 | 20.77 | 37.97% |
| FIS rejilla 2 mfs | 4416.33 | 20.45 | 38.93% |
| FIS rejilla 3 mfs | 4477.43 | 20.73 | 38.08% |
| FIS agrupamiento | 4485.63 | 20.77 | 37.97% |

Tabla 8.21. Errores RMSE en la predicción de la potencia eléctrica media entregada en la subestación y mejora con respecto al modelo persistente para horizontes de predicción de 48.5 a 72 horas.

En la Figura 8.22 están representados los valores reales de la potencia eléctrica media entregada en la subestación y el valor de predicción obtenido con el modelo de la red neuronal modular que mejores resultados ha ofrecido, para los últimos 10 días del

grupo de testeo. Las predicciones del modelo corresponden a horizontes de predicción entre las 48.5 horas y las 72 horas: las predicciones para las 00:30 horas de cada día tienen un horizonte de predicción de 48.5 horas, mientras que las predicciones para las 00:00 horas de cada día tienen un horizonte de predicción de 72 horas.

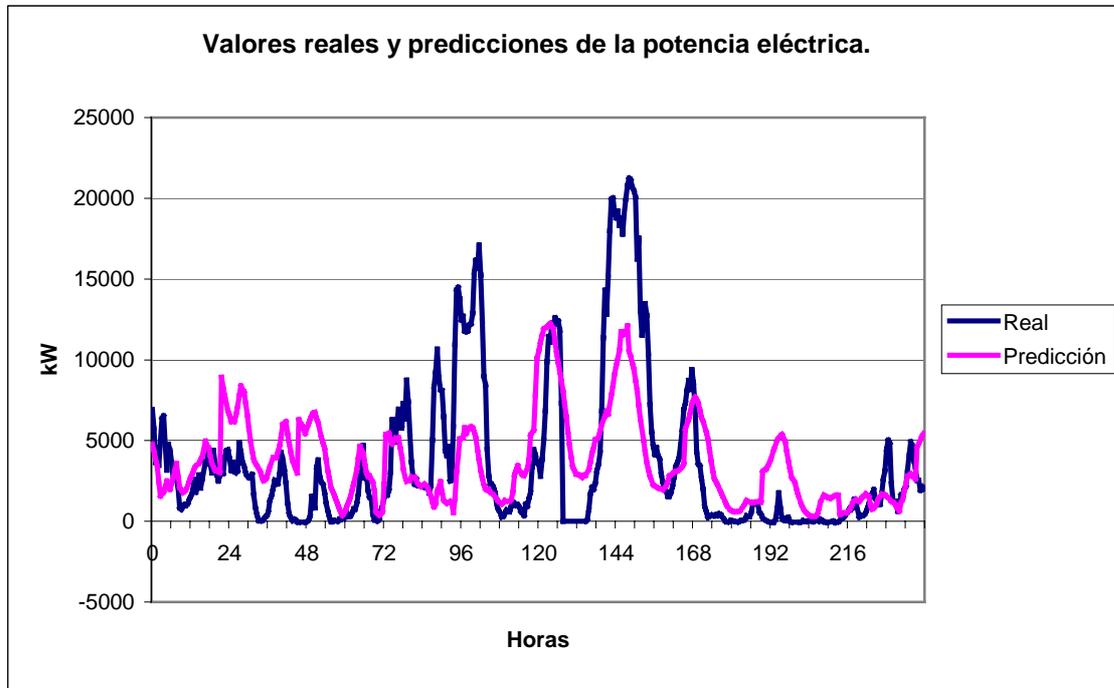


Figura 8.22. Valores reales de potencia eléctrica media entregada en la subestación y predicción obtenida con la red neuronal Modular para horizontes de predicción entre 48.5 y 72 horas.

8.5.2.4. Resultados para horizontes de predicción de 0.5 a 72 horas.

Con el objeto de apreciar gráficamente los resultados obtenidos para todos los horizontes de predicción considerados, en la Figura 8.23 se representa el error RMSE correspondiente a todos los horizontes de predicción (de 0.5 a 72 horas) para el modelo con mejor comportamiento en los tres días. Las predicciones con horizonte entre 0.5 y 24 horas están realizadas utilizando la primera red neuronal de Elman (apartado 8.5.2.1 de este documento). Las predicciones con horizonte entre 24.5 y 48 horas están realizadas usando la segunda red neuronal de Elman (apartado 8.5.2.2 de este documento). Las predicciones con horizontes entre 48.5 y 72 horas están realizadas con la red neuronal modular (apartado 8.5.2.3 de este documento). Para comparar los errores obtenidos con los modelos seleccionados, se han representado también en la Figura 8.23 los errores RMSE en la predicción de la potencia media entregada en la subestación obtenidos usando el modelo persistente.

El error máximo de predicción aparece para un horizonte de predicción de 56.5 horas, siendo este error de 5902.75 kW. El error RMSE mínimo es de 2274.70 kW y aparece con un horizonte de predicción de 21 horas. El aspecto discontinuo (con múltiples saltos) de la curva resultante de la unión de los errores RMSE de las predicciones obtenidas utilizando los modelos seleccionados, se debe al número limitado de datos de testeo para cada horizonte de predicción (30 datos para cada horizonte).

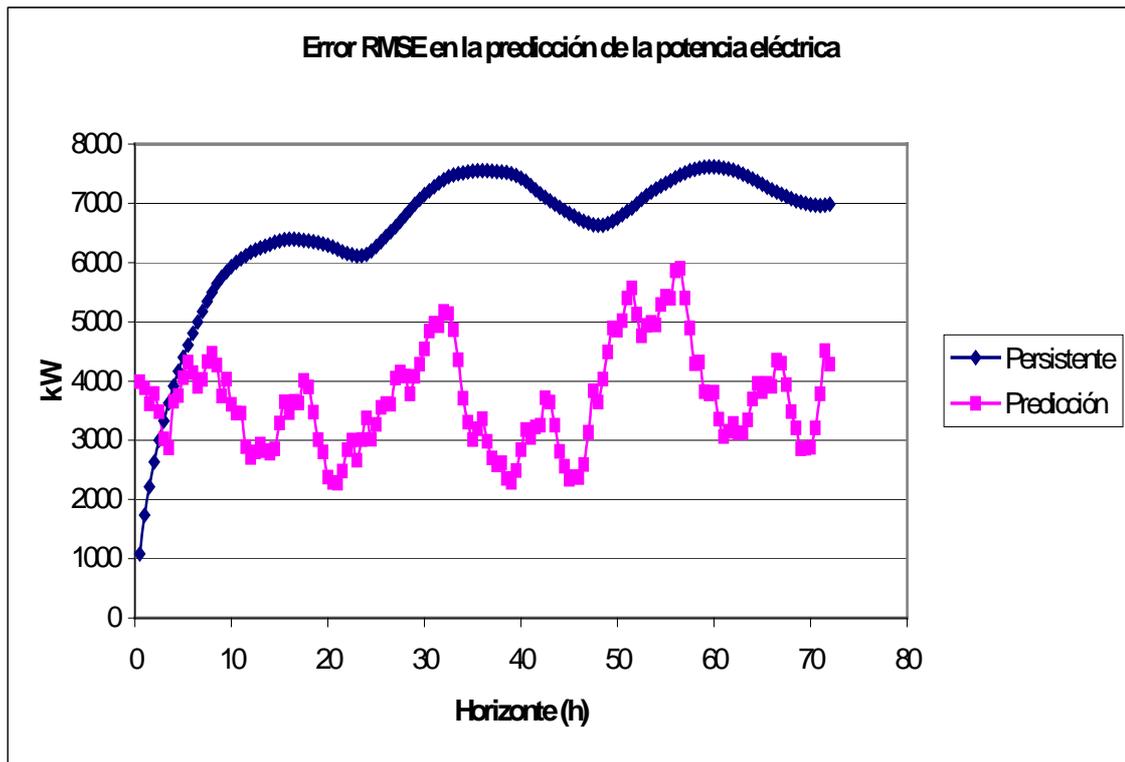


Figura 8.23. Errores RMSE en la predicción de la potencia eléctrica media suministrada a la subestación para horizontes de predicción entre 0.5 y 72 horas.

En la Figura 8.23 se puede apreciar que los errores RMSE de predicción obtenidos con los modelos seleccionados son menores que los errores obtenidos con el modelo persistente para todos los horizontes de predicción, excepto para horizontes por debajo de las 3 horas, en los que el modelo persistente sigue siendo mejor. Esto nos obligó a buscar una alternativa a los modelos seleccionados para realizar las predicciones en el muy corto plazo, inicialmente para horizontes de predicción por debajo de las tres horas, aunque, como posteriormente se detallará, puede ampliarse en función de la hora en la que se realiza la predicción.

8.5.4. Mejora de las predicciones de la potencia eléctrica a muy corto plazo.

8.5.4.1. Aplicación de filtros Kalman.

Los modelos de predicción de la segunda familia (los que usan como variables de entrada valores de predicción de modelos atmosféricos), según la clasificación hecha en el Capítulo 3 de este documento, presentan como principal ventaja frente a los modelos de la primera familia (los que usan como variables de entrada valores previos de las series temporales) su menor error de predicción con horizontes elevados (por encima de unas pocas horas). Por el contrario, para horizontes de predicción muy bajos, los modelos de la primera familia ofrecen resultados con menor error de predicción que los modelos de la segunda familia.

Una mejora para los resultados de los modelos de predicción de la segunda familia consiste en mejorar la calidad de las predicciones suministradas por el modelo atmosférico. Para horizontes de predicción muy pequeños esta mejora puede realizarse una vez obtenidos los valores de predicción, aplicando técnicas estadísticas sobre los mismos.

Ya en los modelos atmosféricos se utilizan técnicas de reducción de los errores en las predicciones numéricas que realizan. La primera técnica utilizada en la reducción de errores de modelos atmosféricos fue la técnica MOS (model output statistics) [Glahn72]. Esta técnica consiste en reducir los errores de predicción de las variables meteorológicas mediante la aplicación de regresión lineal múltiple. Así en el artículo que da origen a la técnica MOS, sus autores ponen como ejemplo, entre otros, el valor de la velocidad media del viento en superficie a las 12 horas en la ciudad de St. Louis. Para esta variable, velocidad media del viento en superficie a las 12 horas, las posibles variables explicativas son el valor de predicción obtenido para esa misma variable por el modelo atmosférico para esa hora y para horas precedentes, así como los valores obtenidos para otras variables (temperatura, velocidad del viento a la altura correspondiente a la presión de 1000 mb, etc.) para esa misma hora u horas precedentes, o incluso pueden utilizarse como variables explicativas los valores de medidas (no de predicciones) de variables meteorológicas en las horas precedentes.

En concreto, en [Glahn72] el valor de predicción para la velocidad media del viento viene determinado por la ecuación (8.27), donde \hat{S}^{12} representa el valor de predicción para la velocidad media del viento en superficie a las 12 horas tras la

aplicación de la técnica MOS, S_0^{12} representa el valor de predicción para la velocidad del viento a las 12 horas para la altura correspondiente a una presión de 1000 mb obtenida del modelo atmosférico, S^{07} representa el valor de la velocidad media del viento en superficie medido a las 7 horas, V_0^{18} corresponde al valor de predicción para la componente v (norte-sur) del viento para las 18 horas para la altura correspondiente a una presión de 1000 mb obtenido del modelo atmosférico, y U_0^{12} el de la componente u (este-oeste) del viento para la altura correspondiente a una presión de 1000 mb para las 12 horas obtenida también con el modelo atmosférico.

$$\hat{S}^{12} = 1.576 + 0.239 S_0^{12} + 0.175 S^{07} - 0.04 V_0^{18} + 0.027 U_0^{12} \quad (8.27)$$

La técnica MOS resulta muy sencilla de aplicar, pero tiene el inconveniente de precisar una gran cantidad de datos obtenidos del modelo atmosférico (e incluso de medidas reales) para poder aplicar la regresión lineal múltiple y poder reducir los errores en la predicción de las variables (velocidad y dirección media del viento) que resulten de mayor interés para la predicción a corto plazo de la potencia eléctrica generada en parques eólicos.

Dentro de los datos disponibles para el desarrollo de los modelos, los valores de predicción obtenidos del modelo atmosférico se limitan a los valores medios de la velocidad y dirección del viento, de la temperatura y de la presión, todos ellos para una altura de 10 m y para intervalos regulares de 30 minutos. Además no se dispone de medidas fiables de esas variables (se dispone de las medidas de las estaciones meteorológicas del parque, pero no fiables). Esto imposibilita el desarrollo de modelos para la reducción de los errores de predicción de los valores de velocidad y dirección media del viento obtenidos del modelo atmosférico MM5.

Una técnica diferente a la MOS que contribuye a la reducción de los errores de predicción de los modelos atmosféricos (requiriendo una menor cantidad de datos), es la de aplicación de un filtro Kalman [Kalman60]. Esta técnica ha sido usada para corregir los errores sistemáticos de los modelos atmosféricos en la predicción de la temperatura en superficie en [Homleid95], [Galanis02]; y en la reducción de los errores en la predicción de la velocidad media del viento [Crochet04], [Doeswijk05] y [Galanis06]. En los cuatro primeros trabajos citados, el filtro Kalman pretende determinar el sesgo de las predicciones del modelo atmosférico. La ecuación (8.28) muestra un ejemplo del valor de predicción, tras aplicar el filtro Kalman, de la velocidad media del viento, donde $\hat{v}(t)$ es el valor “filtrado” de esa variable, $v(t)$ es el valor de predicción para la

velocidad media del viento obtenida del modelo atmosférico y $x(t)$ es el sesgo para ese instante calculado por el filtro Kalman.

$$\hat{v}(t) = v(t) - x(t) \quad (8.28)$$

En el último trabajo publicado, [Galanis06], los autores utilizan el filtro Kalman para determinar los coeficientes de un polinomio de grado $n-1$ que determina el valor del sesgo en función del valor de la predicción obtenida del modelo atmosférico. Así, la velocidad media del viento viene indicada en la ecuación (8.29), donde el sesgo del valor proporcionado por el modelo atmosférico se calcula mediante el polinomio de grado $n-1$, cuyos coeficientes, variables con el tiempo, se calculan por medio del filtro Kalman.

$$\hat{v}(t) = v(t) - \left(a_0(t) + a_1(t)v(t) + \dots + a_{n-1}(t)[v(t)]^{n-1} \right) \quad (8.29)$$

El inconveniente que presentan los filtros Kalman, en el contexto de los datos disponibles para el desarrollo de los modelos de predicción a corto plazo de la potencia eléctrica generada en el parque eólico, es que precisan de valores medidos (reales) de las variables que son filtradas para actualizar los parámetros del filtro. Ya se ha comentado en las anteriores páginas que las medidas de las variables meteorológicas disponibles en el parque no eran fiables. En este sentido conviene desarrollar un modelo de predicción que sea lo más independiente posible de medidas en el propio parque eólico.

Los sistemas SCADA instalados en un parque eólico pueden sufrir ocasionalmente desajustes, de tal forma que los datos registrados en algún momento no son fidedignos. Así, el suministrador de los datos “Serie V”, que son los usados en este caso, manifestó que únicamente los datos correspondientes a la potencia entregada en la subestación eran fiables.

Por ello, se ha optado por actualizar los valores del filtro Kalman, no con el dato de medida real de la velocidad media del viento en el parque eólico, sino con una estimación de la velocidad media del viento obtenida a partir de la energía eléctrica vertida por el parque eólico en la subestación de conexión a la red eléctrica. Para dicha estimación se consideraron las simplificaciones señaladas a continuación.

1. En la serie temporal de potencia eléctrica media vertida en la subestación han estado contribuyendo prácticamente todos los aerogeneradores del parque (si

hubiese habido paradas de algún aerogenerador, no hay información sobre las mismas).

2. La potencia generada por el parque eólico depende de la velocidad media del viento (no existe información fidedigna sobre la dirección media real del viento). La curva de potencia eléctrica total generada en el parque (potencia eléctrica generada en función de la velocidad del viento) es aproximadamente proporcional a la curva de potencia del modelo de aerogenerador instalado en el parque eólico: la curva de potencia (estimación de la curva) del parque eólico se obtiene, para cada valor de velocidad, multiplicando por 12 (existen 12 aerogeneradores idénticos) el valor correspondiente del eje vertical de potencias eléctricas de la curva del aerogenerador.
3. Los valores negativos de potencia eléctrica vertida corresponden a consumos del propio parque eólico, consecuencia del consumo de los elementos que lo componen. En la mayoría de los casos (sobre todo en aquellos con un mayor valor absoluto), la potencia eléctrica generada en el parque es nula. Para aquellos casos en que el consumo es bajo, puede haber generación de potencia en el parque, pero no ser suficiente para cubrir su propia demanda de energía. Aun así, se consideró nula la velocidad media del viento cuando la potencia media vertida en la subestación era nula o negativa.

La potencia eléctrica media generada en el parque eólico es siempre mayor que la vertida en la subestación. Si suponemos unos valores de tensión y factor de potencia aproximadamente constantes en la subestación, la intensidad en la línea eléctrica es proporcional a la potencia entregada en la subestación. Por otra parte, la potencia eléctrica generada en el parque, P_{gen} , es igual a la potencia eléctrica vertida en la subestación, P_{sub} , más las pérdidas en la línea eléctrica, más un consumo del propio parque (supuesto constante), P_c , tal como muestra la ecuación (8.30). Dado que la intensidad es proporcional a la potencia entregada en la subestación, podemos sustituir el último sumando de la ecuación (8.30) para obtener la ecuación (8.31) que representa una estimación de la potencia generada P_{gen} .

$$P_{gen} = P_{sub} + k_1 I^2 + P_c \quad (8.30)$$

$$P_{gen} = P_{sub} + k_2 P_{sub}^2 + P_c \quad (8.31)$$

Con los valores de la serie temporal de potencia eléctrica media vertida en la subestación se determina el valor de la constante k_2 , y se obtiene el consumo del propio parque, P_c , de forma que la potencia generada no sea nunca menor que cero, ni sea, en ningún caso, superior a la potencia instalada del parque más un 2% (es la tolerancia supuesta para la potencia máxima que puede entregar cada aerogenerador).

A partir de esa estimación de la potencia generada en el parque, se determina la velocidad media teórica del viento por medio de la curva de potencia del propio parque.

Para el desarrollo de un sistema de predicción a muy corto plazo de la potencia eléctrica vertida por el parque eólico en la subestación se estudiaron los sistemas de predicción descritos seguidamente.

1. Sistema 1: sistema de predicción basado en una red neuronal utilizando como entradas exclusivamente los últimos valores conocidos (hasta 6 valores) de la potencia eléctrica media vertida en la subestación. Este sistema corresponderá a un modelo de la primera familia de modelos (de acuerdo a la clasificación de modelos de predicción hecha en el Capítulo 3).
2. Sistema 2: sistema de predicción basado en una red neuronal utilizando como entradas los últimos valores conocidos (hasta 2 valores) de la potencia eléctrica media vertida en la subestación, el valor de la velocidad media del viento ajustada por un filtro Kalman (filtrada), con expresión del sesgo obtenida con un polinomio de grado 0 (constante numérica), y los valores del seno y del coseno del ángulo de la dirección del viento suministradas por el modelo atmosférico MM5 para el horizonte de predicción.
3. Sistema 3: sistema de predicción idéntico al del punto 2, pero la velocidad media del viento es ajustada con un filtro Kalman, con expresión del sesgo obtenida con un polinomio de primer grado. El filtro determina los valores de los coeficientes del polinomio, a_0 y a_1 , expresados en la ecuación (8.32), que son variables con el tiempo. En dicha ecuación $v(t)$ representa el valor de predicción, aportado por el modelo atmosférico, para la velocidad media del viento en el horizonte de predicción, y $\hat{v}(t)$ representa el valor filtrado para dicha variable.

$$\hat{v}(t) = v(t) - a_0(t) - a_1(t) v(t) \quad (8.32)$$

4. Sistema 4: sistema idéntico al del punto 2, pero la velocidad media del viento es ajustada con un filtro Kalman con expresión del sesgo obtenida mediante un polinomio de segundo grado. El filtro determina los valores de los coeficientes del polinomio, a_0 , a_1 y a_2 expresados en la ecuación (8.33).

$$\hat{v}(t) = v(t) - a_0(t) - a_1(t) v(t) - a_2(t) v^2(t) \quad (8.33)$$

En el Anexo I de este documento se detalla la metodología seguida para la aplicación del filtro Kalman al valor de predicción de la velocidad media del viento obtenida del modelo atmosférico MM5. La misión de este filtro es la de establecer el valor, para cada instante, de los coeficientes que determinan el valor del sesgo del valor de predicción del modelo atmosférico.

Los cuatro sistemas de predicción basados en redes neuronales fueron optimizados, en un proceso controlado por algoritmo genético, con las mismas características que las expuestas para anteriores ejemplos. Las redes neuronales desarrolladas fueron de los tipos: MLP, FFBP, Modular, RBF y Híbrida (mapa auto-organizado con 20 neuronas en línea seguido de una red neuronal MLP), todas con dos capas ocultas. Los resultados obtenidos con los datos del grupo de testeo, con horizontes de predicción de 30 minutos, están recogidos en la Tabla 8.22. Las Tablas 8.23, 8.24 y 8.25 recogen los resultados de predicción con los datos del grupo de testeo para horizontes de 1 hora, 1 hora y 30 minutos, y 2 horas. Estas tablas incluyen también los errores de predicción RMSE de los modelos persistente, ARMA(1,1) y “nueva referencia”.

Además de los cuatro sistemas indicados, también se desarrolló un sistema de predicción basado en regresión lineal múltiple, siguiendo el sistema descrito en el apartado 7.3.3 de este documento. Los resultados obtenidos con este sistema adicional, aunque fueron superiores a los de las cinco redes neuronales de “Sistema 1”, fueron inferiores a los obtenidos con la mayor parte de las redes neuronales recogidas en “Sistema 2” y “Sistema 3” (el error RMSE en la predicción de la potencia eléctrica media en la subestación para el sistema basado en regresión lineal múltiple era de 1023.14 kW, para un horizonte de media hora, utilizando los dos últimos valores de la potencia eléctrica vertida en la subestación y con dos clusters). Por ello, se optó por no considerarlo en las sucesivas pruebas.

Según se puede apreciar en la Tablas 8.24 y 8.25, los resultados obtenidos, con la determinación del sesgo, mediante un filtro Kalman, en los valores de predicción de

la velocidad media del viento con el modelo atmosférico, no mejora la predicción de la potencia eléctrica vertida en la subestación con respecto al modelo persistente si los horizontes de predicción son de 1.5 horas o superiores. Por el contrario, para horizontes de predicción de 0.5 ó 1 hora, se obtienen mejores resultados aplicando el filtro Kalman, en comparación a los resultados del modelo persistente (recordemos que, para un horizonte de predicción por debajo de las 3 horas, el modelo persistente ofrecía mejores resultados que la red neuronal de Elman).

| Modelo | | Error RMSE (kW) | Error RMSE (%) | mejora |
|------------------|---------|-----------------|----------------|--------|
| Persistente | | 1083.01 | 5.01 | --- |
| ARMA(1,1) | | 1027.15 | 4.76 | 5.16% |
| Nueva referencia | | 1056.60 | 4.89 | 2.44% |
| Sistema 1 | MLP | 1032.34 | 4.78 | 4.68% |
| | FFBP | 1036.47 | 4.80 | 4.30% |
| | Modular | 1030.42 | 4.77 | 4.86% |
| | RBF | 1088.42 | 5.04 | -0.50% |
| | Híbrida | 1108.72 | 5.13 | -2.37% |
| Sistema 2 | MLP | 1007.49 | 4.66 | 6.97% |
| | FFBP | 1009.33 | 4.67 | 6.80% |
| | Modular | 1009.93 | 4.68 | 6.75% |
| | RBF | 1065.87 | 4.93 | 1.58% |
| | Híbrida | 1036.17 | 4.80 | 4.32% |
| Sistema 3 | MLP | 1016.60 | 4.71 | 6.13% |
| | FFBP | 1009.33 | 4.67 | 6.80% |
| | Modular | 1015.63 | 4.70 | 6.22% |
| | RBF | 1051.82 | 4.87 | 2.88% |
| | Híbrida | 1044.42 | 4.84 | 3.56% |
| Sistema 4 | MLP | 1013.42 | 4.69 | 6.43% |
| | FFBP | 1026.40 | 4.75 | 5.23% |
| | Modular | 1030.20 | 4.77 | 4.88% |
| | RBF | 1080.71 | 5.00 | 0.21% |
| | Híbrida | 1038.74 | 4.81 | 4.09% |

Tabla 8.22. Errores RMSE en la predicción de la potencia eléctrica en la subestación y mejora con respecto al modelo persistente para horizontes de predicción de 0.5 horas.

Para un horizonte de predicción de 1.5 horas la red neuronal modular del “Sistema 1” presenta el menor error de predicción, aunque con una ligera ventaja sobre el modelo de nueva referencia. Este último modelo es el de menor error de predicción para un horizonte de predicción de 2 horas, tal como se puede comprobar en la Tabla 8.25.

8.5.4.2. Modelos de redes neuronales de la primera familia de modelos de predicción.

Tras comprobar que el filtro Kalman, aplicado al valor de predicción de la velocidad media del viento suministrada por el modelo atmosférico, no mejora las predicciones de la potencia eléctrica media en la subestación donde vierte la energía el parque eólico si el horizonte de predicción es superior a 1 hora, se desarrollaron nuevos modelos de los denominados en el apartado anterior como “Sistema 1” para horizontes de predicción entre 2.5 y 5 horas. Estos modelos corresponden a redes neuronales que utilizan como entradas los últimos 6 valores conocidos de la potencia eléctrica media vertida en la subestación. Dado que en las pruebas realizadas para horizontes de predicción menores, los modelos de redes neuronales RBF y Híbrido, ofrecían peores resultados (en media) que los otros tres tipos de redes neuronales (MLP, FFBP y Modular), se optó por desarrollar exclusivamente esos tres tipos de modelos. Los errores de predicción para estos modelos, con los datos del grupo de testeo, se encuentran recogidos en la Tabla 8.26. También se han incluido en la tabla los errores de predicción de los modelos persistente, ARMA(1,1) y “nueva referencia”.

8.5.4.3. Sistemas avanzados para la predicción.

Después de completar las pruebas para la mejora de las predicciones a muy corto plazo, podemos seleccionar el modelo que presenta un menor error de predicción para cada uno de los horizontes de predicción. Seguidamente se indica el mejor modelo para cada intervalo de horizonte de predicción en el muy corto plazo.

- Para horizonte de predicción de 0.5 ó 1 hora, el mejor modelo es una red neuronal que utiliza como entradas los últimos valores conocidos (hasta 2 valores) de la potencia eléctrica media vertida en la subestación, el valor de la velocidad media del viento ajustada por un filtro Kalman con expresión del sesgo obtenida con un polinomio de grado 0 (constante numérica), y los valores del seno y del coseno del ángulo de la dirección del viento suministradas por el modelo atmosférico MM5 para el horizonte de predicción.

En las pruebas realizadas, la mejor red para un horizonte de 0.5 horas era una red neuronal MLP, mientras que para una hora de horizonte de predicción la mejor red neuronal era una FFBP. Ya que para el horizonte de predicción de 0.5 horas la diferencia de errores RMSE en la predicción de la potencia eléctrica vertida en la subestación para la red neuronal MLP y la red neuronal FFBP era casi despreciable

(exactamente el 0.18%), podría utilizarse el mismo tipo de red neuronal, la FFBP, para las predicciones con horizonte 0.5 ó 1 hora.

| Modelo | | Error RMSE (kW) | Error RMSE (%) | mejora |
|------------------|---------|-----------------|----------------|--------|
| Persistente | | 1738.31 | 8.05 | --- |
| ARMA(1,1) | | 1676.47 | 7.76 | 3.55% |
| Nueva referencia | | 1677.95 | 7.77 | 3.47% |
| Sistema 1 | MLP | 1687.02 | 7.81 | 2.95% |
| | FFBP | 1686.90 | 7.81 | 2.96% |
| | Modular | 1708.47 | 7.91 | 1.72% |
| | RBF | 1702.12 | 7.88 | 2.08% |
| | Híbrida | 1703.36 | 7.89 | 2.01% |
| Sistema 2 | MLP | 1645.86 | 7.62 | 5.32% |
| | FFBP | 1635.80 | 7.57 | 5.90% |
| | Modular | 1641.17 | 7.60 | 5.59% |
| | RBF | 1665.81 | 7.71 | 4.17% |
| | Híbrida | 1643.21 | 7.61 | 5.47% |
| Sistema 3 | MLP | 1647.37 | 7.63 | 5.23% |
| | FFBP | 1643.16 | 7.61 | 5.47% |
| | Modular | 1659.17 | 7.68 | 4.55% |
| | RBF | 1670.67 | 7.73 | 3.89% |
| | Híbrida | 1655.88 | 7.67 | 4.74% |
| Sistema 4 | MLP | 1698.43 | 7.86 | 2.29% |
| | FFBP | 1703.65 | 7.89 | 1.99% |
| | Modular | 1696.07 | 7.85 | 2.43% |
| | RBF | 1715.20 | 7.94 | 1.33% |
| | Híbrida | 1692.95 | 7.84 | 2.61% |

Tabla 8.23. Errores RMSE en la predicción de la potencia eléctrica en la subestación y mejora con respecto al modelo persistente para horizontes de predicción de 1 hora.

- Para un horizonte de predicción de 1.5 horas el modelo que proporciona un menor error de predicción es una red neuronal modular (no usa la salida del filtro Kalman).
- Para horizontes de predicción entre 2 y 3 horas, ambos inclusive, el mejor modelo de predicción es el modelo “nueva referencia”. Este modelo utiliza exclusivamente los valores pasados de la potencia eléctrica vertida en la subestación para realizar las predicciones.
- Para horizontes de predicción entre las 3.5 y las 24 horas, el mejor modelo de predicción es la red neuronal de Elman descrita en el apartado 8.5.2.1 de este documento. En lo sucesivo se la denominará red neuronal Elman 1. Éste sería el

mejor modelo a partir de ese horizonte si las predicciones del modelo atmosférico, que le sirven de entradas, estuvieran disponibles y actualizadas en todo instante, lo que, como veremos más adelante, no es el caso.

| Modelo | | Error RMSE (kW) | Error RMSE (%) | mejora |
|------------------|---------|-----------------|----------------|---------|
| Persistente | | 2217.59 | 10.27 | --- |
| ARMA(1,1) | | 2127.12 | 9.85 | 4.08% |
| Nueva referencia | | 2115.64 | 9.79 | 4.60% |
| Sistema 1 | MLP | 2129.64 | 9.86 | 3.97% |
| | FFBP | 2139.76 | 9.91 | 3.51% |
| | Modular | 2115.53 | 9.79 | 4.61% |
| | RBF | 2165.69 | 10.03 | 2.34% |
| | Híbrida | 2129.94 | 9.86 | 3.95% |
| Sistema 2 | MLP | 2717.32 | 12.58 | -22.53% |
| | FFBP | 2731.90 | 12.65 | -23.19% |
| | Modular | 2716.20 | 12.57 | -22.48% |
| | RBF | 2909.72 | 13.47 | -31.21% |
| | Híbrida | 2790.42 | 12.92 | -25.83% |
| Sistema 3 | MLP | 2951.61 | 13.66 | -33.10% |
| | FFBP | 2884.13 | 13.35 | -30.06% |
| | Modular | 2893.06 | 13.39 | -30.46% |
| | RBF | 2960.01 | 13.70 | -33.48% |
| | Híbrida | 2910.15 | 13.47 | -31.23% |
| Sistema 4 | MLP | 2754.56 | 12.75 | -24.21% |
| | FFBP | 2870.59 | 13.29 | -29.45% |
| | Modular | 2882.96 | 13.35 | -30.00% |
| | RBF | 2832.93 | 13.12 | -27.75% |
| | Híbrida | 2748.62 | 12.73 | -23.95% |

Tabla 8.24. Errores RMSE en la predicción de la potencia eléctrica en la subestación y mejora con respecto al modelo persistente para horizontes de predicción de 1.5 horas.

- Para horizontes de predicción entre las 24.5 y 48 horas, el mejor modelo de predicción es la red neuronal de Elman descrita en el apartado 8.5.2.2 de este documento. En lo sucesivo se la denominará red neuronal Elman 2.
- Para horizontes de predicción entre las 48.5 y 72 horas, el mejor modelo de predicción es la red neuronal modular descrita en el apartado 8.5.2.3 de este documento. En lo sucesivo se la denominará red neuronal Modular.

| Modelo | | Error RMSE (kW) | Error RMSE (%) | mejora |
|------------------|---------|-----------------|----------------|---------|
| Persistente | | 2631.31 | 12.18 | --- |
| ARMA(1,1) | | 2503.99 | 11.59 | 4.84% |
| Nueva referencia | | 2486.21 | 11.51 | 5.51% |
| Sistema 1 | MLP | 2493.87 | 11.55 | 5.22% |
| | FFBP | 2503.76 | 11.59 | 4.85% |
| | Modular | 2500.37 | 11.58 | 4.98% |
| | RBF | 2530.04 | 11.71 | 3.85% |
| | Híbrida | 2525.43 | 11.69 | 4.02% |
| Sistema 2 | MLP | 3112.44 | 14.41 | -18.28% |
| | FFBP | 3028.53 | 14.02 | -15.10% |
| | Modular | 3149.39 | 14.58 | -19.69% |
| | RBF | 3190.00 | 14.77 | -21.23% |
| | Híbrida | 3082.93 | 14.27 | -17.16% |
| Sistema 3 | MLP | 3204.25 | 14.83 | -21.77% |
| | FFBP | 3333.14 | 15.43 | -26.67% |
| | Modular | 3336.54 | 15.45 | -26.80% |
| | RBF | 3374.85 | 15.62 | -28.26% |
| | Híbrida | 3392.76 | 15.71 | -28.94% |
| Sistema 4 | MLP | 3252.58 | 15.06 | -23.61% |
| | FFBP | 3130.82 | 14.49 | -18.98% |
| | Modular | 3173.99 | 14.69 | -20.62% |
| | RBF | 3104.46 | 14.37 | -17.98% |
| | Híbrida | 3239.04 | 15.00 | -23.10% |

Tabla 8.25. Errores RMSE en la predicción de la potencia eléctrica en la subestación y mejora con respecto al modelo persistente para horizontes de predicción de 2 horas.

La Figura 8.24 muestra gráficamente el modelo que ofrece los mejores resultados preliminares en la predicción de la potencia eléctrica media vertida a la red por el parque eólico para horizontes por debajo de 24 horas.

Todavía es posible mejorar ligeramente las predicciones, en especial en el muy corto plazo. En los sistemas de predicción con series temporales, una solución para mejorar las predicciones obtenidas con los modelos desarrollados consiste en construir un modelo de regresión lineal múltiple que utilice como variables explicativas las predicciones de los dos o tres modelos que mejores resultados han dado en la predicción. La idea que subyace bajo esta regresión lineal múltiple es la mejora de las predicciones por la compensación de errores entre las predicciones de los diferentes modelos.

| Horizonte | Modelo | Error RMSE (kW) | Error RMSE (%) | mejora |
|-----------|------------------|--------------------|-------------------|--------|
| 2.5 horas | Persistente | 2999.84 | 13.89 | --- |
| | ARMA(1,1) | 2830.88 | 13.11 | 5.63% |
| | Nueva referencia | 2810.15 | 13.01 | 6.32% |
| | MLP | 2817.78 | 13.05 | 6.07% |
| | FFBP | 2844.75 | 13.17 | 5.17% |
| 3 horas | Modular | 2850.95 | 13.20 | 4.96% |
| | Persistente | 3328.10 | 15.41 | --- |
| | ARMA(1,1) | 3110.44 | 14.40 | 6.54% |
| | Nueva referencia | 3091.03 | 14.31 | 7.12% |
| | MLP | 3099.09 | 14.35 | 6.88% |
| 3.5 horas | FFBP | 3113.78 | 14.42 | 6.44% |
| | Modular | 3115.44 | 14.42 | 6.39% |
| | Persistente | 3630.82 | 16.81 | --- |
| | ARMA(1,1) | 3359.10 | 15.55 | 7.48% |
| | Nueva referencia | 3340.89 | 15.47 | 7.99% |
| 4 horas | MLP | 3384.69 | 15.67 | 6.78% |
| | FFBP | 3335.88 | 15.44 | 8.12% |
| | Modular | 3308.36 | 15.32 | 8.88% |
| | Persistente | 3911.81 | 18.11 | --- |
| | ARMA(1,1) | 3582.46 | 16.59 | 8.42% |
| 4.5 horas | Nueva referencia | 3563.16 | 16.50 | 8.91% |
| | MLP | 3695.90 | 17.11 | 5.52% |
| | FFBP | 3695.90 | 17.11 | 5.52% |
| | Modular | 3574.89 | 16.55 | 8.61% |
| | Persistente | 4168.84 | 19.30 | --- |
| 5 horas | ARMA(1,1) | 3779.61 | 17.50 | 9.34% |
| | Nueva referencia | 3758.98 | 17.40 | 9.83% |
| | MLP | 3787.56 | 17.54 | 9.15% |
| | FFBP | 3771.66 | 17.46 | 9.53% |
| | Modular | 3836.41 | 17.76 | 7.97% |
| 5 horas | Persistente | 4400.46 | 20.37 | --- |
| | ARMA(1,1) | 3950.37 | 18.29 | 10.23% |
| | Nueva referencia | 3925.95 | 18.18 | 10.78% |
| | MLP | 3791.61 | 17.55 | 13.84% |
| | FFBP | 3721.88 | 17.23 | 15.42% |
| 5 horas | Modular | 3796.00 | 17.57 | 13.74% |

Tabla 8.26. Errores RMSE en la predicción de la potencia eléctrica en la subestación y mejora con respecto al modelo persistente para horizontes de predicción entre 2.5 y 5 horas.

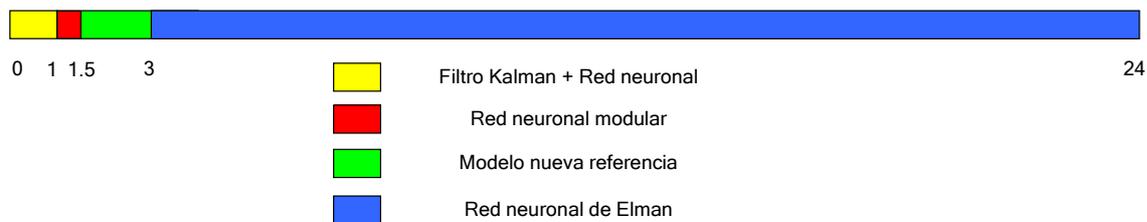


Figura 8.24. Modelos con mejores resultados en la predicción de la potencia eléctrica en la subestación para horizontes de predicción por debajo de 24 horas.

Mejores resultados que los obtenidos con un modelo de regresión lineal múltiple se pueden obtener con un sistema de inferencia difusa del tipo TSK de primer orden, donde la parte consecuente de las reglas responde a un modelo de regresión lineal múltiple. Así, el valor de la salida de un sistema de inferencia difusa TSK representa la media ponderada de las salidas de varios subsistemas de regresión lineal múltiple (tantos como reglas), construidos para ciertos valores locales de las variables de entrada del sistema (los valores locales que intervienen en cada subsistema vienen determinados por el tipo de partición del espacio de entrada).

Se construyó, para cada uno de los horizontes de predicción de menor valor, un sistema de inferencia difusa del tipo TSK de primer orden que utilizaba como entradas las predicciones de los tres modelos con menor error RMSE en la predicción de la potencia eléctrica media vertida en la subestación para los datos de los grupos de entrenamiento y validación. Las entradas de estos modelos de inferencia difusa (salidas de los modelos de predicción desarrollados anteriormente) escogidas para cada uno de los horizontes, así como los errores RMSE obtenidos con los datos del grupo de testeo para estos sistemas de inferencia difusa están recogidas en la Tabla 8.27. En dicha tabla, la “mejora respecto a mejor modelo” representa el porcentaje de mejora, para el error RMSE, con respecto al modelo de menor error de los tres cuyas predicciones constituyen las entradas del sistema de inferencia difusa; la “mejora” representa el porcentaje de mejora, para el error RMSE, con respecto al modelo persistente. Una condición que se impuso, tras las primeras pruebas, fue la de que los tres modelos escogidos, cuyas salidas constituyen las entradas del sistema de inferencia difusa, fueran de “tecnologías diferentes” (por ejemplo, no se admiten dos redes neuronales, incluso siendo distintas, que tengan por entrada la salida del filtro Kalman).

Los sistemas de inferencia difusa fueron construidos con la técnica del agrupamiento substractivo de tal manera que el número de clusters (con los datos de los grupos de entrenamiento y validación) fuera 2. Se escogieron los parámetros del

algoritmo substractivo (modificando exclusivamente el rango de influencia) de tal manera que correspondiesen al valor de frontera entre la obtención de 2 y 3 clusters. Las familias de pertenencia fueron gaussianas. Las posiciones de los clusters (centros de las familias de pertenencia), su varianza y los parámetros de la parte consecuente fueron obtenidos con la función “genfis2” [Jang95].

A partir de un horizonte de predicción de 5 horas, el error de las predicciones conseguidas con el sistema de inferencia difusa correspondiente es mayor que el del mejor de los modelos que utiliza como entradas, por lo que no mejora la calidad de las predicciones (tal y como se observa para el horizonte de predicción de 5 horas en la Tabla 8.27). El mejor modelo para ese horizonte de predicción de 5 horas es la red neuronal FFBP que usa como entradas los últimos valores conocidos de la serie temporal de potencia eléctrica media en la subestación. Concretamente estos valores son $p(t-2.5)$, $p(t-2)$, $p(t-1.5)$, $p(t-1)$, $p(t-0.5)$ y $p(t)$, y la predicción que realiza corresponde a $p(t+5)$.

Con un horizonte de predicción por encima de las 5 horas los errores de predicción obtenidos con la red neuronal de Elman 1 son mucho menores que cualquiera de los modelos utilizados como entradas para los sistemas de inferencia difusa (modelo ARMA(1,1), modelo “nueva referencia”, etc.), por lo que no se obtiene mejora alguna al combinarla con los otros modelos. Esto sería más acusado en el caso de que las predicciones del modelo atmosférico se actualizasen cada 6 horas (de hecho puede hacerse), como se analizará en el próximo apartado.

8.5.5. Nuevo sistema global para la predicción.

En la propuesta del nuevo sistema global de predicción (sistema híbrido global), como consecuencia del análisis de los resultados de predicción de la potencia eléctrica media en la subestación de conexión del parque eólico con la red eléctrica, utilizando los modelos de predicción estudiados, se va a distinguir entre la disponibilidad de predicciones del modelo atmosférico MM5 para cada 24 horas o para cada 6 horas. En todo caso, la propuesta de cada uno de los modelos de predicción (pertenecientes al nuevo sistema global) corresponde al mejor modelo, de los analizados en las páginas anteriores, que obtiene los resultados óptimos para cada horizonte de predicción, en función de cual sea el momento en que se realizan las predicciones.

| Horizonte | Modelos | Error RMSE (kW) | Mejora respecto mejor modelo | Error RMSE (%) | mejora |
|-----------|-----------------------|--------------------|------------------------------------|-------------------|--------|
| 0.5 horas | Kalman orden 0 + FFBP | | | | |
| | ARMA(1,1) | 1002.62 | 0.48% | 4.64% | 7.42% |
| | Nueva referencia | | | | |
| 1 hora | Kalman orden 0 + FFBP | | | | |
| | ARMA(1,1) | 1616.89 | 1.16% | 7.49% | 6.98% |
| | Nueva referencia | | | | |
| 1.5 horas | Modular | | | | |
| | ARMA(1,1) | 1688.63 | 1.45% | 7.82% | 23.85% |
| | Nueva referencia | | | | |
| 2 horas | MLP | | | | |
| | ARMA(1,1) | 2138.34 | 0.94% | 9.99% | 18.74% |
| | Nueva referencia | | | | |
| 2.5 horas | MLP | | | | |
| | ARMA(1,1) | 2516.27 | 0.58% | 11.65% | 16.12% |
| | Nueva referencia | | | | |
| 3 horas | MLP | | | | |
| | ARMA(1,1) | 2864.93 | 0.51% | 13.26% | 13.92% |
| | Nueva referencia | | | | |
| 3.5 horas | Modular | | | | |
| | ARMA(1,1) | 3111.14 | 0.41% | 14.40% | 14.31% |
| | Nueva referencia | | | | |
| 4 horas | Modular | | | | |
| | ARMA(1,1) | 3375.28 | 0.27% | 15.62% | 13.72% |
| | Nueva referencia | | | | |
| 4.5 horas | MLP | | | | |
| | ARMA(1,1) | 3616.76 | 0.97% | 16.74% | 13.24% |
| | Nueva referencia | | | | |
| 5 horas | FFBP | | | | |
| | ARMA(1,1) | 3807.32 | -0.02% | 17.63% | 13.48% |
| | Nueva referencia | | | | |

Tabla 8.27. Mejora de las predicciones a muy corto plazo obtenidas con los sistemas de inferencia difusa.

8.5.5.1. Nuevo sistema global de predicción propuesto con predicciones del modelo MM5 cada 24 horas.

Este es el caso realmente estudiado en los trabajos de investigación, ya que únicamente se dispone de predicciones del modelo MM5 realizadas con los datos de la atmósfera asimilados a las 0 horas GMT. En estas condiciones, el modelo que hay que

usar para calcular los valores de predicción cambia según el instante en que se realicen estas predicciones. La Tabla 8.28 recoge el mejor sistema en función del instante en que se realizan las predicciones (sólo se ha considerado instantes correspondientes a los 0 minutos de cada hora referidos a hora GMT).

En la Tabla 8.28 están señaladas las horas en las que conviene utilizar cada uno de los modelos presentados en las páginas anteriores. Así, por ejemplo, si la predicción se realiza a las 2 horas GMT, son los modelos de inferencia difusa (FIS) recogidos en la Tabla 8.27 los que conviene utilizar con horizontes de predicción de 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5 y 5 horas. Para horizontes de predicción de 5.5, 6 y 6.5 horas (de las 7.5 a las 8.5 horas GMT) conviene usar el modelo “nueva referencia” ya que es el que proporciona un menor error de predicción en esos momentos. A partir de un horizonte de predicción de 7 horas (a partir de las 9 horas GMT) el modelo con menor error es la red neuronal Elman 2, que era el modelo con mejores resultados para la predicción con horizontes de predicción de 24.5 a 48 horas (apartado 8.5.2.2 de este documento). Debe tenerse en cuenta que, actualmente, hasta poco después de las 5 horas GMT no se dispone de nuevas predicciones del modelo atmosférico MM5, con lo que la mejor predicción posible para las siguientes horas corresponde a las predicciones realizadas el día anterior con la red neuronal Elman 2, utilizada para realizar predicciones para el día siguiente. Para las predicciones a partir de las 24 horas, el mejor modelo lo constituye la red neuronal Modular (apartado 8.5.2.3 de este documento) con las predicciones que realizó el día anterior.

Poco después de las 5 horas GMT, en la actualidad ya pueden estar disponibles las últimas predicciones del modelo atmosférico MM5, por lo que a partir de ese momento las predicciones realizadas con la red neuronal Elman 1, que ofrecía menores errores de predicción con horizontes por debajo de las 24 horas, empieza a aparecer como mejor modelo de predicción.

8.5.5.2. Nuevo sistema global de predicción propuesto con predicciones del modelo MM5 cada 6 horas.

Los resultados de los modelos de predicción de escala global, de los que parte el modelo de meso-escala MM5 para realizar sus predicciones, son actualizados cada 6 horas. Tal es el caso del modelo GFS del NCEP norteamericano. Los resultados, con las predicciones para una rejilla con 1° de resolución espacial y para las próximas 180 horas (en intervalos de tres horas), están disponibles unas 4 horas y media después de las 0, 6,

12 y 18 horas GMT. Un equipo (cluster) de potentes computadores personales podría ejecutar, en menos de 1 hora, las rutinas del modelo MM5 y tener las predicciones para las próximas 48 horas correspondientes a la producción eléctrica del parque eólico.

Dado que no se dispone de las predicciones del modelo MM5 con los datos de la atmósfera para las 6, 12 y 18 horas GMT, en los siguientes párrafos se va a suponer que los errores de predicción para los modelos que se desarrollasen, utilizando esas predicciones del modelo MM5, son similares a los obtenidos (utilizando las predicciones del modelo MM5 con los datos de la atmósfera para las 0 horas GMT) para los mismos horizontes de predicción. De todas formas, es lógico suponer que la precisión de las predicciones del modelo atmosférico es mayor cuanto menor sea su antigüedad, por lo que sólo se van a considerar las predicciones obtenidas con el último grupo de datos disponibles de predicciones del modelo MM5.

| Predicción (hora GMT) | FIS | Nueva referencia | Elman 1 | Elman 2 | Modular |
|--------------------------|-------------|---------------------|-------------------|-------------------|------------------|
| 0 | Hasta 5h | 5.5h | | A partir de 6h | A partir de 24 h |
| 1 | Hasta 6h | 6.5h | | A partir de 7h | A partir de 24 h |
| 2 | Hasta 7h | 7.5-8.5h | | A partir de 9h | A partir de 24 h |
| 3 | Hasta 8h | 8.5-9h | | A partir de 9.5h | A partir de 24 h |
| 4 | Hasta 9h | | | A partir de 9.5h | A partir de 24 h |
| 5 | Hasta 10h | | | A partir de 10.5h | A partir de 24 h |
| 6 | Hasta 10h | | A partir de 10.5h | A partir de 24h | A partir de 48 h |
| 7 | Hasta 11h | | A partir de 11.5h | A partir de 24h | A partir de 48 h |
| 8 | Hasta 11.5h | | A partir de 12h | A partir de 24h | A partir de 48 h |
| 9 | Hasta 12h | | A partir de 12.5h | A partir de 24h | A partir de 48 h |
| 10 | Hasta 12.5h | | A partir de 13h | A partir de 24h | A partir de 48 h |
| 11 | Hasta 13.5h | | A partir de 14h | A partir de 24h | A partir de 48 h |
| 12 | Hasta 16h | | A partir de 16.5h | A partir de 24h | A partir de 48 h |
| 13 | Hasta 17.5h | | A partir de 18h | A partir de 24h | A partir de 48 h |
| 14 | Hasta 18.5h | | A partir de 19h | A partir de 24h | A partir de 48 h |
| 15 | Hasta 19h | | A partir de 19.5h | A partir de 24h | A partir de 48 h |
| 16 | Hasta 19h | | A partir de 19.5h | A partir de 24h | A partir de 48 h |
| 17 | Hasta 19.5h | | A partir de 20h | A partir de 24h | A partir de 48 h |
| 18 | Hasta 20h | | A partir de 20.5h | A partir de 24h | A partir de 48 h |
| 19 | Hasta 21h | | A partir de 21.5h | A partir de 24h | A partir de 48 h |
| 20 | Hasta 23h | | A partir de 23.5h | A partir de 24h | A partir de 48 h |
| 21 | Hasta 24h | | | A partir de 24.5h | A partir de 48 h |
| 22 | Hasta 3h | 3.5-4h | | A partir de 28.5h | A partir de 48 h |
| 23 | Hasta 4h | 4.5h | | A partir de 29h | A partir de 48 h |

Tabla 8.28. Modelos a utilizar para realizar las predicciones de la potencia eléctrica en la subestación en función de la hora en que se realizan las predicciones.

La selección del mejor modelo de predicción (con menores errores) depende del instante en el que se requiere la predicción. En las siguientes páginas vamos a determinar (para distintos momentos de realización de las predicciones) el modelo de predicción que logra un menor error de predicción. Estos modelos son los sistemas de inferencia difusa (en lo sucesivo FIS) desarrollados para horizontes de predicción hasta un máximo de 5 horas, la red neuronal Elman 1, la red neuronal Elman 2 y la red neuronal Modular. En lo sucesivo, todas las indicaciones horarias se refieren a horas GMT.

1. Predicción realizada a las 0 horas (o a las 6, 12 o 18 horas).

Si las predicciones se realizan a las 0 horas, se dispone ya de las predicciones del modelo atmosférico MM5 realizadas con el estado de la atmósfera a las 18 horas del día anterior (en lo sucesivo predicción del modelo MM5 a las 18 horas del día anterior). La Figura 8.25 representa los errores RMSE en la predicción de la potencia eléctrica media en la subestación, con los datos del grupo de testeo, para los modelos FIS y para la red neuronal Elman 1 con una antigüedad de 6 horas, ya que fueron obtenidos con las predicciones del modelo atmosférico MM5 con datos de las 18 horas del día anterior (Elman 18h). Como se puede apreciar, para cuatro horas y media después, los resultados de la red neuronal Elman 1 con las predicciones del modelo atmosférico MM5 (con datos de las 18 horas del día anterior) presentan un menor error que los modelos FIS.

Para los siguientes horizontes de predicción (por encima de las 4.5 horas) conviene usar la red neuronal Elman 1 utilizando las predicciones del modelo atmosférico MM5, con datos de las 18 horas del día anterior, para las predicciones hasta las 18 horas del día actual. Para predicciones entre las 18.5 horas del día actual y las 18 horas del día siguiente el modelo que conviene utilizar es la red neuronal Elman 2, con las predicciones del modelo atmosférico MM5 con datos de las 18 horas del día anterior. Por último para predicciones a partir de las 18.5 horas del día siguiente y hasta las 18 horas de dos días después, el modelo que ofrece mejores resultados es la red neuronal Modular, utilizando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior.

Para el caso de realizar estas predicciones a las 6, 12 ó 18 horas sólo habrá que sumar 6, 12 y 18 horas sobre las referencias temporales señaladas en los dos

párrafos anteriores. Por ejemplo, si las predicciones se realizan a las 18 horas, los últimos datos disponibles del modelo atmosférico MM5 son los de las 12 horas de ese mismo día. Las predicciones para las cuatro primeras horas (hasta las 22 horas) se realizarán con los modelos FIS, las predicciones entre las 22.5 horas y las 12 horas del día siguiente con la red neuronal Elman 1, usando las predicciones del modelo atmosférico MM5 con los datos de las 12 horas del día actual. Las predicciones entre las 12.5 horas del día siguiente y las 12 horas de dos días después, se realizaran usando la red neuronal Elman 2 y las predicciones del modelo atmosférico MM5 con los datos de las 12 horas del día actual. Por último, para las predicciones entre las 12.5 horas de dos días después y las 12 horas de tres días después, conviene utilizar la red neuronal Modular y las predicciones del modelo atmosférico MM5 con los datos de las 12 horas del día actual.

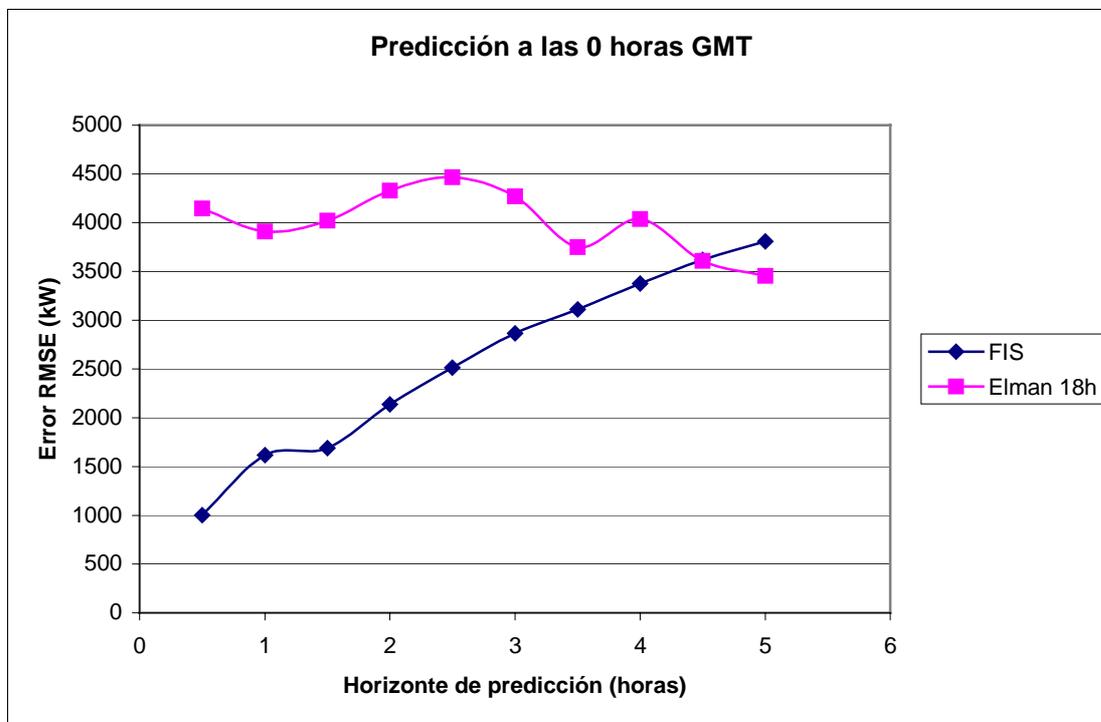


Figura 8.25. Errores RMSE de predicción de la potencia eléctrica media en la subestación para predicciones realizadas a las 0 horas GMT.

2. Predicción realizada a la hora 1 (o a las 7, 13 o 19 horas).

En este caso, a la hora 1 se dispone de las predicciones del modelo atmosférico MM5 de las 18 horas del día anterior. La Figura 8.26 representa los errores RMSE en la predicción de la potencia eléctrica media en la subestación, con los datos del grupo de testeo, para los modelos FIS y para la red neuronal Elman 1 con una

antigüedad de 7 horas, puesto que fueron realizadas con las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior (Elman 18h). Aquí, como en el caso anterior, los errores con la red neuronal Elman 1 con las predicciones del modelo MM5, realizadas con los datos a las 18 horas del día anterior, superan a los modelos FIS para un horizonte de predicción de 4.5 horas.

Para los siguientes horizontes de predicción conviene usar la red neuronal Elman 1, con las predicciones del modelo atmosférico MM5 (utilizando los datos de las 18 horas del día anterior para las predicciones), hasta las 18 horas del día actual. Para predicciones entre las 18.5 horas del día actual y las 18 horas del día siguiente, el modelo idóneo es la red neuronal Elman 2 usando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior. Por último, para predicciones a partir de las 18.5 horas del día siguiente y hasta las 18 horas de dos días después, el modelo que ofrece mejores resultados es la red neuronal Modular, empleando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior.

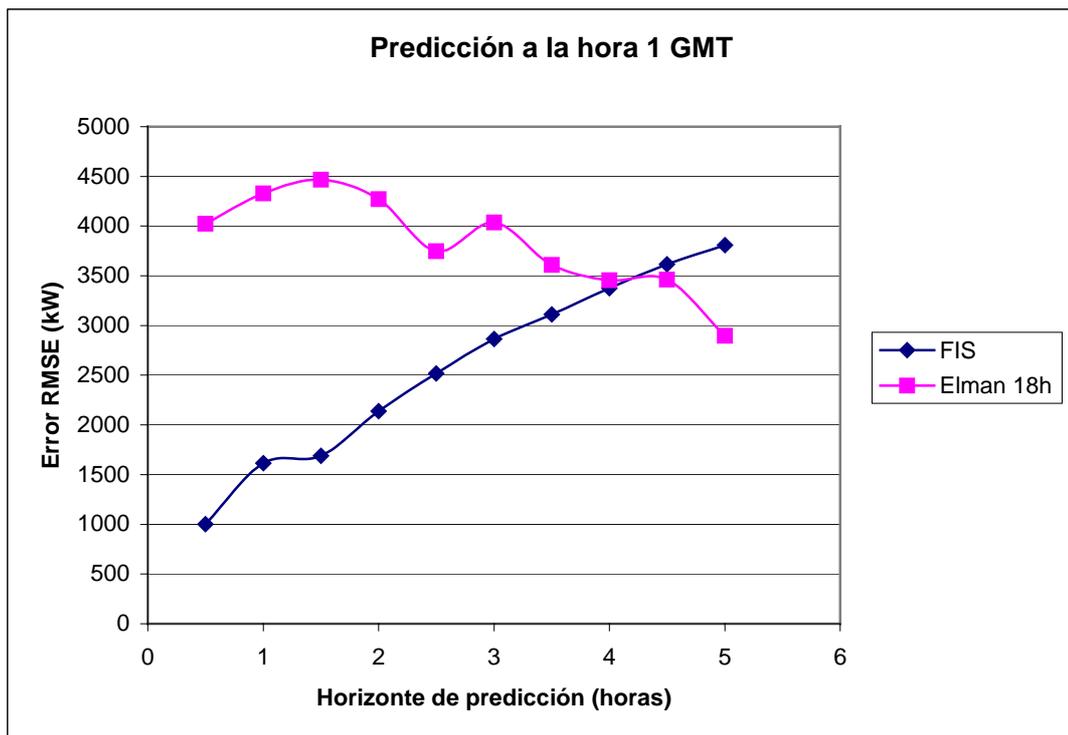


Figura 8.26. Errores RMSE de predicción de la potencia eléctrica media en la subestación para predicciones realizadas a la 1 hora GMT.

Para el caso de realizar estas predicciones a las 7, 13 ó 19 horas sólo habrá que sumar 6, 12 y 18 horas sobre las referencias temporales indicadas en los dos

párrafos anteriores.

3. Predicción realizada a las 2 horas (o a las 8, 14 o 20 horas).

En este caso, a las 2 horas se dispone de las predicciones del modelo atmosférico MM5 de las 18 horas del día anterior. La Figura 8.27 representa los errores RMSE en la predicción de la potencia eléctrica media en la subestación, con los datos del grupo de testeo, para los modelos FIS y para la red neuronal Elman 1 con una antigüedad de 8 horas, puesto que fueron realizadas usando las predicciones del modelo atmosférico MM5 con los datos atmosféricos de las 18 horas del día anterior (Elman 18h). En este caso los errores obtenidos con la red neuronal Elman 1, utilizando las predicciones del modelo MM5 realizadas con los datos a las 18 horas del día anterior, superan a los modelos FIS para un horizonte de predicción de 4 horas.

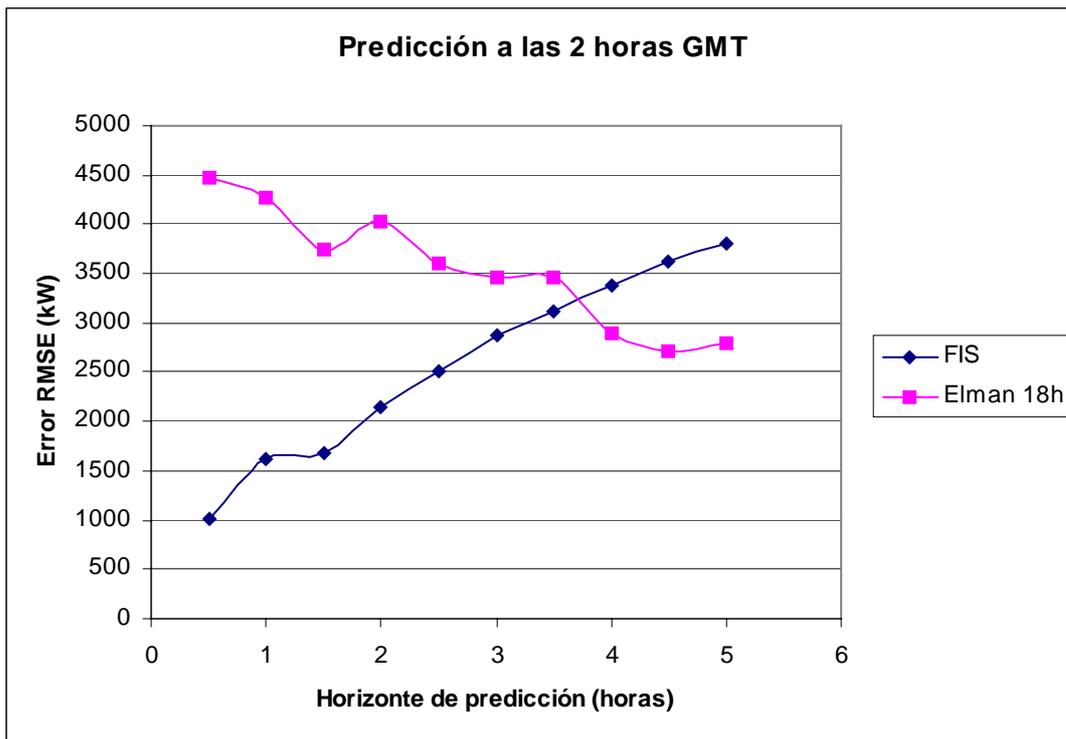


Figura 8.27. Errores RMSE de predicción de la potencia eléctrica media en la subestación para predicciones realizadas a las 2 horas GMT.

Para los siguientes horizontes de predicción conviene usar la red neuronal Elman 1 (haciendo uso de las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior), para las predicciones hasta las 18 horas del día actual.

Para predicciones entre las 18.5 horas del día actual y las 18 horas del día siguiente, el modelo más adecuado es la red neuronal Elman 2, empleando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior. Por último para predicciones a partir de las 18.5 horas del día siguiente y hasta las 18 horas de dos días después, el modelo que ofrece mejores resultados es la red neuronal Modular con las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior.

Para el caso de realizar estas predicciones a las 8, 14 ó 20 horas sólo habrá que sumar 6, 12 y 18 horas sobre las referencias temporales señaladas en los dos párrafos anteriores.

4. Predicción realizada a las 3 horas (o a las 9, 15 o 21 horas).

En este caso, a las 3 horas se dispone de las predicciones del modelo atmosférico MM5 de las 18 horas del día anterior. La Figura 8.28 representa los errores RMSE en la predicción de la potencia eléctrica media en la subestación, con los datos del grupo de testeo, para los modelos FIS y para la red neuronal Elman 1 con una antigüedad de 9 horas, puesto que fueron realizadas con las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior (Elman 18h). En este caso, los errores con la red neuronal Elman 1, empleando las predicciones del modelo MM5 realizadas con los datos a las 18 horas del día anterior, superan a los modelos FIS para un horizonte de predicción de 3.5 horas.

Para los siguientes horizontes de predicción, es idóneo usar la red neuronal Elman 1 (empleando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior), para las predicciones hasta las 18 horas del día actual. Para predicciones entre las 18.5 horas del día actual y las 18 horas del día siguiente, el modelo adecuado es la red neuronal Elman 2, usando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior. Por último, para predicciones a partir de las 18.5 horas del día siguiente y hasta las 18 horas de dos días después, el modelo que ofrece mejores resultados es la red neuronal Modular, usando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior.

Para el caso de realizar estas predicciones a las 9, 15 ó 21 horas sólo habrá que sumar 6, 12 y 18 horas sobre las referencias temporales señaladas en los dos

párrafos anteriores.

5. Predicción realizada a las 4 horas (o a las 10, 16 o 22 horas).

En este caso, a las 4 horas se dispone de las predicciones del modelo atmosférico MM5 de las 18 horas del día anterior. La Figura 8.29 representa los errores RMSE en la predicción de la potencia eléctrica media en la subestación, con los datos del grupo de testeo, para los modelos FIS y para la red neuronal Elman 1, con una antigüedad de 10 horas, puesto que fueron realizadas usando las predicciones del modelo atmosférico MM5 con los datos atmosféricos de las 18 horas del día anterior (Elman 18h). Aquí, los errores de predicción de la red neuronal Elman 1, empleando las predicciones del modelo MM5 realizadas con los datos a las 18 horas del día anterior, superan a los modelos FIS para un horizonte de predicción de 3 horas.

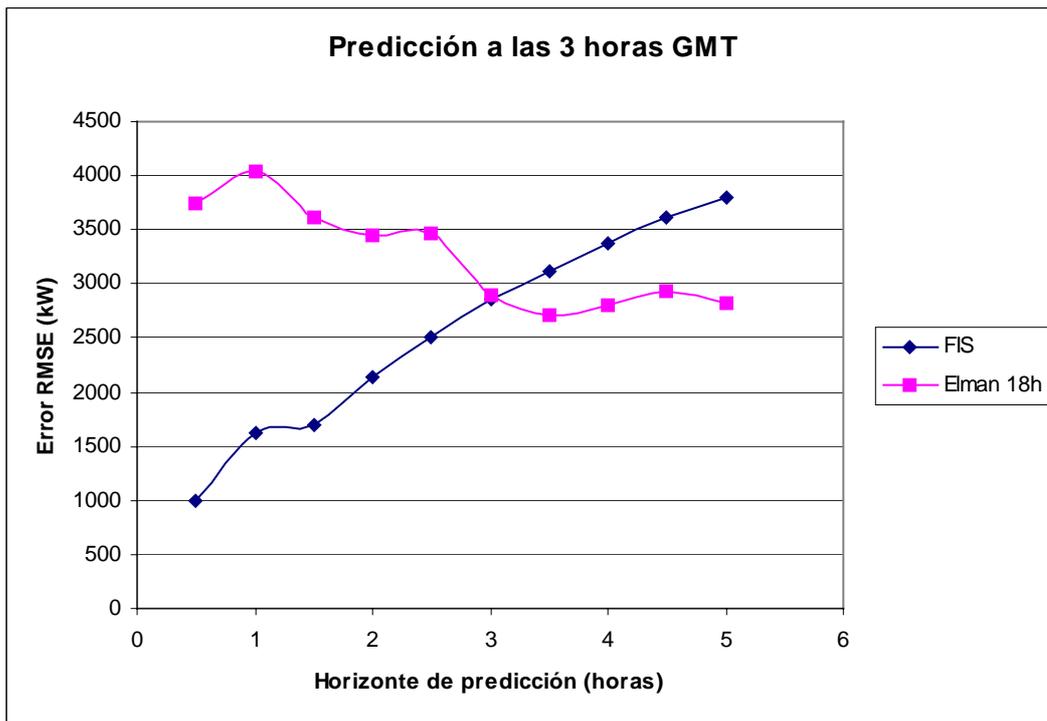


Figura 8.28. Errores RMSE de predicción de la potencia eléctrica media en la subestación para predicciones realizadas a las 3 horas GMT.

Para los siguientes horizontes de predicción, es idóneo usar la red neuronal Elman 1 (empleando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior), para las predicciones hasta las 18 horas del día actual. Para predicciones entre las 18.5 horas del día actual y las 18 horas del día siguiente, el

modelo adecuado es la red neuronal Elman 2, usando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior. Por último para predicciones a partir de las 18.5 horas del día siguiente y hasta las 18 horas de dos días después el modelo que ofrece mejores resultados es la red neuronal Modular, usando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior.

Para el caso de realizar estas predicciones a las 10, 16 ó 22 horas sólo habrá que sumar 6, 12 y 18 horas sobre las referencias temporales señaladas en los dos párrafos anteriores.

6. Predicción realizada a las 5 horas (o a las 11, 17 o 23 horas).

En este caso, a las 5 horas se dispone de las predicciones del modelo atmosférico MM5 de las 18 horas del día anterior (todavía no están procesadas completamente las correspondientes a los datos de las 0 horas del mismo día, salvo que se dispusiera de un número elevado de ordenadores para ejecutar en paralelo el programa que implementa el modelo MM5). La Figura 8.30 representa los errores RMSE en la predicción de la potencia eléctrica media en la subestación, con los datos del grupo de testeo, para los modelos FIS y para la red neuronal Elman 1 con una antigüedad de 11 horas, puesto que fueron realizadas con las predicciones del modelo atmosférico MM5 empleando los datos de las 18 horas del día anterior (Elman 18h). Aquí, como en el caso anterior, los errores con la red neuronal Elman 1, usando las predicciones del modelo MM5 realizadas con los datos a las 18 horas del día anterior, superan a los modelos FIS para un horizonte de predicción de 3 horas.

Para los siguientes horizontes de predicción, es idóneo usar la red neuronal Elman 1 (empleando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior), para las predicciones hasta las 18 horas del día actual. Para predicciones entre las 18.5 horas del día actual y las 18 horas del día siguiente, el modelo adecuado es la red neuronal Elman 2, usando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior. Por último para predicciones a partir de las 18.5 horas del día siguiente y hasta las 18 horas de dos días después el modelo que ofrece mejores resultados es la red neuronal Modular, usando las predicciones del modelo atmosférico MM5 con los datos de las 18 horas del día anterior.

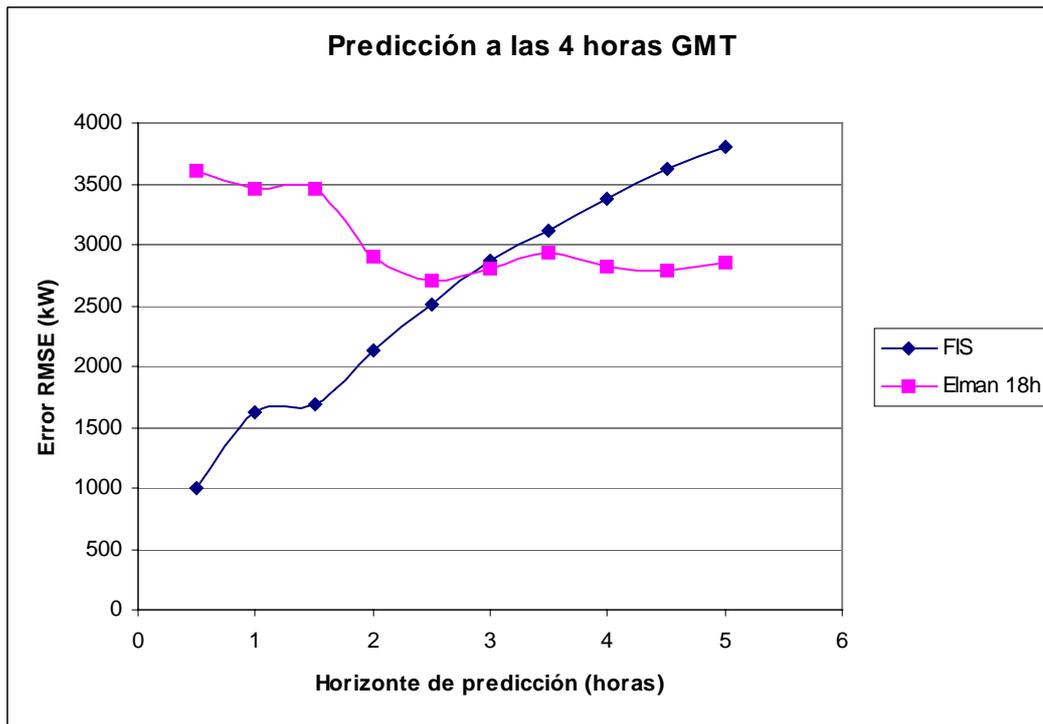


Figura 8.29. Errores RMSE de predicción de la potencia eléctrica media en la subestación para predicciones realizadas a las 4 horas GMT.

Para el caso de realizar estas predicciones a las 11, 17 ó 23 horas sólo habrá que sumar 6, 12 y 18 horas sobre las referencias temporales señaladas en los dos párrafos anteriores.

8.5.6. Aplicación del nuevo sistema global de predicción.

Uno de los intereses principales que suscita una herramienta para la predicción producción eléctrica en el gestor de un parque eólico, es su posible aplicación para la elaboración de ofertas de energía en el mercado eléctrico. Como vimos en el Capítulo 2, estas ofertas se realizan en forma de “paquetes” de energía eléctrica sobre una base horaria, es decir, que se oferta la energía eléctrica que se espera producir hora a hora. Así, la variable que resulta interesante predecir es la potencia eléctrica media horaria, hora a hora, que suministra el parque eólico al sistema eléctrico.

En España, el mercado diario eléctrico centra el interés de los agentes que intervienen en el mismo, ya que en él se realizan habitualmente las transacciones de compra y venta de energía eléctrica. Además, para poder participar en las sesiones del mercado intradiario, es condición haber participado antes en el mercado diario.

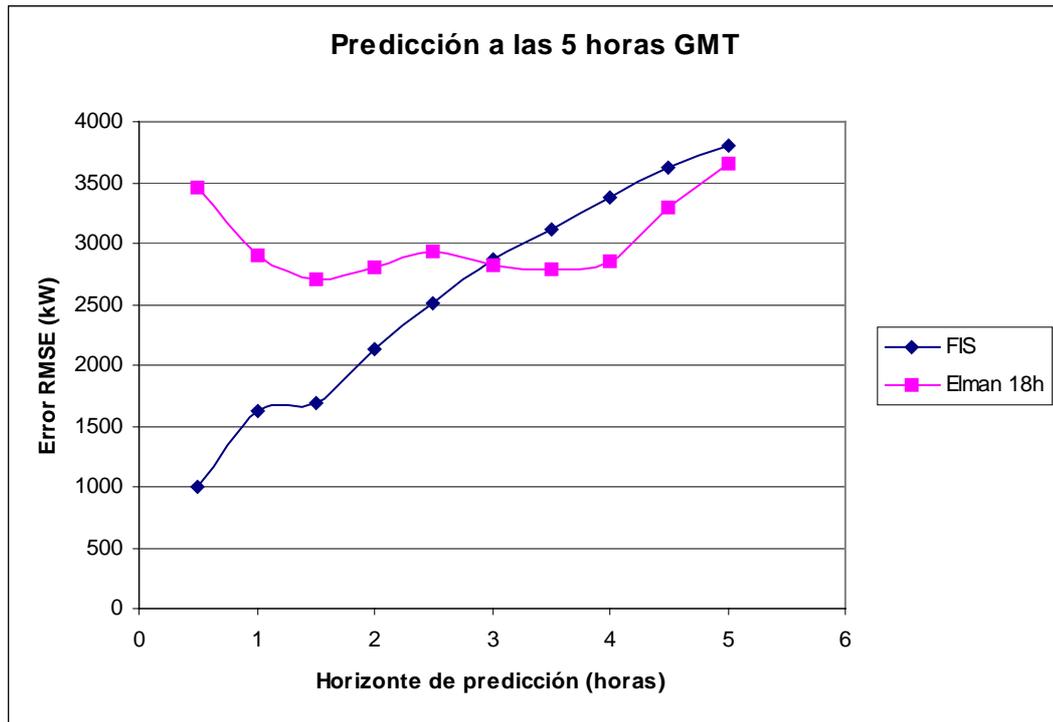


Figura 8.30. Errores RMSE de predicción de la potencia eléctrica media en la subestación para predicciones realizadas a las 5 horas GMT.

Las ofertas en el mercado diario, sobre la base de la potencia media horaria hora a hora para el día siguiente, se deben presentar antes de las 10 horas. En el sistema de predicción propuesto se realizan las predicciones en intervalos de media hora, de forma que habrá que obtener la media de las predicciones correspondientes a dos intervalos consecutivos con objeto de obtener la predicción para toda una hora. Ahora bien, las predicciones con los modelos analizados están centradas en el valor del horizonte de predicción, esto es, la predicción obtenida para las 6.5 horas corresponde a la potencia eléctrica media entre las 6:15 y las 6:45. Para obtener el valor de potencia eléctrica media horaria sólo habrá que aplicar la ecuación (8.34), donde $\hat{p}_h(x)$ representa la predicción de potencia eléctrica media en la subestación el x intervalo horario (entre las $x-1$ y x horas), y donde $p(x)$ representa el valor de predicción con el modelo propuesto para la hora x .

$$\hat{p}_h(x) = 0.25 \cdot p(x-1) + 0.5 \cdot p(x-0.5) + 0.25 \cdot p(x) \quad (8.34)$$

Dado que las ofertas deben ser presentadas antes de las 10 horas (hora civil), lo que corresponde a las 8 ó 9 horas GMT, según sea horario de verano u horario de

invierno, la última predicción disponible del modelo atmosférico MM5 es la de las 0 horas GMT. La Tabla 8.29 recoge los errores RMSE en la predicción de la potencia eléctrica media horaria en la subestación, tanto en kW, como su porcentaje con respecto a la potencia eléctrica instalada en el parque eólico, utilizando el modelo propuesto para ese horizonte de predicción: la red neuronal Elman 2 con las predicciones del modelo atmosférico MM5 de las 0 horas GMT del día anterior.

| Hora | Error RMSE (kW) | Error RMSE (%) | Hora | Error RMSE (kW) | Error RMSE (%) |
|------|--------------------|-------------------|------|--------------------|-------------------|
| 1 | 3077.05 | 14.25% | 13 | 3004.36 | 13.91% |
| 2 | 3500.63 | 16.21% | 14 | 2620.39 | 12.13% |
| 3 | 3720.83 | 17.23% | 15 | 2404.54 | 11.13% |
| 4 | 4106.54 | 19.01% | 16 | 2525.82 | 11.69% |
| 5 | 3929.69 | 18.19% | 17 | 3062.85 | 14.18% |
| 6 | 4296.11 | 19.89% | 18 | 3187.78 | 14.76% |
| 7 | 4798.85 | 22.22% | 19 | 3581.08 | 16.58% |
| 8 | 5003.15 | 23.16% | 20 | 3240.69 | 15.00% |
| 9 | 5078.74 | 23.51% | 21 | 2569.95 | 11.90% |
| 10 | 4321.20 | 20.01% | 22 | 2371.43 | 10.98% |
| 11 | 3332.57 | 15.43% | 23 | 2672.59 | 12.37% |
| 12 | 3191.10 | 14.77% | 24 | 3608.62 | 16.71% |

Tabla 8.29. Errores RMSE de predicción de la potencia eléctrica media horaria en la subestación para predicciones realizadas entre las 5 y las 8 horas GMT del día anterior.

El error RMSE medio para la predicción de la potencia eléctrica media horaria (en relación con las ofertas al mercado eléctrico) es de 4029.7 kW, lo que representa el 18.65% de la potencia eléctrica instalada en el parque eólico.

En cuanto a la participación en los mercados intradiarios, la Tabla 8.30 recoge los momentos finales para presentación de ofertas, las últimas predicciones disponibles del modelo atmosférico MM5 para ese momento y los modelos de predicción propuestos para los intervalos horarios cubiertos por cada sesión de mercado intradiario. Se puede observar que para los mercados intradiarios 3 y 6 no va a existir nuevas ofertas, ya que estas serían coincidentes con las de las sesiones 2 y 3 respectivamente (se trata de los mismos modelos de predicción, con los mismos datos, por lo que las predicciones serían iguales).

8.5.6.1. Valor económico de la predicción.

Como caso ilustrativo del extraordinario interés económico actual de la predicción a corto plazo de las potencias eléctricas en parques eólicos, se va a calcular el beneficio bruto obtenido por la empresa propietaria de un parque eólico, al participar

en el mercado eléctrico español, en los días correspondientes a los datos del grupo de testeo (un total de 700 horas, de las 23 horas del día 9 de julio de 2005, hasta las 2 horas del día 8 de agosto de 2005). Este valor se va a comparar con el beneficio bruto obtenido si esta empresa no hubiese acudido al mercado eléctrico y, por tanto, hubiese cedido la energía producida a una empresa distribuidora y fuera retribuida en forma de tarifa regulada. No se van a tener en cuenta impuestos, ni complementos por energía reactiva, ni los costes de la regulación secundaria y terciaria.

| Intradiario | Hora final | Ultima predicción MM5 disponible | Modelo |
|-------------|------------|----------------------------------|---|
| 1 | 17:45 | 6h GMT mismo día | Elman 1 hasta las 6 horas GMT del día siguiente Elman 2 desde las 6 horas GMT hasta las 24 horas del día siguiente |
| 2 | 21:45 | 12h GMT mismo día | Elman 1 hasta las 12 horas GMT del día siguiente Elman 2 desde las 12 horas GMT hasta las 24 horas del día siguiente |
| 3 | 1:45 | 18h GMT día anterior | Elman 1 hasta las 18 horas GMT del día siguiente Elman 2 desde las 18 horas GMT hasta las 24 horas del día siguiente |
| 4 | 4:45 | 18h GMT día anterior | Elman 1 hasta las 18 horas GMT del día siguiente Elman 2 desde las 18 horas GMT hasta las 24 horas del día siguiente |
| 5 | 8:45 | 0h GMT mismo día | Elman 1 hasta las 24 horas del día |
| 6 | 12:45 | 0h GMT mismo día | Elman 1 hasta las 24 horas del día |

Tabla 8.30. Modelos idóneos para las predicciones de la potencia eléctrica media horaria en la subestación, y para participar en sesiones del mercado intradiario.

A) Tarifa regulada.

En el caso de que el titular del parque eólico hubiese decidido mantener su instalación en la opción a, del apartado 1 del artículo 22 del Real Decreto 436/2004 (venta de la electricidad a la empresa distribuidora de energía eléctrica), el precio de venta de esa energía estaría expresado en forma de tarifa regulada, única para todos los periodos de programación (para todas las horas). Esta tarifa regulada consiste en un porcentaje (para los primeros años es del 90%) de la tarifa eléctrica media o de referencia, que se publica en el real decreto que establece la tarifa eléctrica cada año.

El titular del parque debe comunicar a la empresa distribuidora la previsión de la energía que va a ceder a la red en cada uno de los periodos de programación del mercado de producción de energía eléctrica, es decir, las predicciones de generación para cada uno de los 24 periodos horarios de cada día con un mínimo de 30 horas de anticipación al inicio de dicho día. El modelo a utilizar, dentro de

los desarrollados, corresponde a la red neuronal Modular (apartado 8.5.2.3 de este documento), dado que se trata de realizar las predicciones para pasado mañana.

El RD 2392/2004 establece la tarifa eléctrica para el año 2005, año al que pertenecen los datos disponibles y utilizados en el desarrollo de los modelos de predicción. El citado real decreto fija la tarifa media o de referencia en 7.3304 céntimos de euro por kWh. El RD 436/2004, vigente en agosto de 2005, estipula una tolerancia del 20% para la producción prevista, de forma que si la producción real entra dentro de la banda de la producción prevista $\pm 20\%$ no existirá penalización por desvío; en cambio, si existirá penalización en el caso de que la producción real, hora a hora, supere la banda de tolerancia anterior. La penalización por desvío para cada hora, siempre que el desvío sea superior al 20% citado, será del 10% de la tarifa regulada multiplicada por el valor absoluto del desvío.

En la Tabla 8.31 se recoge el valor de los ingresos económicos brutos obtenidos con el modelo de predicción propuesto (red neuronal Modular), en el caso de que el propietario del parque eólico hubiera elegido ceder la producción eléctrica del mismo a una empresa distribuidora y no realizase correcciones para las sesiones del mercado intradiario.

| Modelo | Ingresos brutos | % desvíos | Diferencia |
|-------------------------|------------------------|------------------|-------------------|
| Propuesto (red Modular) | 178259.82 € | 7.57% | |
| media 1 | 167285.14 € | 13.26% | -10974.68 € |
| media 2 | 168398.15 € | 12.68% | -9861.67 € |
| persistente 1 | 173314.16 € | 10.14% | -4945.66 € |
| persistente 2 | 169767.51 € | 11.97% | -8492.31 € |
| nula | 172976.58 € | 10.31% | -5283.24 € |
| nueva referencia | 172299.01 € | 10.66% | -5960.81 € |
| ARMA (1,1) | 171425.16 € | 11.12% | -6834.66 € |

Tabla 8.31. Ingresos brutos por presentar predicciones a la empresa distribuidora, porcentaje que representa el coste de los desvíos y su diferencia con respecto al modelo propuesto.

En la Tabla 8.31 se han recogido los ingresos brutos obtenidos por la comunicación a la empresa distribuidora de previsiones de generación eléctrica con los valores de predicción obtenidos usando el sistema global de predicción. También la Tabla 10.31 recoge el porcentaje que representa el coste de los desvíos frente a los ingresos brutos y su diferencia con respecto al modelo propuesto (red Modular). Dentro de este conjunto algunos son triviales (propios de

quien no ha desarrollado ningún modelo de predicción) como son el “media 1”, “media 2”, “persistente 1”, “persistente 2” y “nula”.

El modelo “media 1” asigna como valor de predicción, para cada hora, el valor medio de la potencia media horaria de los datos de los grupos de entrenamiento y validación (en este caso el valor de la potencia media horaria de los cuatro primeros meses del grupo de datos utilizado en el desarrollo de los modelos).

El modelo “media 2” asigna como valor de predicción, para cada hora, el valor medio de la potencia media horaria de las 24 últimas horas en el momento de realizar la predicción (las 18 horas) y se toma como valor de predicción para todas las horas de dos días más tarde.

El modelo “persistente 1” asigna como valor de predicción para cada hora la potencia media horaria de la última hora con valor conocido (la hora entre las 17 y las 18 horas), realizándose la predicción con ese valor para las 24 horas de dos días después.

El modelo “persistente 2” asigna como valor de predicción, para cada hora, la potencia media horaria del último día con valores conocidos (el día anterior al de realización de la predicción), y dando esos valores como predicción para dos días después.

El modelo “nula” asigna un valor de predicción nulo para cada hora. Este modelo solamente da una idea de los ingresos obtenidos por la venta de la energía eléctrica producida aun aplicándose a todas las horas la máxima penalización por desvíos entre lo programado y lo real.

En la Tabla 8.31 también se presentan los ingresos brutos obtenidos al utilizar como ofertas de energía eléctrica las predicciones de los modelos nueva referencia y ARMA(1,1). Como se puede observar, en todos los casos, los ingresos brutos obtenidos con el sistema global de predicción propuesto son mayores a los conseguidos con los otros modelos. Evidentemente, las diferencias justifican claramente la inversión de los gestores de parques eólicos en sistemas de predicción de la potencia eléctrica media generada en el parque para ser comunicadas como previsiones a las empresas distribuidoras en el caso de haber optado por la cesión de la energía eléctrica producida a una de ellas.

B) Mercado eléctrico

En el caso de que el titular del parque eólico hubiese decidido acoger su instalación a la opción b, del apartado 1 del artículo 22 del Real Decreto 436/2004, debería realizar la venta de la electricidad libremente al mercado eléctrico. Supondremos que el titular del parque eólico ha realizado las ofertas en el mercado diario con los valores de energía eléctrica producida en cada hora iguales a las predicciones calculadas por el sistema global de predicción propuesto y ofertadas a precio cero (para asegurar su venta). El modelo utilizado en este caso sería la red neuronal Elman 2 con las predicciones de velocidad y dirección medias del viento del modelo MM5 para el día siguiente (horizontes de predicción de 24.5 a 48 horas).

Tras el cálculo de los desvíos producidos en cada uno de los periodos de programación horarios, y la aplicación de la correspondiente penalización (para este caso se ha considerado un 10% del precio marginal para la hora correspondiente, de acuerdo al cálculo de garantías extraordinarias del apartado 23.6.5.3 de Reglas de funcionamiento del mercado diario e intradiario de producción de energía eléctrica, BOE 26 de febrero de 2003, vigente en el periodo evaluado), y aplicar la prima y el incentivo vigentes en el periodo evaluado, los ingresos brutos ascienden a 266301.12 € lo cual representa 88041.3 € más que habiendo escogido la opción de no ofertar la energía eléctrica producida al mercado. Para esta estimación de ingresos brutos no se han tenido en cuenta el coste de la regulación secundaria ni terciaria, ni de la gestión de los desvíos.

En la Tabla 8.32 se recoge el valor de los ingresos económicos brutos obtenidos con el modelo de predicción propuesto, para ofertas en el mercado diario y con otros modelos de predicción simples. También se han recogido los ingresos brutos obtenidos por la presentación de ofertas al mercado eléctrico con los valores de predicción obtenidos con un conjunto de modelos. Esos modelos incluyen los de “media 1”, “media 2”, “persistente 1”, “persistente 2” y “nula”, ya presentados para el caso de venta de la energía producida a una empresa distribuidora. Evidentemente en este caso los modelos “media” y “persistente” consideran los últimos valores conocidos en el momento de realizar la predicción. Para todos los modelos las predicciones son para el día siguiente al del momento de realizar la predicción (entre las 9 y 10 horas), con objeto de presentar las ofertas antes de la hora de cierre del mercado diario.

| Modelo | Ingresos brutos | % desvíos | Diferencia |
|-------------------------|------------------------|------------------|-------------------|
| Propuesto (red Elman 2) | 266301.12 € | 3.61% | |
| media 1 | 253376.46 € | 8.90% | -12924.67 € |
| media 2 | 255731.74 € | 7.90% | -10569.38 € |
| persistente 1 | 259569.50 € | 6.30% | -6731.62 € |
| persistente 2 | 256419.84 € | 7.61% | -9881.28 € |
| nula | 259277.96 € | 6.42% | -7023.16 € |
| nueva referencia | 258606.74 € | 6.70% | -7694.38 € |
| ARMA (1,1) | 259033.65 € | 6.52% | -7267.47 € |

Tabla 8.32. Ingresos brutos por presentar ofertas en el mercado diario, porcentaje que representan las penalizaciones por desvíos y su diferencia con respecto al modelo propuesto.

Como se puede observar, en todos los casos, los ingresos con el modelo propuesto (red Elman 2) son mayores a los conseguidos con los otros modelos. Los ingresos obtenidos justifican el interés de los gestores de parques eólicos en la utilización de sistemas de predicción de la potencia eléctrica media generada en un parque eólico para ser utilizada en la elaboración de ofertas al mercado eléctrico. Por otra parte, también se justifica (al menos con los precios marginales para el periodo evaluado) la selección de la opción de venta de la energía eléctrica producida al mercado eléctrico frente a la opción de venta a una empresa distribuidora (la diferencia en los ingresos es mayor a 88000 € a favor de la primera opción).

Si en lugar de considerar las reglas de mercado vigentes en el periodo evaluado (un mes de verano de 2005), hubiésemos considerado las reglas vigentes a finales de 2006, la diferencia entre las dos opciones se hubiese reducido a unos 70000 €, al haber aumentado la penalización por desvíos para las ofertas presentadas al mercado eléctrico (un 27% del precio marginal para cada hora, frente al 10% anteriormente considerado). Pero en este caso también aumenta la diferencia de ingresos obtenidos con el modelo de predicción propuesto frente a los otros modelos estudiados para la presentación de ofertas al mercado.

8.6. CONCLUSIONES

Se han presentado cinco problemas diferentes de predicción, en el corto plazo, de la producción de energía eléctrica en parques eólicos. Los problemas difieren significativamente entre sí debido a la diferente disponibilidad de datos para realizar las predicciones.

En el primer problema planteado únicamente se dispone de la serie temporal de

velocidades medias horarias del viento para un punto situado en el terreno abarcado por un parque eólico. A partir de esa serie temporal, se ha desarrollado un modelo capaz de predecir, en primer lugar, la velocidad media horaria del viento para la siguiente hora. Después, con la predicción de la velocidad media horaria del viento, se determina la potencia eléctrica media generada para la siguiente hora en el parque eólico.

Este primer problema corresponde a una situación real de un parque eólico de nueva instalación, del cual únicamente se dispone de la serie temporal de velocidad media horaria del viento obtenida por una estación de medida, instalada para la evaluación del propio recurso eólico en esa localización.

El segundo problema planteado corresponde a la predicción de la producción total de energía eléctrica diaria en un parque eólico. El modelo de predicción propuesto está basado en la técnica de series temporales fuzzy. Con esta técnica se recoge información de tipo lingüístico para mejorar la calidad de la predicción. La única información utilizada en la creación del modelo de predicción ha sido la serie temporal de la energía eléctrica diaria producida y la información meteorológica, para el día siguiente, en la región en que se encuentra el parque eólico (información no numérica, sino lingüística).

Este segundo problema corresponde a una situación real de un parque eólico, cuando los datos disponibles son mínimos para el desarrollo de modelos de predicción.

El tercer problema planteado corresponde a la predicción de la potencia eléctrica media generada en un parque eólico para un horizonte de unas pocas horas. La información disponible corresponde a la serie temporal de potencia eléctrica media horaria generada en el mismo. Se ha creado un nuevo modelo, basado en modelos locales de regresión lineal múltiple, que ofrece mejores resultados que los métodos habituales de predicción basados en técnicas de “soft computing” (redes neuronales o sistemas de inferencia difusa), cuando los datos disponibles son escasos y la posibilidad de entrenamiento o aprendizaje de los modelos de predicción es limitada.

Este tercer problema planteado aparece cuando la serie temporal de potencia eléctrica media horaria disponible se limita a unos centenares de registros. Esta situación puede ocurrir, por ejemplo, en los primeros días de funcionamiento de un parque eólico.

El cuarto problema planteado corresponde a la predicción de la potencia

eléctrica media horaria generada, en el corto plazo, en un parque eólico, cuando existen predicciones de velocidad y dirección media horaria del viento proporcionadas por un modelo atmosférico. La limitación aparece en la serie temporal de predicciones del modelo atmosférico, cuando comprende sólo los datos de unas decenas de días, lo que impide desarrollar modelos que precisen aprendizaje a partir de esas predicciones atmosféricas (redes neuronales y sistemas neuro-difusos). Aun así, ha sido posible desarrollar estos modelos entrenándolos con valores reales medidos en el parque eólico, y realizar las predicciones de producción eléctrica utilizando las predicciones del modelo atmosférico. Además la variable de predicción ha sido el factor de capacidad, lo cual contribuye a que el propio gestor del parque mejore la calidad de la predicción de la potencia eléctrica media horaria.

El cuarto problema corresponde a un parque eólico para el que se comienzan a elaborar algunas predicciones de variables atmosféricas. Se dispone de información de variables medidas en el propio parque y registrada en los sistemas SCADA del mismo (potencia eléctrica media generada, velocidad y dirección media del viento, etc.) desde mucho antes de comenzar las predicciones con el modelo atmosférico.

El quinto problema planteado corresponde al de la predicción de la potencia eléctrica media vertida por un parque eólico en la red eléctrica, cuando se dispone de información de la propia serie de potencia eléctrica media y de predicciones de un modelo atmosférico. Se ha desarrollado un nuevo sistema global de predicción y se han determinado los instantes en que han de utilizarse los modelos (que integran aquel sistema global de predicción) para obtener los menores errores posibles. Los horizontes de predicción máximos han sido de 72 horas.

Este quinto problema se refiere también a situaciones habituales en las que el gestor de un parque eólico desea suministrar, bien a la empresa distribuidora, o bien al mercado eléctrico, sus predicciones de producción eléctrica para el corto plazo, y tiene disponibles predicciones de un modelo atmosférico. Respecto al cuarto problema antes indicado, en este quinto problema se dispone de una serie temporal suficiente de predicciones del modelo atmosférico para ser utilizadas como variables de entrada de los modelos de predicción.

Junto a los resultados del quinto problema, se ha presentado la evaluación del valor económico de las predicciones, para el caso haber realizado las ofertas al mercado eléctrico usando los valores de predicción obtenidos con el nuevo sistema global de predicción (opción de venta al mercado de la energía eléctrica producida), o para el caso

de venta a empresa distribuidora. En ambos casos, los beneficios brutos obtenidos con la venta de la energía eléctrica producida por el parque eólico son mayores que los que se obtendrían en el caso de no contar con las predicciones. Este ejemplo ilustra el interés de la predicción de la producción eléctrica (bajo un punto de vista económico) en parques eólicos mediante modelos de predicción como los creados en nuestros trabajos de investigación.

En los cinco problemas se han desarrollado un conjunto de modelos de predicción, correspondiendo una parte de ellos a los presentados en los artículos publicados en la literatura internacional. En todos los casos, los modelos originales desarrollados y propuestos en los trabajos de investigación de este documento presentan características significativamente diferenciadoras (respecto a los publicados) y logran un menor error de predicción.

9. CONCLUSIONES

En este capítulo se presentan las conclusiones generales de la tesis doctoral en cuanto a los nuevos modelos y sistemas de predicción creados y al análisis de los resultados logrados. También se detallan las aportaciones originales realizadas en los trabajos de investigación. Se completa este capítulo señalando futuras líneas de investigación que se pueden llevar a cabo a partir de los modelos y de los resultados de investigación obtenidos.

9.1. RESUMEN DE LA INVESTIGACIÓN Y CONCLUSIONES

En el Capítulo 1 se ha realizado una breve introducción al tema central de la tesis doctoral y se ha descrito el objetivo principal de los trabajos de investigación realizados: el desarrollo de modelos avanzados para la predicción, a corto plazo, de la producción eléctrica en parques eólicos. También se han enumerado los objetivos específicos perseguidos con dichos trabajos de investigación.

En el Capítulo 2 se ha descrito el fenómeno meteorológico del viento, sus causas y la potencia eléctrica que se puede extraer de él. Se ha comentado la evolución del aprovechamiento de la energía eólica que ha realizado el hombre, hasta llegar a la situación actual, en la que se está a punto de alcanzar los 160 GW de potencia eléctrica instalada a nivel mundial sólo en parques eólicos, correspondiendo a España el 12.5% de esa cantidad. Se han analizado las repercusiones de la integración de los parques eólicos en la Red Eléctrica, centrándonos particularmente en el caso español con análisis de las regulaciones introducidas por las autoridades en los últimos años. Se ha puesto de relieve la importancia de la predicción de la producción eléctrica en parques eólicos en el corto plazo, tanto por imposición normativa como por interés económico, comentando además varios estudios publicados donde se analiza el valor económico de la predicción.

En el Capítulo 3 se ha realizado una revisión bibliográfica de los modelos de predicción de la producción eléctrica en parques eólicos, indicando sus limitaciones. En general, los modelos publicados están “focalizados” para un intervalo concreto de su horizonte máximo de predicción. Se ha realizado una clasificación de los modelos publicados en la literatura internacional en función de si requieren, o no, predicciones

previas obtenidas con modelos atmosféricos. Se ha puesto de manifiesto que los modelos de predicción de la producción eléctrica basados en series temporales presentan mejores valores de predicción, que los basados en predicciones previas de modelos atmosféricos, para horizontes de predicción en el muy corto plazo. También se ha concluido que todos los trabajos publicados utilizan el modelo persistente como modelo con el que comparar las predicciones. No se ha publicado la descripción de ningún modelo con un horizonte máximo de predicción por encima de las 48 horas. Sólo en dos artículos publicados se han utilizado modelos atmosféricos no dependientes de institutos meteorológicos. En uno de esos artículos (fruto de los trabajos de investigación recogidos en este documento) se ha utilizado un modelo atmosférico (modelo MM5) cuyos valores de predicción son de libre acceso (sin coste económico), dado que dicho modelo atmosférico no pertenece a ningún instituto meteorológico nacional.

En el Capítulo 4 se han presentado las características de los modelos numéricos de predicción meteorológica (modelos atmosféricos). Los modelos atmosféricos más adecuados, que obtienen predicciones de variables meteorológicas de interés para la predicción en el corto plazo de la producción eléctrica en un parque eólico, son los modelos de meso-escala, que combinan una alta resolución espacial con un alcance o validez de las predicciones dentro del corto plazo. El modelo atmosférico MM5, cuyas predicciones han sido utilizadas en las experiencias recogidas en el Capítulo 8, corresponde a un modelo de meso-escala de dominio público, que goza de crédito dentro de la comunidad científica especializada. Cuenta además con una ventaja, de aplicación práctica relevante: puede implementarse, a coste cero, en sistemas informáticos de uso común, pudiendo ofrecer predicciones de las variables meteorológicas con poco esfuerzo de programación (incluso podrían instalarlo los propietarios o gestores de los parques eólicos).

En el Capítulo 5 se han mostrado las técnicas usadas para el desarrollo de modelos de predicción presentados en los trabajos publicados en la literatura internacional. Estas técnicas también han sido utilizadas para el desarrollo de los nuevos modelos de predicción creados en los trabajos de investigación de este documento. Las técnicas presentadas son las de series temporales (Box-Jenkins), las redes neuronales artificiales y los sistemas de inferencia difusa. Para cada una de estas técnicas se analizan sus ventajas y desventajas, y su aplicación en el diseño de modelos de predicción, a corto plazo, de la producción eléctrica en parques eólicos.

En el Capítulo 6 se ha presentado una técnica de optimización utilizada en el

desarrollo de los modelos de predicción constituidos por redes neuronales artificiales y por sistemas de inferencia difusa. Esta técnica está basada en algoritmos genéticos. En este capítulo se describen los parámetros de los modelos de predicción, creados con redes neuronales, cuyos valores son optimizados, así como un original método para la optimización de sistemas de inferencia difusa mediante un nuevo algoritmo genético.

En el Capítulo 7 se han descrito los datos disponibles y utilizados para la realización de los trabajos de investigación recogidos en este documento, así como los modelos originales de predicción que se han desarrollado para obtener, a partir de esos datos, los mejores valores de predicción para la producción eléctrica de parques eólicos en el corto plazo. Se presentan dichos modelos originales, pertenecientes a las dos familias de modelos de predicción presentadas en el Capítulo 3. Con este conjunto de nuevos modelos se realiza la predicción de producción eléctrica, incluso utilizando una mínima cantidad de datos sobre el comportamiento histórico del parque eólico. La información disponible ha sido de diverso tipo y la cantidad de datos ha sido variable: así se ha utilizado una serie temporal de velocidad media horaria del viento; o bien una serie temporal de energía eléctrica diaria producida y predicciones lingüísticas para la velocidad media del viento (flojo, fuerte, etc.) para el día siguiente; o bien una serie temporal de la potencia eléctrica media horaria producida en el parque eólico; o bien una serie temporal de la potencia eléctrica media producida y predicciones de un modelo atmosférico para el emplazamiento del parque eólico.

En el Capítulo 8 se han presentados los resultados obtenidos, con la aplicación de los modelos presentados en el Capítulo 7, para la predicción de la producción eléctrica, en el corto plazo, de parques eólicos. En todos los casos se han desarrollado un conjunto de modelos diferentes para realizar la predicción (parte de ellos como los presentados en la literatura internacional), siendo los resultados obtenidos con los modelos originales propuestos superiores a los obtenidos con los otros modelos publicados.

Para el caso planteado con mayor información disponible (serie temporal con potencia eléctrica media entregada en la subestación y predicciones completas de un modelo atmosférico), se ha creado un original sistema global de predicción (compuesto por diferentes modelos de predicción) que cubre un horizonte de predicción de hasta 72 horas. También se ha logrado determinar el modelo idóneo (de entre los que integran el sistema global) que conviene utilizar para cada horizonte de predicción y, además, en función del instante en el que se realiza la predicción. Por último, se ha ilustrado la determinación del valor económico de la predicción de la potencia eléctrica,

mostrándose el incremento del beneficio económico obtenido por el gestor del parque eólico, en la venta de la energía eléctrica producida en dicho parque eólico, bajo opciones de venta posibles en España.

9.2. PRINCIPALES APORTACIONES DE LA TESIS

Como principales aportaciones originales de los trabajos de investigación, merecen destacarse las indicadas a continuación.

1. Se han desarrollado un conjunto de nuevos modelos de predicción de la producción eléctrica en un parque eólico cualquiera, adaptables a los datos disponibles del propio parque. Así, se han creado modelos capaces de realizar predicciones en el corto plazo utilizando uno de los grupos de datos, referidos al parque eólico, reseñados a continuación.
 - Serie temporal de velocidad media horaria del viento para el emplazamiento del parque eólico.
 - Serie temporal de energía eléctrica diaria generada en el parque eólico, acompañada, o no, de información lingüística (que recoge incertidumbres) de la velocidad media del viento para el día siguiente y para la región donde se encuentra el parque eólico.
 - Serie temporal de potencia eléctrica media horaria.
 - Serie temporal de la potencia eléctrica media horaria y serie temporal de predicciones de variables meteorológicas de un modelo atmosférico para el emplazamiento del parque eólico.
2. Se han desarrollado modelos originales, que constituyen un nuevo sistema global de predicción, con un horizonte de hasta 72 horas, ampliando así en 24 horas los modelos publicados en la literatura internacional, y también cubriendo necesidades de predicción de la producción eléctrica en parques eólicos dentro del contexto de la legislación española.

Un sistema con este horizonte de predicción es fundamental para poder

realizar las previsiones de generación de energía eléctrica sobre una base horaria, ya que el propietario del parque eólico tendrá que disponer de las mismas de acuerdo a los términos señalados en la legislación española. Ésta puede obligar al propietario del parque eólico a presentar las previsiones de energía generada, sobre una base horaria, con una anticipación mínima de 30 horas al inicio del día correspondiente a las previsiones.

Para este sofisticado sistema global de predicción, compuesto por un amplio conjunto de modelos y que cubre un horizonte de predicción de hasta 72 horas, se ha determinado, en función del momento en que se realiza la predicción (y del horizonte de predicción), el modelo óptimo que logra los mejores resultados de predicción de la potencia eléctrica generada en el parque eólico. Así, el sistema global incluye modelos específicamente desarrollados para realizar predicciones en el muy corto plazo, modelos para realizar predicciones que sirvan para presentar ofertas de venta de la energía eléctrica producida en el mercado diario o en las sesiones del mercado intradiario, o modelos para las realizar predicciones con una anticipación mínima de 30 horas al inicio del día correspondiente a las previsiones, que comentábamos en el párrafo anterior.

3. Se ha creado un nuevo modelo de predicción para la velocidad media horaria del viento, en el emplazamiento de un parque eólico, basado en la predicción de los coeficientes futuros de la transformada discreta de Fourier, aplicada sobre los últimos 24 valores conocidos de la serie temporal de la velocidad media horaria. Los resultados obtenidos superan a los conseguidos con otros modelos usuales de predicción.
4. Se ha aplicado de forma original una técnica de series temporales fuzzy, FTS, en la predicción de la energía eléctrica diaria generada en un parque eólico. Sobre la base de la información (en forma lingüística) suministrada por los medios de comunicación, referente a las características generales del viento (flojo, fuerte, etc.) previsto para el día siguiente en la región donde se encuentra el parque eólico, se ha conseguido mejorar la calidad de la predicción (reducción del error de predicción) con respecto a otros modelos de predicción que no utilizan tal información lingüística. También esta nueva aplicación de FTS permite recoger las incertidumbres asociadas a las predicciones meteorológicas lo que redundará positivamente en la predicción de la potencia eléctrica generada en parques eólicos.

5. Se ha creado nuevo sistema de predicción de series temporales, aplicado a la predicción de la potencia eléctrica media horaria, que combina óptimamente las salidas de varios módulos de regresión lineal múltiple. Los coeficientes de ponderación utilizados para combinar dichas salidas, así como los datos utilizados en la construcción de cada uno de los módulos de regresión lineal, son determinados por la aplicación del agrupamiento substractivo (sobre los valores del vector de entradas). A la hora de construir los módulos de regresión lineal, el aumento del peso relativo, para aquellos valores del vector de entradas más cercanos al centro del cluster, permite reducir el error de predicción. Los resultados de predicción de la potencia eléctrica media horaria obtenidos mediante el sistema creado, superan a los obtenidos con los modelos de predicción más usuales.
6. Para aquellos casos en que se comienzan a realizar las primeras predicciones utilizando modelos atmosféricos para un parque eólico, y se dispone de series temporales históricas de la potencia eléctrica media horaria y de los valores medidos de velocidad y dirección media horaria del viento, se ha desarrollado un nuevo modelo de predicción, basado en redes neuronales artificiales, que permite entrenar las redes neuronales exclusivamente con los valores de las mencionadas series temporales históricas disponibles. Para predicciones de la potencia eléctrica media horaria, se usan además los valores de predicción del modelo atmosférico, a pesar de no haber sido utilizados en el entrenamiento de las redes neuronales.
7. Se ha propuesto una nueva variable como objetivo de la predicción a corto plazo: el factor de capacidad. El uso de esta novedosa variable permite reducir el error de predicción, ya que tiene en cuenta, exclusivamente, los aerogeneradores que están disponibles en cada momento.
8. Para las predicciones a muy corto plazo se ha diseñado un original sistema, cuyas predicciones de potencia eléctrica media generada superan a las de los modelos publicados hasta el momento. Se basa en la combinación óptima de las predicciones de los tres modelos (de entre un significativo conjunto de modelos ensayados) que mejores resultados han ofrecido con los datos utilizados en su ajuste o entrenamiento. La combinación de las salidas de estos tres modelos se realiza por medio de un adecuado sistema neuro-difuso (sistema de inferencia difusa en el que se ajusta el valor de sus parámetros

mediante un proceso de entrenamiento, como en una red neuronal).

9. Se ha desarrollado un nuevo filtro Kalman para la mejora de las predicciones de velocidad media del viento suministradas por un modelo atmosférico, de forma que la salida de dicho filtro sea usada como entrada de una red neuronal, mejorando significativamente la predicción de la potencia eléctrica media. Este nuevo sistema de predicción de la velocidad media del viento, mejora significativamente los errores de predicción de la potencia eléctrica media obtenida con modelos de la segunda familia (los que emplean predicciones de un modelo atmosférico) para horizontes de predicción de hasta 1 hora. La actualización de los parámetros del filtro sólo requiere valores de la potencia eléctrica media medidos en la subestación de conexión del parque eólico a la red eléctrica.
10. Se ha realizado un profundo análisis y una amplia selección de modelos de redes neuronales, utilizándose en todos ellos un proceso de optimización basado en algoritmo genético, con selección óptima de entradas, del número de neuronas en capas ocultas, de valores de los parámetros del algoritmo de entrenamiento, del número de unidades de retardo, del número de neuronas de función de base radial y del número de componentes principales, cuando fuera aplicable, todo ello con objeto de seleccionar la red neuronal idónea que logra un menor error de predicción. Ello ha permitido determinar las estructuras de las redes neuronales que resultan óptimas para la predicción de la producción eléctrica en parques eólicos.
11. El sistema global de predicción creado utiliza un mínimo de variables explicativas, entre ellas las predicciones de variables meteorológicas obtenidas con un modelo atmosférico de libre distribución (sin contratos para el suministro de las predicciones y, por tanto, sin coste adicional).

9.3. FUTURAS LÍNEAS DE INVESTIGACIÓN

Son varias las futuras líneas de investigación que se abren tras los trabajos de investigación recogidos en este documento. Algunas de estas líneas son continuación natural de los trabajos realizados hasta el momento. Otras líneas se centran en desarrollar modelos similares a los propuestos en otros ámbitos de la generación de energía eléctrica con fuentes renovables.

Una línea que se pretende investigar en un futuro consiste en el desarrollo de nuevos modelos de predicción en el corto plazo de la energía eléctrica producida en un parque eólico. El horizonte de predicción de los nuevos modelos debe cubrir, por lo menos, todo el día siguiente (debido al interés que tiene para el mercado eléctrico). Estos nuevos modelos pretenden usar como variable explicativa el valor medio del cubo de la velocidad del viento, que se obtendría de las predicciones de un modelo atmosférico, para la localización del parque eólico. Es muy posible que los errores de predicción de la potencia eléctrica media generada en el parque eólico se reduzcan utilizando esta nueva variable explicativa.

Otras líneas de investigación que se pretenden abordar, como continuación de los trabajos realizados, son las indicadas seguidamente.

1. Mejora de los modelos de predicción con horizontes de predicción por encima de las 5 horas: se pretende utilizar como posibles variables explicativas, en los nuevos modelos a desarrollar, los valores de predicción de la velocidad y dirección media del viento para instantes anteriores y posteriores al correspondiente al horizonte de predicción. Con esto se trata de hacer que los modelos puedan recoger la información de relaciones históricas entre valores de predicción y potencia eléctrica vertida en la subestación, permitiendo mejorar los errores en la predicción de la producción eléctrica debidos al “error de fase” de las previsiones del modelo atmosférico utilizado.
2. Mejora de los modelos de predicción a muy corto plazo: si el modelo atmosférico suministra nuevas predicciones cada seis horas, entonces pueden mejorarse las predicciones, a muy corto plazo, de potencia eléctrica vertida en la subestación utilizando como entradas, en los sistemas de inferencia difusa, las predicciones de dicha potencia eléctrica obtenidas con los últimos valores disponibles de predicción del modelo atmosférico. Obviamente, unas predicciones del modelo atmosférico más recientes pueden reducir los errores de predicción de la producción eléctrica. Por ejemplo, los valores de predicción de producción eléctrica en el parque eólico para las últimas horas de un día presentarán menor error si se han realizado con los valores de predicción de variables meteorológicas suministradas por el modelo atmosférico con la asimilación de datos de las 12 horas, en lugar de haberlas realizado con las suministradas por el modelo atmosférico con la asimilación de datos de las 0 horas.

3. Mejora de los modelos de predicción a corto plazo: en los trabajos publicados en la literatura internacional sobre predicción de energía eléctrica generada en parques eólicos a nivel regional o nacional, [Focken02], [Gastón06], se observa que el error en la predicción de la energía eléctrica generada por un conjunto de parques es menor que la de la suma de los errores de predicción para cada uno de los parques eólicos considerados individualmente. En este sentido, posiblemente se mejoraría la predicción de la energía eléctrica generada en un parque eólico si se realiza como suma de las predicciones de generación de cada uno de los aerogeneradores que componen el parque.

Se podría utilizar un modelo atmosférico para realizar las predicciones de viento en los puntos de una rejilla de alta resolución espacial y, a partir de esas predicciones, obtener las correspondientes a las localizaciones de los aerogeneradores.

4. Mejora del filtro Kalman: el filtro Kalman implementado para la mejora del valor de predicción de la velocidad media del viento (para su posterior utilización en modelos de predicción de la potencia media vertida en la subestación con horizontes de 0.5 ó 1 horas), utiliza una estimación de la velocidad media del viento obtenida a partir de la potencia media en la subestación. Los resultados de predicción de los modelos pueden mejorarse prescindiendo de la estimación y utilizando valores medidos, que sean fiables, de velocidad y dirección media del viento en instantes anteriores al de realizar la predicción.
5. Exploración de alternativas en la combinación de los modelos para predicciones en el muy corto plazo: la solución propuesta para las predicciones en el muy corto plazo ha consistido en nuevo sistema de predicción que logra la combinación idónea de las predicciones de los tres mejores modelos de predicción por medio de un sistema de inferencia difusa. El sistema de inferencia difusa fue construido con los datos de entrenamiento y validación, es decir, con el conjunto de la historia pasada. Esta es una buena solución, pero pueden obtenerse mejores resultados si el sistema de inferencia difusa sólo es construido con la parte más reciente de los datos históricos, es decir, con los x últimos datos, siendo x el valor de la ventana temporal (por ejemplo, x días) que minimiza los errores de predicción. De esta forma, el nuevo sistema de predicción se adaptaría mejor a los cambios estacionales y proporcionaría

predicciones con menor error.

6. Cambio de la función objetivo en los procesos de optimización de modelos de predicción mediante algoritmo genético: los modelos de redes neuronales y sistemas de inferencia difusa obtenidos, optimizados en un proceso controlado con un algoritmo genético, tenían como función objetivo el indicador RMSE (raíz del error cuadrático medio) de la predicción. La selección de individuos (nuevas redes neuronales) para la siguiente generación venía influenciada por el valor del indicador RMSE obtenido en la predicción de la potencia eléctrica media. Bajo el punto de vista de un gestor de parques eólicos, sería más interesante utilizar como función objetivo a minimizar la de los desvíos que pudieran ser económicamente penalizados en la predicción de la potencia media horaria.

Finalmente, como línea de investigación que se puede abrir en campos paralelos a los trabajos de investigación recogidos en este documento, se puede citar el desarrollo de modelos de predicción de la energía eléctrica generada en “granjas solares” en el corto plazo. El problema es similar al planteado en esta tesis doctoral, y las soluciones pueden ser muy similares. Se podría utilizar un modelo atmosférico para realizar predicciones de la radiación solar en la zona de interés, lo que junto a la serie histórica de la producción eléctrica de la “granja solar”, conduciría a crear modelos de predicción de la potencia eléctrica generada en el corto plazo.

- [Akaike76] H. Akaike, "Canonical correlation analysis of time series and the use of an information criterion", en D.G. Lainiotis y R.K. Mehra (Editores), "System identification: Advances and case studies", Academic Press, New York, 27-96, 1976.
- [Alexiadis98] M.C. Alexiadis, P.S. Dokopoulos, H.S. Sahsamanoglou, I.M. Manousaridis, "Short-term forecasting of wind speed and related electrical power", *Solar Energy*, Vol 63. no. 1, 61-68, 1998.
- [Alexiadis99] M.C. Alexiadis, P.S. Dokopoulos, H.S. Sahsamanoglou, "Wind speed and power forecasting based on spatial correlation models", *IEEE Transactions on Energy Conversion*, Vol. 14, no. 3, 836-837, 1999.
- [Atiya99] A.F. Atiya, S.M. El-Shoura, S.I. Shaheen, M.S. El-Sherif, "A comparison between neural-network forecasting techniques. Case study: river flow forecasting", *IEEE Transactions on Neural Networks*, Vol. 10, no. 2, 402-409, 1999.
- [Armstrong01] J.S. Armstrong (editor), "Principles of forecasting: a handbook for researches and practitioners", Kluwer Academic Publishers, Boston, 2001.
- [Back96] A.D. Back, A.C. Tsoi, "Nonlinear system identification using discrete Laguerre functions", *Journal of Systems Engineering*, Vol. 6, no. 3, 194-207, 1996.
- [Bathurst02] G.N. Bathurst, J. Weatherill, G. Strbac, "Trading wind generation in short term energy markets", *IEEE Transactions on Power Systems*, Vol. 17, no. 3, 782-789, 2002.
- [Beyer99] H.G. Beyer, D. Heinemann, H. Mellinghoff, K. Monninch, H.P. Waldl, "Forecast of regional power output of wind turbines", *Proceedings of the European Wind Energy Conference*, Niza, 4p, 1999.

- [Bishop95] C. Bishop, "Neural Networks for Pattern Recognition", Oxford University Press, Oxford, 1995.
- [Blake95] J. Blake, P. Francino, J.M. Catot, I. Solé, "A comparative study for forecasting using neural networks vs genetically identified Box&Jenkins models", *Neural Computing & Applications*, Vol. 3, no. 3, 139-148, 1995.
- [Box76] G.E.P. Box, G.M. Jenkins, "Time series Analysis – Forecasting and control", Englewood Cliffs, New Jersey, Prentice Hall, 1976.
- [Bossanyi85] E.A. Bossanyi, "Short-term wind prediction using Kalman filters", *Wind Engineering*, Vol. 9, no. 1, 1-8, 1985.
- [Bremnes04] J.B. Bremnes, "Probabilistic wind power forecasts using local quantile regression", *Wind Energy*, Vol. 7, no. 1, 47-54, 2004.
- [Bremnes06] J.B. Bremnes, "A comparison of a few statistical models for making quantile wind power forecasts", *Wind Energy*, Vol. 9, no. 1, 3-11, 2006.
- [Bunn85] D.W. Bunn, E.D. Farmer, "Comparative models for electrical load forecasting", John Wiley & Sons, Nueva York, 1985.
- [Cats96] G. Cats, L. Wolters, "The HIRLAM Project", *IEEE Computational Science and Engineering*, Vol. 5, no.4, 4-7, 1996.
- [Ceylan07] R. Ceylan, Y. Özbay, "Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network", *Expert Systems with Applications*, Vol. 33, 286-295, 2007.
- [Chan83] S.M. Chan, D.C. Powell, M. Yoshimura, D.H. Curtice, "Operations requirements of utilities with wind power generation", *IEEE Transactions on Power Apparatus and Systems*, Vol. 102, no. 9, 2850-2860, 1983.

- [Chen00] S.M. Chen, J.R. Hwang, "Temperature prediction using fuzzy time series", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 30, no. 2, 263-275, 2000.
- [Chiu94] S. Chiu, "Fuzzy model identification based on cluster estimation", *Journal of Intelligent and Fuzzy Systems*, Vol. 2, 267-278, 1994.
- [Contaxis91] G.C. Contaxis, J. Kabouris, "Short term scheduling in a wind/diesel autonomous energy system", *IEEE Transactions on Power Systems*, Vol. 6, no. 3, 1161-1167, 1991.
- [Crochet04] P. Crochet, "Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland", *Meteorological Applications*, Vol. 11, 176-187, 2004.
- [Cybenko89] G.V. Cybenko, "Approximation by superpositions of a sigmoidal function", *Mathematics of Control, Signals and Systems*, Vol. 2, 303-314, 1989.
- [Daley99] R. Daley, "Atmospheric data analysis", Cambridge University Press, Cambridge, 1999.
- [Damousis01] I.G. Damousis, P. Dokopoulos, "A fuzzy expert system for the forecasting of wind speed and power generation in wind farms", *Proceedings of the IEEE International Conference on Power Industry Computer Applications, PICA 01*, 63-69, 2001.
- [Damousis04] I.G. Damousis, M.C. Alexiadis, J.B. Theocharis, P.S. Dokopoulos, "A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation", *IEEE Transactions on Energy Conversion*, Vol. 19, no. 2, 352-361, 2004.
- [DeLurgio98] S.A. DeLurgio, "Forecasting principles and applications", McGraw-Hill, Singapur, 1998.
- [Deng89] J.L. Deng, "Introduction to grey system theory", *Journal of Grey System*, Vol. 1, no. 1, 1-24, 1989.

- [Denison01] D. G. T. Denison, P. Dellaportas, B. K. Mallick, "Wind speed prediction in a complex terrain", *Environmetrics*, Vol. 12, no. 6, 499-515, 2001.
- [DeVries92] A. de Vries, J. Principe, "The gamma model: A new model for temporal processing", *Neural Networks*, Vol. 5, pp 565-576, 1992.
- [Dodge06] D.M. Dodge, "Illustrated history of wind power development", en <http://www.telosnet.com/wind/index.html>, 2006.
- [Doeswijk05] T.G. Doeswijk, K.J. Keesman, "Adaptive weather forecasting using local meteorological information", *Biosystems Engineering*, Vol. 91, no. 4, 421-431, 2005.
- [Drezga98] I. Drezga, S. Rahman, "Input variable selection for ANN-based short-term load forecasting", *IEEE Transactions on Power Systems*, Vol. 13, no. 4, 1238-1244, 1998.
- [Dudhia05] J. Dudhia, D. Gill, K. Manning, W. Wang, C. Bruyere, "PSU/NCAR Mesoscale modelling system. Tutorial class notes and user's guide: MM5 modelling system version 3", National Center for Atmospheric Research, Camp Springs, Maryland, 2005.
- [Elman90] J.L. Elman, "Finding structure in time", *Cognitive Science*, Vol. 14, 179-211, 1990.
- [Eur05] Comisión europea, investigación comunitaria, "European Wind Energy at the dawn of the 21st century. Research funded under the Fifth Framework Programme", 2005.
- [El-Fouly06] T.H.M. El-Fouly, E.F. El-Saadany, M.M.A. Salama, "Grey predictor for wind energy conversion systems output power prediction", *IEEE Transactions on Power Systems*, Vol. 21, no. 3, 1450-1452, 2006.
- [Ferreyra05] A. Ferreyra, "Redes Neuronales Difusas para modelado via agrupamiento en-línea: Aplicación a un condensador de aspiración", Tesis doctoral, Instituto Politécnico Nacional, México, 2005.

- [Flores05] P. Flores, A. Tapia and G. Tapia, "Application of a control algorithm for wind speed prediction and active power generation", *Renewable Energy*, Vol.30, no.4, 523-536, 2005.
- [Focken01] U. Focken, M. Lange, H.P. Waldl, "Previento - A wind power prediction system with an innovative upscaling algorithm", Proceedings of the European Wind Energy Conference, EWEC 01, 4p, 2001.
- [Focken02] U. Focken, M. Lange, K. Mönnich, H.P. Waldl, H.G. Beyer, A. Luig, "Short-term prediction of the aggregated power output of wind farms - a statistical analysis of the reduction of the prediction error by spatial smoothing effects", *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 90, no. 3, 231-246, 2002.
- [Galanis02] G. Galanis, M. Anadranistakis, "A one-dimensional Kalman filter for the correction of near surface temperature forecasts", *Meteorological Applications*, Vol. 9, 437-441, 2002.
- [Galanis06] G. Galanis, P. Louka, P. Katsafados, I. Pytharoulis, G. Kallos, "Applications of Kalman filters based on non-linear functions to numerical weather predictions", *Annales Geophysicae*, Vol. 24, 2451-2460, 2006.
- [Gastón06] M. Gastón, L. Frías, I. Martí, "Exploring the limits of wind farm grouping for prediction error compensation", Proceedings European Wind Energy Conference, Atenas, 5p, 2006.
- [Giebel01] G. Giebel, L. Landberg, T.S. Nielsen, H. Madsen, "The Zephyr-Project. The next generation prediction system", Proceedings of Global Wind Power Conference, Paris, 4p, 2001.
- [Giebel03a] G. Giebel, L. Landberg, G. Kariniotakis, R. Brownsword, "State-of-the-art on methods and software tools for short-term prediction of wind energy production", Proceedings European Wind Energy Conference, Madrid, 8p, 2003.

- [Giebel03b] G. Giebel, I. Landberg, J. Badger, K. Sattler, H.A. Nielsen, H. Madsen, "Using ensemble forecasting for wind power", Proceedings European Wind Energy Conference, Madrid, 4p, 2003.
- [Glahn72] H.R. Glahn, D.A. Lowry, "The use of model output statistics (MOS) in objective weather forecasting", *Journal of Applied Meteorology*, Vol. 11, 1203-1211, 1972.
- [Goldberg89] D.E. Goldberg, "Genetic algorithms in search, optimization and machine learning", Addison-Wesley, Reading, Massachusetts, 1989.
- [Gutierrez04] J.M. Gutierrez, R. Cano, A.S. Cofiño, C.M. Sordo, "Redes probabilísticas y neuronales en las ciencias atmosféricas", Ministerio de Medio Ambiente, Madrid, 2004.
- [Hadjili02] M.L. Hadjili, V. Wertz, "Takagi-Sugeno fuzzy modeling incorporating input variables selection", *IEEE Transactions on Fuzzy Logic*, Vol. 10, no. 6, 728-742, 2002.
- [Härdle90] W. Härdle, "Applied nonparametric regression", Cambridge University Press, Cambridge, 1990.
- [Hartnell00] G. Hartnell, L. Landberg, "Wind on the system. Grid integration of wind power", *Renewable Energy World*, Vol. 3, no. 2, 60-71, 2000.
- [Hatziaargyriou01] N. Hatziaargyriou, G. Contaxis, M. Matos, J.A. Pecos Lopes, M.H. Vasconcelos, G. Karioniotakis, D. Mayer, J. Halliday, G. Dutton, P. Dokopoulus, A. Bakirtzis, J. Stefanakis, A. Gigantidou, P. O'donnell, D. McCoy, M.J. Fernández, J.M.S. Cotrim, A.P. Figueira, "Preliminary results from the more advanced control advice project for secure operation of isolated power systems with increased renewable energy penetration and storage", Proceedings of IEEE Power Technology Conference, PowerTech 2001, Oporto, 6p, 2001.
- [Haykin94] S. Haykin, "Neural networks: a comprehensive foundation", Macmillan Publishing, New York, 1994.

- [Ho92] K.L. Ho, Y.Y. Hsu, C.C. Yang, "Short term load forecasting using a multilayer neural network with an adaptive learning algorithm", *IEEE Transactions on Power Systems*, Vol. 7, no. 1, 141-149, 1992.
- [Holland92] J.H. Holland, "Adaptation in natural and artificial systems", Massachusetts Institute of Technology Press, Cambridge, Massachusetts, 1992.
- [Homleid95] M. Homleid, "Diurnal corrections of short-term surface temperature forecasts using the Kalman filter", *Weather and Forecasting*, Vol. 10, no. 4, 689-707, 1995.
- [Hsu91] Y.Y. Hsu, C.C. Yang, "Design of artificial neural networks for short-term load forecasting II. Multilayer feedforward networks for peak load and valley load forecasting", *IEE Proceedings - Generation, Transmission and Distribution*, Vol. 138, no. 5, 414-418, 1991.
- [Hunt01] K. Hunt, G.P. Nason, "Wind speed modelling and short-term prediction using wavelets", *Wind Engineering Journal*, Vol. 25, no. 1, 55-61, 2001.
- [Hwang98] J.R. Hwang, S.M. Chen, C.H. Lee, "Handling forecasting problems using fuzzy time series", *Fuzzy Sets and Systems*, Vol. 100, no. 2, 217-228, 1998.
- [Idae05] Instituto para la Diversificación y Ahorro de la Energía, "Plan de energías renovables para España", Ministerio de Industria, Turismo y Comercio, Madrid, 2005.
- [Jang93a] J.S.R. Jang, "ANFIS: adaptive network-based fuzzy inference systems", *IEEE Transaction on Systems, Man and Cybernetics*, Vol. 23, no. 3, 665-685, 1993.
- [Jang93b] J.S.R. Jang, C.T. Sun, "Functional equivalence between radial basis function networks and fuzzy inference systems", *IEEE Transactions on Neural Networks*, Vol. 4, no. 1, 156-159, 1993.

- [Jang95] J.S.R. Jang, N. Gulley, "The Fuzzy Toolbox for use with MATLAB", The MathWorks, Inc., Natick, Massachusetts, 1995.
- [Jang97] J.S.R. Jang, C.T. Sun, E. Mizutani, "Neuro-fuzzy and soft computing", Prentice Hall, Upper Saddle River, 1997.
- [Joensen97] A. Joensen, H. Madsen, T.S. Nielsen, "Non-parametric statistical methods for wind power prediction", Proceedings of European Wind Energy Conference, Dublin, 788-792, 1997.
- [Joensen99] A. Joensen, L. Landberg, H. Madsen, "A new measure-correlate-predict approach for resource assessment", Proceedings European Wind Energy Conference, Niza, 157-160, 1999.
- [Jolliffe02] I.T. Jolliffe, "Principal component analysis", Springer Verlag, Nueva York, 2002.
- [Jørgensen02a] J.U. Jørgensen, C. Moehrlen, E. McKeogh, "A new generation operational on- and offshore numerical prediction system", Proceedings World Wind Energy Conference, Berlín, 4p, 2002.
- [Jørgensen02b] J.U. Jørgensen, C. Moehrlen, B. Ó Gallachóir, K. Sattler, E.J. McKeogh, "HIRPOM: Description of an operational numerical wind power prediction model for large scale integration of on- and offshore wind power in Denmark", Proceedings Global Wind Power Conference, Paris, 5p, 2002.
- [Kamal97] L. Kamal, Y.Z. Jafri, "Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan", *Solar Energy*, Vol. 61, no. 1, 23-32, 1997.
- [Kalman60] R.E. Kalman, "A new approach to linear filtering and predictions problems", *Transactions of the ASME—Journal of Basic Engineering*, Vol. 82, 35-45, 1960.
- [Kariniotakis96] G.N. Kariniotakis, G.S. Stavrakakis, E.F. Nogaret, "Wind power forecasting using advanced neural networks models", *IEEE Transactions on Energy Conversion*, Vol. 11, no.4, 762-767, 1996.

- [Kariniotakis06] G.N. Kariniotakis et al., “What performance can be expected by short-term wind power prediction models depending on site characteristics?”, Proceedings European Wind Energy Conference, Atenas, 4p, 2006.
- [Kecman01] V. Kecman, “Learning and soft computing”, MIT Press, Cambridge, Massachusetts, 2001.
- [Kohonen95] T. Kohonen, “Self organizing maps”, Springer Verlag, New York, 1995.
- [Landberg94] L. Landberg, S.J. Watson, “Short-term prediction of local wind conditions”, *Boundary Layer Meteorology*, Vol. 70, 171-195, 1994.
- [Landberg97a] L. Landberg, “A model to predict the power output from wind farms”, Proceedings of Windpower Conference, Austin, Texas, 443-451, 1997.
- [Landberg97b] L. Landberg, “A mathematical look at a physical power prediction model”, Proceedings of Windpower Conference, Austin, Texas, 629-636, 1997.
- [Landberg 99] L. Landberg, “Short-term prediction of the power production from wind farms”, *Journal of Wind Engineering and Industrial Aerodynamics*, Vol 80, 207-220, 1999.
- [Landberg01] L. Landberg, “Short-term prediction of local wind conditions”, *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 89, no. 3-4, 235-245, 2001.
- [Landberg02] L. Landberg, G. Giebel, L. Myllerup, J. Badger, H. Madsen, T.S. Nielsen, “Poor man's ensemble forecasting for error estimation”, Proceedings American Wind Energy Association Conference, Portland, 8p, 2002.

- [Landberg03] L.Landberg, G. Giebel, H.A. Nielsen, T.S. Nielsen, H. Madsen, "Short-term prediction - An overview", *Wind Energy*, Vol. 6, no. 3, 273-280, 2003.
- [Larson06] K.A. Larson, K. Westrick, "Short-term wind forecasting using off-site observations", *Wind Energy*, Vol. 9, no. 1, 55-62, 2006.
- [Lee92] K.Y. Lee, Y.T. Cha, J.H. Park, "Short-term load forecasting using an artificial neural network", *IEEE Transactions on Power Systems*, Vol. 7, no. 1, 124-132, 1992.
- [Leng06] G. Leng, T.M. McGinnity, G. Prasad, "Design for self-organizing fuzzy neural networks based on genetic algorithms", *IEEE Transactions on Fuzzy Systems*, Vol. 14, no. 6, 755-766, 2006.
- [Linkens99] D.A. Linkens, M.Y. Chen, "Input selection and partition validation for fuzzy modelling using neural network", *Fuzzy Stes and Systems*, Vol. 107, 299-308, 1999.
- [Ljung99] L. Ljung, "System identification: theory for the user", Prentice-Hall, Upper Saddle River, N.J., 1999.
- [Madsen05] H. Madsen, P. Pinson, G. Kariniotakis, H.A. Nielsen, T.S. Nielsen, "Standardizing the performance evaluation of short-term wind prediction models", *Wind Engineering*, Vol. 29, no. 6, 475-489, 2005.
- [Makridakis98] S. Makridakis, S.C. Wheelwright, R.J. Hyndman, "Forecasting. methods and applications", John Wiley and Sons, Hoboken (NJ), 1998.
- [Makridakis00] S. Makridakis, M. Hibon, " The M3-competition: results, conclusions and implications", *International Journal of Forecasting*, Vol. 16, no. 4, 451-476, 2000.
- [Maqsood04] I. Maqsood, M.R. Khan, A. Abraham, "An ensemble of neural networks for weather forecasting", *Neural Computing & Applications*, Vol. 13, no. 2, 112-122, 2004.

- [Maqsood05] I. Maqsood, M. Khan, G. Huang, R. Abdalla, “Application of soft computing models to hourly weather analysis in southern Saskatchewan, Canada”, *Engineering Applications of Artificial Intelligence*, Vol. 18, no. 1, 115-125, 2005.
- [Márquez07] J. Márquez, J. Usaola, “Combining hydro-generation and wind energy biddings and operation on electricity spot markets”, *Electric Power Systems Research*, Vol. 77, no. 5, 393-400, 2007.
- [Martí06] I. Martí *et al*, “Evaluation of advanced wind power forecasting models – results of the Anemos project”, Proceedings European Wind Energy Conference, Atenas, 29p, 2006.
- [Milligan95] M.R. Milligan, A.H. Miller, F. Chapman, “Estimating the economic value of wind forecasting to utilities”, Proceedings of Windpower Conference, 10p, 1995.
- [Moehrlen02] C. Moehrlen, J.U. Jørgensen, E.J. McKeogh, “Power predictions in complex terrain with an operational numerical weather prediction model in Ireland including ensemble forecasting”, Proceedings World Wind Energy Conference, Berlín, 4p, 2002.
- [Mohammed95] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, A. Azeem, J. Farah, C. Drake, “Practical experiences with an adaptive neural network short-term load forecasting system”, *IEEE Transactions on Power Systems*, Vol. 10, no. 1, 254-265, 1995.
- [Mohandes98] M.A. Mohandes, S. Rehman, T.O. Halawani, “A neural networks approach for wind speed prediction”, *Renewable Energy*, Vol. 13, no. 3, 345-354, 1998.
- [Mörhlen04] C. Mörhlen, “Uncertainty in wind energy forecasting”, tesis doctoral, Universidad de Cork, Irlanda, 2004.
- [More03] A. More, M. C. Deo, “Forecasting wind with neural networks”, *Marine Structures*, Vol. 16, no. 1, 35-49, 2003.

- [Neg02] Neg-Micon, "A new millennium", *Powerful News*, Vol. 1, no. 3, 2002.
- [Nielsen98] T.S. Nielsen, A. Joensen, H. Madsen, "A new reference for wind power forecasting", *Wind Energy*, Vol. 1, no. 1, 29-34, 1998.
- [Nielsen01] T.S. Nielsen, H. Madsen, H.A. Nielsen, L. Landberg, G. Giebel, "Zephyr - the prediction models", Proceedings European Wind Energy Conference, Copenhagen, 4p, 2001.
- [Nielsen02a] H.A. Nielsen, T.S. Nielsen, H. Madsen, "On on-line systems for short-term forecasting for energy systems", Proceedings Operations Research Conference, Vienna, 7p, 2002.
- [Nielsen02b] T.S. Nielsen, H. Madsen, H. Aa. Nielsen, L. Landberg, G. Giebel, "Prediction of regional wind power", Proceedings of Global Wind Power Conference, Paris, 5p, 2002.
- [Nielsen06a] H.A Nielsen, H. Madsen, T.S. Nielsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts", *Wind Energy*, Vol. 9, no. 1, 95-108, 2006.
- [Nielsen06b] T.S. Nielsen, H. Madsen, H.A. Nielsen, P. Pinson, G. Kariniotakis, N. Siebert, I. Marti, M. Lange, U. Focken, L. von Bremen, P. Louka, G. Kallos, G. Galanis, "Short-term wind power forecasting using advanced statistical methods", Proceeding European Wind Energy Conference, Atenas, 9p, 2006.
- [Oja89] E. Oja, "Neural networks, principal components and subspaces", *International Journal of Neural Systems*, Vol. 1, 61-68, 1989.
- [Orlanski75] I. Orlanski, "A rational subdivision of scales for atmospheric processes", *Bulletin of the Americal Meteological Society*, Vol. 56, 527-530, 1975.
- [Park89] J. Park, I. Sandberg, "Universal approximation using radial basis function networks", *Neural Computation*, Vol. 3, 303-314, 1989.

- [Park91] D.C. Park, M.A El-Sahrkawi, R.J. Marks, L.E. Atlas, M.J. Damborg, “Electric load forecasting using an artificial neural network”, *IEEE Transactions on Power Systems*, Vol. 6, no. 2, 442-449, 1991.
- [Peng92] T.M. Peng, N.F. Hubele, G.G. Karady, “Advancement in the application of neural networks for short-term load forecasting”, *IEEE Transactions on Power Systems*, Vol. 7, no. 1, 250-257, 1992.
- [Perivolaris06] Y.G. Perivolaris, A.N. Vougiouka, V.V. Alafouzos, D.G. Mourikis, V.P. Zagorakis, K.G. Rados, D.S. Barkouta, A. Zervos, Q. Wang, “Coupling of a mesoscale atmospheric prediction system with a CFD microclimatic model for production forecasting of wind farms in complex terrain: test case in the island of Evia”, Proceedings European Wind Energy Conference, Atenas, 10p, 2006.
- [Pinson03] P. Pinson, G.N. Kariniotakis, “Wind power forecasting using fuzzy neural networks enhanced with on-line prediction risk assessment”, Proceedings IEEE Power Technology Conference (PowerTech), Bolonia, 64-71, 2003.
- [Pinson04] P. Pinson, G.N. Kariniotakis, “On-line assessment of prediction risk for wind power production forecasts”, *Wind Energy*, Vol. 7, no. 2, 119-132, 2004.
- [Poggi03] P. Poggi, M. Muselli, G. Notton, C. Cristofari, A. Louche, “Forecasting and simulating wind speed in Corsica by using an autoregressive model”, *Energy Conversion and Management*, Vol. 14, no. 20, 3177-3196, 2003.
- [Potter06] C.W. Potter, M. Negnevitsky, “Very short-term wind forecasting for tasmanian power generation”, *IEEE Transactions on Power Systems*, Vol. 21, no. 2, 965-972, 2006.
- [Practical05] Practical Action, The Schumacher Centre for Technology & Development, “Wind for electricity generation”, en http://practicalaction.org/docs/technical_information_service/wind_electricity_generation.pdf, 2005.

- [Principe00] J.C. Principe, N.R. Euliano, W.C. Lefebvre, "Neural and adaptive systems: fundamentals through simulations", John Wiley & Sons, New York, 2000.
- [Ramírez01] I.J. Ramírez-Rosado, L.A. Fernández-Jiménez, "A new model for short-term wind electric power forecasting", Proceedings International Conference on Modelling, Identification and Control, Innsbruck, Austria, 66-69, 2001.
- [Ramírez03] I.J. Ramírez-Rosado, L.A. Fernández-Jiménez, "Next-day wind park electric energy generation forecasting using fuzzy time-series", Proceedings International Conference on Modelling, Identification and Control, Innsbruck, Austria, 237-240, 2003.
- [Ramírez04a] I.J. Ramírez-Rosado, L.A. Fernández-Jiménez, "An advanced model for short-term forecasting of mean wind speed and wind electric power", *Control and Intelligent Systems*, Vol. 32, no. 1, 21-26, 2004.
- [Ramírez04b] I.J. Ramírez-Rosado, L.A. Fernández-Jiménez, "A new electric power forecasting model for wind farms", Proceedings International Conference on Modelling, Identification and Control, Grindelwald, Suiza, 398-401, 2004.
- [Ramírez06] I.J. Ramírez-Rosado, L.A. Fernández-Jiménez, C. Monteiro, "Artificial neural network models for wind power short-term forecasting using weather predictions", Proceedings International Conference on Modelling, Identification and Control, Innsbruck, Austria, 128-132, 2006.
- [Roulston01] M.S. Roulston, D.T. Kaplan, J. Hardenberg, L.A. Smith, "Value of the ECMWF ensemble prediction system for forecasting wind energy production", Proceedings of the European Wind Energy Conference, EWEC 01, 4p, 2001.
- [Roulston03] M. S. Roulston, D. T. Kaplan, J. Hardenberg, L. A. Smith, "Using medium-range weather forecasts to improve the value of wind energy production", *Renewable Energy*, Vol. 28, no. 4, 585-602, 2003.

- [Salmon99] Salmon, J. L. Walmsley, "A two-site correlation model for wind speed, direction and energy estimates", *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 79, no. 3, 233-268, 1999.
- [Sakawa02] M. Sakawa, "Genetic algorithms and fuzzy multiobjective optimization", Kluwer Academic Publishers, Norwell, Massachusetts, 2002.
- [Sánchez06] I. Sánchez, "Short-term prediction of wind energy production", *International Journal of Forecasting*, Vol. 22, no. 1, 43-56, 2006.
- [Sanger89] T. Sanger, "Optimal unsupervised learning in a single layer linear feedforward neural network", *Neural Networks*, Vol. 12, 459-473, 1989.
- [Schwartz98] M.N. Schwartz, B.H. Bailey, "Wind forecasting objectives for utility schedulers and energy traders", *Proceedings of Windpower Conference*, 5p, 1998.
- [Schlink98] U. Schlink, G. Tetzlaff, "Wind speed forecasting from 1 to 30 minutes", *Theoretical and Applied Climatology*, Vol. 60, 191-198, 1998.
- [Sfetsos99] A. Sfetsos, A.H. Coonick, "Times series forecasting of mean hourly wind speed and power", *Jornadas Hispano-Lusas*, Lisboa, 75-82, 1999.
- [Sfetsos00] A. Sfetsos, "A comparison of various forecasting techniques applied to mean hourly wind speed time series", *Renewable Energy*, Vol. 21, no. 1, 23-35, 2000.
- [Sfetsos02] A. Sfetsos, "A novel approach for the forecasting of mean hourly wind speed time series", *Renewable Energy*, Vol. 27, no. 2, 163-174, 2002.
- [Sideratos07] G. Sideratos, N.D. Hatzigiorgiou, "An advanced statistical method for wind power forecasting", *IEEE Transactions on Power Systems*, Vol. 22, no. 1, 258-265, 2007.

- [Sindelar04] R. Sindelar, R. Babuska, "Input selection for nonlinear regression models", *IEEE Transactions on Fuzzy Logic*, Vol. 12, no. 5, 688-696, 2004.
- [Song93] Q. Song, B.S. Chissom, "Fuzzy time series and its models", *Fuzzy Sets and Systems*, Vol. 54, no. 3, 269-277, 1993.
- [Song97] Q. Song, R.P. Leland, B.S. Chissom, "Fuzzy stochastic fuzzy time series and its models", *Fuzzy Sets and Systems*, Vol. 88, no. 3, 333-341, 1997.
- [Takagi85] T. Takagi, M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 15, no. 1, 116-132, 1985.
- [Tambke05] J. Tambke, M. Lange, U. Focken, J.O. Wolff, J.A.T. Bye, "Forecasting offshore wind speeds above the North Sea", *Wind Energy*, Vol. 8, no. 1, 3-16, 2005.
- [Torres05] J.L. Torres, A. García, M. de Blas, A. de Francisco, "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)", *Solar Energy*, Vol. 79, no. 1, 65-77, 2005.
- [Troen91] I. Troen, E.L. Petersen, "European wind atlas", RisØ National Laboratory, RisØ, Dinamarca, 1991.
- [Undén02] P. Undén et al, "Hirnam-5 scientific documentation", http://hirnam.org/open/publications/SciDoc_Dec2002.pdf, 2002.
- [Usaola04] J. Usaola, O. Ravelo, G. González, F. Soto, M.C Dávila, B. Díaz-Guerra, "Benefits for wind energy in electricity markets from using short term wind power prediction tools; a simulation study", *Wind Engineering*, Vol. 28, no. 1, 119-128, 2004.
- [Watson94] S.J. Watson, L. Landberg, J.A. Halliday, "Application of wind speed forecasting to the integration of wind energy into a large scale power

- system”, *IEE Proceedings Generation, Transmission and Distribution*, Vol. 141, no. 4, 357–362, 1994.
- [Watson99] S.J. Watson, G. Giebel, A. Joensen, “The economic value of accurate wind power forecasting to utilities”, *Proceedings de EWEC 99*, European Wind Energy Conference, 4p, 1999.
- [Wei90] W. Wei, “Time series analysis: univariate and multivariate methods”, Addison-Wesley, Reading, Massachusetts, 1990.
- [Werbos90] P. Werbos, “Backpropagation through time: What it does and how to do it”, *Proceedings of the IEEE*, Vol. 78, no. 10, 1550-1560, 1990.
- [Weigend96] A.S. Weigend, “Time series analysis and prediction using gated experts with application to energy demand forecast”, *Applied Artificial Intelligence*, Vol. 10, 583-624, 1996.
- [World07] World Wind Energy Association, “New world record in wind power capacity: 14,9 GW added in 2006 –worldwide capacity at 73,9 GW”, en http://www.wwindea.org/home/images/stories/pdfs/pr_statistics_2006_290107.pdf, 2007.
- [Xiong02] N. Xiong, “A hybrid approach to input selection for complex processes”, *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans*, Vol. 32, no. 4, 532-536, 2002.
- [Yao99] X. Yao, “Evolving artificial neural networks”, *Proceedings of the IEEE*, Vol. 87, no. 9, 1423-1447, 1999.
- [Yu98] K. Yu, M.C. Jones, “Local linear quantile regression”, *Journal of the American Statistical Association*, Vol. 93, 228-238, 1998.
- [Zadeh65] L.A. Zadeh, “Fuzzy sets”, *Information and control*, Vol. 8, 338-353, 1965.

ANEXO. IMPLEMENTACIÓN DE UN NUEVO FILTRO KALMAN

En este Anexo se detallan las características y aplicación del nuevo filtro Kalman para la obtención de unos mejores valores de la velocidad media del viento en el parque eólico, que se usan en la predicción de la producción eléctrica en el muy corto plazo. Inicialmente describimos las ecuaciones que definen el filtro Kalman, para después comentar la aplicación del nuevo filtro al cálculo de la velocidad media del viento para los siguientes periodos de la serie temporal de esa velocidad.

I. El filtro Kalman.

El filtro Kalman [Kalman60] es básicamente un algoritmo recursivo que produce una estimación de los valores de una serie temporal de valores no observables, usando los valores de otras series temporales de variables observables y relacionadas con la primera. Esta estimación de valores no observables puede ir acompañada de la estimación de los parámetros del modelo teórico que genera esos valores no observables. Las estimaciones de la variable no observable, así como las de los parámetros de su modelo teórico, son actualizadas a intervalos regulares de tiempo, coincidentes con la disponibilidad de nuevos valores de las variables observables. El filtro Kalman obtiene un nuevo valor para la variable no observable, combinando el filtrado con una estimación, con máxima verosimilitud, de los parámetros del modelo.

La base matemática para la formulación del filtro de Kalman se encuentra en la notación utilizada en ingeniería de control denominada el espacio de los estados. El modelo que relaciona variables y tiempo parte de la suposición de una o más variables, o parámetros variables, con los cuales podemos describir completamente el estado en que se encuentra algún sistema en un momento determinado. Estas son las variables de estado, y el conjunto de las mismas se conoce como vector de estado. Una o varias de estas variables de estado suelen ser las variables no observables, mientras que en su lugar observamos (podemos medir) otra u otras variables que son combinación lineal de las variables de estado más un error de predicción.

Dos son los bloques matemáticos fundamentales para un filtro Kalman: la ecuación de medida y la ecuación de proceso o transición. La ecuación de medida relaciona las variables observables con las variables de estado, además de incluir un

término de error. La ecuación (A.1) representa esta ecuación de medida, en la que \mathbf{Z}_t representa el vector de variables observables en el instante t (se suponen que son más de una las variables observables, en otro caso sería un escalar, el de la salida observable); \mathbf{C}_t es una matriz numérica conocida en el instante t ; \mathbf{X}_t corresponde al vector de variables de estado en el instante t ; y \mathbf{v}_t representa un vector de error de medición (se le denomina ruido de medida) también para el instante t . La condición temporal que cumplen los términos del error de medición es que tienen valor medio nulo y varianza conocida; en concreto se supone que siguen una distribución normal de media cero y varianza \mathbf{R} . A la ecuación de medida se le pueden añadir variables exógenas.

$$\mathbf{Z}_t = \mathbf{C}_t \mathbf{X}_t + \mathbf{v}_t \quad (\text{A.1})$$

La ecuación de transición es la que nos proporciona la evolución en el tiempo del vector de estado. Esta ecuación está representada en (A.2), donde la matriz \mathbf{A} relaciona el estado (vector de estado) en el instante actual con su valor futuro; \mathbf{U} es el vector de entradas externas; \mathbf{B} es la matriz que relaciona las variables de estado en el instante siguiente con las entradas externas; \mathbf{G} la matriz que relaciona el vector de estado en el instante siguiente con el ruido actual; y \mathbf{w}_{t-1} representa un ruido blanco (se le denomina ruido del sistema) con distribución de probabilidad normal de valor medio cero y varianza \mathbf{Q} .

$$\mathbf{X}_t = \mathbf{A}_{t-1} \mathbf{X}_{t-1} + \mathbf{B}_{t-1} \mathbf{U}_{t-1} + \mathbf{G}_{t-1} \mathbf{w}_{t-1} \quad (\text{A.2})$$

En muchos casos prácticos, las matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{G} , \mathbf{Q} y \mathbf{R} se consideran constantes. En el caso general, que es el que se comenta en estos párrafos, los elementos de las matrices \mathbf{A} , \mathbf{B} y \mathbf{G} se considerarán variables con el tiempo.

El filtro Kalman estima el valor del vector de estado en cualquier instante t con el mínimo valor de la matriz de covarianza del error, \mathbf{P} . El funcionamiento del filtro recuerda al de un sistema de control con realimentación: el filtro Kalman estima primero el estado del proceso y después obtiene una realimentación en la forma de observaciones (afectadas con ruido blanco). Así, a partir de la ecuación de transición (A.2) es posible proyectar en el tiempo el estado actual y la estimación de la covarianza del error, y obtener una estimación del estado (vector de estado) *a priori* para el siguiente instante de tiempo. Con las observaciones (medidas) se produce la realimentación y las estimaciones *a priori* se pueden mejorar para obtener las estimaciones *a posteriori*.

La Figura A.1 representa el proceso iterativo de cálculo que representa el filtro Kalman, cuyo objetivo final es la mejor estimación posible para el valor del vector de estado en el instante t , $\hat{\mathbf{X}}_{t|t}$. El recuadro de la izquierda de la Figura A.1 representa la ecuación de transición, con la estimación *a priori* del estado para el instante siguiente. El significado de $\hat{\mathbf{X}}_{t|t-1}$ es el de la estimación del vector de estado para el instante t , pero calculada en el instante $t-1$; representa, por tanto, una estimación *a priori*. Además de la estimación del estado, se realiza la estimación, también *a priori*, de la matriz de covarianza de error, $\mathbf{P}_{t|t-1}$. El significado de \mathbf{A}_{t-1}^T es el de la matriz transpuesta de la matriz \mathbf{A} en el instante $t-1$.

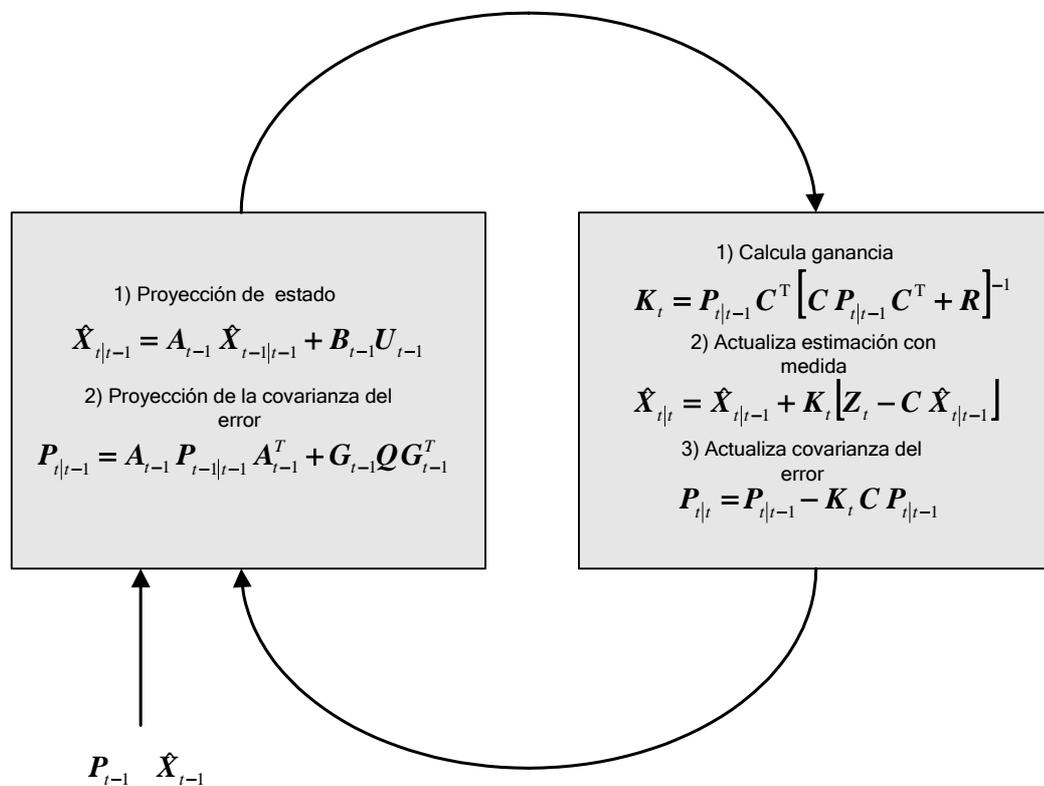


Figura A.1. Diagrama de flujo del algoritmo discreto del filtro Kalman.

El recuadro de la derecha de la Figura A.1 representa la realimentación, incluyendo la ecuación de medida y el cálculo del factor de actualización (ganancia de Kalman) de los valores calculados anteriormente *a priori*. Así, el vector \mathbf{K}_t representa la ganancia de Kalman (el valor de la ganancia puede ser distinto para cada una de las variables de estado; de ahí que sea un vector con igual número de filas que el vector de estado). La ganancia de Kalman multiplica al error de medición en la ecuación central

de este recuadro derecho (el vector entre corchetes representa la diferencia entre las variables observables y su valor teórico, de acuerdo a la estimación *a priori* del vector de estado). En la última ecuación, se calcula el valor *a posteriori* de la matriz de covarianza del error, $\mathbf{P}_{t|t}$.

Para poder usar iterativamente el filtro Kalman y que éste vaya ajustando a cada nuevo instante de tiempo el vector de estado, hacen falta unos valores iniciales de estimación del vector de estado y de la matriz de covarianza de error (representados en la esquina inferior izquierda de la Figura A.1).

En el filtro Kalman las matrices \mathbf{P} , \mathbf{Q} y \mathbf{R} juegan un importante papel, ya que aparecen en tres de las cinco ecuaciones que permiten implementar el filtro. Para la matriz \mathbf{P} se suele tomar un valor inicial elevado, con valores exclusivamente en la diagonal principal. La matriz \mathbf{R} está relacionada con el ruido de medida, cuyos únicos valores no nulos son los de la diagonal principal y representan la varianza en la medida de cada aparato o sensor. Más complicada de calcular resulta la matriz \mathbf{Q} , relacionada con el ruido del sistema. El valor de los elementos de esta matriz se pueden calcular a partir de datos históricos del sistema.

La principal ventaja que aporta el filtro Kalman es la de proporcionar una estimación *a priori* del estado del sistema. Para el problema de predicción, a corto plazo, de la producción eléctrica en parques eólicos, el filtro Kalman puede proporcionar una valiosa información si hacemos que el estado del sistema tenga cierta relación con el error de predicción.

II. Aplicación del nuevo filtro Kalman en la mejora de la predicción de velocidad media del viento.

En [Doeswijk05] se aplica un filtro Kalman en la predicción de la velocidad media del viento en superficie, de tal forma que la misión del filtro es proporcionar una estimación del error sistemático introducido en las predicciones del modelo atmosférico. Los autores consideran ese error sistemático como una variable independiente del valor de las predicciones. La matriz de covarianzas \mathbf{Q} es calculada a partir de la base de datos histórica de predicciones del modelo atmosférico y valores reales medidos (utilizan los datos de un año). La matriz de covarianzas \mathbf{R} toma un valor igual a la estimación que hacen de varianza en las medidas realizadas con un anemómetro. También detallan como las matrices \mathbf{B} y \mathbf{G} cambian con el tiempo, de tal

forma que cuando se actualizan las predicciones del modelo atmosférico son matrices identidad, mientras que en cualquier otro instante su valor es nulo.

En [Galanis06] aplican una novedosa variación para la estimación del error sistemático de un modelo atmosférico en la predicción de la velocidad del viento y de la temperatura en superficie del terreno. El principal inconveniente es que no especifican a qué serie temporal de velocidad del viento o temperatura se refieren (diaria o horaria); pero al usar las mismas ecuaciones que en [Galanis02], donde se pretendía calcular el error sistemático en la predicción de la temperatura máxima y mínima diaria, se puede entender que en [Galanis06] se refieren a valores medios diarios. Aun así se aportan dos ideas valiosas. La primera es que el error sistemático puede ser dependiente del valor de predicción. Esta dependencia es modelizada mediante un polinomio de grado $n-1$ (ver ecuación (8.29)). La segunda idea es un método más sencillo que el propuesto en [Doeswijk05] para el cálculo de las matrices de covarianzas \mathbf{Q} y \mathbf{R} . Éstas son calculadas a partir de los k últimos datos disponibles. Así, calculan la matriz \mathbf{Q} a partir de las diferencias entre vectores de estado en instantes correlativos, tal como indica la ecuación (A.3). Análogamente hacen con la matriz de covarianza \mathbf{R} , como expresa la ecuación (A.4), donde \mathbf{Z} representa el error sistemático (diferencia entre el valor real y el valor de predicción del modelo atmosférico). En ambos cálculos toman como últimos datos disponibles los de la última semana ($k = 6$).

$$\mathbf{Q} = \frac{1}{k} \sum_{i=0}^k \left(\mathbf{X}_{t-i} - \mathbf{X}_{t-i-1} - \left(\frac{\sum_{i=0}^k (\mathbf{X}_{t-i} - \mathbf{X}_{t-i-1})}{k+1} \right) \right)^2 \quad (\text{A.3})$$

$$\mathbf{R} = \frac{1}{k} \sum_{i=0}^k \left(\mathbf{Z}_{t-i} - \mathbf{C}_{t-i} \mathbf{X}_{t-i} - \left(\frac{\sum_{i=0}^k (\mathbf{Z}_{t-i} - \mathbf{C}_{t-i} \mathbf{X}_{t-i})}{k+1} \right) \right)^2 \quad (\text{A.4})$$

Es justamente ese periodo tomado (días), el que nos permite suponer que las series temporales usadas son de valores medios diarios. En [Galanis02] usan exactamente las mismas expresiones.

A partir de estas dos ideas, el nuevo filtro Kalman creado se ha aplicado para la mejora de las predicciones de velocidad media del viento obtenidas del modelo

atmosférico MM5, que hemos usado en la predicción, a muy corto plazo, de la potencia media entregada en la subestación de conexión del parque eólico. Las características conseguidas son:

1. El error sistemático (sesgo) en la predicción de la velocidad media del viento suministrada por el modelo atmosférico MM5 es de tipo no lineal, siendo su valor dependiente del valor de predicción aportado por el modelo MM5. Esta dependencia se modeliza mediante un polinomio de grado $n-1$, tal como muestra la ecuación (A.5). Los n coeficientes de este polinomio formarán los elementos del vector de estado, cuyo valor será actualizado por el filtro Kalman, proporcionando una estimación *a priori* de su valor para el siguiente intervalo de tiempo (en este caso media hora).

$$\hat{v}(t) = v(t) - \left(a_0(t) + a_1(t) v(t) + \dots + a_{n-1}(t) [v(t)]^{n-1} \right) \quad (\text{A.5})$$

Así, el vector de estado es el representado en la ecuación (A.6).

$$\mathbf{X}_t = \begin{bmatrix} a_0(t) \\ a_1(t) \\ \vdots \\ a_{n-1}(t) \end{bmatrix} \quad (\text{A.6})$$

2. La matriz de variables observables queda reducida a una única variable: el sesgo en la determinación de la velocidad media del viento en el parque eólico. Este valor se estima a partir de la potencia media entregada en la subestación y el valor de predicción del modelo atmosférico MM5. La matriz de covarianzas de error de medida, \mathbf{R} , queda reducida a un único escalar (un solo elemento en la matriz). La determinación del valor de este elemento de forma teórica resulta prácticamente imposible, ya que la velocidad media en realidad no es medida directamente, sino a través de una medida indirecta con otra variable (potencia media entregada en la subestación) con la que tiene una relación no lineal.

Para el cálculo de la matriz de covarianza del error de medida, \mathbf{R} , se utiliza la ecuación (A.4) tomando un valor de k de 6, pero en este caso no referido a los valores de los últimos días como en [Galanis06], sino referido a los valores de las últimas 6 medias horas.

Lo mismo sucede en el cálculo de la matriz de covarianza del error del sistema, \mathbf{Q} , cuyo valor es calculado por medio de la ecuación (A.3) tomando también un valor de $k = 6$.

3. Dado que las predicciones del modelo atmosférico MM5 se renuevan cada 24 horas (se podrían renovar cada 6, pero no están disponibles más que los datos correspondientes a las previsiones con el estado de la atmósfera a las 0 horas GMT), se inicializa el filtro en cuanto están disponibles las nuevas predicciones. Si tenemos en cuenta que los resultados del modelo global GFS, que corresponden a las entradas del modelo MM5, no empiezan a estar disponibles hasta pasadas las 3:30 horas GMT, y que el modelo atmosférico MM5 necesita un tiempo adicional para realizar sus predicciones, resulta que hasta las 4:30 horas GMT no se dispone de las previsiones para todo el día. En ese momento (4:30 horas GMT) el modelo MM5 ha aportado predicciones para instantes ya pasados, concretamente para las 0, 0:30, 1, 1:30, 2, 2:30, 3, 3:30, 4 y 4:30 horas GMT. Para esos mismos instantes ya se dispone de la potencia media entregada en la subestación por parte del parque eólico, a partir de la cual se puede realizar una estimación de la velocidad media del viento. Es decir, cuando están disponibles las nuevas predicciones del MM5, que forman una serie temporal a intervalos de 0.5 horas, ya se conocen los valores “medidos” para la variable velocidad del viento para los 10 primeros elementos de la serie. La utilización del filtro Kalman aporta el valor del vector de estado para el elemento 11, que nos permite a su vez determinar el sesgo de la predicción para ese valor y mejorar así el valor de predicción ofrecido por el modelo atmosférico MM5.

Al ser sólo una la variable observada (sesgo en la velocidad media del viento), la matriz \mathbf{C} tiene una única fila cuyo primer elemento es 1, tal como expresa la ecuación (A.7). El número total de columnas es n , igual al número de coeficientes del polinomio que expresa el sesgo de la predicción del modelo atmosférico. El valor de v_t corresponde al valor de la velocidad del viento, medida indirectamente por medio de la potencia media entregada en la subestación en el instante t .

$$\mathbf{C}_t = \left[1 \quad v_t \quad v_t^2 \quad \dots \quad v_t^{n-1} \right] \quad (\text{A.7})$$

Para las 7 primeras medias horas de cada día se toma el valor de la \mathbf{Q} constante

de tal forma que es una matriz identidad con n filas y columnas. A partir de la octava media hora (3:30 horas GMT), el valor de la matriz \mathbf{Q} se calcula de acuerdo a la ecuación (A.3) con un valor de $k = 6$. Debe tenerse en cuenta que $t = 0$ para las 0:00 horas GMT de cada día.

El valor inicial de la matriz \mathbf{P} (para $t = 0$) es el de una matriz con todos los elementos nulos, excepto los de la diagonal principal que se hacen igual a 4, lo que equivale a darle muy poca credibilidad a los primeros valores del vector de estado (las variaciones en las primeras iteraciones pueden ser grandes).

Para el cálculo del valor del único elemento de la matriz \mathbf{R} se utilizó la ecuación (A.4) con un valor de $k = 6$, a partir de la octava media hora de cada día. Para los valores de las primeras 7 medias horas (desde las 0:00 a las 3:00 horas) se utilizó la expresión de la ecuación (A.8), con valores de t de 0 a 6 (en unidades de medias horas), para el cálculo del único elemento de la matriz, r .

$$r = -0.0159t^3 + 0.3438t^2 - 2.4707t + 6 \quad (\text{A.8})$$

La expresión de la ecuación (A.8) se escogió de tal forma que para el primer valor de cada día ($t = 0$, es decir, para las 0:00 horas GMT) el valor de este elemento fuera 6, lo cual es un valor relativamente muy alto, y que fuera decayendo de tal forma que para $t = 7$, el valor obtenido fuera 0.1. La Figura A.2 representa la variación del valor de r , único elemento de la matriz del error de medida para las 7 primeras medias horas de cada día. Se optó por tomar estos valores de tal forma que para los primeros instantes de cada día tuviese un valor relativamente alto que fuera disminuyendo para quedarse en los valores considerados constantes en otros trabajos publicados, como en [Doeswijk05]. A partir de esa 7^a media hora el valor de r cambiará de acuerdo al valor de la ecuación (A.4) como se ha comentado.

Para la obtención de los valores iniciales del vector de estado, \mathbf{X}_t , se realizaron unas pruebas preliminares considerándolos nulos. En estas pruebas se calculó el valor medio de los elementos de este vector para las últimas 10 medias horas de cada día y esos fueron los valores que se tomaron como iniciales para el resto de las pruebas.

La matriz \mathbf{A} se ha tomado constante e igual a la matriz identidad, al igual que la matriz \mathbf{G} . La matriz \mathbf{B} es nula, ya que no hay entradas exógenas.

El vector de ganancia de Kalman es un vector con n filas, cuyos valores van a ir actualizándose en cada iteración (para cada nueva media hora).

Tras el cálculo *a priori* de los valores de los vectores de estado, $\mathbf{X}_{t/t-1}$ y de matriz de covarianzas, $\mathbf{P}_{t/t-1}$, se calcula el vector de ganancia de Kalman y los valores *a posteriori* (una vez obtenido el último valor de velocidad media del viento medido indirectamente). El diagrama de flujo responde al representado en la Figura A.1.

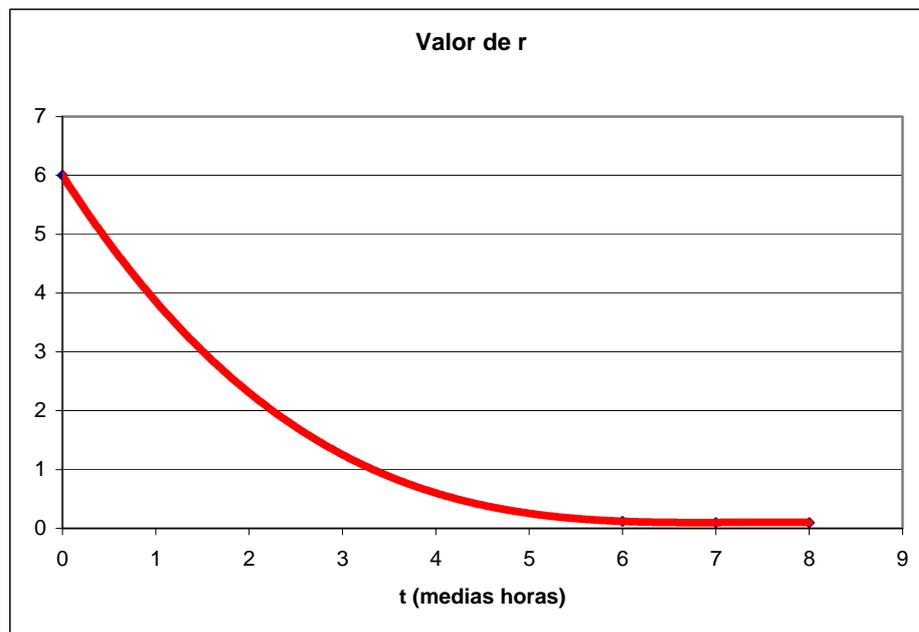


Figura A.2. Variación del elemento de la matriz de covarianza del error de medida para las 7 primeras medias horas de cada día.

El primer valor que el filtro podrá ajustar será el de la 11ª primera media hora. Para este valor el filtro ya ha tenido 4 iteraciones en que a podido ajustar mejor los valores del vector de estado, matriz de covarianzas, etc., independientemente de los valores iniciales tomados.

Se implementó el algoritmo descrito en los párrafos anteriores en forma de programa informático, de tal forma que se pudiera escoger el valor de n , es decir, del número de elementos del vector de estado (número de coeficientes del polinomio que expresa el sesgo de las predicciones del modelo atmosférico MM5). Se utilizaron los valores calculados *a priori* del vector de estado para la determinación del valor de la velocidad media del viento “filtrada”, es decir, restando al valor de predicción del

modelo MM5 el valor del sesgo, de acuerdo a la ecuación (A.5) en la que los coeficientes corresponden al valor *a priori*.

Cuando el horizonte de predicción es diferente a media hora, supongamos que es T medias horas, para la obtención de la velocidad media del viento “filtrada”, se toman los valores *a priori* del vector de estado obtenidos T medias horas antes.