



UNIVERSIDAD DE LA RIOJA

TESIS DOCTORAL

Título
Nuevos modelos de predicción a corto plazo de la generación eléctrica en plantas basadas en energía solar fotovoltaica
Autor/es
María Sonia Terreros Olarte
Director/es
Luis Alfredo Fernández Jiménez
Facultad
Escuela Técnica Superior de Ingeniería Industrial
Titulación
Departamento
Ingeniería Eléctrica
Curso Académico



Nuevos modelos de predicción a corto plazo de la generación eléctrica en plantas basadas en energía solar fotovoltaica, tesis doctoral de María Sonia Terreros Olarte, dirigida por Luis Alfredo Fernández Jiménez (publicada por la Universidad de La Rioja), se difunde bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported. Permisos que vayan más allá de lo cubierto por esta licencia pueden solicitarse a los titulares del copyright.



UNIVERSIDAD DE LA RIOJA

Departamento de Ingeniería Eléctrica

TESIS DOCTORAL

Doctorado en Ingeniería Eléctrica, Matemáticas y
Computación

**Nuevos modelos de predicción a corto plazo
de la generación eléctrica en plantas basadas
en energía solar fotovoltaica**

Autora: D.^a María Sonia Terreros Olarte

Director: Dr. D. Luis Alfredo Fernández Jiménez

Logroño, Mayo de 2017

*Dedico esta tesis a todos los que
contribuyeron a mi formación y,
especialmente, a mi familia.*

Prólogo

Actualmente, a nivel mundial existe una elevada inversión en la explotación de fuentes de energía renovable. Sin embargo, debido a la elevada volatilidad de algunas de ellas como la solar, todavía se producen diversos problemas para la gestión en las redes y en los mercados de electricidad.

Como la inserción de la fuente renovable solar en el sistema eléctrico tiende a aumentar, cada vez es más importante efectuar la predicción de lo que van a generar las centrales eléctricas basadas en esta fuente. Así surge la necesidad de mejorar los sistemas de predicción actuales, con un alto nivel de precisión a fin de minimizar los impactos causados.

La presente tesis doctoral, desarrollada en el Laboratorio de Planificación, Operación y Control de Sistemas de Energía Eléctrica, del Departamento de Ingeniería Eléctrica de la Universidad de La Rioja, tiene como objetivo principal el desarrollo de nuevos modelos de predicción a corto plazo de la potencia media horaria generada en una planta fotovoltaica, basándose en técnicas de minería de datos aplicadas a la base de datos histórica de producción de la planta.

Para evaluar el rendimiento de los modelos desarrollados, se ha procedido a la realización de pruebas en la obtención de predicciones a corto plazo para una instalación fotovoltaica conectada a red, y se ha establecido una comparación con otros modelos publicados en la literatura internacional. Los resultados de las pruebas han sido satisfactorios, obteniendo resultados superiores a los de otros métodos de predicción.

Agradecimientos

Aprovecho la ocasión, para expresar mi agradecimiento, a todos aquellos que han hecho posible la finalización de esta tesis doctoral. En especial al director de la misma, el profesor Dr. D. Luis Alfredo Fernández Jiménez, por todo lo que he aprendido con sus clases, por su continuo apoyo, orientación y consejo en el trabajo realizado. A los integrantes del Grupo de Investigación de “Planificación, Operación y Control de Sistemas de Energía Eléctrica” de la Universidad de La Rioja, con los que he tenido la posibilidad de compartir muchas horas en el laboratorio, y me han dado la oportunidad de disfrutar de su calidad humana, que tantas veces me han demostrado. A todos mis profesores, porque todos ellos de manera directa o indirecta, han colaborado en mi formación académica. Y por último a mi familia, por su comprensión y apoyo. A todos ellos les estoy francamente agradecida.

María Sonia Terreros

Resumen

El importante aumento en el número de instalaciones generadoras de energía eléctrica basadas en fuentes renovables (fundamentalmente solares fotovoltaicas) experimentado en todo el mundo desarrollado, el aumento de la capacidad de las nuevas instalaciones, junto con la obligatoriedad de su participación en los mercados eléctricos, ha impulsado la necesidad de desarrollar modelos de predicción a corto plazo (hasta 72 horas) de la generación eléctrica en estas plantas. La precisión de estos modelos de predicción tiene, así, importantes repercusiones económicas (por su uso para preparar ofertas al mercado eléctrico) como de seguridad de la red (para la determinación de congestiones, establecimiento de reservas y gestión de otras plantas).

El objetivo de los trabajos de investigación de esta tesis es el desarrollo de nuevos modelos de predicción a corto plazo de la energía eléctrica media horaria generada en una planta generadora basada en energía solar fotovoltaica. Los modelos utilizan como entradas variables astronómicas, valores de predicción para las variables meteorológicas y valores pasados de producción eléctrica. Proporcionan como salida la potencia eléctrica media horaria generada para horizontes de predicción a corto plazo. Además uno de los modelos desarrollado en la presente tesis, a pesar de proporcionar un punto de predicción de potencia esperado, proporciona también una predicción de la incertidumbre asociada al valor de predicción. Esta herramienta es de gran importancia, ya que facilita una más completa información de predicción permitiendo, así, mejorar la toma de decisiones. La aplicación de este último modelo proporciona la probabilidad de que la potencia que se generará en el horizonte de predicción considerado se encuentre en un intervalo u otro del rango total de potencia generable.

Los modelos desarrollados utilizan técnicas de minería de datos: árboles de decisión y sistemas de reglas, redes neuronales artificiales, máquinas de vectores soporte, sistemas basados en lógica borrosa y algoritmos evolutivos, y poseen características distintas a las de otros publicados en la literatura internacional. Proporcionan gran cantidad de información a sus usuarios, permitiendo mejorar el rendimiento de los modelos de predicción de potencia y apoyo en su toma de decisiones. Además, la incorporación de la incertidumbre asociada a cada valor de predicción puntual puede permitir al usuario evaluar el riesgo inherente al uso de dicho valor de predicción.

Los resultados computacionales obtenidos de los ensayos efectuados con los nuevos modelos desarrollados, muestran menores errores y/o mayores horizontes de predicción que los de otros modelos publicados en la literatura internacional.

Summary

The significant increase in the number of power generation facilities based on renewable sources (mainly solar photovoltaic) experienced all over the developed world, the increase in the capacity of the new facilities, with the mandatory participation in electricity markets, has promoted the need of short-term forecasting models (horizons up to 72 hours) of the electric power generation in these plants. The accuracy of these forecasting models has important economic implications (due to its use in preparing offers to the electricity markets) and network security implications (determination of network congestions, schedule of power reserves and management of other power plants).

The goal of the research works of this thesis is the development of new short-term forecasting models of hourly average electric power generation in a solar photovoltaic power plant. The models use astronomical input variables, forecasts of weather variables and past values of electric power production. They provide as output the hourly average electric power generation for short-term forecasting horizons. In addition, one of the models developed for the present thesis, in spite of providing an expected power prediction point, also provides the uncertainty associated with that value. This characteristic is very important, since it facilitates a more complete forecasting information allowing, thus, to improve the decision making related to the forecast. The use of this last model provides the probability that the generated power in forecasting horizon considered is in each one of the power intervals in which the power capacity of the plant has been divided.

The developed models use data mining techniques: decision trees, rule-based systems, artificial neural networks, support vector machines, systems based on fuzzy logic and evolutionary algorithms, and have different characteristics from those published in the international literature. They provide a great amount of information to their users, allowing to improve the forecasts and support their decision making processes. In addition, the inclusion of the uncertainty associated with the point forecast may allow to the user assessing the risk inherent to the use of that point forecast.

The computational results obtained from the tests carried out with the new developed models show lower errors and/or larger forecasting horizons than those of other models published in the international literature.

Índice general

Prólogo	v
Resumen	vii
Summary	ix
Índice general	xi
Índice de figuras	xvii
Índice de tablas	xxv
Acrónimos y terminología	xxix

I INTRODUCCIÓN	1
1. Justificación, objetivos y estructura	3
1.1. Justificación	3
1.2. Objetivos	3
1.2.1. Objetivo general	3
1.2.2. Objetivos específicos	4
1.3. Estructura y contenido	4
2. Energías renovables. Energía solar fotovoltaica	9
2.1. Energías renovables	9
2.2. La Energía Solar Fotovoltaica	10
2.2.1. Evolución histórica	10
2.2.2. Ventajas de la utilización de la energía solar	11
2.2.3. Inconvenientes de la utilización de la energía solar	12
2.2.4. Sistemas fotovoltaicos conectados a la red	12
3. Modelos de predicción de la potencia eléctrica	13
3.1. Primeros trabajos	13
3.2. Evolución histórica en los últimos años	15
3.3. Conclusiones	41

4. Astronomía terrestre y solar	43
4.1. Nociones generales	43
4.2. Las referencias locales	45
4.2.1. La referencia horizontal (RH) o azimutal	45
4.2.2. La referencia ecuatorial (RE) o horaria	46
4.2.3. Paso de coordenadas horizontales a horarias	47
4.2.4. Determinación de la declinación	48
4.3. Ángulo de incidencia de los rayos solares directos	49
4.3.1. Ángulo de incidencia en la RH	49
4.3.2. Ángulo de incidencia en la RE	50
4.4. Las coordenadas temporales	50
5. Predicción numérica del tiempo	51
5.1. Introducción	51
5.2. Inicio	51
5.3. Rejillas de integración	52
5.3.1. Coordenadas horizontales	53
5.3.2. Coordenadas verticales	53
5.3.3. Aumento de resolución	56
5.4. Procesos de ejecución	57
5.4.1. Pre-proceso	57
5.4.2. Proceso	58
5.4.3. Post-proceso	59
5.5. Tipos de modelos numéricos del tiempo	59
5.5.1. Modelos globales	60
5.5.2. Modelos regionales	61
5.5.3. Modelos Mesoescalares	62
5.6. Pronóstico de las variables meteorológicas	63
5.6.1. Descarga de datos globales	64
5.6.2. Etapa de preprocesamiento	64
5.6.3. Etapa de procesamiento	66
5.6.4. Etapa de postprocesamiento	67
5.6.5. Variables meteorológicas pronosticadas	68
5.6.6. Valor medio horario de las variables	70
6. Radiación solar	73
6.1. Algunos conceptos	73
6.2. Componentes de la radiación solar	73
6.3. La atmósfera y sus diversos componentes	75
6.4. Irradiancia	76
6.4.1. Irradiancia extra-atmosférica	76
6.4.2. Irradiancia directa total	77
6.4.3. Irradiancia global y difusa en la superficie horizontal	81
6.4.4. Irradiancia difusa sobre un plano de inclinación β	81
6.4.5. Irradiancia del albedo	86
6.4.6. Irradiancia global total sobre un plano de inclinación β	86

6.5. Efectos del ángulo de incidencia y del polvo	87
6.5.1. Irradiancias “efectivas”	88

II TÉCNICAS DE MINERÍA DE DATOS Y PROBABILIDAD 91

7. Árboles de decisión y sistemas de reglas	93
7.1. Sistemas por partición: árboles de decisión	93
7.1.1. Particiones posibles	93
7.1.2. Criterio de selección de particiones	94
7.2. Sistemas de aprendizaje de reglas por cobertura	94
7.3. Poda y reestructuración	95
7.4. Algoritmo M5P	96
7.5. Método M5Rules	98
8. Redes neuronales artificiales	99
8.1. Neurona artificial	99
8.1.1. Elementos	99
8.2. Red neuronal artificial (RNA)	101
8.2.1. Conceptos	101
8.2.2. Fases de operación	101
8.3. Adalina	102
8.3.1. Regla LMS	103
8.4. El perceptrón multicapa (MLP)	105
8.4.1. El MLP como aproximador universal de funciones	105
8.4.2. Aprendizaje por retropropagación de errores (BP)	107
8.4.3. Otros algoritmos	108
8.5. Funciones de base radial (RBF)	110
8.5.1. Descripción matemática del modelo	110
8.5.2. Aprendizaje en las RBF	112
9. Máquinas de vectores soporte	115
9.1. SVM lineal con margen máximo	115
9.1.1. Formulación original de las SVM	115
9.1.2. Formulación dual de las SVM	116
9.1.3. Cálculo de b	117
9.2. SVM con margen máximo en el espacio de características	118
10. Sistemas basados en lógica borrosa	121
10.1. Lógica borrosa	121
10.2. Controlador borroso	127
10.3. Sistemas neuro-difusos	127
11. Algoritmos evolutivos	133
11.1. Esquema general	133
11.2. Algoritmos genéticos	134

12. Probabilidad	139
12.1. Probabilidad básica	139
12.2. Esperanza matemática, varianza y desviación estándar	141
12.3. La distribución normal	142
12.4. Variables aleatorias y distribución de probabilidad	144
 III EVALUACIÓN DE LOS MODELOS	 147
13. Técnicas de evaluación	149
13.1. Diferentes técnicas de evaluación	149
13.1.1. Evaluación mediante validación cruzada	149
13.1.2. Evaluación con entrenamiento, validación y test	149
13.1.3. Evaluación de los modelos en los trabajos de esta tesis	151
13.1.4. Elección de los modelos en los trabajos de esta tesis	153
13.2. Definiciones de error de predicción	154
13.3. Definiciones de medidas de error	155
13.4. Comparación entre modelos de predicción	155
 IV NUEVOS MODELOS DE PREDICCIÓN, RESULTADOS COMPUTACIONALES Y CONCLUSIONES	 159
14. Modelos con técnicas de minería de datos	161
14.1. Variables o atributos disponibles	161
14.1.1. Potencia eléctrica	161
14.1.2. Atributos astronómicos	161
14.1.3. Atributos con predicciones meteorológicas	162
14.1.4. Atributos calculados	163
14.2. Selección de las variables o atributos	165
14.3. Normalización de las variables o atributos	166
14.4. Modelos con técnicas de minería de datos	167
14.4.1. Árboles de decisión y sistemas de reglas	167
14.4.2. Redes neuronales artificiales	168
14.4.3. Máquinas de vectores soporte	169
14.4.4. Sistemas de inferencia neuro-difusos	170
14.4.5. Algoritmos genéticos y redes neuronales artificiales	171
 15. Modelo con evaluación de incertidumbre	 173
15.1. Base de conocimiento	173
15.2. Algoritmo genético	175
15.2.1. Representación de los individuos	175
15.2.2. Creación de una población inicial y evolución	178
15.2.3. Operadores genéticos	178
15.2.4. Resultado	181
15.3. Formulación	181
15.3.1. Mecanismo basado en minería de datos (MDM)	182

15.3.2. Intervalos de potencia	186
15.3.3. Matriz de probabilidad (PM)	186
15.4. Herramientas del modelo	188
15.4.1. Predicción de la incertidumbre	188
15.4.2. Predicción de un punto de potencia	190
16.Resultados computacionales	191
16.1. Variables o atributos seleccionados	191
16.1.1. Atributos en los modelos para días con cielo claro	191
16.1.2. Atributos en los modelos para cualquier tipo de día	192
16.2. Notación utilizada	194
16.2.1. Notación en los modelos para días con cielo claro	195
16.2.2. Notación en los modelos para cualquier tipo de día	196
16.3. Resultados de los modelos para días con cielo claro	198
16.3.1. Horizonte de predicción de 0 a 24 horas	199
16.3.2. Horizonte de predicción de 25 a 48 horas	204
16.3.3. Horizonte de predicción de 49 a 72 horas	210
16.3.4. Horizonte de predicción mayor a 72 horas	216
16.4. Resultados de los modelos para cualquier tipo de día	222
16.4.1. Horizonte de predicción de 1 a 8 horas. Reglas y árboles	222
16.4.2. Horizonte de predicción de 1 a 8 horas. RNAs	225
16.4.3. Horizonte de predicción de 1 a 8 horas. SVM	240
16.4.4. Horizonte de predicción de 1 a 8 horas. ANFIS	243
16.4.5. Horizonte de predicción de 1 a 8 horas. Mejores modelos	243
16.4.6. Horizonte de predicción de 0 a 24 horas	272
16.4.7. Horizonte de predicción de 25 a 48 horas	276
16.4.8. Horizonte de predicción de 49 a 72 horas	281
16.5. Resultados con incertidumbre de la predicción	285
16.5.1. Horizonte de predicción de 0 a 24 horas	285
16.5.2. Horizonte de predicción de 25 a 48 horas	285
16.5.3. Horizonte de predicción de 49 a 72 horas	286
17.Conclusiones	289
17.1. Conclusiones	289
17.2. Principales aportaciones de la tesis	291
17.3. Futuras líneas de investigación	295
A. Artículos científicos	299
Bibliografía	301

Índice de figuras

1.1.	Fases del proceso de descubrimiento de conocimiento en bases de datos. . .	5
2.1.	Esquema de la estructura de un sistema solar fotovoltaico conectado a la red.	12
4.1.	P: Perihelio, A: Afelio. Hemisferio Norte: Ep:Equinoccio de primavera, Sv: Solsticio de verano, Eo: Equinoccio de otoño, Si: Solsticio de invierno. . . .	44
4.2.	La eclíptica.	44
4.3.	Referencia horizontal.	45
4.4.	Referencia ecuatorial.	47
4.5.	Ángulo de incidencia de los rayos solares directos.	49
5.1.	Representación simbólica de discretización de la atmósfera.	53
5.2.	Distribuciones espaciales posibles de las variables dependientes en una rejilla cuadrada.	54
5.3.	Representación esquemática de las coordenadas σ	55
5.4.	Representación esquemática de las coordenadas sigma y ETA.	56
5.5.	Representación esquemática de las coordenadas híbridas utilizadas por el GFS.	57
5.6.	Tipos de observaciones meteorológicas.	58
5.7.	Rejilla global de 2.5° de resolución horizontal utilizada por modelos de circulación general sobre todo el globo; el tamaño de la rejilla es $144 \times 73 = 10512$ puntos.	60
5.8.	Representación gráfica de un modelo regional anidado en un modelo global.	62
5.9.	Rejilla de un modelo global, un modelo regional y un modelo mesoescalar.	63
5.10.	Diagrama de flujo del modelo WRF-NMM.	65
5.11.	Definición de los dominios (WRF Domain Wizard).	66
5.12.	Diferentes géneros de nubes y los niveles aproximados que ocupan en la troposfera.	70
6.1.	Distintos componentes de la radiación solar.	74
6.2.	Estructura general de la atmósfera.	75
7.1.	Ejemplo de poda. Algunos nodos inferiores son eliminados.	95
8.1.	Modelo genérico de neurona artificial.	100
8.2.	Modelo de neurona estándar.	103

8.3. Superficie de error $E(w)$ en el espacio de pesos w , y descenso por el gradiente hacia un mínimo (local).	104
8.4. Arquitectura del MLP de tres capas.	106
8.5. Arquitectura del RBF.	111
8.6. Respuesta localizada de las neuronas ocultas en el RBF (nodos gaussianos). Los puntos representan patrones en el espacio de las entradas, que en su mayoría se agrupan en torno a tres centros.	112
9.1. Margen blando de pérdida para un SVM lineal.	116
10.1. Función de pertenencia gaussiana.	122
10.2. Estructura de un controlador borroso o FLC.	127
10.3. (a) Un modelo difuso Sugeno de primer orden con dos entradas y dos reglas. (b) Arquitectura ANFIS equivalente.	129
11.1. Comportamiento de la búsqueda.	134
11.2. Algoritmo genético básico.	135
11.3. Cruce monopunto.	137
11.4. Cruce de dos puntos.	137
11.5. Mutación.	138
12.1. Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 < \sigma_2$	143
13.1. Evaluación mediante validación cruzada.	150
13.2. Evaluación de una hipótesis utilizando dos subconjuntos de datos.	151
15.1. Diagrama del algoritmo genético del modelo.	176
15.2. Ejemplo de un individuo binario, con dos entradas utilizadas de cinco disponibles.	177
15.3. Ruleta de Holland con 1275 elementos, para una población de 50 individuos ordenados de menor a mayor error.	179
15.4. Se elimina el grupo de 16 bits, correspondiente a la entrada i	181
15.5. Se añade el grupo de 16 bits, correspondiente a la entrada i	182
15.6. Representación gráfica de una función gaussiana local.	184
16.1. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C24, para horizontes de predicción entre 0 y 24 horas con los datos de entrenamiento y test.	200
16.2. Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 20 reglas, para horizontes de predicción entre 0 y 24 horas.	203
16.3. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 20 reglas, para horizontes de predicción entre 0 y 24 horas.	203
16.4. Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 20 reglas, para horizontes de predicción entre 0 y 24 horas con datos de test.	203

16.5. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C48, para horizontes de predicción entre 25 y 48 horas con los datos de entrenamiento y test.	205
16.6. Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 22 reglas, para horizontes de predicción entre 25 y 48 horas.	209
16.7. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 22 reglas, para horizontes de predicción entre 25 y 48 horas. . .	209
16.8. Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 22 reglas, para horizontes de predicción entre 25 y 48 horas con datos de test.	209
16.9. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C72, para horizontes de predicción entre 49 y 72 horas con los datos de entrenamiento y test.	211
16.10. Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 21 reglas, para horizontes de predicción entre 49 y 72 horas.	215
16.11. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 21 reglas, para horizontes de predicción entre 49 y 72 horas. . .	215
16.12. Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 21 reglas, para horizontes de predicción entre 49 y 72 horas con datos de test.	215
16.13. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C, para horizontes de predicción mayores a 72 horas con los datos de entrenamiento y test.	217
16.14. Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 18 reglas, para horizontes de predicción mayores a 72 horas.	221
16.15. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 18 reglas, para horizontes de predicción mayores a 72 horas. . .	221
16.16. Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 18 reglas, para horizontes de predicción mayores a 72 horas con datos de test.	221
16.17. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en M5Rules y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	223

16.18. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en M5P y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	224
16.19. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-GD y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	226
16.20. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-momentum(0.2) y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	227
16.21. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-CGD y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	228
16.22. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-CGD y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de test. . .	229
16.23. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-BFGS y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	230
16.24. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-BFGS y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de test. . .	231
16.25. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-CGD y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	232
16.26. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-CGD y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de test. . .	233
16.27. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-BFGS y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	234
16.28. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-BFGS y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de test. . .	235
16.29. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en RNA y desarrollados con los atributos Gm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	236
16.30. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en RNA y desarrollados con los atributos Hm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	237
16.31. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en MLP-BFGS y desarrollados con los atributos Gm y Hm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	238

16.32. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en MLP-BFGS y desarrollados con los atributos Gm y Hm, para horizontes de predicción de 1 a 8 horas con los datos de test.	239
16.33. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en SVM desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	241
16.34. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en SVM desarrollados con los atributos Gm y Hm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	242
16.35. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	244
16.36. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de validación.	245
16.37. Valores ponderados del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento y validación.	246
16.38. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de test.	247
16.39. Diagramas de dispersión con valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el sistema ANFIS de 19 reglas usando los atributos H1, para el horizonte de predicción de 1 hora.	248
16.40. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con el sistema ANFIS de 19 reglas usando los atributos H1, para el horizonte de predicción de 1 hora.	248
16.41. Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el sistema ANFIS de 19 reglas usando los atributos H1, para el horizonte de predicción de 1 hora con datos de test.	248
16.42. Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	260
16.43. Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de test.	261

16.44. Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos Hm y Gm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.	262
16.45. Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos Hm y Gm, para horizontes de predicción de 1 a 8 horas con los datos de test.	263
16.46. Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de entrenamiento.	266
16.47. Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de test.	267
16.48. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de entrenamiento.	268
16.49. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de test.	269
16.50. Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con datos de test.	270
16.51. Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con datos de test.	271
16.52. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria obtenidos con modelos basados en M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos R24, para horizontes de predicción entre 0 y 24 horas con los datos de entrenamiento.	273
16.53. Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-BFGS con 16 neuronas, para horizontes de predicción entre 0 y 24 horas.	275
16.54. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con la red neuronal RBF-BFGS con 16 neuronas, para horizontes de predicción entre 0 y 24 horas.	275
16.55. Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-BFGS con 16 neuronas, para horizontes de predicción entre 0 y 24 horas con datos de test.	275

16.56. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria obtenidos con modelos basados en M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos R48, para horizontes de predicción entre 25 y 48 horas con los datos de entrenamiento.	277
16.57. Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-CGD con 18 neuronas, para horizontes de predicción entre 25 y 48 horas. .	280
16.58. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con la red neuronal RBF-CGD con 18 neuronas, para horizontes de predicción entre 25 y 48 horas. .	280
16.59. Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-CGD con 18 neuronas, para horizontes de predicción entre 25 y 48 horas con datos de test.	280
16.60. Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria obtenidos con modelos basados en M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos R72, para horizontes de predicción entre 49 y 72 horas con los datos de entrenamiento.	282
16.61. Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo basado en la técnica M5P con 10 instancias, para horizontes de predicción entre 49 y 72 horas.	284
16.62. Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con el modelo basado en la técnica M5P con 10 instancias, para horizontes de predicción entre 49 y 72 horas.	284
16.63. Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo basado en la técnica M5P con 10 instancias, para horizontes de predicción entre 49 y 72 horas con datos de test.	284
16.64. Distribución de probabilidad, correspondiente a la incertidumbre asociada al valor puntual de predicción de la potencia eléctrica media horaria, para algunas horas (tiempo solar) del día 16 de mayo de 2012, desarrolladas para un horizonte de predicción entre 25 y 48 horas.	287

Índice de tablas

3.1. Breve resumen de la estimación de la radiación solar.	13
3.2. Resumen de características de los modelos de predicción de la potencia eléctrica.	39
5.1. Escalas atmosféricas y fenómenos asociados.	59
5.2. Dimensión y resolución de los dominios que se han utilizado.	67
5.3. Clasificación internacional de nubes.	69
5.4. La escala de Beaufort.	71
6.1. Gases en proporciones estables y variables (% en volumen).	76
6.2. Significado de los símbolos de las irradiancias.	77
6.3. Valores típicos de β_A	79
6.4. Significado de los símbolos de las entradas en los modelos de irradiancia horaria para cielo claro.	82
6.5. Entradas en los modelos de irradiancia horaria para cielo claro.	83
6.6. Modelos isótropos y modelos pseudo-isótropos.	84
6.7. Modelos anisótropos.	86
6.8. Albedo medio de algunas superficies.	87
6.9. Valores recomendados para los parámetros implicados en el cálculo de las pérdidas angulares.	88
7.1. Criterios de partición.	94
8.1. Modificación de los pesos en la iteración t	110
9.1. Funciones núcleo comúnmente utilizadas en \mathbb{R}^d	119
10.1. Operaciones entre conjuntos borrosos.	124
10.2. Dos pasos en el proceso de aprendizaje híbrido para ANFIS.	132
11.1. Individuo genético binario.	135
12.1. Función de probabilidad conjunta para X y Y.	145
13.1. Descripción de los subconjuntos de datos.	152
13.2. Notaciones.	154
14.1. Atributos con valores astronómicos.	162

14.2. Atributos con predicciones meteorológicas para los horizontes de predicción entre 0 y 24, 25 y 48, 49 y 72 horas.	162
14.3. Atributos calculados.	163
15.1. Representación de una r -tupla de transición.	174
15.2. Para calcular número de intervalos - 1.	177
15.3. Para calcular el valor de σ_i	178
15.4. Representación de una base de datos histórica con las variables meteorológicas de entradas disponibles.	183
15.5. Representación de un registro j con las variables de entradas disponibles. .	185
15.6. Representación de la matriz PPM_{t+m} de pseudo-probabilidades, para el instante futuro $t + m$, con n intervalos de potencia.	187
15.7. Función $f_{XY,t+m}(x, y)$ de probabilidad conjunta de transiciones de potencia.	188
15.8. Funciones de probabilidad marginal de X , para el instante futuro $t + m + 1$.	189
15.9. Funciones de probabilidad marginal de Y , para el instante futuro $t + m$. . .	189
16.1. Atributos seleccionados para modelos de días con cielo claro para horizontes de predicción entre 0 y 24, 25 y 48, 49 y 72 horas, y mayores a 72 horas. . .	191
16.2. Atributos seleccionados para modelos de cualquier tipo de día con horizontes de predicción entre 1 y 8 horas.	193
16.3. Atributos seleccionados para modelos de cualquier tipo de día con horizontes de predicción entre 0 y 24, 25 y 48, 49 y 72 horas.	194
16.4. Significado de los títulos en los resultados de los modelos de días claros. . .	195
16.5. Notación en los modelos de días claros para horizontes de predicción entre 0 y 24, 25 y 48, 49 y 72 horas, y mayores a 72 horas.	195
16.6. Significado de los títulos en los resultados de los modelos de cualquier tipo de día.	196
16.7. Notación en los modelos de cualquier tipo de día para horizontes de predicción de 1 a 8 horas, entre 0 y 24, 25 y 48, 49 y 72 horas.	197
16.8. Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 0 y 24 horas.	202
16.9. Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción entre 0 y 24 horas.	202
16.10. Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 25 y 48 horas.	208
16.11. Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción entre 25 y 48 horas.	208

16.12. Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 49 y 72 horas.	214
16.13. Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción entre 49 y 72 horas.	214
16.14. Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción mayores a 72 horas.	220
16.15. Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción mayores a 72 horas.	220
16.16. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 1 hora.	252
16.17. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 2 horas.	253
16.18. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 3 horas.	254
16.19. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 4 horas.	255
16.20. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 5 horas.	256
16.21. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 6 horas.	257
16.22. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 7 horas.	258
16.23. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 8 horas.	259

16.24. Coeficientes de determinación R^2 en días de cualquier tipo con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos basados en redes neuronales, para horizontes de predicción de 1 a 8 horas.	264
16.25. Modelos seleccionados para cada horizonte de predicción de 1 a 8 horas. .	265
16.26. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 0 y 24 horas.	274
16.27. Coeficientes de determinación R^2 con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos basados en redes neuronales, para horizontes de predicción entre 0 y 24 horas.	274
16.28. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 25 y 48 horas.	279
16.29. Coeficientes de determinación R^2 con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos basados en redes neuronales, para horizontes de predicción entre 25 y 48 horas.	279
16.30. Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 49 y 72 horas.	283

Acrónimos y terminología

Término	Significado
AE	Algoritmo evolutivo
AG	Algoritmo genético
ALADIN	Aire Limitée, Adaptation Dynamique, Développement InterNational
AnEn	Método de conjunto análogo
ANCFIS	Adaptive Neuro Complex Fuzzy Inferential System
ANFIS	Adaptive Neuro-Fuzzy Inference Systems
APVF	Modelo de predicción de potencia fotovoltaica analítico
AR	Modelo autorregresivo
ARB	Árbol de regresión binario
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Modelo auto-regresivo y de media móvil
ARPEGE	Action de Recherche Petite Échelle Grande Échelle
ARPS	Advanced Regional Prediction System
ARW	Advanced Research WRF
ARX	Modelo de pronóstico autorregresivo con entradas exógenas
BA	Bagging
BFGS	Método BFGS
BO	Boosting
BP	Aprendizaje por retropropagación de errores
BSS	Brier Skill Score
C	Atributos en modelos para días con cielo claro (horizonte mayor a 72 horas)
CAP	Teoría de conjuntos aproximados
CAPS	Center for Analysis and Prediction of Storms
CFRACH	Cobertura de nubes a nivel alto (entre 0 y 1)
CFRACL	Cobertura de nubes a nivel bajo (entre 0 y 1)
CFRACM	Cobertura de nubes a nivel medio (entre 0 y 1)
CFS	Correlation-based Feature Selection
CGD	Gradiente conjugado descendente
Claros_E	Datos horarios de los días claros del 01-01-2010 al 31-12-2011
Claros_En	Datos horarios de los días claros del 01-01-2010 al 31-12-2011 (75 %)
Claros_P	Ponderación del RMSE con: Claros_En (75 %) y Claros_V (25 %)
Claros_Pond	Ponderación del RMSE con: Claros_E (66.66 %) y Claros_T (33.33 %)
Claros_T	Datos horarios de los días claros del 01-01-2012 al 28-12-2012
Claros_V	Datos horarios de los días claros del 01-01-2010 al 31-12-2011 (25 %)

Término	Significado
CMA	Cadena de Markov avanzada
CNRS	Laboratoire d'Aérodynamique
COAMPS	Coupled Ocean/Atmosphere Mesoscale Prediction System
C24	Atributos en modelos para días con cielo claro (horizonte 24 horas)
C48	Atributos en modelos para días con cielo claro (horizonte 48 horas)
C72	Atributos en modelos para días con cielo claro (horizonte 72 horas)
DMAPE	Porcentaje de error absoluto medio diario
DR	Desviación relativa
E	Subconjunto de datos horarios del 01-01-2010 al 31-12-2011
EC	Criterio de evaluación considerado
ECMWF	European Centre for Medium-Range Weather Forecasts
EEMD	Ensemble empirical mode decomposition
ELEC	Técnica eléctrica
ELM	Extreme Learning Machine
EM	Algoritmo expectación maximización
EMOS	Ensemble Model Output Statistics
En	Subconjunto de datos horarios del 01-01-2010 al 31-12-2011 (75 %)
EPS	Ensemble Prediction System
ER	Error relativo
ERM	Error relativo medio
EP	Error en tanto por ciento perdido (missing rate error)
EST	Técnica estadística
ET	Ecuación del tiempo
FA	Red borrosa ARTMAP (Adaptive Resonance Theory Mapping)
FCM	Algoritmo Fuzzy C-means
FF	Algoritmo firefly
FIS	Fuzzy Inference Systems
FISI	Técnica física
FK	Filtro de Kalman
FLC	Fuzzy Logic Controllers
FTDNN	Focused Time-Delay Neural Network
GCOS	Global Climate Observing System
GD	Descenso por el gradiente
GDAS	Global Data Assimilation System
GFS	Global Forecast System
GHI	Irradiancia global sobre el plano horizontal
GLSSVM	Group Least Square Support Vector Machine
GM	Modelo gris
GMDH	Group Method of Data Handling
Gm	Atributos en modelos para cualquier tipo de día, $m \in \{1, \dots, 8\}$ horas
GS	Gauss-Seidel
GSi	Gridpoint Statistical Interpolation
HIRLAM	High Resolution Limited Area Model
HISIMI	Modelo HIstorical SIimilar MIning

Término	Significado
Hm	Atributos en modelos para cualquier tipo de día, $m \in \{1, \dots, 8\}$ horas
IA	Inteligencia Artificial
IFS	Integrated Forecast System
IMF	Intrinsic Mode Function
ISAC	Institute of Atmospheric Sciences and Climate
KKT	Karush-Kuhn-Tucker
KNN	K vecinos más próximos
kWh	Abreviatura de kilovatio-hora
kWp	Kilovatio pico. Máxima potencia generada bajo condiciones estándar
L	Longitud
LAMs	Modelos de área limitada
LB	Lógica borrosa
LS-SVM	Least Square Support Vector Machine
M	Masa de aire
MAE	Error absoluto medio
MAG	Modelo aditivo generalizado
MAPE	Porcentaje de error absoluto medio
MARE	Error relativo absoluto medio
MASS	Mesoscale Atmospheric Simulation System
MC	Simulación de Monte-Carlo
MC2	Mesoscale Compressible Community Model
MDM	Mecanismo basado en minería de datos
ME	Error medio
Meso-NH	Mesoscale no-Hidrostatic Atmospheric Simulation System
MINC	Mínimos cuadrados
MM5	5th generation Mesoscale Model
ML	Modelo lineal
MLP	Perceptrón multicapa
MPVF	Modelo basado en una MLP con dos capas ocultas y aprendizaje BP
MSE	Error cuadrático medio
MSE_0	Error cuadrático medio por el modelo promedio global
M5P	Algoritmo M5P
M5Rules	Método M5Rules
m(0.2)	Método de optimización descenso por el gradiente con momentum = 0.2
NARX	Modelo autorregresivo no lineal con entradas exógenas
NB	Neuro-borroso
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NCO	Paquete netCDF Operators
NMAE	Error absoluto medio normalizado
NMAPE	Porcentaje de error absoluto medio normalizado
NME	Error medio normalizado
NMM	NonHydrostatic Mesoscale Model
NRMSE	Raíz cuadrada del error cuadrático medio normalizado

Término	Significado
NWP	Predicciones numéricas meteorológicas
OMM	Organización meteorológica mundial
P	Ponderación del RMSE de los subconjuntos: En (75 %) y V (25 %)
PDF	Función de densidad de probabilidad
PeEn	Conjunto de persistencia
PEP	Modelo de verificación de probabilidad de error pequeño
PM	Matriz de probabilidad
POM	Modelo polinomial multivariante
Pond	Ponderación del RMSE de los subconjuntos: E (66.66 %) y T (33.33 %)
PPM	Matriz pseudo-probabilidades
PSHLTR	Presión atmosférica a 2 <i>m</i> de altura (<i>Pa</i>)
PSO	Optimización de enjambre de partículas (Particle Swarm Optimization)
QR	Técnica de regresión cuantil
QS	Humedad relativa en superficie (entre 0 y 1)
QSHLTR	Humedad potencial relativa en superficie (entre 0 y 1)
Q10	Humedad relativa a 10 <i>m</i> (entre 0 y 1)
R	Coefficiente de correlación lineal
RA	Razonamiento análogo
RAD-POT	Radiación-potencia
RAMS	Regional Atmospheric Modeling System model
RBF	Función de base radial
R^2	Coefficiente de determinación
RE	Referencia ecuatorial o horaria
RG	Regresión general
RH	Referencia horizontal o azimutal
RF	Random forests
Rm	Atributos en modelos para cualquier tipo de día, $m \in \{1, \dots, 8\}$ horas
RNA	Red neuronal artificial
RNAG	Red neuronal artificial generalizada
RNAMG	Red neuronal retardada en el tiempo con memoria gamma
RNAR	Red neuronal recursiva
RMSE	Raíz cuadrada del error cuadrático medio
RSWIN	Radiación de onda corta sobre superficie horizontal (W/m^2)
RSWINC	Radiación de onda corta sobre superficie horizontal de día claro (W/m^2)
R24	Atributos en modelos para cualquier tipo de día (horas entre 0 y 24)
R48	Atributos en modelos para cualquier tipo de día (horas entre 25 y 48)
R72	Atributos en modelos para cualquier tipo de día (horas entre 49 y 72)
SARIMA	Modelo ARIMA con estacionalidad multiplicativa
SE	Método de suavizado exponencial
SDE	Desviación estándar de los errores
SFCR	Sistemas fotovoltaicos conectados a red
SOM	Redes auto-organizativas (Self-Organizing Maps)
SPF	Predicciones probabilísticas de potencia solar
ST	Series temporales

Término	Significado
SVM	Máquinas de vectores soporte
SVR	Máquinas de vectores soporte para regresión
T	Subconjunto de datos horarios del 01-01-2012 al 28-12-2012
TDNN	Red neuronal artificial retardada en el tiempo
THS	Temperatura en superficie ($^{\circ}K$)
TH10	Temperatura a 10 <i>m</i> ($^{\circ}K$)
TIC	Coefficiente Theil IC
TL	Tiempo legal
TSHLTR	Temperatura potencial en superficie ($^{\circ}K$)
TSK	Takagi-Sugeno-Kang
TSM	Tiempo solar medio en el meridiano de Greenwich
TSM(L)	Tiempo solar medio de un lugar cualquiera de longitud L
TU	Tiempo universal
UA	Unidad astronómica o distancia media entre la Tierra y el Sol
U10	Componente <i>u</i> del viento a 10 <i>m</i> de altura
V	Subconjunto de datos horarios del 01-01-2010 al 31-12-2011 (25 %)
VAR	Varianza del error
VD	The ensemble variance déficit
VEL10	Velocidad del viento a 10 <i>m</i> de altura
V10	Componente <i>v</i> del viento a 10 <i>m</i> de altura
WT	Transformada wavelet
WRF	Weather Research and Forecast Model

Parte I

INTRODUCCIÓN

Capítulos:

1. Justificación, objetivos, estructura y contenido de la tesis
2. Energías renovables. Energía solar fotovoltaica
3. Estado actual de la predicción de la potencia eléctrica media generada en una planta fotovoltaica
4. Astronomía terrestre y solar
5. Predicción numérica del tiempo
6. Radiación solar

Capítulo 1

Justificación, objetivos, estructura y contenido de la tesis

1.1. Justificación

La predicción a corto plazo (hasta 72 horas) de la energía producida en una planta eléctrica tiene importantes repercusiones en su funcionamiento económico y fiable. Además, desde la puesta en marcha de los mercados eléctricos, el gestor de cada planta generadora que participe en los mismos debe suministrar al mercado los valores de energía que va a generar en cada una de las horas del día siguiente. En el caso de una planta eléctrica convencional no hay problema, ya que el valor de la producción de la planta se puede ajustar a voluntad hasta capacidad de la misma, pero en el caso de plantas basadas en energías renovables, donde no hay capacidad de regulación por parte del gestor, hay que aportar unos valores de predicción. Esto, unido al importante incremento de plantas generadoras basadas en energías renovables instaladas en los últimos años, hace que los sistemas de predicción a corto plazo de la energía generada en este tipo de plantas sean fundamentales para la gestión del propio sistema eléctrico.

1.2. Objetivos

1.2.1. Objetivo general

El objetivo general de la tesis es el de desarrollar nuevos modelos avanzados de predicción a corto plazo de la generación eléctrica en plantas basadas en energía solar fotovoltaica, con el fin de lograr, con los datos históricos que se disponen de la producción eléctrica y de las predicciones de las variables meteorológicas, las mejores predicciones de potencia eléctrica media generada en cada una de las horas incluidas en el horizonte de predicción.

1.2.2. Objetivos específicos

Además del objetivo general, se plantean cinco objetivos específicos:

- Desarrollar modelos de predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica con horizontes de predicción de hasta 72 horas. Se utilizarán técnicas incluidas dentro de la minería de datos, algunas de las cuales no se han aplicado hasta ahora en este tipo de modelos de predicción.
- Crear sistemas de predicción más económicos de la producción eléctrica: los datos utilizados para obtener las predicciones serán valores de variables medidas en la misma planta o públicos y accesibles sin ningún coste.
- Producir nuevas herramientas para la evaluación de la incertidumbre en las predicciones.
- Analizar las predicciones de producción eléctrica logradas con los nuevos modelos desarrollados.
- Determinar el intervalo temporal a partir del cual el uso de predicciones meteorológicas, como variables de entrada en los modelos, mejora las predicciones de la potencia media generada.

1.3. Estructura y contenido

En los trabajos de investigación de esta tesis se han seguido las fases del proceso de extracción de conocimiento representadas en el esquema de la Figura 1.1. La tesis se organiza en cuatro partes, que incluyen diecisiete capítulos, y un apéndice.

En la parte I se introduce la parte más general de la tesis:

- En el capítulo 1 se presenta la justificación, los objetivos, la estructura y el contenido de la tesis.
- En el capítulo 2 se describe brevemente la energía solar fotovoltaica, su evolución histórica, sus ventajas e inconvenientes.
- En el capítulo 3 se analiza el estado actual de los modelos de predicción a corto plazo de la potencia eléctrica media, generada en plantas basadas en energía solar fotovoltaica. Se hace un breve resumen de los diferentes modelos publicados en artículos de la literatura internacional.
- En el capítulo 4 se centra en las nociones generales de astronomía terrestre y solar: las referencias locales, el ángulo de incidencia de los rayos solares directos, y las coordenadas temporales.
- En el capítulo 5 se hace un breve repaso de la predicción numérica del tiempo atmosférico, de su evolución histórica, de las rejillas de integración, de los procesos de ejecución, y de los tipos de modelos usados actualmente. Se describe el cálculo de la predicción de las distintas variables meteorológicas obtenidas, utilizando el modelo mesoescalar WRF-NMM.

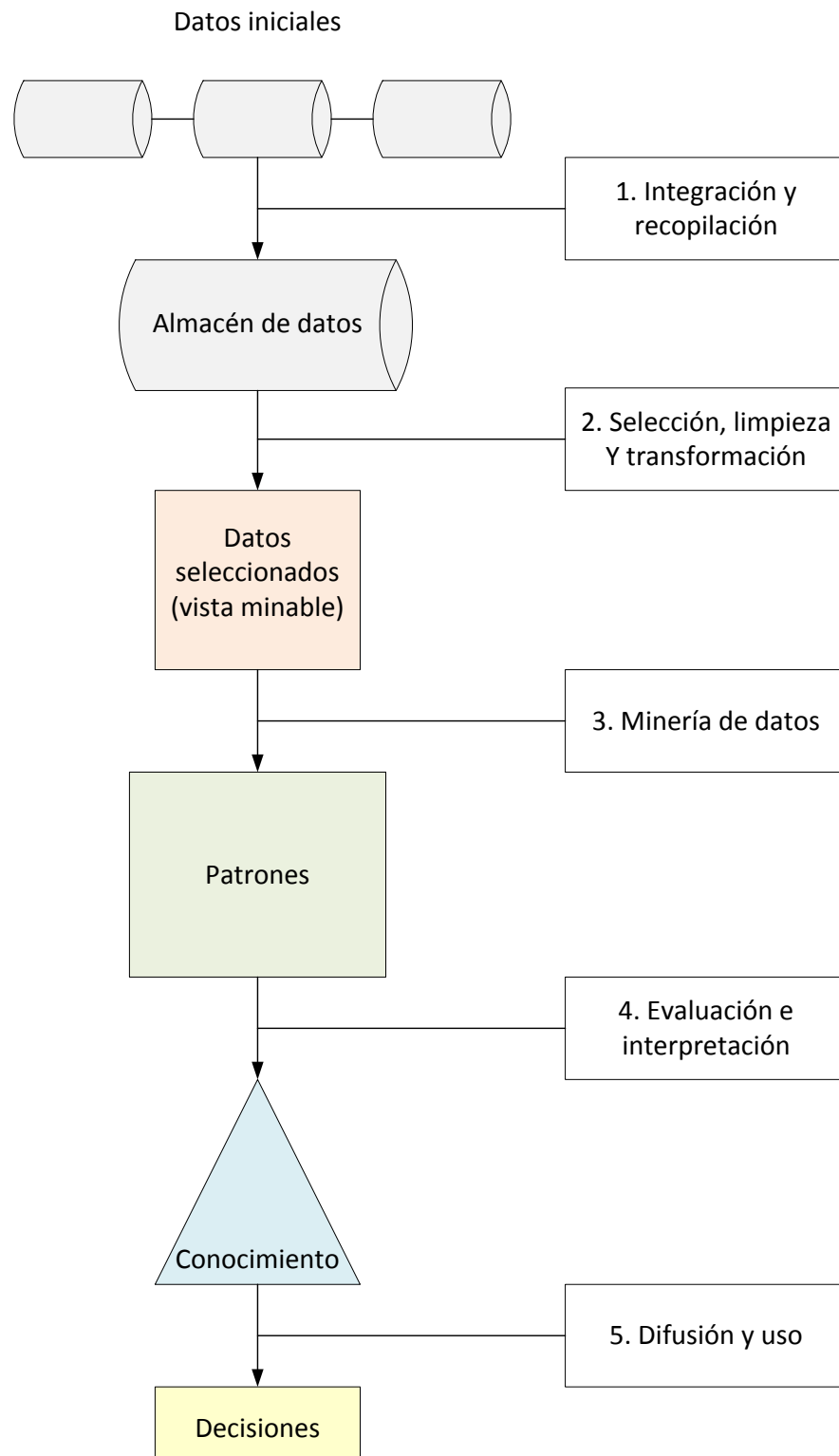


Figura 1.1: Fases del proceso de descubrimiento de conocimiento en bases de datos (adaptado de [Hernández 04]).

- En el capítulo 6 se introduce el concepto de la radiación solar, sus componentes y modelos.

En la parte II se muestran las bases de las técnicas de minería de datos y probabilidad, que se utilizan en esta tesis, para la creación de los modelos de predicción de la potencia eléctrica media horaria:

- En el capítulo 7 se exponen las técnicas y métodos para el aprendizaje basados en modelos de árboles y sistemas de reglas.
- En el capítulo 8 se describen los aspectos básicos de las redes neuronales. Se estudian algunos modelos de redes neuronales artificiales: el perceptrón multicapa entrenado mediante el algoritmo de retropropagación de errores o con algunas variantes, y las redes neuronales de función de base radial.
- En el capítulo 9 se presentan los fundamentos teóricos de las máquinas de vectores soporte. Éstas son interdisciplinarias, utilizan técnicas de optimización, estadística y análisis funcional.
- En el capítulo 10 se centra en los sistemas basados en lógica borrosa, se describen sus bases conceptuales y la estructura de un sistema de inferencia borroso.
- En el capítulo 11 se aporta una introducción a los algoritmos evolutivos, y en especial al funcionamiento básico de los algoritmos genéticos.
- En el capítulo 12 se proporciona un breve repaso a los principales conceptos de probabilidad, que se utilizan en uno de los modelos desarrollados en los trabajos de investigación de esta tesis.

En la parte III se utilizan diferentes técnicas de evaluación de los modelos para estimar la proporción de error:

- En el capítulo 13 se explican las técnicas que se utilizan en los trabajos de investigación de esta tesis, para la evaluación de los modelos de predicción desarrollados.

En la parte IV se presentan los nuevos modelos de predicción desarrollados, sus resultados computacionales, las conclusiones y principales aportaciones de la tesis, así como varias líneas de continuación:

- En el capítulo 14 se describen los nuevos modelos desarrollados de predicción de potencia eléctrica media horaria generada en una planta fotovoltaica, para un horizonte de predicción a corto plazo, utilizando técnicas de minería de datos: árboles de decisión, sistemas de reglas, redes neuronales artificiales, máquinas de vectores soporte y sistemas basados en lógica borrosa; muchos de los modelos han sido optimizados con algoritmos genéticos.
- En el capítulo 15 se especifica la formulación del nuevo modelo desarrollado de predicción con evaluación de incertidumbre, se explican todas sus características y se describen todas sus etapas.

- En el capítulo 16 se exponen los mejores resultados computacionales, obtenidos de los ensayos efectuados con los nuevos modelos desarrollados. Todos los modelos expuestos, poseen unas características distintas a las de otros publicados en la literatura internacional y, además, los mejores consiguen un menor error en sus predicciones.
- En el capítulo 17 se explican todas las conclusiones relevantes extraídas de la elaboración de los trabajos de investigación y las principales aportaciones de esta tesis. Además, se proponen varias líneas de continuación a partir de los trabajos de investigación desarrollados.

En el apéndice A se hace referencia a los artículos científicos publicados en la literatura internacional, que son en parte fruto del trabajo de investigación realizado en esta tesis doctoral.

Capítulo 2

Energías renovables. Energía solar fotovoltaica

2.1. Energías renovables

Solo las fuentes de energía que explotan recursos renovables, como el viento, el agua, la biomasa, la geotermia, o el sol, son inagotables a escala humana. Además, las energías renovables no contribuyen a incrementar el CO_2 y otros gases causantes del efecto invernadero y el calentamiento global, así que su aprovechamiento es sostenible y básico para conseguir la seguridad de abastecimiento [Romero 10]. Entre las energías renovables aprovechables actualmente merecen destacarse [Sanz 09]:

- Hidráulica:
 - Se obtiene de la caída del agua desde cierta altura a un nivel inferior, lo que provoca el movimiento de ruedas hidráulicas o turbinas.
 - Su desarrollo requiere una gran inversión (pantanos, presas, etc.).
- Eólica:
 - Se debe a la energía cinética del aire, la potencia que se obtiene es directamente proporcional al cubo de la velocidad del viento.
 - La electricidad se genera a través de unos aerogeneradores con un tamaño medio de 600–5000 kW con áreas de barrido de palas de hasta 65 metros de diámetro.
 - Hay dos tipos de instalaciones: aisladas y concentradas en forma de parques eólicos.
- Solar:
 - En instalaciones de aprovechamiento de la energía solar térmica se utilizan colectores para concentrar la luz solar sobre un fluido y utilizar directamente la energía calorífica absorbida en forma de calor o utilizarla para producir energía eléctrica.

- En instalaciones fotovoltaicas se produce directamente energía eléctrica a través de unas células fotovoltaicas construidas por un material cristalino semiconductor, fundamentalmente silicio.

2.2. La Energía Solar Fotovoltaica

2.2.1. Evolución histórica

Algunas fechas clave en el desarrollo de la energía solar fotovoltaica son [Romero 10]:

- En 1839: E. Becquerel descubre el efecto fotovoltaico (proceso de producción de corriente en un material sólido con la ayuda de la luz solar).
- En 1870: H. Hertz produce celdas fotovoltaicas que convertían la luz en electricidad con una eficiencia del 1 al 2 %.
- En 1873: W. Smith descubre la fotoconductividad de selenio.
- En 1877: W. G. Adams y R. E. Day construyen la primera celda de selenio.
- En 1904: A. Einstein publica su trabajo acerca del efecto fotovoltaico.
- En 1921: A. Einstein recibe el premio Nobel por sus teorías sobre el efecto fotovoltaico.
- En 1951: Se desarrolla el procedimiento Czochralski, que permitió generar cristales de silicio de alta pureza. En este mismo año se produce una celda de germanio mono-cristalino.
- En 1954: D. M. Chaplin, C. S. Fuller y G. L. Pearson, de los laboratorios Bell, producen la primera célula de silicio, con una eficiencia del 4,5 %.
- En 1955: se le asigna a la industria americana la tarea de producir elementos solares fotovoltaicos para aplicaciones espaciales.
- En 1957: Hoffman Electronic alcanza el 8 % de rendimiento en sus células.
- En 1958: puesta en órbita el primer satélite alimentado parcialmente con energía solar.
- En 1959: Hoffman Electronic alcanza el 10 % de rendimiento en sus células comerciales.
- En 1962: se lanza el primer satélite comercial de telecomunicaciones, con una potencia instalada en paneles fotovoltaicos de 14 W.
- En 1963: Sharp consigue una forma práctica de producir módulos de silicio; en Japón se instala un sistema de 242 W en un faro.
- En 1964: se lanza la nave espacial Nimbus con 470 W de potencia instalada en paneles fotovoltaicos.

- En 1973: la crisis del petróleo incrementa el interés en las energías renovables. La producción mundial de células alcanza los 100 kW.
- En 1974: se fundan las primeras compañías de energía solar.
- En 1975: las aplicaciones terrestres superan a las aplicaciones espaciales.
- En 1977: la potencia total instalada de energía solar fotovoltaica supera los 500 kW.
- En 1981: un avión abastecido por energía solar fotovoltaica, realiza su primer vuelo.
- En 1983: la potencia instalada a nivel mundial supera los 21,3 MW.
- En 1994: se celebra la primera Conferencia Mundial Fotovoltaica en Hawai.
- En 2000: Alemania y Japón despuntan a nivel mundial como productores y, en el caso de Alemania, como consumidores.
- En 2004: se producen e instalan 1.000 MW en generadores fotovoltaicos en todo el mundo.
- En 2007: se duplica la producción de módulos fotovoltaicos (más de 2.000 MW).
- En la actualidad: la mayoría de las células solares se basan en el silicio. Los módulos fotovoltaicos tienen rendimientos del 7 al 17 %, son altamente fiables y su vida media supera los 30 años.

2.2.2. Ventajas de la utilización de la energía solar

Como ventajas que proporciona la utilización de la energía solar, pueden citarse las siguientes:

- Tiene un inmenso potencial, al no tener límite la energía solar que se puede captar. Su presencia en toda la superficie terrestre permite la ubicación de sistemas fotovoltaicos en cualquier parte [Romero 10].
- Se requiere una infraestructura muy simple, comparada con otros sistemas de generación eléctrica [Jutglar 12].
- No consume combustible, ni produce contaminación ambiental y además es silenciosa. Es resistente a condiciones climáticas extremas: granizo, viento, temperatura y humedad [Sanz 09].
- La vida útil de los generadores fotovoltaicos es elevada y además exige un mantenimiento mínimo en tiempo, costes y especialización profesional [Romero 10].
- Se puede aumentar la potencia instalada mediante la incorporación de nuevos módulos [Sanz 09].
- La producción máxima se alcanza normalmente a mediodía, justo cuando hay altos valores de demanda de energía [Romero 10].
- Es una solución muy adecuada para suministrar electricidad a áreas remotas o rurales, aisladas de la red eléctrica [Romero 10].

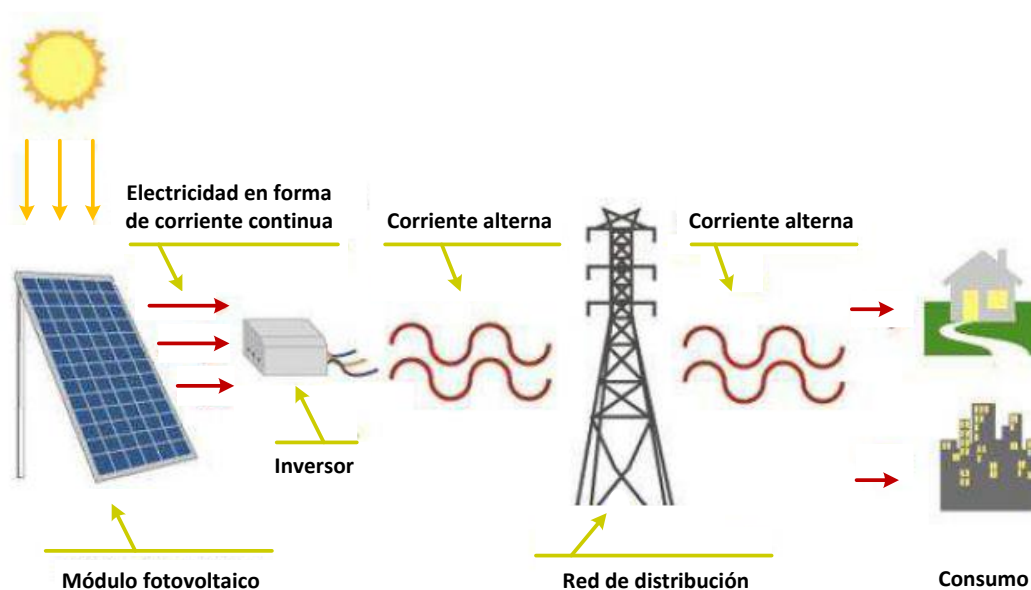


Figura 2.1: *Esquema de la estructura de un sistema solar fotovoltaico conectado a la red [Vilchez 09].*

2.2.3. Inconvenientes de la utilización de la energía solar

Los inconvenientes más resaltables de la utilización de la energía solar son los siguientes:

- Los rendimientos de los sistemas aún son bajos y la instalación completa no es muy barata [Sanz 09].
- Es una generación discontinua de electricidad debido a la propia naturaleza de la fuente de energía (día-noche) [Romero 10].
- En sistemas aislados se requiere una mayor capacidad de almacenamiento de energía para abastecer normalmente consumos nocturnos [Romero 10].
- Este tipo de energía siempre estará sometida a la aleatoriedad de las condiciones atmosféricas [Sanz 09].

2.2.4. Sistemas fotovoltaicos conectados a la red

Los sistemas fotovoltaicos conectados a la red (SFCR) son aquellos cuya instalación está motivada por la inyección de energía eléctrica a la red, ya sea con ánimo de venta de la producción eléctrica de nuestro sistema fotovoltaico o como apoyo a la red eléctrica [Sanz 09].

Están compuestos por un generador fotovoltaico que se encuentra conectado a la red eléctrica convencional a través de un inversor, transfiriéndose energía a la red [Sanz 09], (ver Figura 2.1).

Capítulo 3

Estado actual de la predicción de la potencia eléctrica media generada en una planta fotovoltaica

En este capítulo, se presenta el estado actual de los modelos de predicción a corto plazo de la potencia eléctrica media publicados en la literatura internacional.

3.1. Primeros trabajos

En los primeros trabajos se trata de predecir la radiación solar, no la generación eléctrica. Se desarrollaron modelos basados en técnicas de inteligencia artificial para aplicaciones fotovoltaicas [Mellit 08]. En la Tabla 3.1 se recogen los modelos de predicción de la radiación solar más relevantes publicados en la literatura científica.

Tabla 3.1: *Breve resumen de la estimación de la radiación solar (adaptado de [Mellit 08]).*

Referencia	Técnica	Tema
[Elizondo 94]	RNA	Predicción de la radiación solar diaria
[Williams 94]	RNA	Estimación de la radiación solar diaria
[Williams 96]	RNA	Estimación de la radiación solar diaria
[Alawi 98]	RNA	Radiación solar en áreas no cubiertas por medidas directas con instrumentos
[Guessoum 98]	RNA	Predicción de la radiación solar diaria
[Mohandes 98]	RNA	Predicción de datos de radiación solar global media mensual
[Sen 98]	LB	Estimación de la radiación solar
[Hammer 99]	EST	Predicción de la radiación solar
[Hontoria 99]	RNA	Generador de series de radiación solar
[Kemmoku 99]	RNA	Predicción de la insolación diaria

(Sigue en la página siguiente.)

Continuación de la tabla

Referencia	Técnica	Tema
[Zufiria 99]	RNA	Generador de series de radiación solar
[Hontoria 00]	RNA	Generación del mapa de radiación solar
[Mihalakakou 00]	RNA	Simulación de la radiación solar total
[Mohandes 00]	RNA	Media mensual de valores diarios de radiación solar global en superficies horizontales
[Sfetsos 00]	RNA	Pronóstico de la radiación solar
[Hontoria 01]	RNA	Generación de mapas de radiación solar
[Kalogirou 01]	RNA	Revisión de modelos basados en IA para sistemas de energía renovable
[Dorvio 02]	RNA	Estimación de la radiación solar y índ. de claridad
[Hontoria 02]	RNA	Generador de series artificiales de radiación solar
[Kalogirou 02]	RNA	Predicción de máxima irradiación
[Safi 02]	EST	Predicción de la radiación solar global diaria
[Tymvios 02]	RNA	Energía solar total sobre el nivel horizontal
[Gomez 03]	LB	Modelo de radiación solar
[Reddy 03]	RNA	Estimación diaria y mensual de la radiación solar
[Mellit 04]	RNA	Estimación de la radiación solar de la luz del sol
[Cao 05]	WT-RNA	Estimación de la radiación solar total diaria
[Hontoria 05]	RNA	Generador de series artificiales de radiación solar
[Iqdour 05]	LB	Predicción de la radiación solar global
[Lopez 05]	RNA	Estimación de la irradiación directa
[Mellit 05]	RNA	Predicción de la radiación solar diaria
[Sozen 05]	RNA	Irradiación solar mensual media
[Tymvios 05]	RNA	Estimación de la radiación solar global
[Zarzalejo 05]	RNA	Estimación de la radiación horaria de satélite
[Alam 06]	RNA	Cálculo de la radiación directa
[Cao 06]	WT-RNA	Estimación de la radiación solar total diaria
[Mellit 06a]	WT-RNA	Estimación de la radiación solar total diaria
[Mellit 06b]	WT-RNA	Predicción de ST aplicadas a la radiación solar
[Elminir 07]	RNA	Predicción de la fracción difusa K_D
[Gabriel 07]	RNA	Eficacia de la luz solar con cielo claro
[Kapanis 07]	EST	Predicción de la radiación solar global horaria
[Mellit 07]	NB	Estimación de la radiación solar
[Hocaoğlu 08]	RNA	Predicción de la radiación solar horaria
[Mubiru 08]	RNA	Estimación de la radiación solar mensual media
[Zervas 08]	RNA	Predicción de la irradiancia solar global diaria
[Kapanis 10]	EST	Predicción de la radiación solar global horaria
[Martín 10]	AR, RNA, NB	Predicción de la irradiancia solar global diaria
[Mellit 10]	RA	Predicción de la irradiancia solar horaria global, directa y difusa
[Paoli 10]	RNA	Predicción de radiación solar diaria

(Sigue en la página siguiente.)

Continuación de la tabla

Referencia	Técnica	Tema
[Voyant 11]	RNA	Predicción de la radiación global diaria

3.2. Evolución histórica en los últimos años

A continuación, se describen brevemente las características más relevantes de los modelos de predicción de la generación eléctrica en una planta fotovoltaica publicados en la literatura científica en los últimos años.

En 2009:

- En [Bacher 09] se predice la potencia eléctrica media horaria para el horizonte hasta 36 horas. Se utilizan:
 - Observaciones cada 15 min. de la potencia de 21 sistemas fotovoltaicos con un promedio de 2769 W, ubicados en Dinamarca.
 - Predicciones de la irradiancia global proporcionadas por el Instituto Meteorológico de Dinamarca. Se utiliza el modelo HIRLAM.
 - Período de estudio: año 2006.

Metodología: se obtiene una normalización estadística de la potencia generada usando un modelo de cielo claro, que se encuentra utilizando técnicas de suavizado. Después, se calculan las predicciones de la potencia utilizando modelos de series temporales lineales AR y ARX.

Resultados: se indica que para las predicciones con horizonte hasta 2 horas, las entradas más importantes son las observaciones disponibles de la potencia, mientras que para horizontes mayores, las entradas más importantes son los valores previstos de variables meteorológicas. Para evaluar los modelos se utiliza el RMSE y se comparan con el de un modelo de referencia. El modelo ARX es el que mejores predicciones obtiene.

- En [Li 09] se predice la potencia eléctrica media cada 15 min. para el horizonte de un día de antelación. Se utilizan:
 - Datos de potencia en intervalos de 15 min. de un sistema fotovoltaico conectado a red con 5.6 kWp.
 - Período de estudio: de 2007 hasta 2009.

Metodología: se establecen dos “modelos grises” (*Grey forecast model*), del tipo GM(1,1). En el primero se aplican los datos de 8:00 a 11:15, y en el segundo de 12:15 a 14:15. A continuación, se calcula la matriz de transferencia de probabilidad de la desviación relativa mediante el análisis de la cadena de Markov avanzada.

Los resultados de las predicciones de los GM(1,1) se modifican basándose en la desviación relativa y en su probabilidad.

Resultados: los resultados de los modelos se comparan con la desviación relativa (DR), y se indica que el modelo de la cadena Grey-Markov avanzada es más preciso.

En 2010:

- En [Huang 10] se predice la potencia eléctrica media cada 15 min. para el horizonte de un día de antelación. Se utilizan:
 - Datos de potencia en intervalos de tiempo de 15 min. de una planta fotovoltaica de 1 MW.
 - Datos de irradiancia y temperatura, medidos cada 15 min. en la planta fotovoltaica.
 - Predicciones cada 15 min. de la irradiancia solar, la temperatura del aire, el coeficiente de nubosidad, la humedad, proporcionados por un modelo de mesoescala.
 - Posición del sol.
 - Período de estudio: 1 año de datos.

Metodología: se estudian dos modelos. Un modelo de red neuronal, con una capa oculta, que utiliza el algoritmo de retropropagación de errores o BP (*back-propagation*) y como entradas las predicciones meteorológicas. Un modelo físico que se obtiene por combinación de la posición del sol y el modelo de la célula fotovoltaica. Para comprobar la validez del modelo físico y el impacto de las predicciones de la irradiancia y de la temperatura, utiliza estas variables como entradas del modelo.

Resultados: para la comparación de los resultados obtenidos con los diferentes modelos, se utiliza el NRMSE. El modelo que da menor error para todo el año es el basado en red neuronal. Por otra parte, se encontró que el principal origen de los errores de predicción proviene de la no exactitud de la información de las predicciones meteorológicas. Además, los datos de irradiancia en tiempo real se utilizan como entradas del modelo, mejorando aún más el rendimiento de la predicción.

En 2011:

- En [Wang 11] se predice la potencia eléctrica media horaria durante 12 horas, entre las 7:00 am. y las 6:00 pm. Se utilizan:
 - Datos de potencia suministrados por una instalación fotovoltaica.
 - Período de estudio: del 12 al 17 de abril.

Metodología: se combinan un modelo gris y un modelo de red neuronal que entrena las neuronas de las capas ocultas con el algoritmo de retropropagación de errores (BP). Se toma la salida de potencia de un sistema fotovoltaico y se construye un modelo GM(1,1) eligiendo muestras adecuadas. Luego, se construye el modelo de

red neuronal BP, utilizando series de error residual entre los valores ajustados y los valores reales. Finalmente, se modifica el valor GM(1,1) acorde al valor de la salida de la red neuronal.

Resultados: para comparar los resultados obtenidos con el modelo GM, y la combinación del modelo GM con una RNA-BP, se elige como criterio de evaluación el error relativo medio (ERM). Los resultados muestran que la aplicación de la combinación aumenta la precisión del modelo de predicción.

En 2012:

- En [Fernández 12a] se predice la potencia eléctrica media horaria para el horizonte de 1 a 39 horas. Se utilizan:
 - Datos de la potencia horaria de una planta fotovoltaica con una capacidad de 36 kWp, en La Rioja, España.
 - Predicciones de: flujo de calor sensible de la superficie, flujo de calor latente de la superficie, radiación de onda corta hacia la superficie, radiación de onda larga hacia la superficie, radiación de onda corta saliente, radiación de onda larga saliente y temperatura. Estos valores se obtienen con el modelo de predicción numérica meteorológica de mesoscala MM5 para horizontes de 0 a 48 horas en intervalos de 15 min.
 - Momento de año y del día correspondiente al horizonte de predicción.
 - Período de estudio: del 2 de junio de 2007 hasta el 25 de mayo de 2008.

Metodología: se lleva a cabo un proceso de optimización (con algoritmo genético) de diferentes tipos de RNA seleccionadas, que son el perceptrón multicapa (MLP), la red recurrente Elman, la de función de base radial (RBF), y las retardadas en el tiempo (TDNN). Además, se testean otras familias diferentes de modelos: persistencia, ARIMA, K vecinos más próximos (KNN) y sistemas de inferencia neuro-difusos (ANFIS).

Resultados: el rendimiento de los modelos se evalúan con RMSE, MAE y ME. Se concluye que el mejor modelo de predicción es el que se nombra como MLP 2A, es una RNA MLP, con dos capas ocultas.

- En [Fernández 12b] se predice la potencia eléctrica media horaria para el horizonte hasta 48 horas. Se utilizan:
 - Datos históricos de potencia horaria de una planta fotovoltaica conectada a red, con capacidad de 36 kWp, localizada en La Rioja, España.
 - Predicciones de la radiación de onda corta hacia la superficie y de la temperatura media horaria, obtenidas con un modelo de mesoescala.
 - Período de estudio: 1 año.

Metodología: se basa en un sistema difuso *Takagi-Sugeno-Kang* (TSK), que se entrena con la metodología ANFIS [Jang 93]. Las variables de entrada (escogidas de

entre las disponibles), y los valores de los parámetros necesarios para el algoritmo agrupamiento substractivo utilizado para reducir el número de reglas del sistema difuso TSK, se seleccionan con un proceso de optimización controlado por un algoritmo genético (AG). Se utilizan 50 generaciones con 50 individuos, las tasas de cruce y mutación son de 90 % y 1 % respectivamente, y se aplica elitismo.

Resultados: para la evaluación del modelo se usa la raíz cuadrada del error cuadrático medio normalizado (NRMSE). El modelo propuesto logra resultados satisfactorios.

- En [Fonseca 12] se predice la potencia eléctrica media horaria para el horizonte de 1 hora. Se utilizan:
 - Datos de medidas horarias de potencia de una planta fotovoltaica de 1 MWp en Japón.
 - Valores previstos para la hora correspondiente al horizonte de predicción y para la hora anterior de temperatura normalizada, humedad relativa y nubosidad (nivel: bajo, medio y alto). Estos valores son proporcionados por el sistema de predicción de tiempo de la Agencia meteorológica de Japón.
 - Datos calculados (para la hora correspondiente al horizonte de predicción y para la hora anterior) de la insolación extraterrestre.
 - Período de estudio: de mayo de 2008 a julio de 2010.

Metodología: se utilizan las máquinas de vectores soporte para regresión (SVR).

Resultados: para evaluar la exactitud de las predicciones se consideran la raíz cuadrada del error cuadrático medio (RMSE), el error absoluto medio (MAE), y el porcentaje de error absoluto medio (MAPE). Los resultados muestran que el uso de SVR y la nubosidad pronosticada numéricamente, producen pronósticos de producción de potencia con un buen nivel de precisión.

- En [Lin 12] se predice la potencia eléctrica media semanal. Se utilizan:
 - Previsiones del tipo de nubosidad obtenida por predicciones meteorológicas.
 - La fórmula [Smith 84] de la relación entre la radiación solar y la nubosidad (se utilizan 19 años de datos meteorológicos de una estación meteorológica).

Metodología: se introduce un modelo que predice la radiación solar horizontal, utilizando la información del tipo de nubosidad obtenida de la predicción meteorológica.

Resultados: se muestra que el modelo puede predecir la salida con relativa exactitud, utiliza los datos de la potencia eléctrica generada en estudios de flujos de cargas en sistemas de distribución.

- En [Mandal 12] se predice la potencia eléctrica media horaria para el horizonte de 12 horas. Se utilizan:
 - Datos de potencia horaria recolectados de un sistema fotovoltaico de 15 kWp, localizado en Oregon.

- Datos históricos horarios de radiación solar y temperatura.
- Período de estudio: 30 días del 2011.

Metodología: se utiliza un método híbrido basado en la combinación de la transformada wavelet (WT) y una red neuronal con funciones de base radial (RBF).

Resultados: para evaluar la exactitud del modelo propuesto WT-RBF se utilizan como criterios el porcentaje de error absoluto medio (MAPE), el error absoluto medio (MAE), y la raíz cuadrada del error cuadrático medio (RMSE). La capacidad de predicción del modelo WT-RBF se confirma comparando sus resultados con una RNA entrenada con el algoritmo de retropropagación de errores (BP), una RNA con RBF y una WT-RNA (BP). Los resultados demuestran que el modelo WT-RBF funciona con mayor precisión.

- En [Pedro 12] se predice la potencia eléctrica media horaria para el horizonte de 1 y 2 horas. Se utilizan:
 - Datos de potencia de una planta fotovoltaica de 1 MWp, en California.
 - Período de estudio: de noviembre 2009 a agosto de 2011.

Metodología: Se estudian los siguientes métodos: persistente, ARIMA, K vecinos más próximos (KNN), redes neuronales artificiales (RNA) unidireccionales con una capa oculta y entrenadas con el algoritmo de retropropagación de errores Levenberg-Marquardt, y redes neuronales artificiales optimizadas por algoritmo genético (AG).

Resultados: la precisión de los modelos se evalúa usando el error absoluto medio (MAE), el error medio (ME), la raíz cuadrada del error cuadrático medio (RMSE), la raíz cuadrada del error cuadrático medio normalizado (NRMSE), y el coeficiente de determinación R^2 .

Los hallazgos muestran que los modelos de predicción basados en RNA obtienen mejores resultados que las otras técnicas de predicción. Se logran mejoras sustanciales con una optimización por un AG de los parámetros de la RNA.

- En [Shi 12] se predice la potencia eléctrica media cada 15 min. con un horizonte de 24 horas. Se utilizan:
 - Datos históricos de salida de potencia cada 15 min. de una instalación fotovoltaica en el sur de China, con capacidad de 20 kWp.
 - Predicciones meteorológicas para el próximo día.
 - Período de estudio: del 13 de enero al 29 de octubre de 2010.

Metodología: se proponen cuatro modelos, basándose en la clasificación de las condiciones del tiempo (días: despejado, nuboso, con niebla y con lluvia) y en el uso de máquinas de vectores soporte (SVM).

Resultados: para evaluar la precisión de los pronósticos se utilizan el error relativo medio (ERM) y la raíz cuadrada del error cuadrático medio (RMSE). Se puede observar que de los cuatro modelos el que mejores resultados obtiene es el desarrollado para días despejados.

En 2013:

- En [Bouzerdoum 13] se predice la potencia eléctrica media horaria para el horizonte de unos pocos días. Se utilizan:
 - Observaciones cada 10 min. de la potencia de una planta solar fotovoltaica de 20 kWp, instalada en Italia.
 - Medidas de datos físicos: irradiancia, temperatura del módulo, y temperatura ambiente.
 - Medidas de datos eléctricos: tensión, intensidad y potencia.
 - Período de estudio: del 29 de enero al 25 de mayo de 2009.

Metodología: se desarrollan tres modelos, uno basado en la técnica SARIMA, otro en SVM, y un tercero de tipo híbrido combinando ambos modelos.

Resultados: el comportamiento de los modelos se evalúa con NRMSE, NME, MPE, y R, y se demuestra que el modelo híbrido se comporta mejor que los modelos SARIMA y SVM.

- En [Bracale 13] se predice la potencia eléctrica media horaria para el horizonte de 1 hora. Se utilizan:
 - Medidas de la potencia media horaria de un sistema fotovoltaico de 75 kWp en Colorado.
 - Medidas de: índice de claridad, temperatura ambiente, humedad relativa, velocidad del viento, y cobertura de nubes. Estos datos son de una estación meteorológica próxima al sistema fotovoltaico.
 - Período de estudio: años 2010 y 2011.

Metodología: se pronostica la función de densidad de probabilidad del índice de claridad horario mediante el modelo de serie temporal autorregresivo Bayesiano; el modelo tiene en cuenta la dependencia de la radiación solar sobre algunas variables meteorológicas, tales como la cobertura de nubes y la humedad. Después, se utiliza el procedimiento de simulación de Monte-Carlo (MC) para evaluar la predicción de la función de densidad de probabilidad (PDF) de la potencia horaria.

Resultados: se usa el valor medio de la función de densidad de probabilidad (PDF) pronosticada para estimar la potencia generada, y para evaluar el modelo se utiliza el error relativo absoluto medio (MARE).

- En [Haque 13] se predice la potencia eléctrica media horaria hasta el horizonte de 12 horas. Se utilizan:
 - Datos de potencia horaria de un sistema fotovoltaico de 15 kWp localizado en Oregon.
 - Datos de radiación solar.
 - Período de estudio: mayo de 2011.

Metodología: se presenta un algoritmo que utiliza una combinación de una técnica de filtrado de datos, basada en la transformada wavelet (WT) y en un modelo de computación basado en una red borrosa ARTMAP (*adaptive resonance theory mapping*, FA), el cuál usa una técnica de optimización basada en el algoritmo *firefly* (FF). Los parámetros de entrada de la red FA son la potencia eléctrica media y la radiación solar de la hora actual y de la misma hora del día anterior.

Resultados: se calculan tres tipos de medidas para evaluar el modelo WT-FF-FA propuesto, el MAPE, el NRMSE y la varianza del error (VAR). Los resultados se comparan con otros modelos de predicción (RNA-BP, RNA-RBF, RNA-RG, FA, WT-RNA-BP, WT-RNA-RBF, WT-RNA-RG, y WT-FA). En todos los casos, el modelo híbrido propuesto supera a los otros modelos.

- En [Kardakos 13] se predice la potencia eléctrica media horaria para las 24 horas siguientes. Se utilizan:
 - Un conjunto de datos que incluye la potencia horaria de cuatro plantas fotovoltaicas, ubicadas en diferentes sitios de Grecia con potencias instaladas entre 0.15 y 5 MWp.
 - Medidas reales de la radiación horaria obtenidas en estaciones meteorológicas cercanas y valores previstos por medio de un modelo de predicción numérica meteorológica no especificado.
 - Período de estudio: del 1 de enero de 2011 al 31 de diciembre de 2012.

Metodología: se abordan dos métodos prácticos, SARIMA y redes neuronales artificiales con una capa oculta. La salida del modelo SARIMA es multiplicada por un factor dependiente de los valores previstos para la radiación solar (lo llama modelo SARIMA modificado). En el caso de las redes neuronales usa como entradas potencias generadas para la misma hora del horizonte de predicción en los tres días anteriores, así como diferencias entre los valores reales y previstos de radiación y potencia en el día anterior para la misma hora que el horizonte de predicción y la hora anterior.

Resultados: con el fin de evaluar el rendimiento de los modelos de predicción propuestos, estos se comparan en términos del NRMSE. Los resultados de la aplicación de todos los modelos muestran que, en general, los modelos que hacen uso de datos de radiación solar, es decir el modelo SARIMA modificado y los modelos RNA, son preferibles, ya que conducen a una mejora considerable en las previsiones en comparación con el modelo de persistencia y el modelo SARIMA puro (que no tienen en cuenta datos de radiación). El modelo SARIMA(3, 1, 2) \times (3, 1, 2)₂₄ modificado tiene el mejor rendimiento para todos sitios en términos de promedio anual.

- En [Li 13] se predice la potencia eléctrica media cada 15 min. Se utilizan:
 - Datos de la potencia cada 15 min de un sistema fotovoltaico conectado a red.
 - Datos de la radiación solar cada 15 min.
 - Período de estudio: un día.

Metodología: se aplica un modelo gris GM(1,1), cuyos valores de predicción son evaluados con un modelo probabilístico que se denomina de pequeño error (PEP).

Resultados: se sugiere que la predicción de potencia generada se haga con la combinación de la prevista con el modelo GM(1,1) y las obtenidas con otros modelos (que no detalla) cuando el modelo PEP proporcione un valor por encima de un umbral.

- En [Mao 13] se predice la potencia eléctrica media horaria para un horizonte de 24 horas. Se utilizan:
 - Datos de potencia horaria de una planta fotovoltaica conectada a red con capacidad total de 100 MWp.
 - Valores previstos de temperatura máxima y tipo de índice del tiempo (soleado 1, nublado 0.77, lluvioso 0.38, nevado 0.35), proporcionados por los servicios meteorológicos.
 - Período de estudio: del 1 de enero al 31 de julio de 2012.

Metodología: se propone un método de predicción combinado basado en *ensemble empirical mode decomposition* (EEMD) y en máquinas de vectores soporte (SVM). Primero, los datos históricos horarios de la potencia fotovoltaica se clasifican según el tipo de tiempo, y los del mismo tipo se separan usando el método EEMD, obteniendo componentes IMF (*Intrinsic Mode Function*) y un residuo. Para cada componente IMF se presenta un modelo SVM. Seleccionando los mejores parámetros y la función kernel, se obtiene el valor de la secuencia de cada componente.

Resultados: se escogen como índices de evaluación el RMSE, el MAPE y el coeficiente *Theil IC* (TIC), para evaluar los resultados. Éstos muestran que el método de combinación EEMD–SVM reduce significativamente el error de predicción en comparación con los métodos SVM y RNA (BP), aunque el índice TIC con el modelo de RNA (BP) es el menor.

- En [Monteiro 13a] se predice la potencia eléctrica media horaria para el horizonte de 1 a 24 horas. Se utilizan:
 - Valores de la potencia horaria de una planta solar fotovoltaica de 2.8 MWp. de capacidad, ubicada en España.
 - Predicciones de la radiación sobre superficie horizontal y de la temperatura proporcionadas por el modelo meteorológico de mesoescala *Weather Research and Forecasting* (WRF).
 - *Seno* y *coseno* de la hora solar para la localización de la planta en el instante correspondiente.
 - Período de estudio: un año.

Metodología: se crea un mecanismo basado en minería de datos (MDM) que selecciona los casos relevantes presentes en la base de conocimiento del modelo (conjunto de datos pasados de predicciones de las variables meteorológicas y potencia generada por la planta). Después de la aplicación de este mecanismo es posible definir una

función de probabilidad conjunta que representa las transiciones de intervalos de potencia. Su estructura final se optimiza usando un algoritmo genético (AG).

Resultados: asociado al modelo se encuentra un conjunto de herramientas. La herramienta de predicción puntual utiliza las funciones de probabilidad conjunta que representan las transiciones de potencia anteriores y posteriores en el instante para el que se pretende efectuar la predicción. Con esta herramienta es posible proporcionar el valor esperado de potencia eléctrica media horaria para un determinado instante de horizonte temporal. La herramienta de predicción de incertidumbre proporciona la incertidumbre asociada a cada predicción puntual obtenida por la herramienta anterior.

Para la evaluación el modelo se utiliza el RMSE, los resultados obtenidos se comparan con los de otros dos modelos: el de persistencia y una red neuronal MLP con una capa oculta optimizada con un algoritmo genético. Se observa que con el modelo optimizado propuesto se obtienen mejores resultados.

- En [Monteiro 13b] se predice la potencia eléctrica media horaria para el horizonte que cubre todas las horas de luz del día siguiente. Se utilizan:
 - Observaciones horarias de la potencia de una planta solar de 36 kWp situada en La Rioja, España.
 - Predicciones de: radiación de onda corta hacia la superficie terrestre, radiación de onda larga hacia la superficie, temperatura en la superficie, flujo de calor sensible de la superficie, flujo de calor latente de la superficie, radiación de onda corta saliente, y radiación de onda larga saliente, pronosticadas todas con un modelo NWP de mesoescala.
 - *Seno* y *coseno* de la fracción transcurrida del año y del día.
 - Período de estudio: un año.

Metodología: se presentan y comparan dos modelos. El modelo de predicción de potencia fotovoltaica analítico (APVF) consistente en un modelado original para ajustar datos de radiación de cielo despejado con un índice de atenuación de irradiación, y combinados con un índice de atenuación de la producción de energía eléctrica. El otro modelo (MPVF) está basado en una red neuronal perceptrón multicapa con dos capas ocultas y aprendizaje por retropropagación de errores (BP). Este modelo fue seleccionado como el mejor entre un amplio conjunto de modelos posibles de diferentes familias (modelos de series temporales, modelos de redes neuronales y modelos neurodifusos).

Resultados: para calcular el rendimiento de los modelos de predicción se utilizan los índices de error NRMSE y NMAE. A pesar de que ambos modelos son muy diferentes, logran resultados muy similares.

- [Singh 13] se predice la potencia eléctrica media para el horizonte de 15 min. Se utilizan:
 - Datos de potencia cada 15 min. de una planta fotovoltaica de 101 kWp, situada en Rajasthan.

- Observaciones cada 15 min. de radiación solar, temperatura ambiente, velocidad del viento y temperatura del módulo.
- Período de estudio: año 2011.

Metodología: se utilizan los modelos: *Adaptive Neuro-Fuzzy Inference* (ANFIS), redes neuronales artificiales con el algoritmo de entrenamiento de retropropagación de errores y redes neuronales generalizadas (RNAG).

Resultados: con la red neuronal generalizada (RNAG) se obtiene menor error RMSE, que con los modelos basados en RNA y ANFIS.

- En [Yan 13] se predice la potencia eléctrica media de una instalación eólica y fotovoltaica cada 15 min. para los próximos 15 min., las próximas cuatro horas, el próximo día, y predicciones medias mensuales. Se utilizan:
 - Datos de potencia de un parque eólico con una potencia instalada de 98 MW y una planta fotovoltaica con 50 MWp situadas, muy próximas, en China.
 - Predicciones del viento y de la radiación solar, proporcionadas por un modelo de mesoescala.
 - Período de estudio: del 1 de marzo de 2011 hasta agosto de 2012.

Metodología: primero, el sistema propuesto hace predicciones numéricas meteorológicas (NWP) con alta resolución, mediante el aprovechamiento de los datos meteorológicos en tiempo real. En segundo lugar, se utiliza una combinación de diferentes modelos estadísticos para lograr predicciones a corto plazo y a muy corto plazo de la generación en aerogeneradores y paneles fotovoltaicos para obtener después las predicciones de todo el parque eólico e instalación fotovoltaica.

Resultados: para la evaluación del error de predicción se adopta una variante del NRMSE. El sistema propuesto predice a la vez la generación eléctrica del parque eólico y planta fotovoltaica con resultados satisfactorios.

- En [Yang 13] se predice la potencia eléctrica cada media hora para el horizonte de 24 horas. Se utilizan:
 - Datos históricos de potencia eléctrica media horaria obtenidos de un sistema fotovoltaico conectado a red cuya capacidad es de 2.2 MWp, en China.
 - Datos de temperatura, humedad, velocidad del viento y radiación solar medidos en una estación situada junto al sistema fotovoltaico.
 - Período de estudio: meses de septiembre y octubre.

Metodología: se utiliza el modelo neuronal de mapas autoorganizados (SOM) para identificar los tipos de datos meteorológicos recogidos por la estación y agrupar las muestras que tienen el mismo estado de la atmósfera. Para cada una de estas muestras se entrena una red neuronal, con el algoritmo de retropropagación de errores (BP) y con aplicación de la transformada wavelet (WT), con el fin de formar modelos de predicción adecuados para cada tipo de tiempo. Los datos de entrada

son los valores de la potencia generada en los cuatro días anteriores con el mismo tipo meteorológico.

Resultados: los criterios de evaluación de la predicción utilizados son el porcentaje de error absoluto medio diario (DMAPE) y la raíz del error cuadrático medio (RMSE). Los resultados experimentales muestran que el reconocimiento de agrupamiento basado en la red neuronal SOM es muy significativo, lo cual mejora la predicción en gran medida.

- En [Yazdanbaksh 13] se predice la potencia eléctrica media para el horizonte de 1 min. Se utilizan:
 - Datos de potencia que se obtienen utilizando el modelo propuesto en [Bellini 09].
 - Datos de irradiancia solar y temperatura medidas cada minuto proporcionadas por la *Lowry Range Solar Station* (LRSS).
 - Período de estudio: 7 días de observaciones.

Metodología: se comparan dos modelos ANFIS y de redes neuronales con funciones de base radial, con uno nuevo basado en lógica difusa compleja: *Adaptive Neuro Complex Fuzzy Inferential System* (ANCFIS).

Resultados: se comparan los tres modelos en términos del RMSE. Se obtiene mejor precisión en la predicción de la potencia de salida con el modelo ANCFIS propuesto.

En 2014:

- En [Al-Messabi 14] se predice la potencia eléctrica media cada 10 min. para un horizonte de 24 horas. Se utilizan:
 - Datos cada 10 min. recogidos de una planta fotovoltaica de 1.76 kWp instalada en Arabia Saudí.
 - Período de estudio: del 2 al 6 de febrero de 2010 y del 7 al 10 de febrero de 2011.

Metodología: se desarrolla una red denominada *focused time-delay neural network* (FTDNN). El problema de la optimización de la estructura de la FTDNN se reduce solo a optimizar el número de unidades de retardo y el número de neuronas en la capa oculta. Este problema se resuelve a través de dos algoritmos de optimización: genético y de enjambre de partículas (PSO).

Resultados: se comparan los modelos FTDNN, FTDNN optimizada por un AG, y optimizada por PSO. Con la PSO se alcanza un mejor RMSE.

- En [Li 14a] se predice la potencia eléctrica media horaria de cada hora hasta las próximas 24 horas. Se utilizan:
 - Datos de potencia generada en un sistema fotovoltaico.
 - Índice del tiempo, radiación solar, y velocidad del viento medidas en una estación meteorológica situada junto al sistema fotovoltaico.

- Período de estudio: dos años de datos (de junio a septiembre).

Metodología: a través de la combinación de lo que denominan análisis del mecanismo de producción de potencia en la planta fotovoltaica se propone un modelo basado en una red neuronal RBF optimizada. Se utilizan como entradas los datos del día anterior.

Resultados: para evaluar el modelo se utiliza el RMSE. Los resultados experimentales demuestran que el modelo propuesto reduce la desviación entre la potencia prevista y la potencia real de manera significativa, y se puede lograr la predicción rápida y precisa.

- En [Li 14b] se predice la potencia eléctrica media horaria para el horizonte a corto plazo. Se utilizan:
 - Datos históricos de la potencia media horaria generada en una planta fotovoltaica.
 - Período de estudio: del 1 al 10 de abril de 2013.

Metodología: se propone la mejora de un modelo de predicción gris GM(1,1) mediante el uso del método de suavizado exponencial (SE), mostrando éste un efecto de suavizado sobre la secuencia de fuerte volatilidad.

Resultados: con el fin de comparar de forma más explícita la precisión de la predicción de los dos modelos, se utiliza la diferencia de los datos reales y previstos. La evaluación de los resultados muestra que el modelo de predicción GM(1,1) mejorado tiene una mayor precisión en la predicción.

- En [Lo Brano 14] se predice la potencia eléctrica media horaria para el horizonte a corto plazo. Se utilizan:
 - Datos históricos cada 30 min. de:
 - La potencia de salida en dos módulos fotovoltaicos situados en Italia.
 - Temperatura de la célula solar, tensión a circuito abierto, corriente de circuito corto de dichos módulos.
 - Medidas en una estación meteorológica próxima de la temperatura del aire, irradiancia solar y velocidad del viento.
 - Período de estudio: base de datos con más de 6000 datos horarios.

Metodología: la predicción de la potencia de salida se obtiene utilizando tres tipos de redes neuronales: un perceptrón multicapa (MLP) con una capa oculta, una red neuronal recursiva (RNAR) y una red neuronal retardada en el tiempo con memoria gamma (RNAMG), todas ellas entrenadas con retropropagación de errores (BP).

Resultados: la evaluación se realiza mediante tres índices: el error medio (ME), el error absoluto medio (MAE), y la desviación estándar (SDE). Los resultados obtenidos confirman la idoneidad de la metodología adaptada para el problema de pronóstico de potencia a corto plazo. Las topologías RNA basadas en MLP fueron las mejores en términos de error de predicción.

- En [Mellit 14] se predice la potencia eléctrica media horaria para el horizonte de 1 hora. Se utilizan:
 - Datos de la potencia de salida en una instalación fotovoltaica de 1 MWp, instalada en Italia.
 - Predicciones de la irradiancia solar proporcionadas por *European Centre for Medium-Range Weather Forecasts* (ECMWF).
 - La temperatura en los paneles de la instalación.
 - Período de estudio: del 1 de enero al 31 de diciembre de 2011.

Metodología: se utilizan los datos previstos de irradiancia solar para clasificar el tipo de día (soleado, parcialmente nublado y nublado), y posteriormente, se construyen tres modelos distintos con redes neuronales artificiales unidireccionales entrenadas con el algoritmo BP para cada uno de esos tipos de día. También se construye un modelo global para el que se toma todos los datos sin clasificar.

Resultados: con el fin de evaluar el rendimiento de los modelos de RNA diseñados, se evalúan en términos de R, RMSE, ME y MAPE. Se puede observar que los modelos RNA-1 (usado para días soleados) y RNA-3 (usado para días nublados) son más precisos que el modelo RNA-global, mientras que el modelo RNA-2 (usado para días parcialmente nublados) y el global dan resultados bastante similares.

- En [Murakami 14] se predice la potencia eléctrica media para el horizonte de: 10–30 min., varias horas y para el próximo día. Se utilizan:
 - Datos históricos de potencia de un sistema fotovoltaico de 3 kWp situado en Tokio.
 - El índice de claridad y la irradiancia.
 - Período de estudio: del 19 de enero al 1 de diciembre de 2012.

Metodología: se investigan métodos de predicción de salida de potencia eléctrica media para varios horizontes de tiempo. Se aplica el método del promedio incluyendo el cálculo teórico de la irradiancia solar, para la predicción de 10–30 min., y la técnica de simulación de Monte-Carlo (MC) para la predicción a medio plazo (de varias horas hasta el día siguiente).

Resultados: se adopta la técnica de simulación de Monte-Carlo para la predicción a medio y largo plazo. También se propone un modelo de simulación que demuestra que la agregación de predicciones para varias instalaciones fotovoltaicas reduce el error en la predicción de la potencia global generada por el conjunto. Se encuentra que el método propuesto puede predecir la potencia en 30 min. con un error promedio menor del 5 %.

- En [Wang 14] se predice la potencia eléctrica media cada 15 min. con un horizonte de 24 horas. Se utilizan:
 - Datos de potencia de muchas plantas fotovoltaicas de la parte del noroeste de China.

- Datos históricos y en tiempo real proporcionados por una estación meteorológica.
- Datos históricos y en tiempo real de la intensidad de la radiación solar en un plano exo-atmosférico.
- Calibración de la transparencia atmosférica eficiente.
- Período de estudio: un año.

Metodología: se examinan dos modelos de predicción de la radiación. En el modelo de cielo despejado se predice la radiación en la superficie mediante, el cálculo del plano de corte de la intensidad instantánea de la radiación solar y el coeficiente de transparencia atmosférica. En el modelo de cielo no despejado, la radiación futura es prevista mediante la búsqueda de la curva cuadrática de la radiación medida en la superficie y el tiempo. Finalmente, se desarrolla un modelo de conversión radiación-potencia para calcular la predicción de la potencia correspondiente a través de la salida del modelo de predicción de radiación a corto plazo.

Resultados: se compara el modelo propuesto con el modelo de predicción de potencia tradicional basado en NWP. Para la evaluación del error se utiliza el RMSE, el MAE, el coeficiente de correlación (CC), y el porcentaje de paso (PP). Se observa que el modelo propuesto basado en patrones meteorológicos tiene predicciones más precisas.

- En [Wu 14] se predice la potencia eléctrica media horaria para el horizonte de 1 hora. Se utilizan:
 - Bases de datos de tres sistemas fotovoltaicos:
 - Valores promedio cada 15 min. de la potencia de una planta fotovoltaica de 45.36 kWp en Malaysia, junto con la temperatura atmosférica, la irradiancia solar, y la temperatura en el módulo, con un total de 9009 registros.
 - Valores promedio cada 60 min. de la potencia de una planta fotovoltaica de 72 kWp. en Taiwán, junto con la temperatura atmosférica, la irradiancia solar, y la temperatura en el módulo, con un total de 6830 registros.
 - Valores promedio cada 10 min. de la potencia de una planta fotovoltaica de 70 kWp. en un edificio de Taiwan, junto con la temperatura atmosférica, la irradiancia solar, y la temperatura en el modulo, con un total de 17280 registros.
 - Predicciones meteorológicas horarias con el tipo de día previsto (soleado, nublado, lluvioso) y el promedio de la cobertura de nubes, proporcionadas por el modelo WRF del *Central Weather Bureau* (CWB) de Taiwan.

Metodología: el proceso de predicción se divide en dos etapas. En la primera, la potencia se predice individualmente por cuatro modelos: ARIMA, SVM, RNA y ANFIS. A continuación, los resultados de predicción se convierten en las entradas del modelo de la segunda etapa. En ésta, se combinan los modelos usando un AG,

que asigna el coeficiente de peso para cada modelo individual, y se forma el modelo híbrido final.

Resultados: para evaluar el rendimiento del modelo de predicción propuesto se utiliza la raíz del error cuadrático medio normalizado (NRMSE). Los resultados demuestran que el modelo de predicción híbrido genera las predicciones más precisas en la mayoría de los casos.

- En [Yan 14] se predice la potencia eléctrica media horaria, para el horizonte de 1 hora. Se utilizan:
 - Datos de potencia de una instalación fotovoltaica en China.
 - Datos de temperatura máxima y mínima diaria, de velocidad del viento media diaria, de presión media del aire y el tipo de día (soleado, nublado, cubierto y otro), proporcionados por una estación meteorológica.
 - Predicciones proporcionadas por el departamento meteorológico.
 - Período de estudio: del 1 de abril del 2014 al 5 de mayo 2014.

Metodología: el modelo se basa en el algoritmo *Fuzzy C-means* (FCM), usa datos históricos meteorológicos para clasificar 6 tipos de días (soleado, nublado, cubierto rara vez, medio cubierto, muy cubierto, lluvia). Luego, se utilizan los datos históricos que tienen similares condiciones meteorológicas con la predicción del día y se construye un modelo de predicción con una red neuronal RBF para cada uno de los seis tipos de días.

Resultados: se escoge el MSE para evaluar el error. Los resultados muestran una buena exactitud predictiva del modelo.

- En [Yang 14] se predice la potencia eléctrica media para el horizonte de 5 min. Se utilizan:
 - Datos de potencia eléctrica media cada minuto de una instalación fotovoltaica, en China.
 - Datos cada min. de irradiancia directa normal, velocidad del viento, temperatura, humedad y presión del aire.
 - Período de estudio: 4 de mayo 2014.

Metodología: se utiliza una red neuronal entrenada mediante el algoritmo de retropropagación de errores (BP) como modelo básico de predicción y se aplica un algoritmo genético para optimizar los pesos. La red neuronal tiene tres capas, las entradas son: valores de irradiancia directa normal del minuto actual y hasta cuatro minutos antes, valor de la potencia actual y hasta cuatro minutos antes, y el valor de la presión media del aire de los últimos cinco minutos.

Resultados: los resultados muestran que la predicción basada en una red neuronal, combinada con el algoritmo genético para optimizar pesos, tiene mayor exactitud que la red neuronal con los pesos generados aleatoriamente.

- En [Zamo 14] se predice la potencia eléctrica media horaria para el horizonte entre 28 y 45 horas. Se utilizan:
 - Observaciones cada 10 min. de la potencia de 28 plantas fotovoltaicas en Francia.
 - Predicciones de irradiación solar, nubosidad, temperatura del aire, velocidad y dirección de viento y presión, proporcionadas por el modelo ARPEGE.
 - Período de estudio: años 2009, 2010 y 2011.

Metodología: se comparan varios modelos, como los basados en persistencia, modelo lineal (ML), árbol de regresión binario (ARB), *bagging* (BA), *boosting* (BO), *random forests* (RF), máquinas de vectores soporte (SVM), y modelo aditivo generalizado (MAG).

Resultados: para elegir el mejor modelo se utilizan el RMSE y el MAE. El RMSE y MAE de todos modelos son por lo menos un tercio menores que los del modelo de persistencia. El modelo basado en *random forests* es más exacto, su RMSE o MAE son casi la mitad que el de persistencia.

- En [Zhong 14] se predice la potencia eléctrica media horaria para el horizonte diario. Se utilizan:
 - Datos históricos de potencia, obtenidos de una planta en China.
 - Predicciones de radiación solar, temperatura, y tipo de día, proporcionadas por un servicio meteorológico.
 - Período de estudio: del 13 al 30 de septiembre del 2013.

Metodología: se diseña un modelo con un algoritmo de optimización basado en la combinación de *Particle Swarm Optimization* (PSO) y una red neuronal MLP. El sistema de predicción se divide en cuatro sub-modelos de acuerdo con el tipo de estación: primavera, verano, otoño e invierno. Además, cada sub-modelo se divide en tres sub-redes de acuerdo con los tipos de día: soleado, nublado, lluvioso.

Resultados: el número de neuronas en las capas ocultas (rango de 3 a 12) se selecciona con el error cuadrático medio (MSE), con el fin de conseguir la red óptima. Los errores relativos de mayor tamaño se observan principalmente durante los días de lluvia. Los resultados experimentales muestran que el modelo de optimización propuesto mejora la precisión de la predicción.

En 2015:

- En [Alessandrini 15] se predice la potencia eléctrica media horaria para el horizonte de 0 a 72 horas. Se utilizan:
 - Datos de la potencia de salida de tres parques fotovoltaicos con 15.42 kWp en total, ubicados en Italia.

- Predicciones de la irradiancia horizontal global, cobertura de nubes, y temperatura del aire a 2 m, obtenidas con el modelo de mesoescala RAMS.
- Ángulo de la elevación del Sol.
- Período de estudio: entre año y medio y dos años, dependiendo del parque.

Metodología: se propone la aplicación de un método de conjunto análogo (AnEn) para generar predicciones probabilísticas de potencia solar (SPF). El AnEn se basa en un conjunto histórico de predicciones meteorológicas proporcionadas por modelos de predicción numérica del tiempo y observaciones de la potencia generada. Los pronósticos se eligen basándose en su similitud con los pronósticos anteriores.

Resultados: el rendimiento del AnEn para SPF se compara con una técnica de regresión cuantil (QR) y un conjunto de persistencia (PeEn). El QR es un método para predicciones probabilísticas que, similarmente al AnEn, está basado en un conjunto de datos históricos. El PeEn es un modelo de persistencia para predicciones probabilísticas, donde se utilizan las 20 medidas de potencia disponibles más recientes del mismo tipo de tiempo para formar un conjunto. La evaluación de los resultados se lleva a cabo evaluando los indicadores importantes de un sistema probabilístico tales como la consistencia estadística, la fiabilidad, la resolución y la habilidad. El AnEn funciona tan bien como el QR para eventos comunes, proporcionando predicciones con fiabilidad similar, resolución y nitidez, mientras que exhibe más habilidad para eventos raros y durante horas con una elevación solar baja.

- En [Caro 15] se predice la potencia eléctrica media horaria para el horizonte de un día. Se utilizan:
 - Datos históricos de la generación de la potencia horaria de 33 plantas fotovoltaicas ubicadas en España.
 - Datos históricos de la predicción de la radiación horaria proporcionadas por un servicio de predicción meteorológica.
 - Período de estudio: 5 meses.

Metodología: se propone un procedimiento basado en el modelo de suavizado exponencial multivariante y en el empleo de la predicción de la radiación solar para el próximo día. El modelo se estima utilizando el algoritmo *Expectation-Maximization* (EM), es un procedimiento iterativo para obtener la estimación de máxima verosimilitud de los parámetros desconocidos de un modelo dado.

Resultados: la comparación métrica que se utiliza es el error relativo (ER), se calcula considerando la potencia eléctrica diaria producida. Los resultados indican que el algoritmo desarrollado es numéricamente preciso y computacionalmente eficiente.

- En [Chu 15] se predice la potencia eléctrica media cada 5 min. para el horizonte de 5, 10 y 15 min. Se utilizan:
 - Datos de la potencia de salida de una estación fotovoltaica de 48 MWp en Nevada.

- Período de estudio: del 1 de noviembre al 5 de diciembre de 2011.

Metodología: se emplea una red neuronal artificial optimizada con un algoritmo genético para mejorar la potencia de salida de tres modelos de predicción: uno determinista basado en técnicas de seguimiento de la nube, ARMA y k vecinos más próximos.

Resultados: el rendimiento de todas las predicciones se evalúan en términos de ME, MAE y RMSE, y se comparan con los obtenidos con el modelo de persistencia de referencia. Se demuestra que el modelo basado en red neuronal optimizada obtiene los mejores resultados.

- En [De Felice 15] se predice la potencia eléctrica media diaria con un horizonte de hasta diez días. Se utilizan:
 - Datos de la potencia de salida de 65 plantas fotovoltaicas con un total de 425 MWp, en Italia.
 - Medidas locales de la radiación solar y de la temperatura del aire.
 - Predicciones de la radiación solar y de la temperatura del aire, proporcionadas por el modelo global IFS.
 - Período de estudio: entre 18 y 24 meses de producción de potencia diaria.

Metodología: se calibran modelos con máquinas de vector soporte (SVM) utilizando observaciones disponibles y luego se aplican los mismos modelos con las variables de predicciones meteorológicas.

Resultados: la variabilidad de la cobertura de nubes afecta en gran medida la producción de la potencia solar. Se observa que durante el verano el error de las predicciones es más bajo y durante el invierno es mayor. Como medida de error se selecciona el RMSE y el porcentaje de error absoluto medio (MAPE).

- En [Filipe 15] se calculan predicciones probabilísticas de la potencia eléctrica media horaria para el horizonte de 48 horas. Se utilizan:
 - Datos de potencia fotovoltaica de una instalación fotovoltaica con 16 kWp, localizada en el norte de Portugal.
 - Temperatura ambiente, irradiancia global sobre el panel, y temperatura de la celda.
 - Predicciones de la velocidad del viento y de la irradiancia horizontal y directa, proporcionadas por el modelo de mesoescala WRF.
 - El ángulo del panel, y el albedo del terreno en torno a la planta fotovoltaica.
 - Período de estudio: de mayo de 2013 hasta noviembre 2014.

Metodología: las predicciones de la potencia se calculan combinando linealmente las predicciones generadas por un modelo autorregresivo (AR) y por dos tipos diferentes de modelos: uno estadístico (*gradient boosting technique*) [Bühlmann 06] y

otro físico [Sumaili 14], que convierten la irradiancia solar en potencia. Finalmente, un modelo de regresión produce un conjunto de predicciones probabilísticas usando como entrada el punto de predicción generada.

Resultados: los pronósticos son evaluados por el MAE y el RMSE normalizado por la potencia nominal. Las predicciones probabilísticas se evalúan utilizando el *Continuous Ranked Probability Score* (CRPS), descrito en [Anastasiades 13].

La combinación de los modelos físico y estadístico ha demostrado un mejor rendimiento, teniendo en cuenta que resulta un RMSE inferior en comparación con los modelos individuales que se ejecutan por separado. Además, la herramienta de pronóstico híbrido propuesta muestra una mejora con respecto al modelo de persistencia y al modelo diurno (utiliza medidas de potencia de los dos últimos días).

- En [Golestaneh 15] se generan predicciones probabilísticas para la generación de la potencia eléctrica media de dos plantas fotovoltaicas para el horizonte entre unos pocos minutos hasta 1 hora. Se utilizan:
 - Datos de potencia cada min. recolectados de dos plantas fotovoltaicas de 433 kWp y 1 kWp, situadas en Singapur.
 - Período de estudio: años 2012 y 2013.

Metodología: se propone un método basado en *Extreme Learning Machine* (ELM) como un modelo de regresión rápido, entrenado de varias formas para obtener las predicciones puntuales y cuartiles de la generación de la potencia eléctrica media horaria.

Resultados: se proporciona la precisión de los pronósticos puntuales separados por estaciones del año, se usa el RMSE y el MAE. La estación de verano presenta la mejor precisión en el pronóstico de potencia.

- En [Li 15] se predice la potencia eléctrica media cada 15 min. para los horizontes de: una semana, un mes, una estación, y un año. Se utilizan:
 - Datos de potencia cada 15 min. de un sistema fotovoltaico de 5.6 kWp.
 - Temperatura del panel solar, temperatura exterior, temperatura interior, humedad y radiación solar.
 - Período de estudio: año 2010.

Metodología: se utiliza la teoría de conjuntos aproximados (CAP) [Pawlak 95]. Los datos redundantes se eliminan de la tabla de decisión con el algoritmo de reducción de Pawlak.

Resultados: las investigaciones demuestran que cuanto mayor es el número de datos de condiciones meteorológicas y de potencia, más precisos son los resultados de predicción de la generación de la potencia basada en los conjuntos aproximados. El cálculo del modelo de predicción por conjuntos aproximados es más sencillo y el promedio del error relativo (ER) de predicción es inferior al que se obtiene con el modelo gris (GM) y el modelo basado en cadenas de Markov.

- En [Rashkovska 15] se predice la potencia eléctrica media horaria para el horizonte de 6 horas. Se utilizan:
 - Datos de potencia cada 2 min. recolectados de una planta fotovoltaica situada en Eslovenia.
 - Medidas cada 2 min. de temperatura del aire, humedad relativa, velocidad y dirección del viento, precipitación, intensidad de la radiación solar en las partes visibles y ultravioleta del espectro, proporcionadas por una estación meteorológica situada en el mismo lugar.
 - Predicciones meteorológicas de radiación solar, temperatura del aire, cobertura de nubes, proporcionadas por la agencia de medio ambiente de Eslovenia utilizando el modelo regional ALADIN.
 - Elevación del sol y el azimut en cada hora.
 - Período de estudio: del 1 de mayo hasta el 14 de noviembre de 2013.

Metodología: se utiliza una red neuronal unidireccional con dos capas ocultas, que se entrena con el algoritmo de retropropagación de Levenberg-Marquardt. Se investigan 4 modelos:

- El modelo 1: tiene como entradas los datos horarios de potencia eléctrica generada y medidas locales de radiación solar, temperatura del aire, de cada una de las cinco horas anteriores.
- El modelo 2: utiliza como entradas las del modelo 1 y la elevación del sol y el azimut.
- El modelo 3: además de las entradas del modelo 2, incluyen predicciones de radiación solar, temperatura del aire y cobertura de nubes (de cada una de las 5 horas anteriores).
- El modelo 4: usa las entradas del modelo 2, además incluye predicciones de radiación solar, temperatura del aire y cobertura de nubes.

Resultados: para determinar el modelo con mejor resultado se utilizan el error cuadrático medio (MSE), el coeficiente de correlación lineal (R), el error absoluto medio (MAE), y el porcentaje de error absoluto medio normalizado (NMAPE). Se puede concluir que para horizontes de predicción más largos, la inclusión de información del ángulo de elevación del sol disminuye significativamente el error global.

- En [Yang 15] se predice la potencia eléctrica media para los horizontes de 1 min. hasta 2 horas. Se utilizan:
 - Datos de potencia de plantas fotovoltaicas localizadas en California y en Colorado.
 - Datos históricos de irradiancia solar de 13 ubicaciones en California, con 1 hora de resolución.
 - Datos históricos de irradiancia solar de cinco lugares cercanos en Colorado, con 1 min. y 1 hora de resolución.

- Período de estudio: año 2010.

Metodología: se propone un modelo de pronóstico autorregresivo con entradas exógenas (ARX), que se puede ampliar con información meteorológica adicional.

Resultados: con el fin de comparar la calidad de las predicciones del modelo, se utilizan dos índices de rendimiento MAE y RMSE. El modelo muestra tener un mejor rendimiento en comparación con el modelo de persistencia convencional.

En 2016:

- En [De Giorgi 16] se predice la potencia eléctrica media horaria para diversos horizontes hasta 24 horas. Se utilizan:
 - Datos de potencia horaria recolectados de un sistema fotovoltaico con una potencia nominal total de 960 kWp, instalado en Italia.
 - Medidas horarias de la temperatura del módulo, la temperatura ambiente, la irradiancia en el plano inclinado y su ángulo de inclinación.
 - Período de estudio: del 5 de marzo de 2012 hasta el 31 de diciembre de 2013.

Metodología: se describen tres modelos de predicción: el LS-SVM (*Least Square Support Vector Machine*), una red neuronal conocida como GMDH (*Group Method of Data Handling*), y un nuevo algoritmo híbrido GLSSVM (*Group Least Square Support Vector Machine*) que se basa en la combinación de los dos primeros modelos. Cada modelo se combina con tres estrategias (Directa, Recursiva y Directa-Recursiva) para cada horizonte de predicción.

Resultados: con el fin de evaluar el rendimiento de los diferentes métodos de predicción se utiliza el error absoluto medio normalizado (NMAE), el error medio normalizado (NME) y la raíz del error cuadrático medio (NRMSE). La salida demuestra que el método GLSSVM con la estrategia Directa-Recursiva permite obtener un error normalizado menor.

- En [Do 16] se predice la potencia eléctrica media horaria para el horizonte de 1 hora. Se utilizan:
 - Datos de potencia cada 5 min. de una planta fotovoltaica de 113 kWp situada en el Caribe.
 - Datos de potencia cada 10 min. de una planta fotovoltaica de 93 kWp situada en Francia.
 - Medidas cada hora de la temperatura del aire y de la cobertura de nubes, en la estación meteorológica de un aeropuerto cercano.
 - Período de estudio: del 1 de enero de 2010 hasta el 31 de diciembre de 2011.

Metodología: se utilizan tres modelos: el modelo de persistencia escalado, un modelo polinomial multivariante y una red neuronal artificial unidireccional, con una capa oculta, que se entrena con la técnica de aprendizaje supervisado usando el algoritmo

Levenberg-Marquardt y cuyas variables de entrada son la cobertura de nubes, la temperatura ambiente en la hora actual y la potencia media producida los últimos 60 min.

Resultados: se utilizan los indicadores NME, NMAE y NRMSE para comparar la precisión de las predicciones. El modelo polinomial multivariante (POM) tiene una mejor calidad en las predicciones, que la red neuronal artificial, siendo los dos más precisos que el modelo de persistencia.

- En [Larson 16] se predice la irradiancia global horizontal y la potencia eléctrica media horaria para el horizonte del próximo día. Se utilizan:
 - Datos de potencia cada 15 min. de dos plantas fotovoltaicas con 1 MWp cada una, ubicadas en California.
 - Datos cada 15 min. de la irradiancia global horizontal, disponibles de piranómetros co-localizados.
 - Período de estudio: cuatro años, del 2011 hasta 2014.

Metodología: se presenta un método que se basa en la optimización por mínimos cuadrados (MINC) de la predicción numérica meteorológica (NWP). Se supone que existe una función lineal $f(x; w) = \sum_{j=0}^n w_j x_j = w^T x$, que puede asignar la irradiancia global horizontal (GHI) a la salida de potencia, donde w es el vector peso, x es el vector de entrada, y n es el número de parámetros del modelo (en este estudio $n = 2$ que son la GHI prevista por uno de los modelos NWP y el ángulo zenital). A continuación, se usan los mínimos cuadrados para encontrar los pesos óptimos w^* , es decir, los pesos que minimizan el error entre la $f(x; w)$ y el valor real de la potencia.

Resultados: se utilizan métricas de error estándar para evaluar el rendimiento de las predicciones de la irradiancia global horizontal (GHI) y de la potencia de salida: el error medio absoluto (MAE), el error medio (ME), la raíz del error cuadrático medio (RMSE), y el RMSE normalizado (NRMSE).

La metodología propuesta reduce el error en los pronósticos de potencia. Si bien la mejora de la precisión de la predicción de la irradiancia es importante para muchas aplicaciones, los resultados sugieren que, para predicciones de potencia para el próximo día, los métodos estadísticos pueden compensar errores sistemáticos en las predicciones de irradiancia y son adecuados para modelar la complejidad de las operaciones de la planta de potencia.

- En [Li 16] se predice la potencia eléctrica media para los horizontes de 15 min., 1 y 24 horas. Se utilizan:
 - Datos de potencia cada 15 min. de una planta fotovoltaica de 6 MWp, localizada en Florida.
 - Datos de temperatura ambiente, velocidad y dirección del viento proporcionadas por los servicios meteorológicos.

- Variables adicionales: el ángulo de elevación del Sol, su ángulo zenital, y el coseno del ángulo de incidencia en el panel.
- Período de estudio: año 2014.

Metodología: se evalúan y comparan dos métodos comunes: vectores soporte para regresión y una red neuronal artificial unidireccional, con 1 sola capa oculta y 26 entradas (varios valores retardados en el tiempo de producciones históricas de potencia, entradas meteorológicas e información del tiempo), entrenada con el algoritmo de retropropagación de errores o BP (*back-propagation*).

Se propone un enfoque jerárquico basado en los algoritmos de aprendizaje automático. En la aproximación del pronóstico jerárquico, se utiliza las herramientas de aprendizaje máquina a un nivel micro para todas predicciones y se evalúan las representaciones a nivel macro para la planta entera.

Resultados: la precisión del modelo se determina utilizando el cálculo de los estadísticos de error como el error medio (ME), el error absoluto medio (MAE), la raíz cuadrada del error cuadrático medio (RMSE), el error relativo medio (ERM), el porcentaje de error absoluto medio (MAPE) y el RMSE relativo. A partir del análisis del error entre los valores reales y previstos de la producción de la potencia, se concluye que la técnica jerárquica supera a los modelos tradicionales.

- En [Rana 16] se predice la potencia eléctrica media para horizontes de 5, 10, 15, ... hasta 60 min., (es decir 12 horizontes). Se utilizan:
 - Datos históricos de potencia recogidos de un sistema fotovoltaico conectado a red, capaz de generar hasta 1.22 MW, instalado en Australia.
 - Medidas históricas de datos meteorológicos: irradiancia solar, temperatura, humedad y velocidad del viento, en el mismo lugar de la planta fotovoltaica.
 - Período de estudio: años 2013 y 2014.

Metodología: se combina un método de selección de variables usando el algoritmo *Correlation-based Feature Selection* (CFS), con redes neuronales y vectores soporte. Se construye un modelo de predicción independiente para cada uno de los 12 horizontes de predicción. Se desarrollan dos tipos de modelos: los modelos univariantes, que utilizan sólo los datos de potencia fotovoltaica anteriores y los modelos multivariantes, que también utilizan los datos meteorológicos anteriores.

Resultados: la exactitud de los modelos de predicción se evalúan usando dos medidas de rendimiento: el error absoluto medio (MAE) y el error relativo medio (ERM). Los resultados de los modelos univariantes son similares a los modelos multivariantes. Por lo tanto, la salida de la potencia para horizontes de predicción a muy corto plazo entre 5 y 60 min. se puede predecir con precisión utilizando sólo los datos de potencia fotovoltaica anteriores, sin información meteorológica. El modelo más preciso fue el conjunto univariante de redes neuronales, prediciendo la salida de la potencia por separado para cada paso del horizonte de predicción.

- En [Soubdhan 16] se predice la radiación solar para los horizontes entre 1 min. y 1 hora, y la potencia eléctrica media para el horizonte de 1 hora. Se utilizan:

- Datos de potencia recolectados cada 5 min. de un sistema fotovoltaico de 32.3 kWp, instalado en Guadeloupe.
- Datos solares medidos en una estación meteorológica cercana con un intervalo de 1 s.
- Medidas horarias de temperatura ambiente y cobertura de nubes, medidos en otra estación también cercana.
- Período de estudio: años 2012 y 2013.

Metodología: se desarrolla un método de predicción basado en el filtro de Kalman (FK), combinado con un procedimiento de estimación de parámetros construido con un modelo autorregresivo (AR), o con un algoritmo Expectación-Maximización (EM). El modelo está diseñado para funcionar con datos univariantes o multivariantes, de acuerdo a su disponibilidad.

Resultados: la precisión de esta técnica se estudia con un conjunto de criterios de rendimiento como la raíz del error cuadrático medio (RMSE), el error medio (ME) y el error absoluto medio (MAE) en sus formas normalizadas.

Se comparan los resultados de las diferentes pruebas realizadas para el horizonte de predicción desde un minuto hasta una hora, con el modelo de persistencia simple. El rendimiento de esta técnica es superior con mucho al modelo de persistencia tradicional.

- En [Sperati 16] se produce predicciones probabilísticas de potencia eléctrica media cada 3 horas, para el horizonte de 0 hasta 72 horas. Se utilizan:
 - Datos de potencia horaria de tres plantas fotovoltaicas con un total de 15.41 kWp, ubicadas en Italia.
 - Medidas horarias de la irradiancia global horizontal y de la temperatura a 2 m. del suelo, disponibles para dos de las plantas.
 - Período de estudio: entre 18 y 26 meses, dependiendo de la planta.

Metodología: se evalúan las predicciones del modelo de predicción numérica meteorológica EPS del *European Centre for Medium-Range Weather Forecasts* (ECMWF). El EPS se basa en la ejecución del modelo meteorológico varias veces, comenzando de condiciones iniciales ligeramente perturbadas, lo que permite obtener un conjunto de predicciones o ensemble. La distribución de los valores obtenidos para las diferentes ejecuciones permite la estimación de la incertidumbre de la predicción. Se aplica una red neuronal unidireccional, con una sola capa oculta, que utiliza como variables de entrada la irradiancia global horizontal, la temperatura a 2 m. del suelo, la cobertura de nubes total, el ángulo *azimut* y el ángulo de elevación del Sol. Estas variables se usan para generar una función de densidad de probabilidad (PDF) de potencia. Luego, se aplican dos métodos estadísticos: *the ensemble variance deficit* (VD) y *the Ensemble Model Output Statistics* (EMOS), para calibrar el conjunto de predicciones de potencia, lo cual influye en la predicción de potencia probabilística final.

Resultados: se evalúan la predicción probabilística con la desviación estándar de los miembros del conjunto o ensemble y con el RMSE normalizado. El valor global de una predicción probabilística [Jolliffe 03] se define por el *Brier Skill Score* (BSS) y el *Relative Operating Characteristic Skill Score* (ROCSS). La evaluación adicional se realiza evaluando el *Continuous Ranked Probability Score* (CRPS) de los métodos de pronóstico [Carney 06]. Se representa el error en tanto por ciento perdido (EP) en diagramas.

Se muestra que el EPS acoplado con VD y EMOS es un método válido para producir predicciones de potencia solar probabilísticas a corto plazo, en diferentes condiciones climáticas. Los métodos VD y EMOS muestran un nivel similar de consistencia estadística y habilidad de pronóstico para cada caso de prueba. En general, el rendimiento de VD y EMOS aparece significativamente mejor que el conjunto de persistencia.

- En [Vaz 16] se predice la potencia eléctrica media para los horizontes de 15 min. hasta un mes. Se utilizan:
 - Datos de potencia cada min. de cinco sistemas fotovoltaicos con un total de 4320 Wp, instalados en Holanda.
 - Datos horarios de la radiación solar y de la temperatura ambiente.
 - Período de estudio: del 1 de agosto de 2012 hasta el 31 de julio de 2013.

Metodología: se implementa un sistema de arquitectura de red neuronal configurado como un modelo autorregresivo no lineal con entradas exógenas (NARX). Se utilizan como entradas datos meteorológicos locales (la radiación solar y temperatura ambiente) y las mediciones de cinco sistemas fotovoltaicos próximos. La red se entrena con el algoritmo Levenberg-Marquardt y se aplican 4 casos con diferentes entradas, para seleccionar la mejor arquitectura de la red neuronal.

Resultados: para cuantificar la calidad de las predicciones se utiliza la raíz del error cuadrático medio normalizado (NRMSE). El modelo red neuronal con estructura NARX supera claramente al modelo de persistencia para horizontes de predicción de más de 15 min.

En la Tabla 3.2 se resumen brevemente las características más relevantes de cada uno de los modelos de predicción de la potencia eléctrica analizados. Para cada modelo se indican las técnicas que se utilizan, su resolución, su horizonte de predicción, y los criterios que se siguen para su evaluación.

Tabla 3.2: *Resumen de características de los modelos de predicción de la potencia eléctrica.*

Referencia	Técnica	Resol.	Horizonte	Evaluación
[Bacher 09]	AR, ARX	1 h	36 h	RMSE
[Li 09]	GM-MARKOV	15 min	1 día	DR
[Huang 10]	RNA, FÍSI	15 min	1 día	NRMSE

(Sigue en la página siguiente.)

Continuación de la tabla

Referencia	Técnica	Resol.	Horizonte	Evaluación
[Wang 11]	GM-RNA	1 h	12 h	ERM
[Fernández 12a]	RNA-AG	1 h	1 a 39 h	MAE, ME, RMSE
[Fernández 12b]	ANFIS-AG	1 h	48 h	NRMSE
[Fonseca 12]	ν -SVR	1 h	1 h	MAE, MAPE, RMSE
[Lin 12]	GS	1 h	7 días	—
[Mandal 12]	WT-RBF	1 h	12 h	MAE, MAPE, RMSE
[Pedro 12]	ARIMA, KNN, RNA, RNA-AG	1 h	1 y 2 h	RMSE, NRMSE, MAE, ME, R^2
[Shi 12]	SVM	15 min	24 h	ERM, RMSE
[Bouzerdoun 13]	SARIMA-SVM	1 h	días	NRMSE, MPE, NME, R
[Bracale 13]	AR, MC	1 h	1 h	MARE
[Haque 13]	WT-FF-FA	1 h	12 h	MAPE, VAR, NRMSE
[Kardakos 13]	SARIMA, RNA	1 h	1 día	NRMSE
[Li 13]	GM	15 min	Corto	PEP
[Mao 13]	EEMD-SVM	1 h	24 h	RMSE, MAPE, TIC
[Monteiro 13a]	MDM-AG	1 h	24 h	RMSE
[Monteiro 13b]	APVF, MPVF	1 h	1 día	NRMSE, NMAE
[Singh 13]	ANFIS, RNA, RNAG	15 min	15 min	RMSE
[Yan 13]	EST	15 min	15 min, 4 h 1 d, 1 mes	NRMSE
[Yang 13]	SOM, WT-RNA	30 min	24 h	DMAPE, RMSE
[Yazdanbaksh 13]	ANCFIS	1 min	1 min	RMSE
[Al-Messabi 14]	FTDNN-PSO	10 min	24, 48 h	RMSE
[Li 14a]	RBF	1 h	24 h	RMSE
[Li 14b]	GM-SE	1 h	24 h	ME
[Lo Brano 14]	MLP, RNAR, RNAMG	1 h	Corto	ME, MAE, SDE
[Mellit 14]	RNA	1 h	1 h	R, RMSE, ME, MAPE
[Murakami 14]	MC	30 min	10–30 min 1 día	ME
[Wang 14]	RAD-POT	15 min	24 h	CC, MAE, RMSE, PP
[Wu 14]	(ARIMA, SVM, RNA, ANFIS)-AG	1 h	1 h	NRMSE

(Sigue en la página siguiente.)

Continuación de la tabla

Referencia	Técnica	Resol.	Horizonte	Evaluación
[Yan 14]	FCM, RBF	1 h	1 h	MSE
[Yang 14]	RNA-AG	1 min	5 min	MSE
[Zamo 14]	ML, BA, BO, RF, ARB, SVM, MAG	1 h	28–45 h	MAE, RMSE
[Zhong 14]	RNA-PSO	1 h	1 día	MSE
[Alessandrini 15]	AnEn	1 h	0 a 72 h	NRMSE, MAE
[Caro 15]	EM	1 h	1 día	ER
[Chu 15]	RNA-AG	5 min	5, 10, 15 min	ME, MAE, RMSE
[De Felice 15]	SVM	1 día	10 días	RMSE, MAPE
[Filipe 15]	AR-ELEC-EST	1 h	48 h	MAE, NRMSE, CRPS
[Golestaneh 15]	ELM	1 min	10 min, 1 h	MAE, RMSE
[Li 15]	CAP	15 min	Largo	ER
[Rashkovska 15]	RNA	1 h	6 h	MAE, NMAPE, MSE, R
[Yang 15]	ARX	1 min, 1 h	1 min, 2 h	MAE, RMSE
[De Giorgi 16]	GLSSVM	1 h	1–24 h	NMAE, NME, NRMSE
[Do 16]	POM, RNA	1 h	1 h	NMAE, NME, NRMSE
[Larson 16]	MINC	1 h	1 d	MAE, ME, RMSE, NRMSE
[Li 16]	RNA, SVR	15 min, 1 h	15 min, 1 h, 24 h	ME, MAE, ERM, RMSE, MAPE
[Rana 16]	(RNA, SVR)-CFS	5 min	5, 10, ..., 60 min	ERM, MAE
[Soubdhan 16]	(AR, EM)-FK	1 h	1 h	NMAE, NME, NRMSE
[Sperati 16]	EPS, RNA, VD, EMOS	3 h	72 h	RMSE, BSS, EP, ROCSS, CRPS
[Vaz 16]	RNA-NARX	15 min	15 min-1 mes	NRMSE

3.3. Conclusiones

Las conclusiones más relevantes que se pueden extraer después de haber analizado la literatura internacional son:

- En los últimos años, el desarrollo de modelos de predicción de la potencia eléctrica generada en una planta fotovoltaica ha sido un campo de investigación muy activo.

Estos modelos, cada vez más precisos, han permitido mejorar la gestión de esas plantas.

- El modelo de referencia más utilizado, para efectuar una comparación con los nuevos modelos de predicción, es el denominado persistente.
- Hay disparidad de criterios sobre el indicativo de error utilizado en la literatura internacional, la mayoría de los autores se decantan por el RMSE.
- Para horizontes de predicción cortos se utilizan, fundamentalmente, medidas en una planta fotovoltaica (potencia y variables meteorológicas: temperatura, irradiancia y velocidad del viento); para horizontes mayores (24 h o más) se suelen utilizar las predicciones de variables meteorológicas, además de los datos históricos de producción de potencia eléctrica media horaria.
- La inclusión de variables calculadas en función del día del año para el que se realiza la predicción (por ejemplo la elevación solar), mejora la precisión de los modelos.
- Los resultados dependen de la planta elegida. Un modelo desarrollado para una planta que obtiene muy buenos resultados de predicción, aplicado a otra planta situada en una localización diferente, puede obtener resultados mucho peores. En general, aquellos modelos de plantas donde se obtienen mayores porcentajes de días claros obtienen mejores resultados.
- En muchos de los trabajos publicados se echa en falta una comparativa de modelos, y también la evaluación de la incertidumbre asociada a la predicción de los modelos. En aquellos casos en que se comparan los resultados obtenidos por diferentes modelos, estos se limitan a dos o tres y, en la mayoría de los casos, de la misma familia.
- Son raros los modelos que predicen hasta el tercer día.
- En modelos a muy corto plazo (unas pocas horas) normalmente no se utilizan predicciones meteorológicas, sin embargo en los modelos con mayor horizonte de predicción si se usan. Pero no hay ningún estudio que determine a partir de que horizonte de predicción mejora el modelo con el uso de las predicciones meteorológicas. Esto puede ser importante de cara a considerar la compra o no de predicciones de variables meteorológicas, ya que, en muchos casos, la obtención de estas predicciones pueden tener un coste económico.
- La inclusión de variables de tipo astronómico, como el ángulo de elevación del Sol, permite reducir los errores de predicción de los modelos.

Capítulo 4

Astronomía terrestre y solar

En este capítulo se recuerdan algunas nociones de astronomía terrestre y solar extraídas de [Chiron 03, Chiron 08], y [Orús 07].

4.1. Nociones generales

- Sol

- La trayectoria diurna que describe el Sol en el cielo varía de forma sensible de un día a otro. Por tanto, para las medidas de radiación, es importante saber localizar en cada instante del día y del año la posición de este astro en la bóveda celeste.
- Cuando la Tierra se encuentra a una distancia del Sol de 1 *UA* (unidad astronómica o distancia media entre la Tierra y el Sol), la radiación solar produce, sobre un plano orientado perpendicularmente a su dirección de propagación y situado en el límite superior de la atmósfera, una irradiancia total $I_o = 1367 \text{ W m}^{-2}$. Este valor se conoce como *constante solar*.

- Tierra

- Gira sobre sí misma alrededor de un eje que atraviesa su superficie en dos puntos denominados polo Norte y polo Sur geográficos.
- Su trayectoria en el espacio es una aproximación de una elipse de Kepler de la cual uno de sus focos es el Sol, ver Figura 4.1. Los dos puntos de la elipse en los que la Tierra está más próxima y más lejana del Sol son los dos extremos del eje mayor: *perihelio* a 0.983 *UA* y *afelio* a 1.1017 *UA*.

- Esfera celeste

- Si se extiende el plano del ecuador de la Tierra hasta el infinito, corta sobre la esfera celeste un gran círculo denominado *ecuador celeste*, que subtiende un ángulo constante de $23^\circ 26.5'$ con el plano de la eclíptica, ver Figura 4.2.

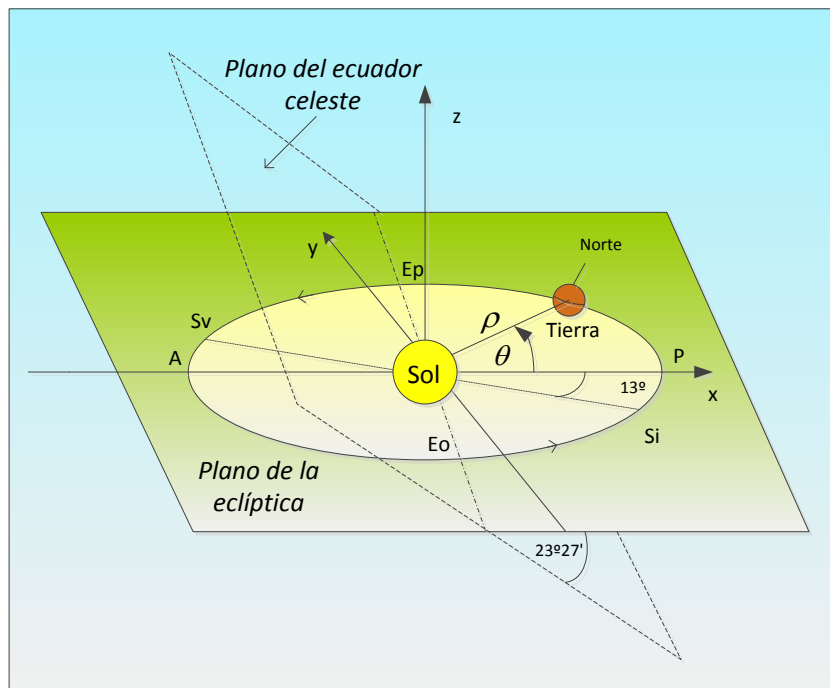


Figura 4.1: P: Perihelio, A: Afelio. Hemisferio Norte: Ep: Equinoccio de primavera, Sv: Solsticio de verano, Eo: Equinoccio de otoño, Si: Solsticio de invierno (adaptado de [Chiron 08]).

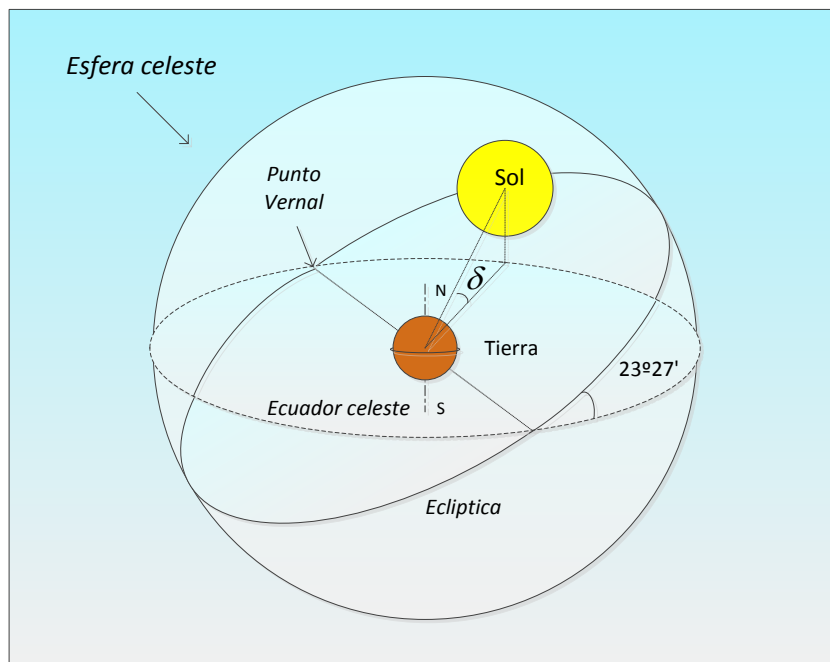


Figura 4.2: La eclíptica (adaptado de [Chiron 08]).

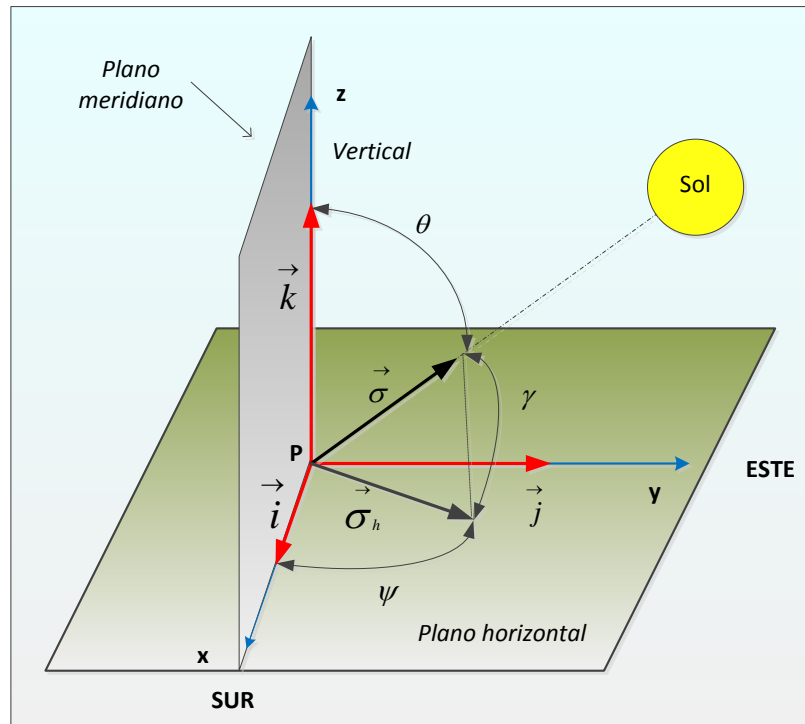


Figura 4.3: Referencia horizontal (adaptado de [Chiron 08]).

- Cuando el Sol se encuentra en una u otra de las dos intersecciones de la eclíptica y del ecuador celeste, la duración del día es igual a la de la noche: son los *equinoccios* (de primavera y otoño). El punto de intersección que corresponde al equinoccio de primavera es el *punto vernal*.

4.2. Las referencias locales

Se usan dos referencias locales que permiten determinar la posición del *Sol* en el cielo a partir de un punto *P* fijo sobre la superficie de la Tierra.

4.2.1. La referencia horizontal (RH) o azimutal

Los tres ejes de esta referencia son (ver Figura 4.3):

- $P\vec{x}$ horizontal y apunta hacia el sur.
- $P\vec{y}$ horizontal y apunta hacia el este.
- $P\vec{z}$ vertical y apunta hacia arriba.

Los vectores unitarios \vec{i} , \vec{j} y \vec{k} , tienen el mismo sentido y dirección que los tres ejes respectivamente. Las dos coordenadas que definen la dirección del Sol son angulares:

- *Azimut* $\psi = (\vec{i}, \vec{\sigma}_h)$ de origen \vec{i} ; $\vec{\sigma}_h$ es la proyección sobre el plano horizontal (Pxy) del vector unitario $\vec{\sigma}$ que apunta hacia el Sol.

Convencionalmente:

- $\psi < 0$ si $\vec{\sigma}_h$ apunta hacia el este.
- $\psi > 0$ si $\vec{\sigma}_h$ apunta hacia el oeste.
- *Ángulo Zenital (solar)* $\theta = (\vec{k}, \vec{\sigma})$ comprendido entre 0 y $\pi/2$; o bien su complementario la *Altura o Elevación* $\gamma = (\vec{\sigma}_h, \vec{\sigma}) = \pi/2 - \theta$.

Se tiene, por tanto:

$$\sigma = \cos \gamma \cos \psi \vec{i} - \cos \gamma \sin \psi \vec{j} + \sin \gamma \vec{k} \quad (4.1)$$

el signo menos del segundo término es por convenio del signo para ψ .

4.2.2. La referencia ecuatorial (RE) o horaria

La RE se deduce por una rotación $\pi/2 - \Lambda$ de RH alrededor del soporte de $P\vec{y}$, que lleva a $P\vec{z}$ a apuntar hacia el norte paralelamente al eje de la Tierra. El ángulo Λ es la latitud del punto P ($\Lambda > 0$ si P está el hemisferio norte, y $\Lambda < 0$ si está en el sur), ver Figura 4.4. Así, $P\vec{x}$ y $P\vec{y}$ se encuentran, en un plano paralelo al plano ecuatorial.

Los tres ejes de referencia se denotan por: \overrightarrow{PX} , \overrightarrow{PY} y \overrightarrow{PZ} .

Los tres vectores unitarios: \vec{I} , \vec{J} y \vec{K} , tienen la misma dirección y sentido que los tres ejes respectivamente.

La dirección del Sol se define mediante dos coordenadas angulares:

- El *ángulo horario* $\omega = (\vec{I}, \vec{\sigma}_e)$ de origen \vec{I} , con $\vec{\sigma}_e$ la proyección de $\vec{\sigma}$ sobre el plano (PXY):
 - $\omega < 0$ si $\vec{\sigma}_e$ apunta hacia el este.
 - $\omega > 0$ si apunta hacia el oeste.
- La *declinación* $\delta = (\vec{\sigma}_e, \vec{\sigma}) \in [-23^\circ 26.5' \text{ y } 23^\circ 26.5']$ de origen $\vec{\sigma}_e$, es:
 - $\delta > 0$ cuando $\vec{\sigma}$ está por encima del plano del ecuador (del 21 de marzo al 23 de septiembre).
 - $\delta < 0$ cuando se encuentra por debajo de este plano (del 23 de septiembre al 21 de marzo).
 - $\delta = 0$ para los equinoccios.

El vector unitario $\vec{\sigma}$ en función de las coordenadas de RE es:

$$\vec{\sigma} = \cos \delta \cos \omega \vec{I} - \cos \delta \sin \omega \vec{J} + \sin \delta \vec{K} \quad (4.2)$$

el signo menos del segundo término es por el convenio del signo sobre ω .

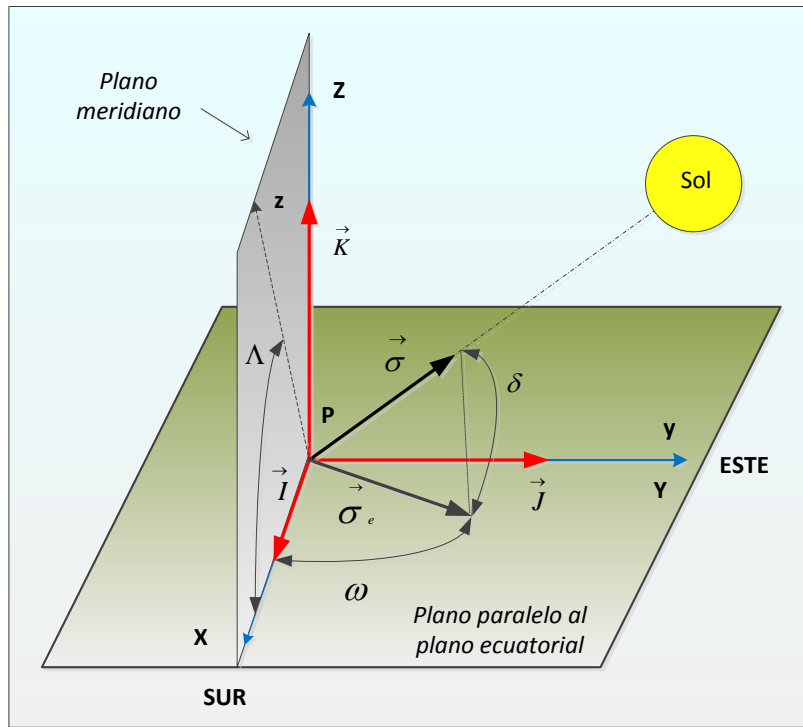


Figura 4.4: Referencia ecuatorial (adaptado de [Chiron 08]).

4.2.3. Paso de coordenadas horizontales a horarias

Los triedros de referencia de los sistemas de coordenadas horizontales y horarias tienen el eje y común y ambos están orientados en sentido retrógrado, por lo que podrá efectuarse el cambio de un sistema al otro por un simple giro alrededor del eje $y \equiv Y$, de ángulo $\pi/2 - \Lambda$ en valor absoluto [Orús 07].

Se recuerda que las matrices que definen un giro de ángulo i son:

- Alrededor del eje x :

$$R_1(i) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos i & \sin i \\ 0 & -\sin i & \cos i \end{pmatrix}$$

- Alrededor del eje y :

$$R_2(i) = \begin{pmatrix} \cos i & 0 & -\sin i \\ 0 & 1 & 0 \\ \sin i & 0 & \cos i \end{pmatrix}$$

- Alrededor del eje z :

$$R_3(i) = \begin{pmatrix} \cos i & \sin i & 0 \\ -\sin i & \cos i & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Estas matrices son ortogonales, por tanto sus inversas coinciden con sus transpuestas:

$$R_h^{-1} = R_h^T \quad (h = 1, 2, 3)$$

y además

$$R_h^T(i) = R_h(-i)$$

Se hace un cambio de base donde $R_2(i)$ tiene las propiedades indicadas en [Orús 07]:

$$\begin{bmatrix} \vec{I} \\ \vec{J} \\ \vec{K} \end{bmatrix} = \begin{pmatrix} \cos i & 0 & -\sin i \\ 0 & 1 & 0 \\ \sin i & 0 & \cos i \end{pmatrix} \cdot \begin{bmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{bmatrix} \quad (4.3)$$

Para pasar de coordenadas horizontales a horarias se toma $i = \Lambda - \pi/2$, ya que el ángulo está contado en sentido contrario al de la orientación del triedro:

$$\begin{aligned} \begin{bmatrix} \vec{I} \\ \vec{J} \\ \vec{K} \end{bmatrix} &= \begin{pmatrix} \cos(\Lambda - \pi/2) & 0 & -\sin(\Lambda - \pi/2) \\ 0 & 1 & 0 \\ \sin(\Lambda - \pi/2) & 0 & \cos(\Lambda - \pi/2) \end{pmatrix} \cdot \begin{bmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{bmatrix} \\ &= \begin{pmatrix} \sin \Lambda & 0 & \cos \Lambda \\ 0 & 1 & 0 \\ -\cos \Lambda & 0 & \sin \Lambda \end{pmatrix} \cdot \begin{bmatrix} \vec{i} \\ \vec{j} \\ \vec{k} \end{bmatrix} \end{aligned} \quad (4.4)$$

Se identifica el vector $\vec{\sigma}$ de (4.1) con (4.2), se reemplazan $\vec{I}, \vec{J}, \vec{K}$ por sus expresiones respectivas en función de $\vec{i}, \vec{j}, \vec{k}$ y se obtienen las siguientes relaciones entre las coordenadas de las dos referencias:

$$\sin \gamma = \sin \Lambda \sin \delta + \cos \omega \cos \Lambda \cos \delta \quad (4.5)$$

$$\sin \psi = \cos \delta \frac{\sin \omega}{\cos \gamma} \quad (4.6)$$

4.2.4. Determinación de la declinación

Se denota J' como el *ángulo del día*, que se deduce de J , *número del día del año* (1 para el 1 de enero y 365 para el 31 de diciembre o 366 si el año es bisiesto):

$$J' = 360^\circ J / 365.25 \quad (4.7)$$

En [Chiron 08] se propone una relación que permite estimar la *declinación del día*, δ , a partir de J' :

$$\sin \delta = 0.3978 \sin[J' - 80.2^\circ + 1.92^\circ \sin(J' - 2.80^\circ)] \quad (4.8)$$

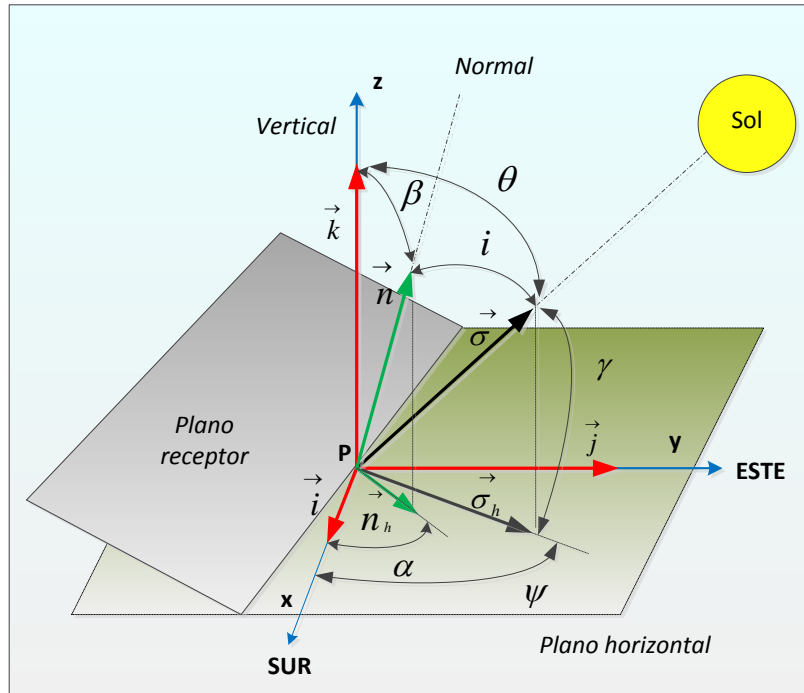


Figura 4.5: Ángulo de incidencia de los rayos solares directos (adaptado de [Chiron 08]).

4.3. Ángulo de incidencia de los rayos solares directos

Es importante conocer el *ángulo de incidencia* que forman los *rayos solares directos* con el plano que les recibe:

- Se denota como \vec{n} el vector unitario normal que parte del plano receptor *por el lado iluminado*, este ángulo de incidencia se define como la magnitud positiva $i = (\vec{n}, \vec{\sigma})$.
- Su valor se obtiene determinando el producto escalar $\vec{n} \cdot \vec{\sigma} = \cos i$ cuya expresión depende de la referencia utilizada.

4.3.1. Ángulo de incidencia en la RH

Para expresar $\cos i$ en esta referencia (ver Figura 4.5), se introduce:

- El *ángulo azimutal* del plano receptor, que es el ángulo $\alpha = (\vec{i}, \vec{n}_h)$ de origen \vec{i} , siendo el vector \vec{n}_h la proyección de \vec{n} sobre el plano horizontal (Pxy);
 - $\alpha < 0$ si \vec{n}_h apunta hacia el este.
 - $\alpha > 0$ si \vec{n}_h apunta hacia el oeste.
- La *inclinación* del plano receptor, es el ángulo $\beta = (\vec{k}, \vec{n})$.

Por consiguiente

$$\vec{n} = \cos \alpha \sin \beta \vec{i} - \sin \alpha \sin \beta \vec{j} + \cos \beta \vec{k} \quad (4.9)$$

de donde se deduce a partir de $\vec{\sigma}$ dada por (4.1), que:

$$\cos i = \sin \beta \cos \gamma \cos(\psi - \alpha) + \cos \beta \sin \gamma \quad (4.10)$$

4.3.2. Ángulo de incidencia en la RE

En esta referencia

$$\begin{aligned} \vec{n} = & (\cos \alpha \sin \beta \sin \Lambda + \cos \beta \cos \Lambda) \vec{I} \\ & - \sin \alpha \sin \beta \vec{J} + (\cos \beta \sin \Lambda - \cos \alpha \sin \beta \cos \Lambda) \vec{K} \end{aligned} \quad (4.11)$$

de donde se deduce por la expresión de $\vec{\sigma}$ dada por (4.2), que:

$$\begin{aligned} \cos i = & \cos \alpha \sin \beta \sin \Lambda \cos \delta \cos \omega + \cos \beta \cos \Lambda \cos \delta \cos \omega \\ & + \sin \alpha \sin \beta \cos \delta \sin \omega + \cos \beta \sin \Lambda \sin \delta - \cos \alpha \sin \beta \cos \Lambda \sin \delta \end{aligned} \quad (4.12)$$

4.4. Las coordenadas temporales

Existen varias coordenadas de tiempo:

- La *ecuación del tiempo* (ET) se expresa en *horas decimales* en función del ángulo del día:

$$ET = -0.128 \sin(J' - 2.80^\circ) - 0.165 \sin(2J' + 19.70^\circ) \quad (4.13)$$

- El *tiempo universal* (TU) es el tiempo solar medio (TSM) en el *meridiano de Greenwich*. Está relacionado con $TSM(L)$ (el TSM de un lugar cualquiera de longitud L), por la expresión:

$$TSM(L) = TU + \frac{L}{15} \quad (4.14)$$

en la que TU y $TSM(L)$ están en horas decimales. La longitud L se cuenta en grados de arco a partir del meridiano de Greenwich y debe ser tomada positiva hacia el este y negativa hacia el oeste.

- El *tiempo legal* (TL) es el TSM de un meridiano tomado como referencia y que atraviesa, el país o sus proximidades. Se tiene:

$$TL = TU + \frac{L_R}{15} + C \quad (4.15)$$

TL y TU son expresados en horas decimales, L_R en grados (en España $L_R = +15^\circ$). La constante C es la *corrección de hora de verano* (en España son 2 horas en verano y 1 en invierno).

- El *ángulo horario* del sol en la RE se calcula de la siguiente manera:

$$\omega = 15 \left[\left(TL + \frac{L - L_R}{15} - C + ET \right) - 12 \right] \quad (4.16)$$

Capítulo 5

Predicción numérica del tiempo

5.1. Introducción

Desde la antigüedad, el ser humano se ha interesado por predecir los fenómenos meteorológicos, debido a la influencia que estos tienen en el desarrollo de la vida en general [Oldani 07].

La predicción se clasifica según el periodo de tiempo que abarca [Fuentes 00]:

- A *corto plazo*: se elabora para las próximas 24 horas, anticipando las condiciones atmosféricas para el segundo y tercer día.
- A *medio plazo*: para un periodo de 3 ó 4 días.
- A *largo plazo*: para periodos superiores a 5 días, pudiendo abarcar periodos mensuales o estacionales.

El método de la *predicción numérica* se sigue a corto y medio plazo. Considera la circulación atmosférica como un problema termodinámico y de mecánica de fluidos. En un momento determinado, en un área geográfica, se toma como punto inicial el análisis del estado de la atmosfera, y a partir de ahí, resolviendo ecuaciones básicas de la física de la misma, se predicen los futuros valores de temperatura, presión, densidad y velocidad del viento. Para ello se utilizan varios modelos atmosféricos, y se requiere el uso de ordenadores con gran capacidad de almacenamiento de datos y con gran velocidad de resolución [Fuentes 00].

5.2. Inicio

Se describe brevemente la evolución histórica [Lezaun 02]:

- En 1903: Vilhelm Bjerknes propuso tratar la evolución de la atmósfera según las leyes de la termodinámica y de la mecánica de fluidos, y defendió que la predicción del tiempo es un problema de fijar valores iniciales, pero para ello hay que resolver un sistema de ecuaciones en derivadas parciales no lineales para las que no se dispone de soluciones analíticas.

- En los años 1910: Lewis Fry Richardson propuso un método aritmético (diferencias finitas) para resolver ecuaciones en derivadas parciales de forma aproximada, y aplicó su método en las ecuaciones de la predicción del tiempo propuestas por Bjerknes. Sus trabajos se publicaron en 1922 en el libro titulado *Weather Prediction by Numerical Process*.
- En 1928: Richard Courant, Kurt Friedrichs y Hans Lewy estudiaron la manera de resolver las ecuaciones en derivadas parciales por “diferencias finitas”, y fijaron de un modo preciso las condiciones de estabilidad que se deben respetar en la discretización.
- En 1939: Carl-Gustaff Rossby dedujo una ecuación cuya solución da la velocidad de propagación de ciertas ondas de longitud larga (denominadas ondas de Rossby) presentes en las corrientes del oeste que circundan la tierra en altura en las latitudes medias. En 1940 hizo varias aproximaciones en las ecuaciones utilizadas por Richardson y consiguió la ecuación de balance de la componente vertical de la vorticidad, la cual traduce el comportamiento de una atmosfera promedio.
- En 1950: Jule Charney, Ragnar Fjörtoft y von Neumann realizaron la primera predicción numérica del tiempo. Para ello eligieron el modelo propuesto por Rossby. Los cálculos numéricos los realizaron en el ENIAC (uno de los primeros de los ordenadores creados por el hombre).
- En 1951: Norman Phillips, introdujo un modelo que tiene en cuenta la estructura vertical de la atmósfera.
- En 1961: Edward Lorenz simuló mediante ordenador el comportamiento de la atmósfera sobre largos periodos de tiempo, y en 1963 publicó el artículo *Deterministic non periodic flows* en el cual introdujo el primer sistema dinámico caótico o *sistema de Lorenz*.
- A mediados de los años 1960, dado que los ordenadores eran más potentes, se volvió a las ecuaciones primitivas de Richardson. Desde entonces los modelos de predicción numérica se han ido haciendo cada vez más complejos y para su resolución se hace uso de un importante entramado matemático.

5.3. Rejillas de integración

Para desarrollar un modelo de predicción numérica del tiempo, el espacio ocupado por la atmósfera se divide mediante una rejilla tridimensional (dos dimensiones horizontales y una vertical, ver Figura 5.1) de manera que en cada nodo debemos calcular los valores de las variables del modelo, de sus derivadas parciales discretizadas y de las fuerzas que intervienen en las ecuaciones del modelo. Como resultado obtenemos en los nodos los valores de las variables básicas del modelo (temperatura, humedad específica y componentes del viento) y de las variables derivadas (precipitación, nubosidad, etc.) previstas para un tiempo futuro [García-Moya 02].

Para un modelo global de circulación general de la atmósfera necesitamos elegir una rejilla que cubra todo el planeta. El inconveniente que tienen estas rejillas es que cuando

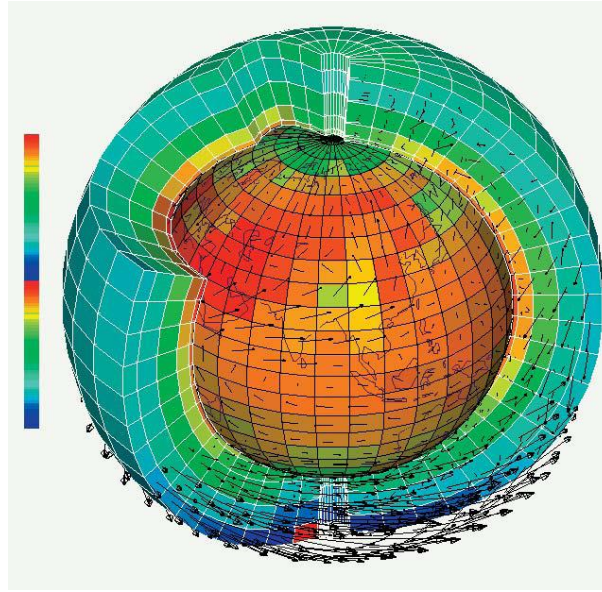


Figura 5.1: *Representación simbólica de discretización de la atmósfera [Lighezzolo 14].*

queremos aumentar la resolución horizontal del modelo debemos aumentar el número de nodos de la rejilla, esto hace que aumente el número de operaciones numéricas que debemos hacer para cubrir un determinado horizonte de predicción (3 o 4 días), y que aumente también el tiempo que necesita el ordenador para hacer la misma predicción, lo que lo hace inviable en un modelo operativo. Una solución a este problema es considerar una rejilla con resolución más alta en la zona de interés de las predicciones. Otra solución es considerar la rejilla de integración en la parte de la superficie de la Tierra que nos interese. El inconveniente de estos modelos es que para poder resolver las ecuaciones del sistema en los puntos de la última fila o columna del área de integración necesitamos saber los valores de las variables fuera del área cubierta por la rejilla. Para ello se usan los valores de las variables obtenidas de un modelo cuya rejilla de integración incluya a la del nuestro [García-Moya 02].

5.3.1. Coordenadas horizontales

Para representar y calcular las cantidades físicas involucradas en los modelos de predicción numérica, en una rejilla cuadrada se consideran cinco distribuciones distintas de las variables de interés, que denominaremos por h , u y v [Arakawa 77], ver Figura 5.2.

5.3.2. Coordenadas verticales

Una descripción adecuada de la estructura vertical de la atmósfera conduce a un mejor pronóstico. Algunas de estas coordenadas son [Lighezzolo 14]:

Coordenadas σ

- Definen su base a nivel suelo.

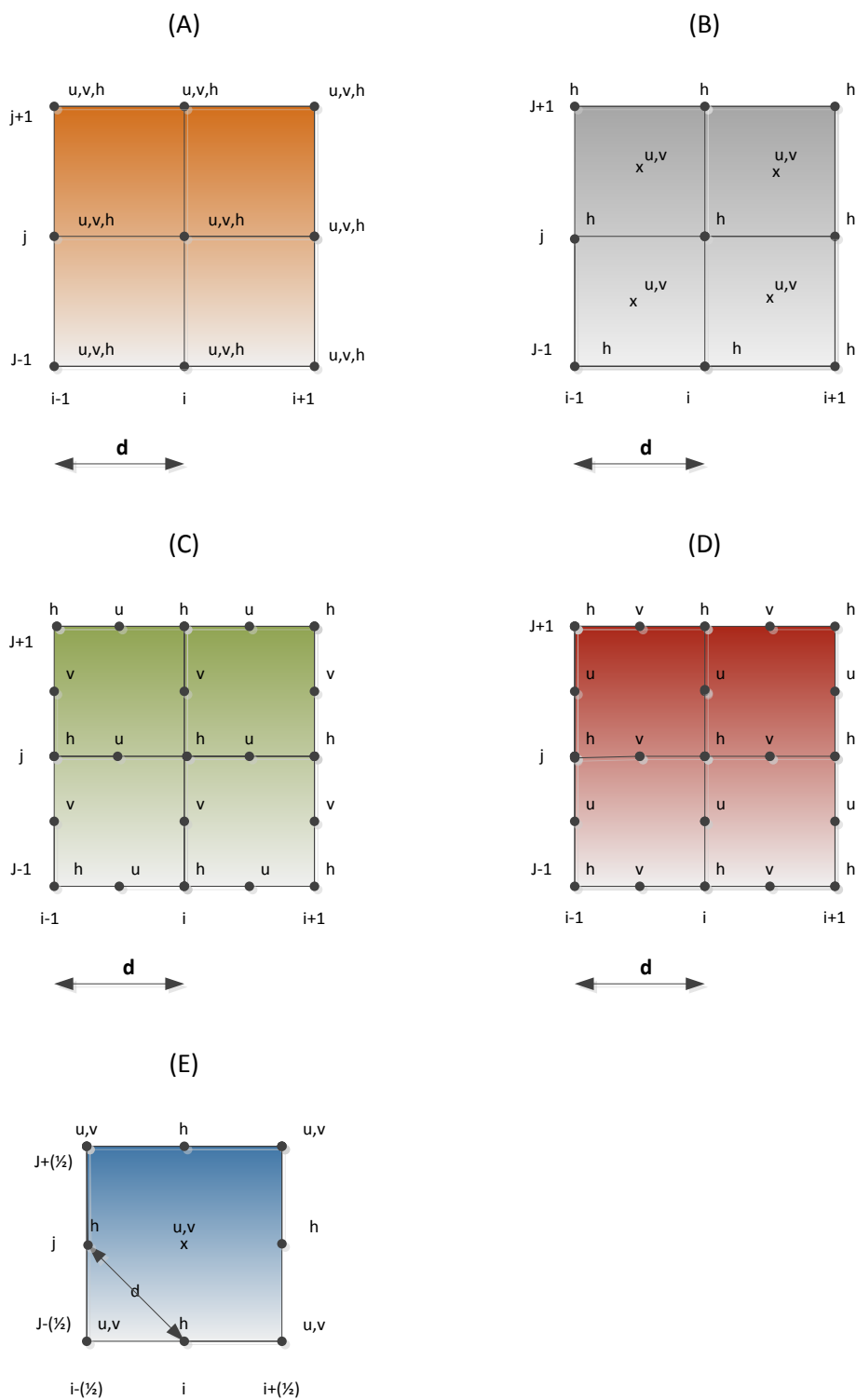


Figura 5.2: Distribuciones espaciales posibles de las variables dependientes en una rejilla cuadrada (adaptado de [Arakawa 77]).

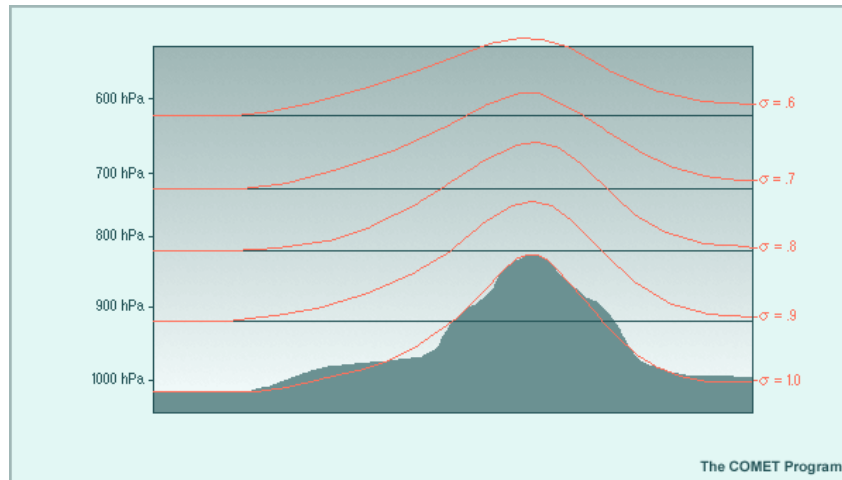


Figura 5.3: Representación esquemática de las coordenadas σ [Lighezzolo 14].

- La superficie sigue el modelo de terreno.
- Se define como:

$$\sigma = p/p_s,$$

donde:

1. p es la presión de un nivel pronosticado dentro del modelo y
 2. p_s es la presión en la superficie de la tierra.
- El valor de σ sobre la superficie es uno.
 - En la capa superior (entre 25 y 1 hPa) el valor de σ es igual a cero.
 - En la Figura 5.3 se muestra una representación grafica.

Coordenadas ETA

- Usa la presión media a nivel del mar.
- Se define como:

$$\eta_s = \frac{p_r(z_s) - p_t}{p_r(z = 0) - p_t},$$

donde:

1. p_t es la presión en el tope del modelo,
 2. $p_r(z = 0)$ es la presión atmosférica media a nivel del mar (MSL = 1013 hPa),
 3. $p_r(z_s)$ es la presión atmosférica al nivel Z_s del terreno.
- El dominio va desde 0 en el tope de la atmósfera hasta 1 a nivel del mar.
 - En la Figura 5.4 se muestra una representación gráfica.

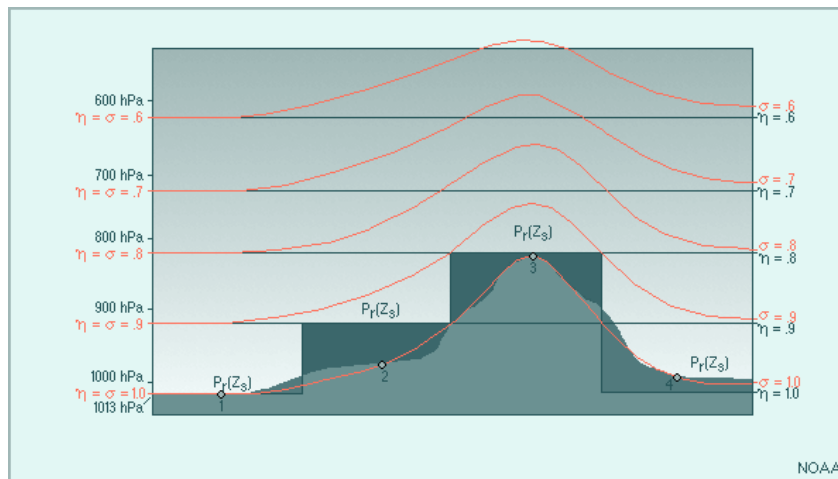


Figura 5.4: Representación esquemática de las coordenadas sigma y ETA [Lighezzolo 14].

Coordenadas híbridas

- Usa coordenadas sigma en las capas inferiores mientras que en las superiores utiliza coordenadas planas, las cuales tienen mejores propiedades numéricas y aumentan la eficiencia y precisión de los cálculos de transferencia radiativa usados en la asimilación de observaciones satelitales de radiancia.
- Esta clase de coordenadas las utiliza el modelo global GFS, ver Figura 5.5.

5.3.3. Aumento de resolución

El problema consiste en “interpolar” las predicciones de un modelo numérico dadas en una cierta rejilla a puntos sub-rejilla (una rejilla de mayor resolución). Este problema aparece cuando la resolución espacial del modelo numérico de circulación atmosférica es insuficiente para una aplicación dada, o cuando se desean obtener predicciones locales. Éste se complica cuando se desean obtener predicciones de variables en superficie cuya salida no proporciona directamente el modelo, sino que han de estimarse a partir de otras variables disponibles [Gutiérrez 04].

Se han propuesto diferentes técnicas para abordar este problema [Gutiérrez 04], utilizando además de las predicciones de los modelos numéricos, las series temporales de observaciones disponibles en el lugar de interés.

Técnicas dinámicas

Incrementan la resolución de la rejilla del modelo atmosférico global, anidando a éste un modelo atmosférico regional o mesoescalar limitado al lugar de interés. De esta manera se puede mejorar la orografía y la parametrización de algunos procesos físicos locales (radiación, cúmulos,...).

A veces, estos métodos numéricos se acoplan con técnicas estadísticas, como los filtros de Kalman [Bergman 85] para eliminar errores sistemáticos del modelo global, adaptando

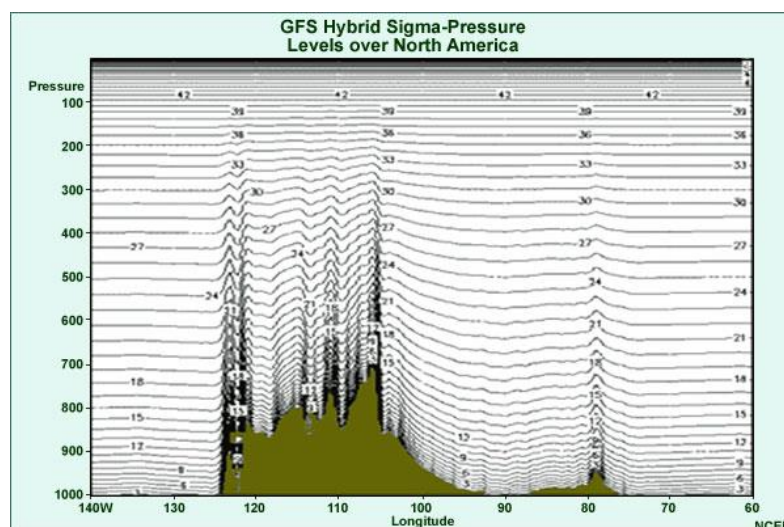


Figura 5.5: *Representación esquemática de las coordenadas híbridas utilizadas por el GFS [Lighezzolo 14].*

dinámicamente las salidas de los modelos numéricos en base a las observaciones disponibles.

Técnicas físicas

Aplican relaciones físicas entre las diferentes variables del modelo y otras variables auxiliares (orografía,...) para estimar el valor de los campos físicos que se desean predecir en una rejilla de alta resolución a partir de las salidas de baja resolución de los modelos.

Técnicas estadísticas

Éstas se aplican cuando se dispone de un conjunto suficiente de datos en registros históricos, de la variable que se desea predecir en la rejilla de alta resolución, o en los puntos locales de interés. En este caso, distintos métodos estadísticos permiten combinar las salidas de los modelos numéricos con esta información.

5.4. Procesos de ejecución

5.4.1. Pre-proceso

Para el establecimiento del punto de partida de la atmósfera se realizan varias tareas [Daley 99, Fernández 08]:

- Determinación de la zona de estudio.
- Resolución de la rejilla.
- Características del suelo.

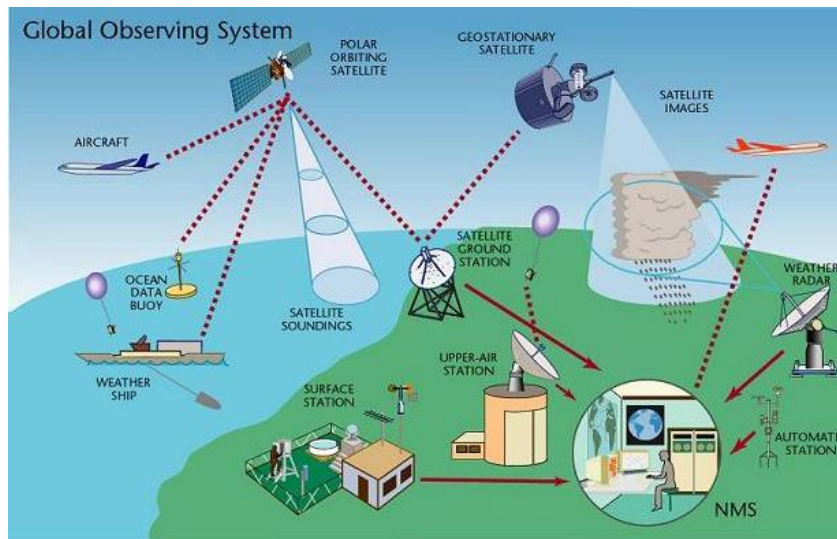


Figura 5.6: Tipos de observaciones meteorológicas [GOS].

- “Asimilación” de los datos atmosféricos disponibles.

Para conocer los valores iniciales de todas las variables en todos y cada uno de los nodos que componen la rejilla del modelo, se utilizan las observaciones meteorológicas que se realizan en todo el mundo [García-Moya 02].

El *Global Climate Observing System* (GCOS) es el sistema coordinado que recopila y comprueba las observaciones de variables atmosféricas y oceánicas (en superficie) a nivel global. Para las observaciones superficiales hay aproximadamente 10000 observatorios que realizan observaciones al menos cada 6 horas, y a menudo cada hora. Además existen unas 1000 estaciones de radiosonda y más de 3000 aviones que realizan observaciones en varios niveles de la atmósfera. Unos 7300 barcos, 600 boyas a la deriva, 300 boyas fijas y 600 plataformas proporcionan observaciones en los océanos [Gutiérrez 04] (ver Figura 5.6).

5.4.2. Proceso

El modelo matemático está formulado por un sistema de ecuaciones diferenciales en derivadas parciales, las cuales traducen las leyes generales de la física que rigen la atmósfera terrestre [Lezaun 02].

Estas ecuaciones son [Ledesma 11]:

- La primera ley de la termodinámica.
- La segunda ley del movimiento de Newton.
- La conservación de la masa que representa la ecuación de continuidad.
- La ley de los gases perfectos.

Tabla 5.1: *Escalas atmosféricas y fenómenos asociados.*

[Orlanski 75]	[Stull 88]	[Thunis 96]	Escala	Duración	Fenómenos Atmosféricos
Macro- α	Macro	Macro- α	10000 km	1 mes	Ondas largas
Macro- β	Macro	Macro- β	10000 km	1 mes	Anticiclones
Meso- α	Meso- α	Macro- γ	2000 km	1 semana	Huracanes
Meso- β	Meso- β	Meso- β	200 km	1 día	Brisas
Meso- γ	Meso- γ	Meso- γ	20 km	1 hora	Tormentas
Micro- α	Micro- α	Meso- δ	2 km	1 hora	Tornados
Micro- β	Micro- β	Micro- β	200 m	30 mín.	Trombas
Micro- γ	Micro- γ	Micro- γ	20 m	1 mín.	Ondas sonoras
Micro- γ	Micro- δ	Micro- δ	2 m	1 seg.	Turbulencia

Puesto que las ecuaciones son no lineales, en la práctica se suele recurrir a distintas aproximaciones que simplifican la resolución numérica, siendo las más usuales la de diferencias finitas, la de elementos finitos y la de esquemas de interpolación [Fernández 08].

Además es necesario conocer los valores de las variables atmosféricas sobre la frontera del dominio atmosférico D considerado. En el caso en el que D sea toda la atmósfera terrestre sólo hay que tener en cuenta sus límites inferior y superior. En el caso en el que esté limitado por una frontera lateral, además habrá que hacer suposiciones más o menos justificadas sobre la evolución de las variables sobre esa frontera [Lezaun 02].

A partir de los valores iniciales en cada nodo de la rejilla se resuelven las ecuaciones resultantes [García-Moya 02].

5.4.3. Post-proceso

Los resultados del modelo son datos con los valores de las variables meteorológicas (temperatura, humedad, viento, presión etc.) que se pronostican para cada uno de los nodos de la rejilla del modelo a lo largo del periodo de predicción [García-Moya 02].

5.5. Tipos de modelos numéricos del tiempo

Los fenómenos que suceden en la atmósfera tienen dimensiones espaciales y temporales muy dispares superponiéndose usualmente unos con otros [Jorba 05]. En la Tabla 5.1 se detalla dicha subdivisión según distintos autores. Cada tipo de modelo caracteriza un conjunto de estos fenómenos, desde fenómenos sinópticos de miles de kilómetros y períodos de días, hasta fenómenos mesoescalares de escalas de kilómetros y períodos de horas [Gutiérrez 04].

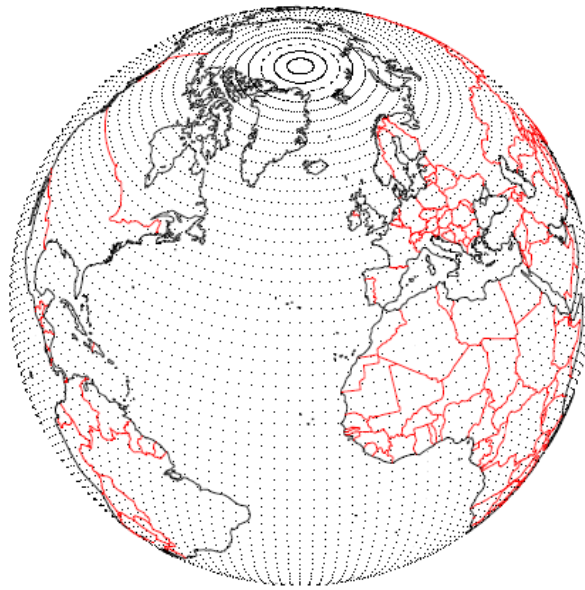


Figura 5.7: *Rejilla global de 2.5° de resolución horizontal utilizada por modelos de circulación general sobre todo el globo; el tamaño de la rejilla es $144 \times 73 = 10512$ puntos [Gutiérrez 04].*

5.5.1. Modelos globales

Son modelos de macro-escala, que simulan el flujo general atmosférico de todo el planeta [Fernández 08]. Se integran sobre todo el globo y su tratamiento numérico se realiza en coordenadas esféricas [Gutiérrez 04].

La resolución horizontal de estos modelos viene caracterizada por el número de modos esféricos que se consideren en el desarrollo de las soluciones; así, un modelo truncado a 144 modos se denomina T144 y tiene una resolución horizontal de 2.5 grados, que corresponden aproximadamente a 250 km en nuestra latitud [Gutiérrez 04] (ver Figura 5.7).

Pueden tener diferente número de niveles para la altura sobre la superficie de la Tierra. Algunos expresan la componente vertical en función de la presión; utilizan las coordenadas sigma y otros utilizan un sistema de niveles verticales fijos [Fernández 08]. Dada su escasa resolución espacial, estos modelos utilizan la aproximación hidrostática de la atmósfera (los movimientos verticales están parametrizados) [Gutiérrez 04].

Se usan principalmente para predicciones a medio plazo, estacionales y climáticas [García-Moya 02]. Algunos de estos modelos se describen brevemente:

- **Unified Model.** Desarrollado por *U. K. Met. Office*. Su resolución horizontal es de aproximadamente 25 km con 70 niveles verticales y genera pronósticos de hasta 144 horas (6 días) dos veces al día [Lighezzolo 14].
- **Integrated Forecast System (IFS).** Desarrollado y mantenido por el *European Center for Medium-Range Weather Forecast* (ECMWF). Su resolución horizontal es de 25 km con 91 niveles verticales y genera pronósticos de hasta 240 horas (10 días) dos veces al día [Lighezzolo 14].

- **Action de Recherche Petite Échelle Grande Échelle (ARPEGE)**. Lo puso en servicio operativo *Météo-France*. Su resolución horizontal es variable: las dimensiones de la rejilla son mínimas sobre Francia y crecen progresivamente hasta sus antípodas, donde son máximas. La resolución horizontal máxima corresponde a una rejilla de 19.1 km de lado y la mínima a una de 234 km. La atmósfera está dividida en 31 capas, su nivel más alto está en los 5 mbar de presión y el más bajo 20.5 m por encima del suelo. Las observaciones para los datos iniciales se realizan intermitentemente, con 4 análisis diarios [Lezaun 02].
- **Global Forecast System (GFS)**. Desarrollado por *National Centers for Environmental Prediction* (NCEP) USA. El modelo corre en dos partes. La primera de ellas tiene alta resolución y pronostica 192 horas (8 días), mientras que la segunda parte corre desde la hora 192 hasta la 384 (16 días) en baja resolución. Horizontalmente la resolución varía desde 35 a 70 km, verticalmente divide la atmósfera en 64 capas y temporalmente produce pronósticos cada 3 horas en las primeras 192 horas y cada 12 horas hasta las 384 horas.

Para inicializar las predicciones el GFS utiliza información que proviene del *Global Data Assimilation System* (GDAS), éste es el componente operativo del *Gridpoint Statistical Interpolation* (GSI). El GSI puede asimilar los siguientes tipos de observaciones: en superficies, radio sondeos, reportes de aeronaves, de boyas, de radar y de satélites. El modelo global GFS tiene sus datos de salida disponibles libremente en un dominio público en internet (a través de servidores NOMADS) [Lighezzolo 14].

5.5.2. Modelos regionales

Se conocen también como modelos de área limitada (LAMs) [Lighezzolo 14]. Son los utilizados en la predicción operativa a corto y medio plazo, donde cada servicio meteorológico se centra en una zona geográfica [Gutiérrez 04].

Necesitan además de las condiciones iniciales, unas condiciones de contorno en las fronteras del dominio, éstas se consiguen interpolando los valores de salida de un modelo global [Fernández 08]. Permiten el uso de una rejilla con mayor resolución horizontal que la de los modelos globales, concentrando los recursos computacionales en la zona de interés. Esto permite resolver explícitamente los fenómenos meteorológicos a pequeña escala que no pueden ser resueltos con una rejilla menos fina como la de los modelos globales [Lighezzolo 14]. El anidamiento de rejillas de tamaño decreciente y resolución creciente y las integraciones se realizan de forma anidada, aprovechando las salidas de una rejilla como condiciones de contorno de la siguiente [Gutiérrez 04] (ver Figura 5.8).

Logran representaciones detalladas de procesos físicos y resuelven topografías complejas [Lighezzolo 14]. En la actualidad existen muchos modelos de área limitada, algunos de ellos se describen brevemente a continuación:

- **High Resolution Limited Area Model (HIRLAM)**. Resultado del trabajo común llevado a cabo por los países escandinavos, Irlanda, Holanda y España [Lezaun 02]. Su resolución horizontal es de 5–15 km, es hidrostático y utiliza coordenadas híbridas en la vertical. El tipo de rejilla utilizado es el Arakawa-C [Lighezzolo 14]. Su alcance temporal es de 72 horas. Presenta el anidamiento, permitiéndole resoluciones espaciales de unos 5 km y alcances de 36 horas [Fernández 08].

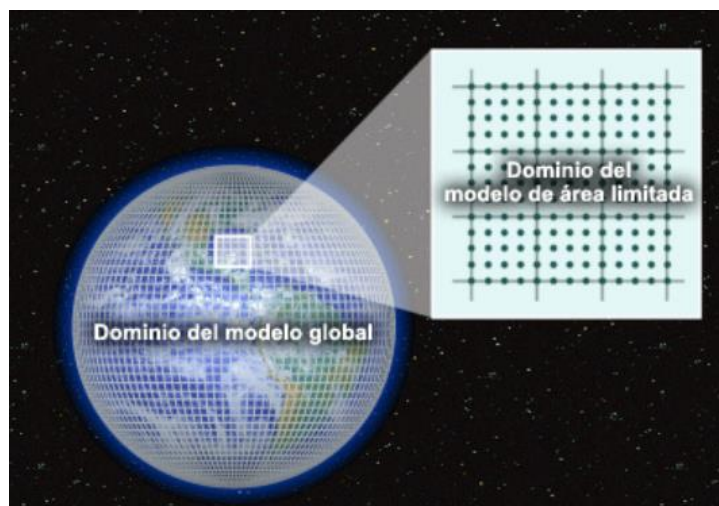


Figura 5.8: Representación gráfica de un modelo regional anidado en un modelo global [Lighezzolo 14].

- **BOLAM**. Desarrollado en el *Institute of Atmospheric Sciences and Climate* (ISAC). Es un modelo no hidrostático. Su resolución horizontal es aproximadamente 11 km en coordenadas geográficas rotadas y 50 niveles de resolución vertical. Las condiciones iniciales pueden ser tomadas del GFS o el IFS [Lighezzolo 14].
- **Aire Limitée, Adaptation Dynamique, Développement InterNational (ALADIN)**. Desarrollado por *Météo-France*. En este modelo el dominio es un cuadrado de 2740 km de lado, en proyección Lambert. Los niveles verticales son los mismos que los de ARPEGE y su resolución horizontal es de 9.9 km . Las condiciones iniciales y de contorno vienen de ARPEGE [Lezaun 02].
- **ETA-12**. Lo anida el NCEP con distintas rejillas de alta resolución (8 km) sobre zonas de interés [Gutiérrez 04].

5.5.3. Modelos Mesoescalares

La resolución de algunos modelos regionales ha aumentado hasta llegar a unos pocos kilómetros, con el objetivo de mejorar la resolución de procesos físicos de pequeña escala [Gutiérrez 04] (ver Figura 5.9). Muchos de éstos incluyen la posibilidad de anidamiento, y además no suelen usar 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 [Fernández 08]. A continuación se citan algunos de estos modelos [Jorba 05]:

- **Mesoscale Atmospheric Simulation System (MASS)**. Comercializado por la compañía norteamericana MESO.
- **Regional Atmospheric Modeling System model (RAMS)**. Desarrollado en la *Colorado State University*.

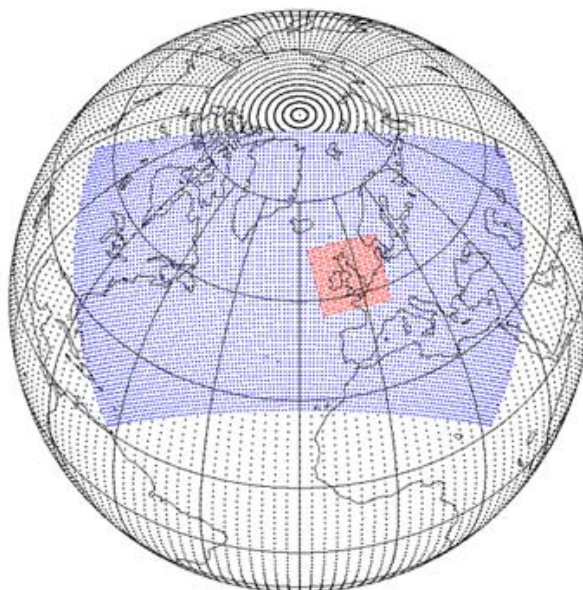


Figura 5.9: *Rejilla de un modelo global, un modelo regional y un modelo mesoescalar [García-Moya 02].*

- **5th generation Mesoscale Model (MM5).** Desarrollado por la *Pennsylvania State University* junto con el *National Center for Atmospheric Research* (NCAR).
- **Mesoscale Compressible Community Model (MC2).** Se trata de una extensión del modelo de área limitada compresible desarrollado por [Tanguay 90] y [Laprise 97] en la *Université du Québec à Montreal* (UQAM).
- **Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS).** Desarrollado por la *Marine Meteorology Division* (MMD) de el *Naval Research Laboratory* (NRL).
- **Mesoscale no-Hidrostatic Atmospheric Simulation System (Meso-NH).** Resultado de la colaboración entre el *Centre National de Recherches Météorologiques (Météo-France)* y el *Laboratoire d'Aérodynamique* (CNRS).
- **Advanced Regional Prediction System (ARPS).** Desarrollado por el *Center for Analysis and Prediction of Storms* (CAPS) de la *University of Oklahoma*.
- **Weather Research and Forecast Model (WRF).** Existen dos módulos principales: el WRF de investigación avanzada (*Advanced Research WRF*, ARW), desarrollado principalmente en la NCAR, y el Modelo de Mesoescala No-Hidrostático (*NonHydrostatic Mesoscale Model*, NMM) desarrollado en el NCEP [Ramírez 12].

5.6. Pronóstico de las variables meteorológicas

Para obtener las predicciones de las variables meteorológicas utilizadas en los trabajos de esta tesis doctoral, se ha utilizado el modelo mesoescalar WRF-NMM versión 3. Éste

soporta una gran variedad de capacidades como simulaciones de datos reales, anidamiento unidireccional y bidireccional, y resoluciones que van desde cientos de metros a miles de kilómetros.

Los factores principales por los que se ha optado por la selección y aplicación de este modelo han sido:

- La disponibilidad de esta herramienta 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 forman este modelo fueron instalados y compilados por primera vez en 2010. Las predicciones meteorológicas usadas en los trabajos de la presente tesis doctoral se obtuvieron con el modelo instalado en una maquina Quad Xeon, 16 Gb. de RAM, bajo el sistema operativo Linux Ubuntu 12.04.
- El uso extendido entre la comunidad científica.
- Su potencia.

A continuación se describen las diferentes etapas (ver Figura 5.10) que se han seguido, para conseguir un pronóstico de 72 horas diario de las diferentes variables meteorológicas.

5.6.1. Descarga de datos globales

En este laboratorio todas las mañanas desde hace varios años, se descargan de los servidores NOMADS del NCEP americano, los ficheros con los resultados del modelo atmosférico GFS $1^\circ \times 1^\circ$, correspondientes a los datos atmosféricos de las 00, 06, 12 y 18 horas GMT, para horizontes de predicción de 0 a 180 horas, en intervalos de 3 horas, de la rejilla que abarca la zona comprendida entre los meridianos: 30° Oeste y 20° Este, y los paralelos: 20° Norte y 60° Norte, para ello se creó un programa que realizaba automáticamente estas descargas.

En esta tesis para establecer las condiciones de inicialización y de frontera e interpolar las variables a los puntos de las rejillas de análisis de cada dominio, se utilizan un subconjunto de estos ficheros, los correspondientes de los datos atmosféricos de las 00 horas, de las fechas comprendidas entre el 1 de enero de 2010 hasta el 28 de noviembre de 2012, para un pronóstico de 72 horas. El formato de estos ficheros es *GRIB2*.

5.6.2. Etapa de preprocesamiento

El sistema de preprocesamiento del WRF (WPS) es un conjunto de tres programas [Janjic 14]:

Geogrid

Define el dominio sobre el que se está haciendo la predicción. Éste está definido en el archivo *namelist.wps* [Ramírez 12].

El modelo simula las variables meteorológicas a partir de la información topográfica que se le suministre, por eso es necesario trabajar con una resolución elevada en el área

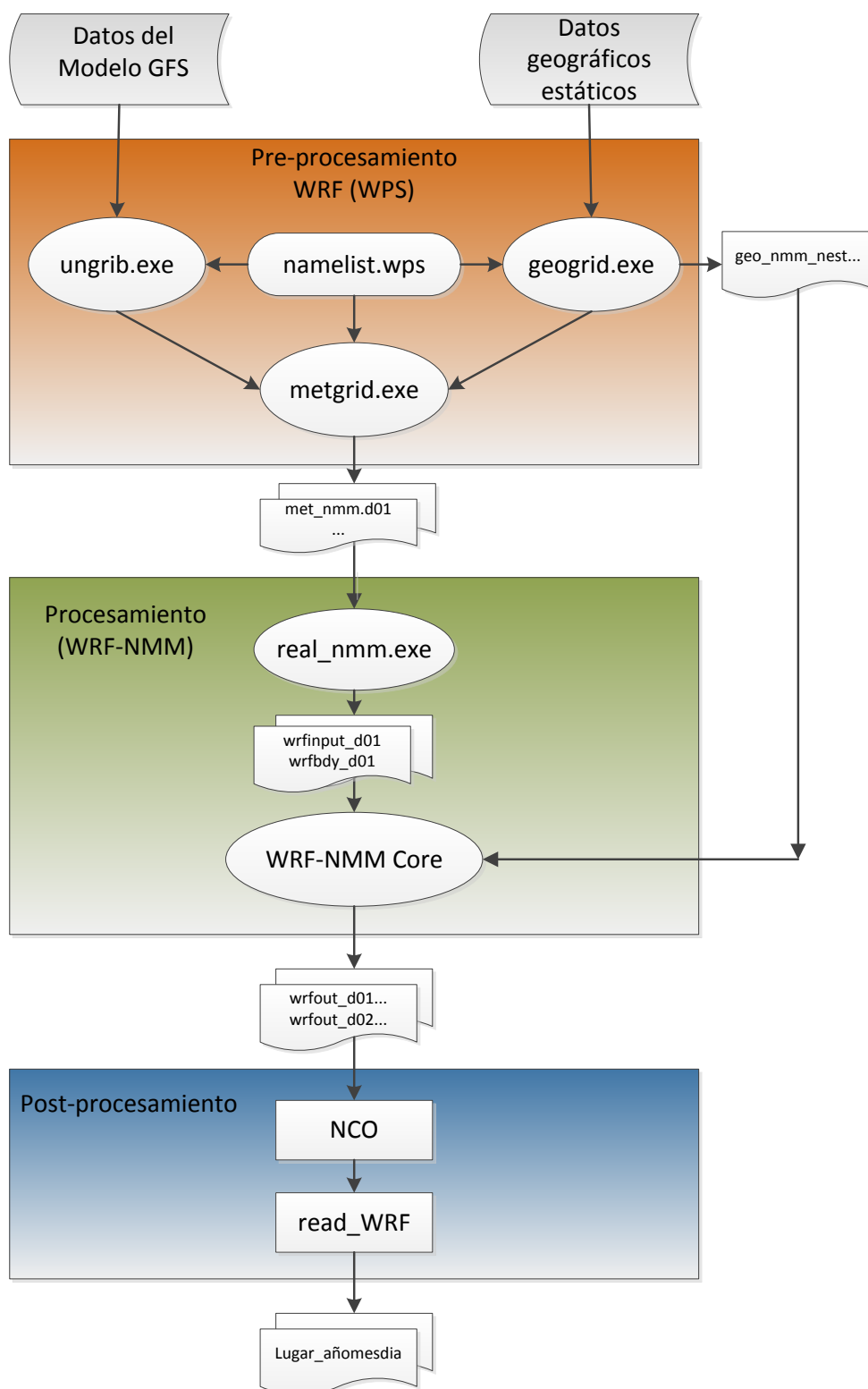


Figura 5.10: Diagrama de flujo del modelo WRF-NMM (adaptado de [Janjic 14]).

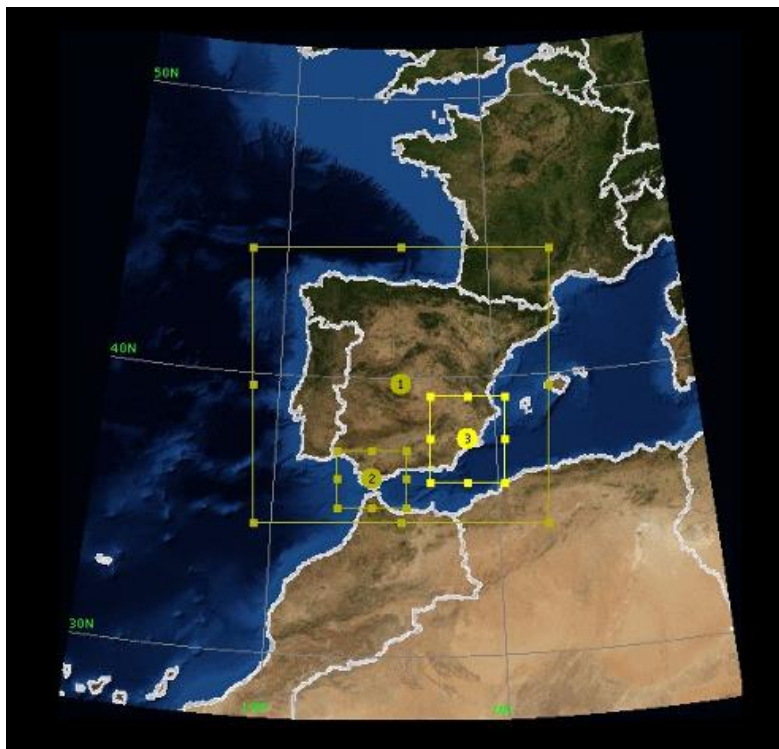


Figura 5.11: Definición de los dominios (WRF Domain Wizard).

geográfica de interés. Para este estudio se han definido dos dominios presentados en la Figura 5.11. El dominio externo (D1) engloba toda la Península Ibérica y el dominio interno (D3) se centra en el sureste de la Península. Sus dimensiones y resoluciones se detallan en la Tabla 5.2.

Una vez editado el archivo *namelist.wps*, se ejecuta el archivo *geogrid.exe* para generar los datos estáticos sobre el dominio [Lighezzolo 14]. Se crean los archivos: *geo_nmm_nest...* [Janjic 14]. En estos archivos está contenida toda la información de las variables meteorológicas correspondientes al dominio, su formato es *Grib* [Lighezzolo 14].

Ungrib

Lee archivos *Grib* y los convierte en archivos más simples con un formato intermedio [Ramírez 12].

Metgrid

Interpola horizontalmente los datos extraídos por Ungrib dentro del dominio definido por Geogrid [Ramírez 12]. Los archivos generados son de la forma: *met_nmm.d01...*

5.6.3. Etapa de procesamiento

En esta etapa se corre el modelo. Primero se edita el archivo *namelist.input* y después se ejecutan los archivos *real.exe* y *wrf.exe* [Lighezzolo 14]:

Tabla 5.2: *Dimensión y resolución de los dominios que se han utilizado.*

Dominio	Zona de estudio	Inicio y Frontera	Ratio	Rejilla horizontal (Puntos)	Distancia media horizontal puntos	Niveles (50 mb mayor altura)	Salida de datos
D1	Península Ibérica	GFS (1°x1°)	5:1	(61x112)	10 km	45	Cada hora
D3	Sureste Península	WRF (D1)	3:1	(45x105)	1100 m	45	Cada media hora

Real_nmm

Interpola verticalmente los archivos generados por *metgrid.exe*. Crea los datos de la condiciones iniciales y condiciones de contorno, además realiza algunos chequeos de consistencia entre los datos. Después de su correcta ejecución, se crean los archivos: *wrfinput_d0x* y *wrfbdy_d0x*, donde *x* respresenta el número del dominio.

WRF-NMM Core

Por último se ejecuta el programa WRF encargado de aplicar las ecuaciones finales del modelo [Ramírez 12]. Se crean los archivos de salida: *wrfout_d01...*, *wrfout_d03...* en formato *netCDF*.

5.6.4. Etapa de postprocesamiento

Las salidas que proporciona este modelo contienen información pronosticada para las 72 horas siguientes, de diferentes variables meteorológicas. Los ficheros que componen estas salidas tienen formato *netCDF*. Para poder analizar y editar esta información existe un gran número de herramientas que permiten visualizar los datos del modelo WRF, en esta investigación se utilizan primero el paquete NCO y después el programa *read_WRF*.

NCO

El paquete NCO (*netCDF Operators*) es un conjunto de programas de libre distribución diseñados para manipular y gestionar ficheros en los formatos típicos utilizados por los programas de predicción numérica meteorológica. Para nuestro caso las acciones que realizamos con comandos y programas del paquete NCO son las siguientes:

- Para cada uno de los días, se obtienen 145 ficheros de salida del WRF (*wrfout_d03_año-mes-día_00:00:00*, *wrfout_d03_año-mes-día_00:00:30*, ..., *wrfout_d03_año-mes-(día+4)_00:00:00*).

- Mediante la orden `ncrcat -v TIMES, THS, TSHLTR, TH10, QS, QSHLTR, Q10, CFRACH, CFRACL, CFRACM, RSWIN, RSWINC, U10, V10, PSHLTR wrfout_d03_*wrfSal.nc` se agrupan los 145 ficheros y se extraen los valores de las variables de interés generando un nuevo fichero con formato *netCDF* que se llama *wrfSal.nc*.
- Este fichero contiene el valor de todas las variables seleccionadas, para todos los puntos de la rejilla del dominio (D3) y para todos los niveles que están definidos en la vertical.

Read_WRF

El programa `Read_WRF` extrae datos correspondientes a variables meteorológicas para un punto geográfico concreto de los ficheros de salida del programa `wrf.exe`. Este programa fue compilado en el equipo utilizado para obtener las predicciones para los trabajos de investigación de esta tesis a partir del código fuente accesible en los servidores del NCEP.

Mediante la orden `read __wrf_nc wrfSal.nc -ts xy X Y THS, TSHLTR, TH10, QS, QSHLTR, Q10, CFRACH, CFRACL, CFRACM, RSWIN, RSWINC, U10, V10, PSHLTR -lev 1 -rot`, para cada uno de los días, se genera un fichero con formato *txt* llamado *Lugar_añomesdia*. Éste contiene los datos pronosticados ordenados temporalmente (de media en media hora desde las 00:00:00 horas hasta las 00:00:00 horas del 4º día siguiente), de cada una de las variables meteorológicas seleccionadas, para el punto de la rejilla con coordenadas (x,y), para el nivel 1 (superficie del terreno), con las coordenadas del viento en sistema rotado (coinciden Norte-Sur y Este-Oeste).

5.6.5. Variables meteorológicas pronosticadas

Las variables meteorológicas obtenidas son las siguientes:

Temperatura

La temperatura representa el grado de calor o frío de una sustancia medido con el termómetro. Es la velocidad media de moléculas y átomos de una sustancia [Ledesma 11]. Las variables correspondientes a temperaturas extraídas de los ficheros de salida del programa WRF, y para el punto geográfico más cercano a la instalación fotovoltaica, son las siguientes:

- THS: temperatura en superficie ($^{\circ}K$).
- TSHLTR: temperatura potencial en superficie ($^{\circ}K$).
- TH10: temperatura a 10 m ($^{\circ}K$).

Humedad relativa

Es la relación entre la humedad absoluta (densidad de vapor) y la humedad absoluta saturada (cantidad de vapor de agua por m^3 que necesita una masa de aire para estar saturada) [Ledesma 11]. Las variables de humedad relativa extraídas son:

Tabla 5.3: *Clasificación internacional de nubes [Ledesma 11].*

Familia	Símbolo	Género	Base media	Tope medio
Nubes altas	Ci	Cirros	6000 m	12000 m
	Cs	Cirrostratos		
	Cc	Cirrocúmulos		
Nubes medias	As	Altostratos	2500 m	6000 m
	Ac	Altocúmulos		
Nubes bajas	St	Estratos	150–600 m	750 m
	Sc	Estratocúmulos	600–1500 m	2400 m
	Ns	Nimbostratos	0–2500 m	6000 m
Nubes de desarrollo vertical	Cu	Cúmulos	300–2400 m	6000 m
	Cb	Cumulonimbos	600–2400 m	12000 m

- QS: humedad relativa en superficie (entre 0 y 1).
- QSHLTR: humedad potencial relativa en superficie (entre 0 y 1).
- Q10: humedad relativa a 10 m (entre 0 y 1).

Nubes

Las nubes son masas de vapor de agua suspendidas en la atmósfera debidas a un proceso de condensación en mayor o menor cuantía, constituyendo un compuesto de gotitas de agua y cristales de hielo [Ledesma 11]. En 1929 se presentó el atlas de nubes vigente actualmente en todo el mundo, se establecieron cuatro categorías especificadas en la Tabla 5.3. Los distintos tipos de nubes se pueden observar en la Figura 5.12. Las variables con datos correspondientes a nubes extraídas son:

- CFRACL: cobertura de nubes a nivel bajo (entre 0 y 1).
- CFRACM: cobertura de nubes a nivel medio (entre 0 y 1).
- CFRACH: cobertura de nubes a nivel alto (entre 0 y 1).

Radiación de onda corta

Radiación emitida por el Sol, (radiación solar) [Ledesma 11]. Las variables de radiación extraídas son:

- RSWIN: radiación de onda corta (W/m^2).
- RSWINC: radiación de onda corta de día claro (W/m^2).

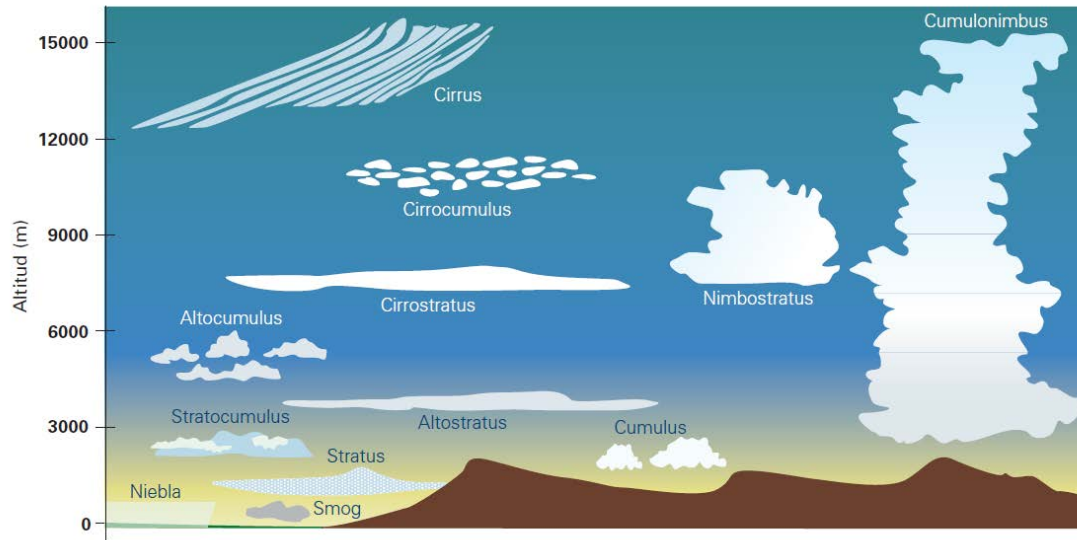


Figura 5.12: *Diferentes géneros de nubes y los niveles aproximados que ocupan en la troposfera [Martín 04].*

Viento

El viento es el aire en movimiento en la superficie terrestre y queda definido por su dirección y su velocidad [Ledesma 11].

El viento se clasifica, por la velocidad y por los efectos que produce, según una escala ideada en 1805 por Beaufort [Oldani 07], ver Tabla 5.4. Las variables extraídas correspondientes a viento son las siguientes:

- U10: componente u del viento a 10 m de altura.
- V10: componente v del viento a 10 m de altura.
- VEL10: velocidad del viento, $\sqrt{(U10)^2 + (V10)^2}$ (m/s).

Presión atmosférica

La presión atmosférica representa el peso de la atmósfera por unidad de superficie [Ledesma 11]. La variable de presión atmosférica extraída es:

- PSHLTR: presión atmosférica a 2 m de altura (Pa).

5.6.6. Valor medio horario de las variables

Los valores de las variables meteorológicas fueron obtenidos para intervalos de media hora (tal y como indica la Tabla 5.2). Para obtener su valor medio horario se procedió de la siguiente manera:

Tabla 5.4: *La escala de Beaufort (adaptado de [Oldani 07]).*

Grado	Velocidad (m/s)	Efectos producidos en el suelo
0. Calma	< 1	No viento, el humo asciende vertical
1. Brisa muy ligera	De 0.6 a 1.7	El humo se inclina
2. Ventolina	De 1.8 a 3.3	Mueve las hojas
3. Brisa muy débil	De 3.4 a 5.2	Agita las hojas
4. Brisa debil, floja	De 5.3 a 7.4	Mueve las ramas pequeñas
5. Brisa moderada	De 7.5 a 9.8	Mueve las ramas gruesas
6. Brisa fresca, fresquito	De 9.9 a 12.4	Se advierte dentro de las casas
7. Frescachón, viento fuerte	De 12.5 a 15.2	Agita ramas y obstaculiza el camino
8. Temporal	De 15.3 a 18.2	Rompe las ramas y agita árboles
9. Temporal fuerte	De 18.3 a 21.5	Levanta las tejas de los tejados
10. Temporal duro	De 21.6 a 25.1	Arranca y rompe árboles
11. Temporal muy duro	De 25.2 a 29	Provoca grandes devastaciones
12. Temporal huracanado	Más de 29	Provoca la destrucción general

- Se creó un programa en Visual Basic, para leer los valores de los ficheros *Lugar_añomesdia* (con formato *txt*) de todos los días ordenados por fecha y guardarlos en un fichero de la hoja de cálculo excel (*xls*).
- El valor medio horario pronosticado de cada variable meteorológica, se calcula sumando el valor pronosticado de esa variable al comienzo de la hora multiplicado por 0.25, más su valor a y media de esa hora multiplicado por 0.50, más su valor al final de esa hora multiplicado por 0.25.
- Por último con estos datos se han creado 3 ficheros *xls* que contienen para todos los días los valores medios horarios de las variables meteorológicas pronosticados para el primer día (primeras 24 horas), para el segundo día (horas entre 25 y 48) y tercer día (horas entre 49 y 72).

Capítulo 6

Radiación solar

La naturaleza intrínseca de la radiación solar y, su carácter aleatorio hace que la radiación solar que llega a la superficie de un generador fotovoltaico solo puedan ser predicciones sobre futuro, que asocian un cierto grado de incertidumbre [Lorenzo 06]. En este capítulo se describen modelos para predecir la irradiancia incidente sobre la superficie de un generador fotovoltaico.

6.1. Algunos conceptos

La comisión de radiación de la *Organización Meteorológica Mundial* (OMM), ha establecido los siguientes conceptos [Ledesma 11]:

- *Radiación*. Es la energía electromagnética propagada o recibida.
- *Irradiancia*. Es la cantidad de energía electromagnética que incide o atraviesa la unidad de superficie en unidad de tiempo.
- *Irradiación*. Es la cantidad total de energía electromagnética recibida por la superficie irradiada durante el tiempo considerado y despedida hacia el espacio.

La irradiación a lo largo de una hora (en Wh/m^2) coincide numéricamente con el valor medio de la irradiancia durante esa hora (en W/m^2) [Lorenzo 06].

6.2. Componentes de la radiación solar

La radiación solar se modifica al atravesar la atmósfera por interacción con sus componentes [Lorenzo 06]:

- Las nubes reflejan radiación.
- El ozono, el oxígeno, el dióxido de carbono y el vapor de agua absorben radiación en algunas longitudes de onda específicas.

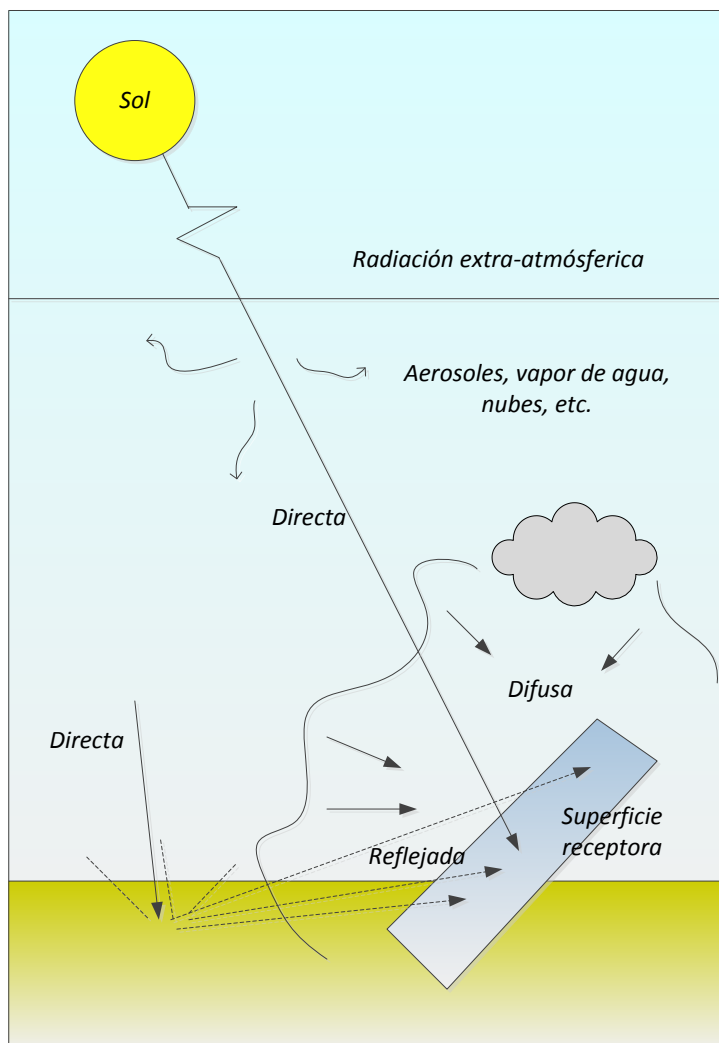


Figura 6.1: *Distintos componentes de la radiación solar (adaptado de [Lorenzo 06]).*

- Las gotas de agua y el polvo en suspensión dispersan radiación.

El resultado de estos procesos es que la radiación solar que incide sobre un receptor situado en la superficie de la Tierra se divide en tres componentes (ver Figura 6.1):

- La *radiación directa*, constituida por los rayos recibidos en línea recta desde el Sol.
- La *radiación difusa*, constituida por los rayos no directos y dispersados por la atmósfera en la dirección del receptor.
- La *radiación del albedo*, procedente del suelo y debida a la radiación incidente en él y reflejada hacia el receptor.

La radiación total (directa + difusa + albedo) incidente en una superficie se denomina *radiación global* [Lorenzo 06].

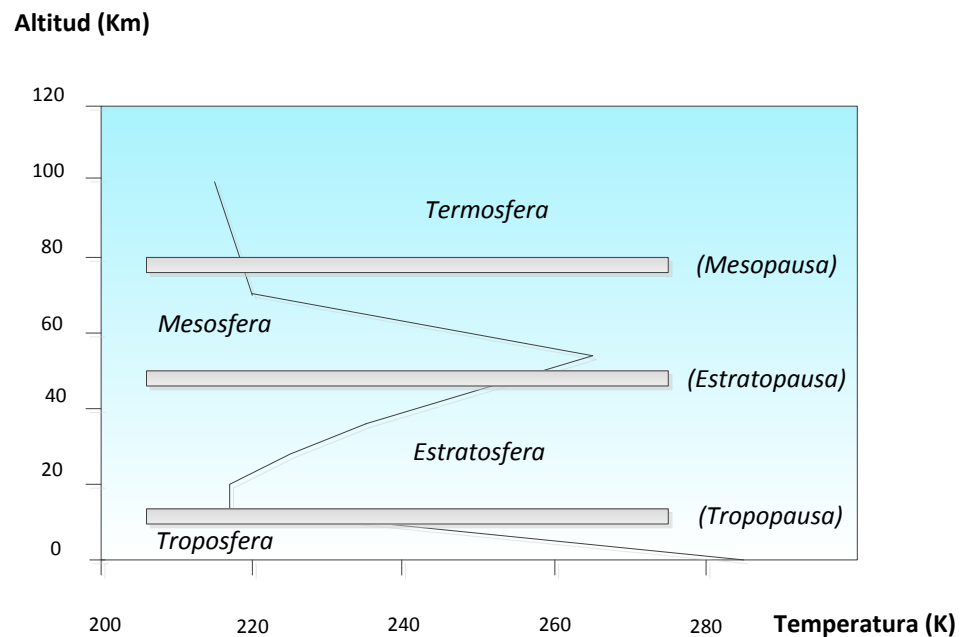


Figura 6.2: Estructura general de la atmósfera (adaptado de [Chiron 08]).

6.3. La atmósfera y sus diversos componentes

En la Figura 6.2 se describe esquemáticamente la estructura general de la atmósfera y se muestra el perfil vertical de su temperatura media en el caso de la atmósfera estándar [Chiron 08].

La radiación solar directa en el curso de su paso por la atmósfera, sufre diversos procesos de atenuación, como consecuencia de las interacciones que se producen entre los fotones y los constituyentes gaseosos, sólidos o líquidos de ésta [Chiron 08]. En la Tabla 6.1 se presentan los gases que se pueden encontrar en la atmósfera en proporciones estables y variables.

Además, numerosos fenómenos producen suspensiones de partículas sólidas o líquidas (*aerosoles*). Estas partículas, cristales o gotas tienen diámetros medios que varían entre $0.01 \mu m$ y algunas decenas de μm . Se trata principalmente de [Chiron 08]:

- Gotas de agua y cristales de hielo que componen las nubes.
- Micrometeoritos.
- Cenizas volcánicas y gotas de ácido sulfúrico inyectadas a la atmósfera por erupciones.
- Polvo elevado por las tormentas.
- Materias orgánicas tales como pólenes.
- Humos de combustión natural.

Tabla 6.1: *Gases en proporciones estables y variables (% en volumen) [Chiron 08].*

Gases			Volumen %
En proporciones estables	Nitrógeno	(N_2)	78.110 ± 0.004
	Oxígeno	(O_2)	20.953 ± 0.001
	Argón	(Ar)	0.934 ± 0.001
	Neón	(Ne)	$(18.180 \pm 0.040) 10^{-4}$
	Helio	(He)	$(5.240 \pm 0.004) 10^{-4}$
	Kriptón	(Kr)	$(1.140 \pm 0.010) 10^{-4}$
	Xenón	(Xe)	$(0.087 \pm 0.001) 10^{-4}$
	Hidrógeno	(H_2)	$0.5 10^{-4}$
	Óxido de nitrógeno	(N_2O)	$(0.500 \pm 0.100) 10^{-4}$
En proporciones variables	Agua	(H_2O)	de 0 a 7
	Anhídrido carbónico	(CO_2)	de 0.01 a ...
	Anhídrido sulfuroso	(SO_2)	de 0 a 0.001
	Ozono	(O_3)	de 0 a 0.00001
	Metano	(CH_4)	de 0.0002 a ...
	Peróxido de nitrógeno	(NO_2)	trazas

- Desechos industriales (minas a cielo abierto...).
- Cristales de sal en el litoral marino.

6.4. Irradiancia

Aunque hay varios mapas del mundo de radiación solar, no son lo suficientemente detallados como para ser utilizados para la determinación de la energía solar disponible en áreas pequeñas. Estas circunstancias han impulsado el desarrollo de procedimientos de cálculo para proporcionar estimaciones de irradiancia para las áreas donde las mediciones no se llevan a cabo, y para situaciones en las que se produjeron lagunas en los registros de medición [Badescu 12]. El significado de los símbolos que se utilizan para las irradiancias (expresadas en W/m^2) se describen en la Tabla 6.2.

6.4.1. Irradiancia extra-atmosférica

La radiación que alcanza la superficie externa de la atmósfera de la Tierra se llama *radiación extra-atmosférica*, y consiste esencialmente en rayos que llegan desde el Sol en línea recta. La radiación extra-atmosférica experimenta variaciones debidas al movimiento aparente del Sol. Estas variaciones son perfectamente predecibles y se pueden calcular teóricamente con meras consideraciones geométricas [Lorenzo 06]. En los trabajos de investigación de esta tesis, la *irradiancia extra-atmosférica* sobre una superficie horizontal

Tabla 6.2: Significado de los símbolos de las irradiancias [Chiron 08].

Símbolo	Significado
I_e	Irradiancia extra-atmosférica
I_o	Irradiancia total “normal” (cuando la distancia Tierra-Sol está 1 UA)
I_n	Irradiancia directa total normal
I_h	Irradiancia directa total horizontal
I_β	Irradiancia directa total sobre un plano de inclinación β
$I_{ef,\beta}$	Irradiancia directa total “efectiva” sobre un plano de inclinación β
D_h	Irradiancia difusa total sobre un plano horizontal
D_{hc}	D_h proveniente de la zona circumsolar
D_{hv}	D_h proveniente del resto de la bóveda del cielo
D_β	Irradiancia difusa total sobre un plano de inclinación β
$D_{ef,\beta}$	Irradiancia difusa “efectiva” sobre un plano de inclinación β
R_β	Irradiancia del albedo sobre un plano de inclinación β
$R_{ef,\beta}$	Irradiancia del albedo “efectiva” sobre un plano de inclinación β
G_h	Irradiancia global total sobre un plano horizontal
G_β	Irradiancia global total sobre un plano de inclinación β
$G_{ef,\beta}$	Irradiancia global total “efectiva” sobre un plano de inclinación β

se calcula basándose en una fórmula que se utiliza en [Monteiro 13b]:

$$I_e = I_o K_d (\sin \delta \sin \Lambda + \cos \delta \cos \Lambda \cos \omega) \quad (6.1)$$

donde:

- $I_o = 1367 \text{ W/m}^2$ es la irradiancia total “normal” cuando la distancia Tierra-Sol está 1 UA [Chiron 08].
- K_d es el factor corrector de esta distancia [Chiron 08]:

$$K_d = 1 + 0.03344 \cos(J' - 2.80^\circ) \quad (6.2)$$

- J' es el *ángulo del día* (ver subsección 4.2.4).
- δ es la *declinación del día* (ver subsección 4.2.4).
- el ángulo Λ es la latitud de un punto P .
- ω es el *ángulo horario* (ver sección 4.4).

6.4.2. Irradiancia directa total

Irradiancia directa total normal

Los cinco procesos que causan la atenuación de la radiación solar directa en la atmósfera se agrupan en [Chiron 08]:

- La difusión, dispersión o “scattering” molecular de Rayleigh.
- La difusión, la refracción y la absorción continua por los aerosoles.
- La absorción selectiva por el ozono.
- La absorción selectiva por el vapor de agua.
- La absorción selectiva por otros gases (O_2 , CO_2 , ...).

En los trabajos de investigación desarrollados en esta tesis se considera uno de los modelos más extendidos, el elaborado por Bird y Hulström [Bird 81] para cielo claro, en el que se admite que la *irradiancia directa total normal*, I_n , es proporcional al producto de cinco transmitancias atmosféricas específicas *totales*, que corresponden a los cinco procesos de extinción respectivos [Chiron 08]:

$$I_n = 0.9751 K_d I_o T_R T_A T_O T_W T_G \quad (6.3)$$

Las transmitancias específicas totales se calculan a partir de las siguientes expresiones [Bird 81, Chiron 08]:

- Transmitancia Rayleigh

La transmitancia atmosférica específica del proceso Rayleigh de atenuación se expresa como:

$$T_R = \exp \left[-0.0903 (M')^{0.84} \left(1 + M' - (M')^{1.01} \right) \right] \quad (6.4)$$

donde M' viene definida por la ecuación (6.12).

- Transmitancia de los aerosoles

La transmitancia atmosférica de los aerosoles se toma como:

$$T_A = \exp \left[-K_A^{0.873} \left(1 + K_A - K_A^{0.7088} \right) M^{0.9108} \right] \quad (6.5)$$

con

$$K_A = 0.2758 K_{A\lambda_1} + 0.35 K_{A\lambda_2}$$

y

$$K_{A\lambda} = \beta_A \lambda^{-\alpha_A}$$

para

$$\lambda_1 = 0.38 \mu m \quad y \quad \lambda_2 = 0.50 \mu m$$

β_A es un parámetro que varia en función de la cantidad de aerosoles en la atmósfera y se le denomina, cuando λ se expresa en μm , *coeficiente de turbiedad de Angström*. En la Tabla 6.3 se pueden observar algunos de sus valores típicos. En los trabajos de investigación de esta tesis, dicho coeficiente se calcula dando distinta prioridad a los diferentes tipos de nubes, un 10 % a las de nivel alto, un 35 % a las de nivel medio, y un 55 % a las de nivel bajo. Para su cálculo se utiliza la siguiente expresión:

$$\beta_A = 0.25 (0.1 CFRACH_24 + 0.35 CFRACM_24 + 0.55 CFRACL_24) + 0.05$$

donde:

Tabla 6.3: Valores típicos de β_A [Chiron 08].

Atmósfera	Visibilidad (km)	β_A
Turbiedad baja (lugar rural)	100	0.05
Turbiedad media (lugar urbano)	25	0.10
Turbiedad alta (lugar industrial)	5	0.30

- CFRACH_24 es la cobertura de nubes a nivel alto (entre 0 y 1) media horaria.
- CFRACM_24 es la cobertura de nubes a nivel medio (entre 0 y 1) media horaria.
- CFRACL_24 es la cobertura de nubes a nivel bajo (entre 0 y 1) media horaria.

Las predicciones para estas variables meteorológicas se obtienen a partir del modelo numérico del tiempo WRF-NMM, y se utilizan las de horizontes entre 0 y 24 horas. El parámetro α_A , se le denomina *exponente de Angström*, depende del valor medio de tamaño de los aerosoles (es próximo a cero cuando las partículas son grandes, y se acerca a cuatro cuando son pequeñas). En los trabajos de investigación de esta tesis se considera $\alpha_A = 2$.

■ Transmitancia del ozono estratosférico

El ozono *estratosférico* natural (no hay que confundirle con el ozono *troposférico*) se encuentra situado en la capa de atmósfera entre 10 y 30 *km* de altitud. Su *espesor reducido* l_O (el espesor que presentaría este gas, si todo él se encontrase en las condiciones normales de temperatura y presión) está comprendido entre 2 y 4.5 *mm*. La transmitancia del ozono se muestra como:

$$T_O = 1 - 0.1611 X_O (1 + 139.48 X_O)^{-0.3035} - 0.002715 X_O (1 + 0.044 X_O + 0.0003 X_O^2)^{-1} \quad (6.6)$$

con $X_O = l_O M$ y M definida en la ecuación (6.11).

Las variaciones de l_O son a la vez temporales y geográficas. En los trabajos de investigación de esta tesis, para calcular el valor de l_O se utiliza una ecuación descrita en [Karavana 13] para el área de Europa, basada en un modelo matemático definido en [Heuklon 79]:

$$l_O = \frac{235 + [150 + 40 \sin(0.9865 (J - 30)) + 20 \sin(3 (L + 20))] \sin^2(1.28 \Lambda)}{100} \quad (6.7)$$

donde:

- J es el *número del día del año* (1 para el 1 de enero y 365 para el 31 de diciembre o 366 si el año es bisiesto).
- Λ es la latitud (en grados).

- L es la longitud (en grados).
- Transmitancia del vapor de agua (gas)

De la masa total de agua en forma gaseosa en la atmósfera, aproximadamente la mitad se encuentra por debajo de los 2000 m . Para cuantificar esta masa de agua se utiliza la *altura de agua condensable o precipitable*, w , que es el espesor total que habría a nivel de suelo si ésta condensase en forma líquida (entre 0.5 y 5 cm). La fórmula de Leckner relaciona w con la presión parcial de vapor y con la temperatura absoluta T del aire a nivel de suelo ($^{\circ}K$). En los trabajos de investigación de esta tesis w se calcula de la siguiente forma:

$$w = 0.493 \frac{(PSHLTR_24/100)}{TH10_24} \quad (6.8)$$

donde:

- PSHLTR_24 es la presión atmosférica media horaria a 2 m de altura (Pa).
- TH10_24 es la temperatura media horaria a 10 m ($^{\circ}K$).

Las predicciones de estas variables meteorológicas se obtienen con el modelo mesoescalar WRF-NMM, y se utilizan las de horizontes de predicción entre 0 y 24 horas. La transmitancia del vapor de agua (gas) se determina por medio de la expresión siguiente:

$$T_W = 1 - 2.4959 w M \left[(1 + 79.034 w M)^{0.6828} + 6.385 w M \right]^{-1} \quad (6.9)$$

donde M viene definida por la ecuación (6.11).

- Transmitancia de otros gases

Los otros gases activos en la absorción selectiva son esencialmente el oxígeno O_2 y el dióxido de carbono CO_2 . Se utiliza la relación siguiente:

$$T_G = \exp \left[-0.0127 (M')^{0.26} \right] \quad (6.10)$$

donde M' viene definida por la ecuación (6.12).

La *masa de aire*, M , se define como la relación entre el camino que recorren los rayos directos del Sol a través de la atmósfera para llegar a la superficie receptora y el que recorrerían en caso de incidencia normal y superficie al nivel del mar [Lorenzo 06]. Se presenta como [Bird 81]:

$$M = \left[\cos \theta + 0.15 (93.885 - \theta)^{-1.25} \right]^{-1} \quad (6.11)$$

donde θ es el ángulo zenital (solar), y se denota

$$M' = \frac{M (PSHLTR_24/100)}{1013} \quad (6.12)$$

Irradiancia directa total horizontal

La *irradiancia directa total horizontal* se expresa como [Chiron 08]:

$$I_h = I_n \sin \gamma = I_n \cos \theta \quad (6.13)$$

donde γ es el ángulo de la altura o elevación del Sol ($\gamma = \pi/2 - \theta$ es el complementario del ángulo zenital solar θ comprendido entre 0 y $\pi/2$).

Irradiancia directa total sobre el plano de inclinación β

La *irradiancia directa total* sobre un plano de inclinación β se expresa como [Chiron 08]:

$$I_\beta = I_n \cos i \quad (6.14)$$

donde i es el *ángulo de incidencia* que forman los *rayos solares directos* con el vector unitario normal \vec{n} que parte del plano receptor *por el lado iluminado*. En los trabajos de investigación de esta tesis, para calcular el *ángulo de incidencia* i , se tiene en cuenta que las placas solares en la planta fotovoltaica de estudio están orientadas al Sur. Es decir, el vector \vec{n}_h que es la proyección de \vec{n} sobre el plano horizontal (Pxy) apunta al Sur.

6.4.3. Irradiancia global y difusa en la superficie horizontal

En [Badescu 12] se revisan 54 modelos para el cálculo de la irradiancia horaria global y difusa en la superficie horizontal, para cielo claro. En la Tabla 6.4 se describe el significado de los símbolos que se utilizan para las entradas de estos modelos, varios de ellos se referencian en la Tabla 6.5.

En los trabajos de investigación de esta tesis la *irradiancia difusa* observada sobre un plano horizontal, D_h , se obtiene como en [Notton 06]:

$$D_h = G_h - I_n \sin \gamma \quad (6.15)$$

donde G_h es la *irradiancia global total* sobre un plano horizontal y γ es el ángulo de la elevación del Sol. Para poder realizar el cálculo de la irradiancia difusa media horaria D_h sobre un plano horizontal, a G_h se le asigna el valor medio horario de la variable meteorológica de radiación de onda corta de día claro (W/m^2), RSWINC_24, pronosticada con el modelo mesoescalar WRF-NMM para horizontes de predicción entre 0 y 24 horas.

6.4.4. Irradiancia difusa sobre un plano de inclinación β

La *irradiancia difusa total* procedente del cielo observada sobre un plano con un ángulo de inclinación β , se designa por D_β , y se calcula teniendo en cuenta distintos modelos geométricos, algunos se muestran en las Tablas 6.6 y 6.7. A continuación se resumen brevemente algunas características de varios de estos modelos [Notton 06, Serrano 10]:

Tabla 6.4: Significado de los símbolos de las entradas en los modelos de irradiancia horaria para cielo claro de la Tabla 6.5 (adaptado de [Badescu 12]).

Símbolo	Significado
θ	Ángulo Zenital (grados)
M	Masa de aire
ρ_s	Albedo del suelo
P	Presión del aire en la superficie (hPa)
T	Temperatura del aire, de bulbo seco ($^{\circ}K$)
Q	Humedad relativa en el aire (%)
V	Velocidad de viento (m/s)
vi	Visibilidad (km)
w	Agua precipitable (cm)
l_o	Longitud vertical de ozono reducido ($atm - cm$)
u_N	Longitud vertical del NO_2 total ($atm - cm$)
α_A	Exponente de la longitud de onda de Angström
β_A	Turbidez de Angström
τ_a	Coefficiente de turbidez de banda ancha de Unsworth-Monteith
T_L	Turbidez de Linke
ϖ	Albedo de dispersión simple del aerosol

Modelos isótropos y modelos pseudo-isótropos

- En el modelo [Liu 62] se supone que la irradiancia difusa es uniforme sobre la bóveda celeste y se aproxima a las condiciones del cielo cubierto. Este modelo subestima la componente difusa sobre un plano inclinado, pero a menudo se utiliza en cálculos de ingeniería porque es fácil de implementar.
- En [Hamilton 85] se indica que la hipótesis de un cielo isótropo para la irradiancia difusa es incorrecta, ya que el cielo del sur parece ser responsable del 63 % de la irradiancia solar difusa total procedente del cielo.
- En [Jimenez 86] se supone que la irradiancia difusa sobre una superficie horizontal es del 20 % de la irradiancia global.
- En [Koronakis 86] se modifica la hipótesis de un cielo difuso isótropo.
- En [Badescu 02] se cambian algunos coeficientes del modelo isótropo y se obtienen resultados aceptables para distintas orientaciones simultáneamente.

Modelos anisótropos

- En el modelo [Bugler 77] se añade un término correctivo al modelo isótropo para tener en cuenta la irradiancia difusa proveniente del área próxima al sol y del resto del cielo según la elevación solar.

Tabla 6.5: *Entradas en los modelos de la irradiancia horaria para cielo claro (adaptado de [Badescu 12]).*

Modelo	θ	M	ρ_s	P	T	Q	V	vi	w	l_o	u_N	α_A	β_A	τ_a	T_L	ϖ
[Sharma 65]	•															
[Dogniaux 76]	•	•		•												•
[Hottel 76]	•															
[Paltridge 76]	•															
[Schulze 76]	•															
[Suckling 77]	•	•	•	•					•							
[Chandra 78]	•															•
[Rodgers 78]	•			•					•					•		
[Watt 78]	•	•	•	•					•	•		•	•			
[Atwater 79]	•		•	•					•					•		
[Barbaro 79]	•	•							•							
[Biga 79]	•															
[Davies 79]	•		•	•	•				•							
[King 79]	•	•	•	•					•				•			
[Bird 80]	•	•	•	•					•	•		•	•			
[Ideriah 81]	•	•		•					•			•	•			
[Perrin 82]	•	•		•					•	•			•			
[Iqbal 83]	•	•	•	•	•				•	•		•	•			
[Kasten 83]	•	•		•												
[Carroll 85]	•	•	•	•					•				•			
[Josefsson 85]	•	•	•	•					•							
[Machler 85]	•			•				•	•							
[Gueymard 89]	•		•	•					•	•		•	•			•
[Bashahu 94]	•	•		•					•				•			
[Nijegorodov 97]	•	•		•	•				•	•		•	•			
[Maxwell 98]	•	•	•	•					•	•				•		
[Muneer 98]	•	•		•					•	•						
[Psiloglou 00]	•	•	•	•					•	•			•			
[Yang 01]	•	•		•					•	•			•			
[Paulescu 03]	•	•		•					•	•			•			
[Gueymard 04]	•		•	•					•				•			
[Ashrae 05]	•															
[Krarti 06]	•				•	•	•									
[Zhang 06]	•				•	•										
[Belcher 07]	•			•		•		•	•							
[Badescu 08]	•	•		•					•	•						
[Gueymard 08]	•		•	•					•	•	•	•	•			•
[Ineichen 08]	•			•					•					•		
[Kambezidis 08]	•	•	•	•					•	•			•			
[Esra 10]	•	•		•											•	
[Janjai 10]	•	•		•					•	•		•	•			

Tabla 6.6: Modelos isótopos y modelos pseudo-isótopos [Notton 06].

Modelo	Irradiancia difusa sobre un plano de inclinación β
[Liu 62]	$D_\beta = (1/2) D_h [1 + \cos \beta]$
[Hamilton 85]	$D_\beta = (1/3) D_h [2 + \cos \beta]$
[Jimenez 86]	$D_\beta = (1/2) 0.2 G_h [1 + \cos \beta]$
[Koronakis 86]	$D_\beta = (1/3) D_h [2 + \cos \beta]$
[Badescu 02]	$D_\beta = (1/4) D_h [3 + \cos (2\beta)]$

- En el modelo [Temps 77] se introducen dos factores en la ecuación propuesta en [Liu 62], que simulan la anisotropía del cielo en condiciones de cielo claro. El primero tiene en cuenta la irradiancia difusa que proviene de la proximidad del sol, y el segundo toma en cuenta el brillo del cielo cerca del horizonte.
- En el modelo [Hay 85] se tiene en cuenta, la anisotropía de la radiancia del cielo claro, y se propone considerar la irradiancia difusa horizontal D_h constituida por dos partes distintas [Chiron 08]:

- D_{hc} debida a un flujo *direccional* proveniente de la zona circumsolar

$$D_{hc} = D_h F_{Hay} \quad (6.16)$$

- D_{hv} debida a un flujo *hemisférico* e *isótopo* correspondiente al resto de la bóveda del cielo

$$D_{hv} = D_h (1 - F_{Hay}) \quad (6.17)$$

donde F_{Hay} es el índice de anisotropía de Hay. Se define como

$$F_{Hay} = \frac{I_n}{K_d I_o} \quad (6.18)$$

La irradiancia horizontal D_{hc} es producida por un flujo direccional de la misma dirección que el sol, y su transposición sobre un plano *normal* a los rayos directos se obtiene dividiendo por $\sin \gamma$; además, por la relación (6.14) se multiplica por $\cos i$ para tener su valor sobre un plano con un ángulo de inclinación β que recibe los rayos directos bajo el ángulo de incidencia i . Así, se puede calcular la irradiancia difusa total observada sobre un plano de inclinación β

$$D_\beta = \frac{\cos i}{\sin \gamma} D_{hc} + \frac{1}{2} (1 + \cos \beta) D_{hv} \quad (6.19)$$

- En el modelo [Klucher 79] se módica el modelo [Temps 77] para simular las condiciones presentes durante los cielos parcialmente nublados incorporando una función de modulación para determinar la magnitud de los efectos de cielo claro.

- El modelo circunsolar [Iqbal 83] se aplica a condiciones de cielo despejado, toda la irradiancia solar se supone que procede de la dirección del sol, y la componente difusa se trata como irradiancia de haz. Con este modelo se sobrestima generalmente la irradiancia difusa del cielo y se utiliza raramente.
- En el modelo [Ma 83] se propone un modelo donde la irradiancia difusa viene separada en dos términos, uno procedente de la zona circunsolar y el otro del resto de la bóveda del cielo, se utiliza el índice de claridad del cielo horario M_T .
- En los modelos [Gueymard 84, Gueymard 86] se considera que la irradiancia para cielos parcialmente nublados es una combinación lineal de valores para cielos cubiertos R_{d1} y para cielos despejados R_{d0} . Se utiliza un término de ponderación de la nubosidad N_G . La irradiancia para cielo claro R_{d0} se calcula como la suma de la componente circunsolar y de unos términos hemisféricos. La irradiancia para un cielo cubierto R_{d1} sólo depende del ángulo de inclinación del plano y de un factor de corrección comprendido entre los valores 1 y 2.
- En el modelo propuesto en [Perez 86, Perez 87, Perez 89] se considera tres zonas del cielo como fuentes diferentes de irradiancia difusa: una región circunsolar con una cierta extensión angular, una banda horizontal con otra extensión angular y el resto del hemisferio celeste. La irradiancia dentro de cada zona es constante e igual a k_3L , k_4L y L , respectivamente [Lorenzo 06].
- En el modelo [Skartveit 86] se señala que para los cielos cubiertos, una parte importante de la irradiancia difusa proviene del zenit. Se introduce un factor de corrección Z (es una función lineal del índice de anisotropía de Hay, F_{Hay}). En la fórmula $S(\omega_i, \Omega_i)$ es la porción de ángulo sólido con obstáculos en el horizonte real.
- En el modelo [Muneer 87, Muneer 90] se distingue la irradiancia difusa sobre la superficie inclinada para las superficies iluminadas por el sol bajo cielo nublado y bajo cielo no nublado. Se utiliza un índice b de distribución de la irradiancia introducido en [Moon 42] para modelar la distribución de luminancia de un cielo cubierto. A nivel mundial, en [Muneer 97] se sugiere un valor promedio de $b = 2.5$ para las superficies iluminadas por el sol bajo un cielo nublado.
- En el modelo [Reindl 90] se agrega un horizonte que pule el término difuso del modelo de Hay. La magnitud del brillo del horizonte se controla por una función de modulación f que se multiplica por el término de corrección del brillo del horizonte que se utiliza en el modelo [Temps 77].

En los trabajos de investigación de esta tesis para estimar la irradiancia difusa D_β sobre un plano de inclinación β y orientación Sur, se hacen pruebas con varios modelos geométricos, tanto isótropos como anisótropos. Tras analizar los resultados de las pruebas, se selecciona uno de los modelos que mejor estima la irradiancia difusa [Notton 06], el elaborado por [Hay 85], que también está descrito en [Chiron 08].

Tabla 6.7: Modelos anisótrpos [Notton 06].

Modelo	Irradiancia difusa sobre un plano de inclinación β
[Bugler 77]	$D_\beta = \left[\left(D_h - 0.05 \frac{I_\beta}{\cos \theta} \right) \frac{1+\cos \beta}{2} \right] + 0.05 I_\beta \cos i$
[Temps 77]	$D_\beta = \frac{1}{2} D_h (1 + \cos \beta) \left[1 + \sin^3 \left(\frac{\beta}{2} \right) \right] (1 + \cos^2 i \sin^3 \theta)$
[Hay 85]	$D_\beta = D_h \left[\frac{\cos i}{\sin \gamma} F_{Hay} + \frac{1}{2} (1 - F_{Hay}) (1 + \cos \beta) \right]$
[Klucher 79]	$D_\beta = D_h \left[0.5 \left(1 + \cos \left(\frac{\beta}{2} \right) \right) \right] \left[1 + F \sin^3 \left(\frac{\beta}{2} \right) \right] \left[1 + F \cos^2 i \sin^3 \theta \right]$ con $F = 1 - \left(\frac{D_h}{G_h} \right)^2$
[Iqbal 83]	$D_\beta = D_h \frac{\cos i}{\cos \theta}$
[Ma 83]	$D_\beta = D_h \left[M_T \frac{\cos i}{\cos \theta} + (1 - M_T) \cos^2 \left(\frac{\beta}{2} \right) \right]$
[Gueymard 84]	$D_\beta = D_h [(1 - N_G) R_{d0} + N_G R_{d1}]$
[Perez 86]	$D_\beta = D_h [(1 + \cos \beta) (1 - k_3) / 2 + k_3 (\cos i / \cos \theta) + k_4 \sin \beta]$
[Skartveit 86]	$D_\beta = D_h \left[F_{Hay} \frac{\cos i}{\cos \theta} + Z \cos \beta + (1 - F_{Hay} - Z) \left(\frac{1+\cos \beta}{2} \right) - S(\omega_i, \Omega_i) \right]$
[Muneer 87]	$D_\beta = D_h \left[\cos^2 (\beta/2) + \frac{2b}{\pi(3+2b)} (\sin \beta - \beta \cos \beta - \pi \sin^2 (\beta/2)) \right]$ $= D_h [T] \text{ (Cielo nublado)}$ $D_\beta = D_h \left[T (1 - F_{Hay}) + F_{Hay} \left(\frac{\cos i}{\cos \theta} \right) \right] \text{ (Cielo no nublado)}$
[Reindl 90]	$D_\beta = D_h \left[(1 - F_{Hay}) \left(\frac{1+\cos \beta}{2} \right) \left(1 + f \sin^3 \left(\frac{\beta}{2} \right) \right) + F_{Hay} \frac{\cos i}{\cos \theta} \right]$

6.4.5. Irradiancia del albedo

El albedo de la superficie terrestre es el porcentaje de radiación que la Tierra devuelve por reflexión [Ledesma 11]. A excepción de los ambientes con mucha nieve, la reflectividad de la mayoría de los suelos es más bien baja, por lo que la contribución del albedo a la radiación que incide sobre los receptores fotovoltaicos suele ser muy pequeña [Lorenzo 06]. Si se supone que el suelo es horizontal, infinito e isotrópico, la *irradiancia del albedo* sobre una superficie inclinada viene dada por [Lorenzo 06, Chiron 08]:

$$R_\beta = \frac{1}{2} (1 - \cos \beta) \rho_s G_h \quad (6.20)$$

donde ρ_s es la reflectividad del suelo. Varía de unos lugares a otros pues depende de las características de la superficie [Ledesma 11]. En la Tabla 6.8 se especifica el albedo ρ_s de algunas superficies. Cuando no se conoce ésta, es habitual suponer que $\rho_s = 0.2$ [Lorenzo 06], como se ha hecho en los trabajos de investigación de esta tesis.

6.4.6. Irradiancia global total sobre un plano de inclinación β

La mayoría de las veces la irradiancia global total es medida sobre un plano horizontal. Para conocer el valor sobre un plano con un ángulo de inclinación β (situado en la misma ubicación), se utiliza una relación de *transposición* de la irradiancia, capaz de expresar D_β y R_β en función de D_h y G_h . La *irradiancia global total* para cielo claro en las zonas sin relieve se puede expresar sobre un plano de inclinación β , que recibe los rayos directos

Tabla 6.8: *Albedo medio de algunas superficies (adaptado de [Ledesma 11]).*

Superficies	Albedo %
Tierra amarilla silícea	35
Desierto	20–28
Suelo arcilloso	20
Arena seca	18
Arena húmeda	9
Humus, tierra vegetal negra	8–14
Suelo sin vegetación	7–20
Cubierto con hierba verde	20–33
Barbecho seco	8–12
Barbecho húmedo	5–7
Copas de pinos y robles	18
Media en trigo y centeno	10–25
Avena y guisantes	20
Nieve limpia, caída recientemente	90
Nieve sucia con polvo o barro	45

del sol bajo un ángulo de incidencia i , como [Chiron 08]:

$$G_{\beta} = I_{\beta} + D_{\beta} + R_{\beta} \quad (6.21)$$

6.5. Efectos del ángulo de incidencia y del polvo

La eficiencia de los módulos fotovoltaicos se ve afectada por la posición relativa de su superficie con respecto al Sol [Lorenzo 06]. El efecto del ángulo de incidencia y del polvo en la respuesta eléctrica de un módulo fotovoltaico se puede estimar aplicando las ecuaciones descritas en [Martín 01, Lorenzo 06] a las diferentes componentes de la irradiancia:

- A la componente directa se le aplica

$$FT_I(i) = 1 - \left[\frac{\exp\left(-\frac{\cos i}{a_r}\right) - \exp\left(-\frac{1}{a_r}\right)}{1 - \exp\left(-\frac{1}{a_r}\right)} \right] \quad (6.22)$$

- A la componente difusa se le aplica

$$\begin{aligned}
 FT_D(\beta) = 1 - \exp \left[-\frac{1}{a_r} \left[c_1 \left(\sin \beta + \frac{\pi - \frac{\beta\pi}{180} - \sin \beta}{1 + \cos \beta} \right) \right. \right. \\
 \left. \left. + c_2 \left(\sin \beta + \frac{\pi - \frac{\beta\pi}{180} - \sin \beta}{1 + \cos \beta} \right)^2 \right] \right] \quad (6.23)
 \end{aligned}$$

Tabla 6.9: Valores recomendados para los parámetros implicados en el cálculo de las pérdidas angulares [Lorenzo 06].

Grado de suciedad	$T_{sucio}(0)/T_{limpio}(0)$	a_r	c_2
Limpio	1	0.17	-0.069
Bajo	0.98	0.20	-0.054
Medio	0.97	0.21	-0.049
Alto	0.92	0.27	-0.023

- A la componente del albedo se le aplica

$$FT_R(\beta) = 1 - \exp \left[-\frac{1}{a_r} \left[c_1 \left(\sin \beta + \frac{\frac{\beta\pi}{180} - \sin \beta}{1 - \cos \beta} \right) + c_2 \left(\sin \beta + \frac{\frac{\beta\pi}{180} - \sin \beta}{1 - \cos \beta} \right)^2 \right] \right] \quad (6.24)$$

donde $c_1 = 4/(3\pi)$. En la Tabla 6.9 se incluyen los valores recomendados para los parámetros que se utilizan en el cálculo de las pérdidas angulares. En los trabajos de investigación de esta tesis se considera el grado de suciedad alto, $T_{sucio}(0)/T_{limpio}(0) = 0.92$, $a_r = 0.27$ y $c_2 = -0.023$.

6.5.1. Irradiancias “efectivas”

Las irradiancias “efectivas” que alcanzan a las células solares de un módulo fotovoltaico se calculan basándose en las fórmulas descritas en [Lorenzo 06]:

Irradiancia directa total “efectiva”

La irradiancia directa total “efectiva” está dada por

$$I_{ef,\beta} = I_\beta \frac{T_{sucio}(0)}{T_{limpio}(0)} FT_I(i) \quad (6.25)$$

donde I_β es la irradiancia directa total sobre un plano de inclinación β (en los trabajos de investigación de esta tesis se calcula mediante la ecuación (6.14)).

Irradiancia difusa “efectiva”

La irradiancia difusa “efectiva” está dada por

$$D_{ef,\beta} = D_\beta \frac{T_{sucio}(0)}{T_{limpio}(0)} FT_D(i) \quad (6.26)$$

donde D_β es la irradiancia difusa sobre un plano de inclinación β (en los trabajos de investigación de esta tesis se calcula mediante la ecuación (6.19)).

Irradiancia del albedo “efectiva”

La irradiancia del albedo “efectiva” está dada por

$$R_{ef,\beta} = R_{\beta} \frac{T_{sucio}(0)}{T_{limpio}(0)} FT_{AL}(i) \quad (6.27)$$

donde R_{β} es la irradiancia del albedo sobre un plano de inclinación β (en los trabajos de investigación de esta tesis se calcula mediante la ecuación (6.20)).

Irradiancia global total “efectiva”

El valor medio horario de la irradiancia global total “efectiva” sobre un plano de inclinación β y orientación Sur, para las condiciones ideales de un cielo *claro*, en los trabajos de investigación de esta tesis se calcula como la suma de:

- La irradiancia directa total “efectiva” (ecuación 6.25).
- La irradiancia difusa “efectiva” (ecuación 6.26).
- La irradiancia del albedo “efectiva” (ecuación 6.27).

Se expresa como:

$$G_{ef,\beta} = I_{ef,\beta} + D_{ef,\beta} + R_{ef,\beta} \quad (6.28)$$

Parte II

TÉCNICAS DE MINERÍA DE DATOS Y PROBABILIDAD

Capítulos:

- 7. Árboles de decisión y sistemas de reglas
- 8. Redes neuronales artificiales
- 9. Máquinas de vectores soporte
- 10. Sistemas basados en lógica borrosa
- 11. Algoritmos evolutivos
- 12. Probabilidad

Capítulo 7

Árboles de decisión y sistemas de reglas

Un árbol de decisión es un conjunto de condiciones organizadas en una estructura jerárquica, de tal manera que la decisión final a tomar se puede determinar siguiendo las condiciones que se cumplen desde la raíz del árbol hasta alguna de sus hojas [Hernández 04]. Obviamente las condiciones se desarrollan en función de los valores de las variables o atributos disponibles.

Los sistemas de reglas son una generalización de los árboles de decisión, pero no se exige exclusión ni exhaustividad en las condiciones de las reglas (puede haber cualquier número de reglas).

La diferencia más importante entre los sistemas de aprendizaje de árboles de decisión y los sistemas de inducción de reglas proposicionales es el tipo de algoritmo que utilizan, que puede ser el de partición o el de cobertura [Hernández 04].

7.1. Sistemas por partición: árboles de decisión

Un árbol de decisión conducirá un ejemplo hasta una y sólo una hoja. El espacio de instancias se va partiendo de arriba abajo, utilizando cada vez una partición disjunta, es decir, cada instancia cumple o no cumple una condición. Estos algoritmos se llaman algoritmos de “divide y vencerás” [Hernández 04].

Un aspecto importante en los sistemas de aprendizaje de árboles de decisión es el *criterio de partición*, ya que una mala elección de la partición generará un árbol peor [Hernández 04]. Existen distintos algoritmos de “partición”: CART [Breiman 84], ID3 [Quinlan 83, Quinlan 86], C4.5 [Quinlan 93], y ASSISTANT [Cestnik 87].

7.1.1. Particiones posibles

Las particiones son un conjunto de condiciones exhaustivas y excluyentes. La mayoría de algoritmos de aprendizaje de árboles de decisión sólo permiten un juego muy limitado de particiones [Hernández 04].

Tabla 7.1: *Criterios de partición [Hernández 04].*

Criterio	$f(p^1, p^2, \dots, p^c)$	Referencias
Error Esperado	$\min(p^1, p^2, \dots, p^c)$	[Breiman 84]
GINI (CART)	$1 - \sum (p_i)^2$	[Breiman 84]
Entropía (gain)	$\sum p_i \log(p_i)$	[Quinlan 93]
DKM	$2 (\prod p_i)^{1/2}$	[Kearns 96]

Si un atributo x_i es numérico y continuo, puede tomar muchos valores diferentes en los ejemplos. Se intentan obtener particiones que separen los ejemplos en intervalos. Las particiones numéricas admitidas son de la forma $(x_i \leq a, x_i > a)$ donde a es una constante numérica elegida entre un conjunto finito de constantes que discriminen los ejemplos vistos. Estas particiones sólo afectan a un atributo de un ejemplo a la vez, es decir, ni relacionan dos atributos del mismo ejemplo, ni dos atributos de distintos ejemplos [Hernández 04].

7.1.2. Criterio de selección de particiones

Si existen n atributos y v valores posibles para cada atributo, el número de particiones que se pueden obtener es de $n \cdot v$. Una vez elegida la partición se continúa hacia abajo la construcción del árbol y no se vuelven a plantear las particiones ya construidas [Hernández 04]. Existen numerosos criterios de partición (ver Tabla 7.1), éstos buscan la partición $part$ con el menor $I(part)$ [Hernández 04]:

$$I(part) = \sum_{j=1}^h p_j f(p_j^1, p_j^2, \dots, p_j^c) \quad (7.1)$$

donde:

- h es el número de nodos hijos de la partición.
- p_j es la probabilidad de “caer” en el nodo j .
- p_j^1 es la proporción de elementos de tipo 1 en el nodo j .
- p_j^2 es la proporción de elementos de tipo 2 en el nodo j .
- p_j^c es la proporción de elementos de tipo c en el nodo j .

7.2. Sistemas de aprendizaje de reglas por cobertura

Los métodos que generan estos sistemas van añadiendo reglas, una detrás de otra. Pero a diferencia de los árboles de decisión, no se sigue con las condiciones complementarias a la utilizada en la regla anterior (exclusión y exhaustividad), sino que se descartan los

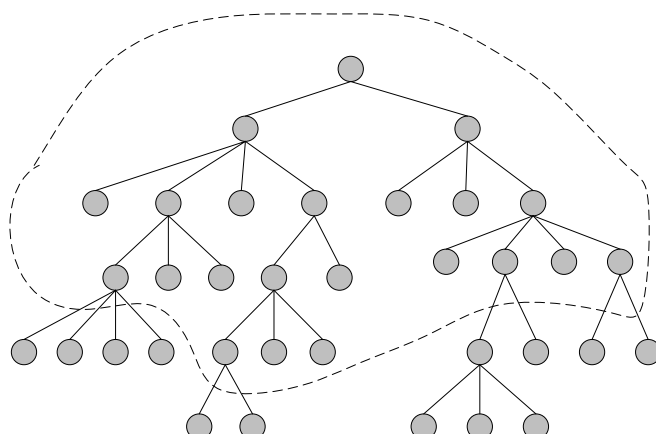


Figura 7.1: *Ejemplo de poda. Algunos nodos inferiores son eliminados [Hernández 04].*

ejemplos ya cubiertos por las reglas ya obtenidas, y con los ejemplos que quedan se empieza de nuevo. Esto hace que puedan aparecer nuevas condiciones que se solapen (o no) con las anteriores [Hernández 04]. Algunos algoritmos de “cobertura” son: AQ [Michalski 83, Michalski 86], CN2 [Clark 89, Clark 91], FOIL [Quinlan 90, Quinlan 93] y el FFOIL [Quinlan 96].

7.3. Poda y reestructuración

Intentar aproximar demasiado hace que el modelo sea demasiado específico y poco general con otros datos no vistos. Para limitar este problema se modifican los algoritmos de aprendizaje eliminando condiciones de las ramas del árbol o de algunas reglas [Hernández 04]. En el caso de los árboles de decisión se puede ver gráficamente en la Figura 7.1.

Se distingue entre métodos de prepoda y pospoda [Hernández 04]:

- **Prepoda:** el proceso se realiza durante la construcción del árbol o el conjunto de reglas. Se trata de determinar el criterio de parada a la hora de seguir especificando una rama o una regla. Los criterios de prepoda pueden estar basados en el número de ejemplos por nodo.
- **Pospoda:** el proceso se realiza después de la construcción del árbol o el conjunto de reglas. En el caso de los árboles de decisión se trata de eliminar nodos de abajo a arriba hasta un cierto límite, y en el caso de los sistemas de reglas de eliminar condiciones con el objetivo de hacer las reglas más generales.

Existen numerosos métodos para podar árboles de decisión [Esposito 97] o para podar reglas de decisión [Cohen 93]; la mayoría de éstos tienen uno o más parámetros que permiten decidir el grado de poda. Además, existen problemas con características especiales que hacen necesario suavizar la poda [Hernández 04].

7.4. Algoritmo M5P

Los árboles modelo [Quinlan 92] tienen una estructura de árbol de decisión convencional, pero utilizan funciones lineales en las hojas [Holmes 99]. La primera aplicación de los árboles modelo, M5, fue más bien abstracta [Quinlan 92], y la idea se reconstruye y mejora en un sistema llamado M5' (M5P) [Wang 97]. Al igual que el aprendizaje del árbol de decisión convencional, M5P construye un árbol mediante el fraccionamiento de los datos en función de los valores de los atributos [Holmes 99].

Construcción del árbol inicial

El criterio de reparto se basa en la desviación estándar de los valores que llegan a un nodo R como una medida del error en un nodo, y se calcula la reducción esperada en el error como un resultado de prueba con cada atributo de ese nodo. Se elige el atributo at que maximiza la reducción del error esperado. La reducción de la desviación estándar (SDR) se calcula mediante la fórmula [Wang 97]:

$$SDR = sd(I) - \sum_i \frac{|I_i|}{|I|} sd(I_i) \quad (7.2)$$

donde:

- I es el conjunto de ejemplos que llegan al nodo.
- $|I|$ es el número de instancias de entrenamiento que llegan al nodo.
- I_1, I_2, \dots son los subconjuntos que resultan de la división del nodo de acuerdo con el atributo seleccionado.
- $|I_i|$ es el número de instancias en el subconjunto I_i .
- La función sd calcula la desviación estándar de los valores de la clase de un conjunto de ejemplos que llegan a un nodo.

La división en M5P cesa cuando los valores de la clase de todas las instancias que llegan a un nodo varían poco, o solo quedan unas pocas instancias [Wang 97].

Simplificación de modelos lineales

Los modelos lineales para un nodo R particular del árbol con k atributos $\{at_1, \dots, at_k\}$, que se dividen por debajo de ese nodo, se calcula como una combinación lineal de esos atributos y se incluye un término constante [Wang 97]:

$$w_0 + w_1 at_1 + w_2 at_2 + \dots + w_k at_k \quad (7.3)$$

Para decidir si quitar o no un atributo, se debe calcular el error esperado antes y después. La fórmula de la estimación del error en posteriores instancias calcula la media del error residual, $e(I, mode)$, producido al clasificar con el modelo creado $mode$ cada instancia del conjunto I [Riaño 16]:

$$e(I, mode) = \frac{1}{|I|} \sum_{i \in I} \|c_i - c(mode, i)\| \quad (7.4)$$

donde:

- $|I|$ es el número de instancias de entrenamiento que llegan al nodo.
- c_i es el valor numérico asociado a la instancia $i \in I$.
- $c(mode, i)$ es el valor numérico que establece el modelo $mode$ de la instancia $i \in I$.

Como esto subestima el error esperado para casos no vistos, para compensarlo se multiplica por el factor $(|I| + \nu)/(|I| - \nu)$ donde $|I|$ es el número de ejemplos de entrenamiento que llegan al nodo y ν es el número de atributos en el modelo que representa el valor de la clase en ese nodo [Wang 97].

Primero se simplifican los modelos lineales de los nodos sucesores de un determinado nodo. Después se realiza una búsqueda en escalada de forma que en cada iteración elimina aquel atributo tal que, si se elimina, el modelo lineal restante minimiza el error [Riaño 16].

Poda del árbol

A la hora de tomar la decisión de si se debe o no podar un subárbol, dado que cada nodo interno del árbol tiene un modelo simplificado lineal y un modelo descrito por su subárbol, es necesario calcular sus errores. Para calcular el error del subárbol, el error esperado de cada rama se combina en un solo valor global para ese nodo y se pondera cada rama mediante la proporción de los ejemplos de entrenamiento que van hacia abajo [Riaño 16].

Se elimina el subárbol de regresión por debajo de cada nodo del árbol de regresión siempre y cuando el error que produce ese subárbol sea mayor o igual que el producido por el del modelo lineal asociado al nodo [Riaño 16].

Suavizado

En la etapa final se utiliza un proceso de suavizado para compensar las discontinuidades bruscas que inevitablemente se producen entre los modelos lineales adyacentes en las hojas del árbol podado, particularmente para algunos modelos construidos a partir de un pequeño número de instancias de entrenamiento [Wang 97]. En el procedimiento de suavizado descrito en [Quinlan 92], primero se utiliza el modelo hoja para calcular el valor pronosticado, y luego se filtra ese valor a lo largo del camino de regreso a la raíz, suavizándose en cada nodo mediante la combinación con el valor pronosticado por el modelo lineal para ese nodo. Se calcula de la siguiente forma [Wang 97]:

$$p' = \frac{|I|p + sq}{|I| + s} \quad (7.5)$$

donde:

- p' es la predicción pasada al siguiente nodo superior.
- p es la predicción pasada a este nodo desde abajo.
- q es el valor predicho por el modelo en ese nodo.

- $|I|$ es el número de instancias de entrenamiento que llegan al nodo por debajo.
- s es una constante (valor por defecto 15).

El suavizado aumenta sustancialmente la precisión de las predicciones [Wang 97].

7.5. Método M5Rules

Se presenta el método M5'Rules (M5Rules) [Holmes 99] para la generación de reglas de árboles modelo, basado en la estrategia básica del algoritmo PART [Frank 98]. En este método se aplica un árbol modelo, utilizando el algoritmo M5P con el conjunto completo de datos de entrenamiento y se genera un árbol podado. A continuación, la mejor hoja conforme a alguna heurística, se convierte en una regla y se descarta el árbol. Todas las instancias cubiertas por esa regla se eliminan del conjunto de datos. El proceso se aplica de forma recursiva a las instancias restantes y termina cuando todas las instancias son cubiertas por una o más reglas. Ésta es la estrategia básica divide y vencerás para las reglas de aprendizaje [Holmes 99]. A diferencia del algoritmo PART [Frank 98], M5Rules construye árboles completos en lugar de árboles parcialmente explorados [Holmes 99]. Este método se centra en la generalización de reglas que utilizan modelos lineales sin suavizar, ya que el árbol del cual se genera una regla, se desecha en cada etapa [Holmes 99].

Para determinar en cada etapa que hoja del árbol es la candidata mejor para añadir al conjunto de reglas, el enfoque más obvio es elegir la hoja que cubre la mayoría de los ejemplos [Frank 98]. Los valores en las hojas del árbol y sobre la consecuente de las reglas son la cobertura y el porcentaje de la raíz del error cuadrático medio, respectivamente, para las instancias que llegan a esas hojas (las que satisfacen esas reglas) [Holmes 99]. El porcentaje de la raíz del error cuadrático medio se calcula de la forma siguiente [Holmes 99]:

$$RMSE \% = \frac{\sqrt{\sum_{i=1}^{N_r} (Y_i - \hat{Y}_i)^2 / N_r}}{\sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2 / N}} \quad (7.6)$$

donde:

- Y_i es el valor real de la clase para el ejemplo i .
- \hat{Y}_i es el valor predicho de la clase para el ejemplo i or el modelo lineal en una hoja.
- N_r es el número de ejemplos cubiertos por la hoja.
- \bar{Y} es la media de los valores de la clase.
- N es el número total de ejemplos.

Capítulo 8

Redes neuronales artificiales

En este capítulo se expone los fundamentos básicos de las redes neuronales artificiales.

8.1. Neurona artificial

Podemos dar como definición que una neurona es un dispositivo simple de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas, proporciona una única salida [Martín 06].

8.1.1. Elementos

Los elementos que constituyen la neurona de etiqueta i son los siguientes [Rumelhart 86a, McClelland 86] (ver Figura 8.1):

- Conjunto de *entradas*, $x_j(t)$.
- *Pesos sinápticos* de la neurona i , w_{ij} que representan la intensidad de interacción entre cada neurona presináptica j y la neurona postsináptica i . Los términos presináptica y postsináptica indican la posición anterior o posterior a la sinapsis o unión entre neuronas.
- *Regla de propagación* $\sigma(w_{ij}, x_j(t))$, que proporciona el valor del potencial postsináptico $h_i(t) = \sigma(w_{ij}, x_j(t))$ de la neurona i en función de sus pesos y entradas.

Se pueden utilizar diferentes funciones, algunos ejemplos son:

- La función de tipo *lineal*

$$h_i(t) = \sum_j w_{ij} x_j \quad (8.1)$$

- La *distancia euclídea*

$$h_i^2(t) = \sum_j (x_j - w_{ij})^2 \quad (8.2)$$

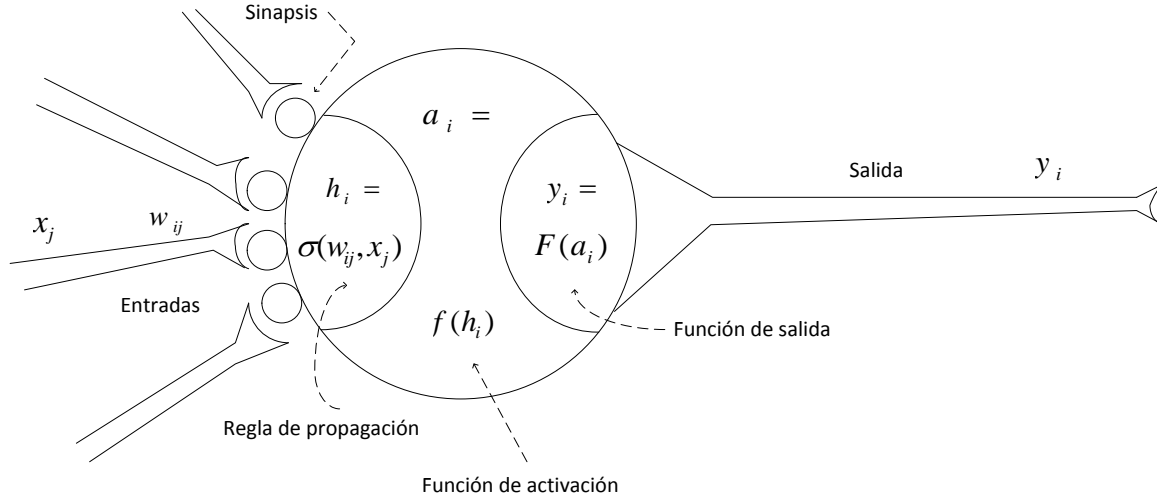


Figura 8.1: *Modelo genérico de neurona artificial [Rumelhart 86a].*

- *Función de activación* $f_i(a_i(t-1), h_i(t))$, que proporciona el estado de activación actual $a_i(t) = f_i(a_i(t-1), h_i(t))$ de la neurona i , en función de su estado anterior $a_i(t-1)$ y de su potencial postsináptico actual.

En muchos modelos de redes neuronales artificiales se considera que el estado actual de la neurona no depende de su estado anterior, sino únicamente del actual

$$a_i(t) = f_i(h_i(t)) \quad (8.3)$$

En ocasiones los algoritmos de aprendizaje requieren que la función de activación sea derivable, las más empleadas son:

- Las funciones de tipo sigmoideo:

$$f(x) = \frac{1}{1 + e^{-x}} \in [0, 1] \quad (8.4)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \tanh(x) \in [-1, 1] \quad (8.5)$$

- La función gaussiana, se utiliza junto con reglas de propagación que involucren el cálculo de cuadrados de distancias (por ejemplo, la euclídea) entre los vectores de entradas y pesos.
- *Función de salida* $F_i(a_i(t))$, proporciona la salida actual de la neurona i en función de su estado de activación

$$y_i(t) = F_i(f_i[a_i(t-1), \sigma_i(w_{ij}, x_j(t))]) \quad (8.6)$$

Frecuentemente la función de salida es simplemente la identidad $F(x) = x$, de modo que el estado de activación de la neurona se considera como la propia salida

$$y_i(t) = F_i(a_i(t)) = a_i(t) \quad (8.7)$$

8.2. Red neuronal artificial (RNA)

Una red neuronal artificial es un grafo dirigido, con las siguientes propiedades [Müller 90]:

1. A cada nodo i se asocia una variable de estado x_i .
2. A cada conexión (i, j) de los nodos i y j se asocia un peso $w_{ij} \in \mathbb{R}$.
3. A cada nodo i se asocia un umbral θ_i .
4. Para cada nodo i se define una función $f_i(x_j, w_{ij}, \theta_i)$, que depende de los pesos de sus conexiones, del umbral y de los estados de los nodos j a él conectados. Esta función proporciona el nuevo estado del nodo.

8.2.1. Conceptos

Las *capas* son unidades estructurales que se agrupan las neuronas. El conjunto de una o más capas forma la *red neuronal*, hay tres tipos de capas [Martín 06]:

- *Capa de entrada*: compuesta por neuronas que reciben datos procedentes del entorno.
- *Capa oculta*: aquella que no tiene una conexión directa con el entorno.
- *Capa de salida*: aquella cuyas neuronas proporcionan la respuesta de la red neuronal.

Se distinguen:

- *Redes monocapa*: compuestas por una única capa de neuronas.
- *Redes multicapa*: sus neuronas se organizan en varias capas.

Según el flujo de datos se puede hablar de:

- *Redes unidireccionales (feedforward)*: la información circula en un único sentido desde las neuronas de entrada hacia las de salida.
- *Redes recurrentes (feedback)*: la información puede circular entre las capas en cualquier sentido, incluido el de salida-entrada.

8.2.2. Fases de operación

Se distinguen dos fases de operación en los sistemas neuronales [Martín 06]:

Fase de aprendizaje

El aprendizaje consiste en determinar un conjunto de pesos sinápticos que permite a la red realizar correctamente el tipo de procesamiento deseado. Cuando se construye un sistema neuronal, se parte de un cierto modelo de neurona y de una determinada arquitectura de red y se establecen los pesos sinápticos iniciales como nulos o aleatorios.

Los pesos sinápticos se modifican siguiendo una cierta regla de aprendizaje, construida a partir de la optimización de una función de error, que mide la eficacia actual de la operación de la red. Se denomina $w_{ij}(t)$ al peso que conecta la neurona presináptica j con la postsináptica i en la iteración t , el algoritmo de aprendizaje, en función de las señales que en el instante t llegan procedentes del entorno, proporciona el valor $\Delta w_{ij}(t)$ que da la modificación que se debe incorporar en dicho peso:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t)$$

El proceso de aprendizaje es *iterativo*, actualizándose los pesos, una y otra vez, hasta que la red neuronal alcanza el rendimiento deseado.

Se definen tres tipos básicos de aprendizaje [Haykin 99]:

- *Aprendizaje supervisado*: se presenta a la red un conjunto de patrones (valores de las entradas), junto con la salida deseada u objetivo, e iterativamente ésta ajusta sus pesos hasta que su salida tiende a ser la deseada, se utiliza para ello información del error que comete en cada paso.
- *Aprendizaje no supervisado*: se presentan a la red multitud de patrones sin adjuntar la respuesta que deseamos. La red, por medio de la regla de aprendizaje, estima la función densidad de probabilidad $p(\vec{x})$, de esta manera se pueden agrupar patrones según su similitud (*clustering*).
- *Aprendizaje híbrido*: coexisten en la red los dos tipos básicos de aprendizaje, el supervisado y el no supervisado, los cuales tienen lugar en distintas capas de neuronas.

Fase de ejecución

Una vez que el sistema ha sido entrenado, los pesos y la estructura quedan fijos, estando la red neuronal dispuesta para procesar datos.

8.3. Adalina

Este modelo [Widrow 60, Widrow 88] utiliza una neurona de respuesta lineal, cuyas entradas pueden ser continuas (ver Figura 8.2). El nodo de la adalina incorpora un parámetro adicional θ_i denominado *bias*, que proporciona un grado de libertad adicional. La ecuación de la adalina queda

$$y_i(t) = f\left(\sum_{j=1}^n w_{ij} x_j - \theta_i\right) = \sum_{j=1}^n w_{ij} x_j - \theta_i, \quad \forall i, \quad 1 \leq i \leq m \quad (8.8)$$

En la adalina se utiliza la regla de aprendizaje de *Widrow-Hoff*, también conocida como regla LMS (*Least Mean Squares*, mínimos cuadrados).

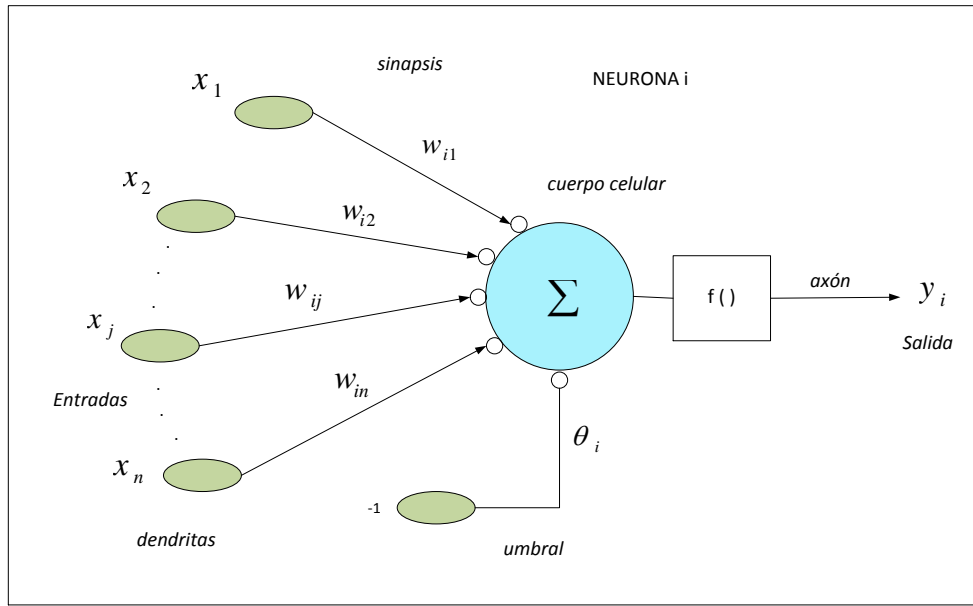


Figura 8.2: *Modelo de neurona estándar (adaptado de [Martín 06]).*

8.3.1. Regla LMS

Aprendizaje como optimización de una función coste

Este método de aprendizaje permite derivar de forma sistemática reglas de aprendizaje para arquitecturas concretas. Se define una función error o coste $E(\cdot)$ que proporciona el error actual E que comete la red neuronal, que depende del conjunto de pesos sinápticos W , $E = E(W)$, $E : \mathbb{R}^n \rightarrow \mathbb{R}$. La representación gráfica de esta función, se puede imaginar como una hipersuperficie con montañas y valles (Figura 8.3). La posición ocupada por un valle se corresponde con una configuración de pesos W^* localmente óptima, al tratarse de un mínimo local de la función error [Martín 06].

El método de optimización (minimización) *descenso por el gradiente* (GD) que se aplica a la función coste proporciona una regla de actualización de pesos, que en función de los patrones de aprendizaje modifica iterativamente los pesos hasta alcanzar el punto óptimo de la red neuronal. Éste se realiza en los siguientes pasos [Martín 06]:

1. Se parte en $t := 0$ de una cierta configuración $W(0)$, y se calcula el sentido de la máxima variación de la función $E(W)$ en $W(0)$, que viene dado por su gradiente en $W(0)$.
2. El sentido de la máxima variación (máximo gradiente) apuntará hacia una colina del paisaje de la hipersuperficie de $E(\cdot)$.
3. Se modifican los parámetros W siguiendo el sentido contrario al indicado por el gradiente de la función error. Así, se lleva a cabo un descenso por la hipersuperficie del error, aproximándose en una cierta cantidad al valle que corresponde a un mínimo local (ver Figura 8.3).

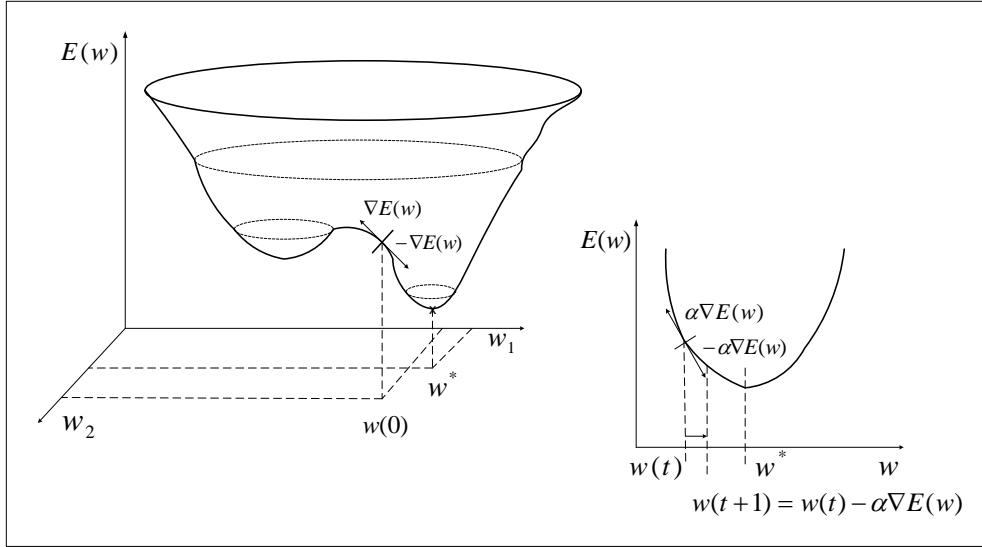


Figura 8.3: Superficie de error $E(w)$ en el espacio de pesos w , y descenso por el gradiente hacia un mínimo (local) [Martín 06].

4. El proceso se itera hasta alcanzarlo (ver [Rumelhart 86a]).
5. Matemáticamente se expresa del siguiente modo:

$$W(t+1) = W(t) - \alpha \nabla E(W) \quad (8.9)$$

donde α (que puede ser diferente para cada peso) indica el tamaño del paso tomado en cada iteración, que idealmente debe ser infinitesimal.

Una actualización de este tipo conduce a un mínimo del funcional de error $E(W)$. Se supone una matriz de pesos $W = \{w_{ij}\}$ y se calcula la variación que en $E(W)$ se produce en la iteración t [Martín 06]:

$$\begin{aligned} \delta(E(w_{ij})) &= \sum_{ij} \frac{\partial E(w_{ij})}{\partial w_{ij}} \delta w_{ij} = \sum_{ij} \frac{\partial E(w_{ij})}{\partial w_{ij}} \left(-\alpha \frac{\partial E(w_{ij})}{\partial w_{ij}} \right) \\ &= -\alpha \sum_{ij} \left(\frac{\partial E(w_{ij})}{\partial w_{ij}} \right)^2 \leq 0 \end{aligned} \quad (8.10)$$

luego la variación en la función error es siempre menor que cero, por lo que siempre disminuye. Mediante este procedimiento se asegura alcanzar un mínimo local de la función, aunque puede no coincidir, en general, con el mínimo global.

Una deducción simple de la regla LMS

Para un conjunto de muestras finitas $(\vec{x}^\mu, \vec{t}^\mu)$, $\mu = 1, \dots, p$, con \vec{x} un conjunto de patrones de entrada y \vec{t} las salidas deseadas, se plantea la siguiente función error [Martín 06]:

$$E[w_{ij}] = \frac{1}{2} \sum_{\mu=1}^p \sum_{i=1}^m (t_i^\mu - y_i^\mu)^2 \quad (8.11)$$

Mediante esta función se obtiene el error cuadrático medio que corresponde a las salidas actuales de la red respecto de los objetivos. El proceso de optimización es el del descenso por el gradiente. Se calcula

$$\frac{\partial E[w_{ij}]}{\partial w_{ij}} = -\frac{1}{2} 2 \sum_{\mu=1}^p (t_i^\mu - y_i^\mu) \frac{dy_i^\mu}{dw_{ij}} = -\sum_{\mu=1}^p (t_i^\mu - y_i^\mu) x_j^\mu \quad (8.12)$$

y el incremento en los pesos queda

$$\Delta w_{ij} = -\alpha \frac{\partial E[w_{ij}]}{\partial w_{ij}} = \alpha \sum_{\mu=1}^p (t_i^\mu - y_i^\mu) x_j^\mu \quad (8.13)$$

esta expresión es la regla LMS.

8.4. El perceptrón multicapa (MLP)

El perceptrón multicapa (MLP) es un modelo de red neuronal unidireccional organizada en capas y con aprendizaje supervisado [Martín 06]. La estructura del perceptrón multicapa o MLP (*Multi-Layer Perceptron*) de tres capas, se presenta en la Figura 8.4, con las entradas, salida, pesos y umbrales de las neuronas definidas de la siguiente forma:

- x_i entradas de la red.
- y_j salidas de la capa oculta.
- z_k salidas de la capa final.
- t_k salidas objetivo.
- w_{ji} pesos de la capa oculta.
- θ_j umbrales de la capa oculta.
- w'_{kj} pesos de la capa de salida.
- θ'_k umbrales de la capa de salida.

8.4.1. El MLP como aproximador universal de funciones

Un MLP de una única capa oculta constituye un aproximador universal de funciones [Funahashi 89, Hornik 89]. Se enuncia el siguiente teorema:

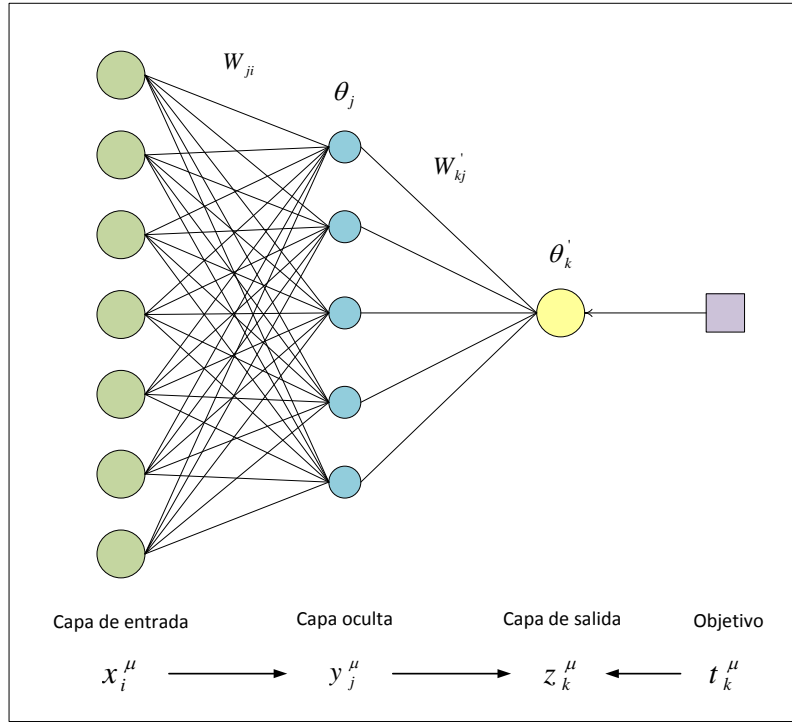


Figura 8.4: *Arquitectura del MLP de tres capas (adaptado de [Martín 06]).*

Teorema 8.4.1.1. [Funahashi 89] Sea $f(x)$ una función no constante, acotada y monótona creciente. Sea K un subconjunto compacto (acotado y cerrado) de \mathbb{R}^n . Sea un número real $\epsilon \in \mathbb{R}$, y sea un entero $c \in \mathbb{Z}$, tal que $c \geq 3$, que fijamos. En estas condiciones, se tiene que:

Cualquier función $\vec{g} : \vec{x} \in K \rightarrow (g_1(\vec{x}), g_2(\vec{x}), \dots, g_m(\vec{x})) \in \mathbb{R}^m$, con $g_l(\vec{x})$ sumables en K , puede ser aproximada en el sentido de la topología L_2 en K por la relación entrada-salida representada por una red neuronal unidireccional (MLP) de c capas ($c - 2$ ocultas), con $f(x)$ como función de transferencia de las neuronas ocultas, y funciones lineales para las de las capas de entrada y de salida.

Es decir: $\forall \epsilon > 0$, \exists un MLP de las características anteriores, que implementa la función

$$\vec{g}' : \vec{x} \in K \rightarrow (g'_1(\vec{x}), g'_2(\vec{x}), \dots, g'_m(\vec{x})) \in \mathbb{R}^m \quad (8.14)$$

de manera que

$$d_{L_2(K)}(\vec{g}, \vec{g}') = \left(\sum_{l=1}^m \int_K |g_l(x_1, \dots, x_n) - g'_l(x_1, \dots, x_n)| d\vec{x} \right)^{1/2} < \epsilon \quad (8.15)$$

Las funciones sigmoideas que se emplean en el MLP cumplen las condiciones exigidas a $f(x)$ [Martín 06].

Este teorema no informa sobre el número exacto de nodos ocultos, solo que hay que colocar los necesarios para aproximar una función determinada [Hornik 89].

8.4.2. Aprendizaje por retropropagación de errores (BP)

Se utiliza el algoritmo de retropropagación de errores o BP (*back-propagation*) para entrenar la red MLP [Rumelhart 86a, Rumelhart 86b, Hecht-Nielsen 90].

Dado un patrón de entrada $(x_1, \dots, x_n)^\mu$, $(\mu = 1, \dots, p)$, la operación global de un MLP de tres capas se expresa del siguiente modo [Martín 06]:

$$z_k^\mu = g\left(\sum_j w'_{kj} y_j^\mu - \theta'_k\right) = g\left(\sum_j w'_{kj} f\left(\sum_i w_{ji} x_i^\mu - \theta_j\right) - \theta'_k\right) \quad (8.16)$$

$g(\cdot)$ es la función de activación de las neuronas de salida, se considera la identidad, y $f(\cdot)$ es la función de activación de las neuronas ocultas, que se toma de tipo sigmoideo (8.4) o (8.5).

Se parte del error cuadrático medio como función coste:

$$E(w_{ji}, \theta_j, w'_{kj}, \theta'_k) = \frac{1}{2} \sum_\mu \sum_k \left(t_k^\mu - g\left(\sum_j w'_{kj} y_j^\mu - \theta'_k\right) \right)^2 \quad (8.17)$$

cuya minimización se realiza mediante descenso por el gradiente (en los trabajos realizados en esta tesis la capa de salida solo tiene una neurona $k = 1$, que corresponde al resultado de predicción); hay un gradiente respecto de los pesos de la capa de salida (w'_{kj}) y otro respecto de los de la oculta (w_{ji})

$$\delta w'_{kj} = -\alpha \frac{\partial E}{\partial w'_{kj}} \quad (8.18)$$

$$\delta w_{ji} = -\alpha \frac{\partial E}{\partial w_{ji}} \quad (8.19)$$

Las expresiones de actualización de los pesos (8.18) y (8.19) se calculan teniendo en cuenta las dependencias funcionales y aplicando la regla de la cadena:

$$\delta w'_{kj} = \alpha \sum_\mu \Delta'_k{}^\mu y_j^\mu, \quad (8.20)$$

$$\text{con } \Delta'_k{}^\mu = \left(t_k^\mu - g(h'_k{}^\mu) \right) \frac{\partial g(h'_k{}^\mu)}{\partial h'_k{}^\mu} \text{ y } h'_k{}^\mu = \sum_j w'_{kj} y_j^\mu - \theta'_k$$

$$\delta w_{ji} = \alpha \sum_\mu \Delta_j^\mu x_i^\mu, \quad (8.21)$$

$$\text{con } \Delta_j^\mu = \left(\sum_k \Delta'_k{}^\mu w'_{kj} \right) \frac{\partial f(h_j^\mu)}{\partial h_j^\mu} \text{ y } h_j^\mu = \sum_i w_{ji} x_i^\mu - \theta_j$$

$h'_k{}^\mu$ y h_j^μ son los potenciales postsinápticos (o locales). En la actualización de los umbrales se considera que el umbral es un caso particular de peso sináptico, cuya entrada es una constante igual a -1.

El procedimiento que se ha seguido en los trabajos realizados en esta tesis para entrenar mediante BP [Martín 06] la arquitectura MLP de tres capas ha sido el siguiente:

1. En $t := 0$ se comienza con pesos y umbrales iniciales aleatorios (valores pequeños positivos y negativos).

2. Para cada patrón de entrada $(x_1, \dots, x_n)^\mu$, $(\mu = 1, \dots, p)$ del conjunto de aprendizaje:
 - a) Se lleva a cabo una fase de ejecución para obtener la respuesta de la red ante el patrón μ -ésimo (8.16).
 - b) Se calcula las señales de error $\Delta'_k{}^\mu$ (8.20) y $\Delta_j{}^\mu$ (8.21).
3. Se calcula el incremento total actual (para todos los patrones) de los pesos y umbrales: $\delta w'_{kj}$ (8.20) y δw_{ji} (8.21).
4. Se actualizan pesos y umbrales.
5. Se calcula el error actual (8.17), $t := t + 1$, y se vuelve al paso 2) si todavía no es satisfactorio.

8.4.3. Otros algoritmos

Para resolver algunos de los inconvenientes del BP se plantean algunas modificaciones, como incluir en el algoritmo un término de inercia. Otro grupo de algoritmos se basan en realizar el descenso por el gradiente utilizando también la información proporcionada por el ritmo de cambio de la pendiente. Los algoritmos de gradientes conjugados y de Newton son ejemplos de ello; cada uno emplea distintas aproximaciones que evitan el cálculo directo de la Hessiana [Martín 06].

Descenso por el gradiente con *momentum* (GD-m)

En [Rumelhart 86b, Martín 06] se incluye en el algoritmo BP un término de inercia (*momentum*), que consiste en añadir al cálculo de la variación de los pesos (8.18) y (8.19) un término adicional proporcional al incremento de la iteración anterior

$$\delta w'_{kj}(t+1) = -\alpha \left. \frac{\partial E}{\partial w'_{kj}} \right|_t + \beta \delta w'_{kj}(t-1) \quad (8.22)$$

$$\delta w_{ji}(t+1) = -\alpha \left. \frac{\partial E}{\partial w_{ji}} \right|_t + \beta \delta w_{ji}(t-1) \quad (8.23)$$

con β un parámetro entre 0 y 1.

Gradiente conjugado descendente (CGD)

Primero se da una estimación inicial $W(0)$ del punto mínimo W^* de $E(W)$, después el algoritmo genera una secuencia de puntos $W(1), W(2), \dots$, y se sigue el siguiente proceso [Charalambous 92]:

1. Se inicializa: $W = W(0)$ (el punto de partida), $D(0) = -\nabla E(W(0))$ (el negativo del vector gradiente en $W(0)$), y $t := 0$ (t es el contador de iteraciones).

2. Se comienza en el punto $W(t)$, y se busca a lo largo de la línea que es paralela a la dirección $D(t)$, para determinar una longitud de paso α_t que disminuirá “suficientemente” el valor de la función de variable única

$$\Psi(\alpha) = E(W(t) + \alpha D(t))$$

relativa a su valor en $\alpha = 0$.

3. Se actualiza la estimación del punto mínimo:

$$W(t+1) = W(t) + \alpha_t D(t)$$

4. La dirección de búsqueda es:

$$D(t+1) = -\nabla E(W(t+1)) + \beta_{t+1} D(t)$$

donde

$$\beta_t = \frac{\nabla E(W(t))^T \nabla E(W(t))}{\nabla E(W(t-1))^T D(t-1)}$$

5. Si $\|\nabla E(W(t+1))\| < \epsilon$ se para; en otro caso $t := t+1$ y se vuelve al paso 3.

El método BFGS

A continuación se describe el algoritmo de cuasi-Newton más popular, el método BFGS, llamado así por sus descubridores Broyden, Fletcher, Goldfarb y Shanno [Nocedal 99]:

Se inicia el punto $W(0)$, la tolerancia de la convergencia $\epsilon > 0$, y la aproximación de la Hessiana inversa $H(0)$;

$t := 0$;

Mientras $\|\nabla E(W(t))\| > \epsilon$;

1. Se calcula la dirección de búsqueda:

$$P(t) = -H(t) \nabla E(W(t));$$

2. Se establece $W(t+1) = W(t) + \gamma_t P(t)$ donde γ_t se calcula a partir de un procedimiento de búsqueda lineal que satisface las condiciones de Wolfe:

$$\begin{aligned} E(W(t) + \gamma_t P(t)) &\leq E(W(t)) + c_1 \gamma_t \nabla E(W(t))^T P(t), \\ \nabla E(W(t) + \gamma_t P(t))^T P(t) &\geq c_2 \nabla E(W(t))^T P(t), \end{aligned}$$

con $0 < c_1 < c_2 < 1$.

3. Se define:

$$S(t) = W(t+1) - W(t) \quad y \quad Q(t) = \nabla E(W(t+1)) - \nabla E(W(t))$$

4. Se calcula:

$$H(t+1) = (I - P(t) S(t) Q(t)^T) H(t) (I - P(t) Q(t) S(t)^T) + P(t) S(t) S(t)^T$$

5. Se actualiza: $t := t+1$;

fin (mientras)

Tabla 8.1: *Modificación de los pesos en la iteración t .*

Algoritmos	$\Delta W(t)$	Referencias
GD	$-\alpha \nabla E(W(t))$	[Martín 06]
GD-m	$-\alpha \nabla E(W(t)) + \beta \Delta W(t-1)$	[Rumelhart 86b, Martín 06]
CGD	$\alpha_t D(t)$	[Charalambous 92]
BFGS	$-\gamma_k H(t) \nabla E(W(t))$	[Nocedal 99]

En la Tabla 8.1 se resume brevemente la modificación de los pesos en la iteración t , $\Delta W(t)$, de los diferentes algoritmos que se utilizan en los trabajos de investigación de esta tesis.

8.5. Funciones de base radial (RBF)

El modelo de funciones de base radial o RBF [Moody 89, Poggio 90] es una red unidireccional para aproximación funcional, pero que puede considerarse de tipo híbrido por incorporar aprendizaje supervisado y no supervisado. Las RBF permiten modelar con relativa facilidad sistemas no lineales arbitrarios [Martín 06].

La arquitectura de una red RBF cuenta con tres capas de neuronas, de entrada, oculta y de salida [Martín 06]:

- Las neuronas de entrada, simplemente envían la información del exterior hacia las neuronas de la capa oculta.
- Las neuronas ocultas operan en base a la distancia que separa el vector de entradas respecto del vector sináptico que cada una almacena (denominado centroide), cantidad a la que aplican una función radial con forma gaussiana (Figura 8.5). Es decir, en el RBF las neuronas son de respuesta localizada, pues sólo responden con una intensidad apreciable cuando el vector de entradas presentado y el centroide de la neurona pertenecen a una zona próxima en el espacio de las entradas.
- Las neuronas de la capa de salida son lineales, calculan la suma ponderada de las salidas que proporciona la capa oculta.

8.5.1. Descripción matemática del modelo

A continuación se describe matemáticamente el modelo [Martín 06]. Se denominan (ver Figura 8.5):

- x_i las entradas de la red.
- y_j las salidas de la capa oculta.

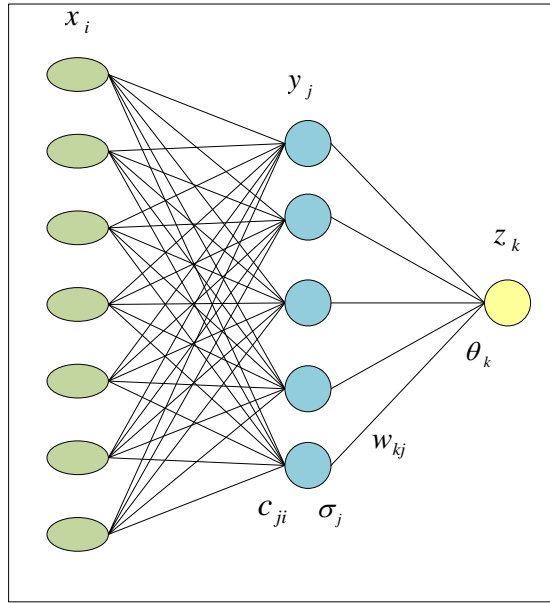


Figura 8.5: *Arquitectura del RBF (adaptado de [Martín 06]).*

- z_k las salidas de la capa final.

Cada neurona j de la capa oculta:

- Almacena un vector c_{ji} , el centroide.
- Calcula la distancia eulídea r_j que separa el vector de entradas x_i de su centroide

$$r_j^2 = \|\vec{x} - \vec{c}_j\|^2 = \sum_i (x_i - c_{ji})^2 \quad (8.24)$$

La salida de la neurona y_j se calcula a partir de una función de activación denominada función radial $\phi(r)$. Una de las más típicas es la función gaussiana

$$\phi(r) = e^{-r^2/2\sigma^2} \quad (8.25)$$

El término *función de base radial* procede precisamente de la simetría radial de estas funciones (el nodo da una salida idéntica para aquellos patrones que distan lo mismo del centroide). El parámetro de normalización σ mide la anchura de la gaussiana, y equivaldría al radio de influencia de la neurona en el espacio de las entradas; a mayor σ la región que la neurona *domina* en torno al centroide es más amplia (ver Figura 8.6).

La salida de la neurona j se escribe:

$$y_j = e^{-r_j^2/2\sigma_j^2} = e^{-\sum_i (x_i - c_{ji})^2 / 2\sigma_j^2} \quad (8.26)$$

Así, si el vector de entradas coincide con el centroide de la neurona j ($\vec{x} = \vec{c}_j$), ésta responde con la máxima salida (la unidad). Si el patrón de entrada es muy diferente del centroide, la respuesta tiende a cero.

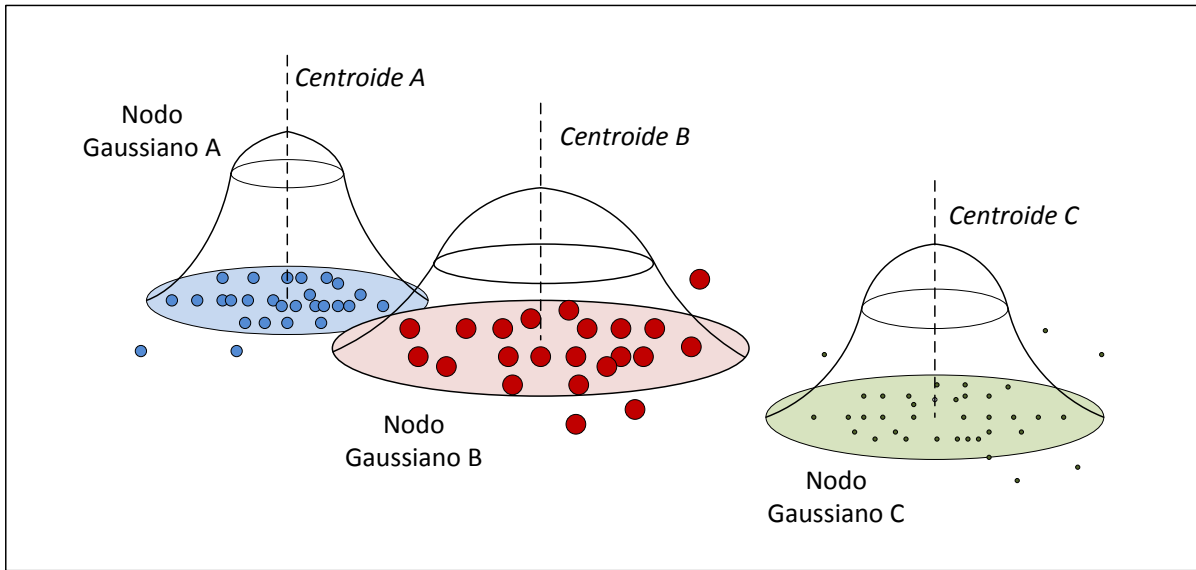


Figura 8.6: *Respuesta localizada de las neuronas ocultas en el RBF (nodos gaussianos). Los puntos representan patrones en el espacio de las entradas, que en su mayoría se agrupan en torno a tres centros (adaptado de [Martín 06]).*

Las salidas de las neuronas ocultas son a su vez las entradas de las neuronas de salida, las cuales calculan su respuesta z_k de la forma

$$z_k = \sum_j w_{kj} y_j + \theta_k = \sum_j w_{kj} \phi(r_j) + \theta_k \quad (8.27)$$

siendo w_{kj} el peso que conecta la neurona oculta j con la salida k , y θ_k un parámetro adicional de la neurona k (umbral o *bias*). En la RBF, cada nodo gaussiano se ocupa de una zona del espacio, y el conjunto de nodos debe cubrir totalmente la zona de interés (ver Figura 8.6).

8.5.2. Aprendizaje en las RBF

Se siguen los siguientes pasos [Martín 06]:

- Como cada nodo radial cubre una parte del espacio de entrada (ver Figura 8.6), se elige el número k de nodos adecuados para que cubran suficientemente dicho espacio.
- Una vez determinado el número de nodos radiales, se trata de encontrar los parámetros de la arquitectura.
- Se emplea un aprendizaje por etapas:
 - En la primera se realiza el entrenamiento de las neuronas ocultas gaussianas.
 - En la segunda se realiza el entrenamiento de las neuronas de salida.

- En el entrenamiento de los nodos gaussianos se determina:
 1. Primero el valor de los centroides c_{ji} de partida, se aplica el algoritmo de las k -medias (ver [Tou 74]). Se toman como centroides los primeros k patrones de aprendizaje.
 2. En cada iteración t se reparten los patrones de aprendizaje \vec{x} entre las k neuronas. Cada patrón se asigna a la neurona de cuyo centroide dista menos.
 3. Se calculan los nuevos centroides de cada neurona como promedio de los patrones de aprendizaje asignados en el paso (2). Se tiene

$$c_{ji} = \frac{1}{n_j} \sum_{\vec{x} \in \text{neurona } j} \vec{x} \quad (8.28)$$

llamando n_j al número de patrones que han correspondido a la neurona j en el reparto.

4. Si los valores de los centroides no han variado respecto de la iteración anterior, el algoritmo ya ha convergido. Si no es así, volver a (2).
- Obtenidos los centroides, se procede al cálculo de los parámetros de escala σ_j de cada neurona:

$$\sigma_j^2 = \frac{1}{N} \sum_{l=1}^N \|\vec{c}_l - \vec{c}_j\|^2 \quad (8.29)$$

En los trabajos de esta tesis se ha tomado $N = 1$, por lo que se tiene en cuenta sólo el nodo más cercano [Wassermann 93].

- Por último, se procede al entrenamiento de las neuronas de salida. Los pesos w_{kj} y los umbrales θ_k se calculan aplicando el algoritmo LMS a la expresión de la salida de la capa final

$$z_k = \sum_j w_{kj} \phi(r_j) + \theta_k \quad (8.30)$$

que queda de la forma

$$w_{kj}(t+1) = w_{kj}(t) + \alpha \sum_{\mu=1}^p (t_k^\mu - z_k^\mu) \phi(r_j)^\mu \quad (8.31)$$

siendo t_k los valores objetivo. Los umbrales se actualizan considerando que se trata de pesos con entradas de valor -1.

Capítulo 9

Máquinas de vectores soporte

En este capítulo se presentan los fundamentos teóricos de las máquinas de vectores soporte (SVM), que se utilizan en los trabajos de investigación recogidos en esta tesis.

9.1. SVM lineal con margen máximo

La SVM lineal con margen máximo es el modelo más sencillo de SVM y el que tiene condiciones de aplicabilidad más restringidas [Hernández 04].

9.1.1. Formulación original de las SVM

A continuación se presenta la idea básica de la formulación original de las SVM [Smola 04]. Dado un conjunto de datos de entrenamiento $\{(x_1, y_1), \dots, (x_\ell, y_\ell)\} \subset \mathcal{X} \times \mathbb{R}$, donde \mathcal{X} denota el espacio de los parámetros de entrada (por ejemplo, $\mathcal{X} = \mathbb{R}^d$), el objetivo en ε -SV para la regresión [Vapnik 95], es encontrar una función $f(x)$ que tenga como máxima desviación ε de los valores reales y_i obtenidos para todos los datos de entrenamiento, y sea al mismo tiempo lo más plana posible.

Primero se describe el caso de las funciones lineales f , que toman la forma

$$f(x) = \langle w, x \rangle + b \text{ con } w \in \mathcal{X}, b \in \mathbb{R} \quad (9.1)$$

donde $\langle \cdot, \cdot \rangle$ denota el producto escalar en \mathcal{X} . La exactitud en el caso (9.1) significa que se busca un w pequeño. Una manera de asegurar esto es minimizar la norma, $\|w\|^2 = \langle w, w \rangle$. Se puede escribir este problema como un problema de optimización convexa:

$$\begin{aligned} &\text{Minimizar} && \frac{1}{2} \|w\|^2 \\ &\text{sujeto a:} && \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon \end{cases} \end{aligned} \quad (9.2)$$

La suposición (9.2) es que tal función f en realidad existe y aproxima todos pares (x_i, y_i) con una precisión ε , es decir, el problema de optimización convexa es factible. A veces, se desea permitir algunos errores. La función de pérdida “margen blando” [Bennett 92]

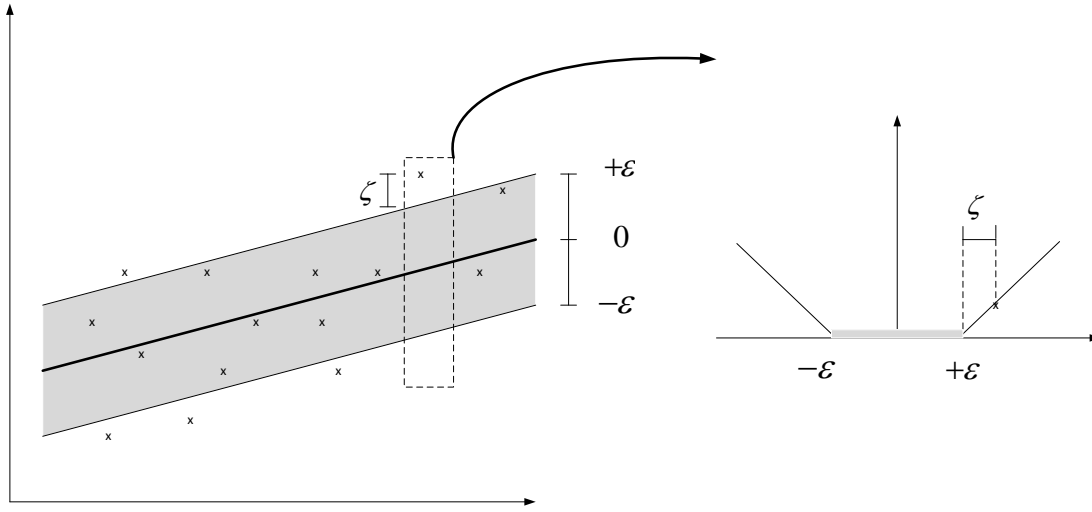


Figura 9.1: *Margen blando de pérdida para un SVM lineal (adaptado de [Smola 04]).*

adaptada a máquinas SV por [Vapnik 95], introduce variables de holgura ζ_i, ζ_i^* para hacer frente a las restricciones no factibles del problema de optimización (9.2):

$$\begin{aligned} &\text{Minimizar} \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\ell} (\zeta_i + \zeta_i^*) \\ &\text{sujeto a:} \quad \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \zeta_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \geq 0 \end{cases} \end{aligned} \quad (9.3)$$

La constante $C > 0$ determina el equilibrio entre la exactitud de f y el valor hasta el cual las desviaciones mayores que ε son toleradas. Esto corresponde a tratar con una función de pérdida ε -insensible, $|\zeta|_\varepsilon$ descrita por:

$$|\zeta|_\varepsilon := \begin{cases} 0 & \text{si } |\zeta| \leq \varepsilon \\ |\zeta| - \varepsilon & \text{en otro caso.} \end{cases} \quad (9.4)$$

La Figura 9.1 representa esta situación gráficamente.

En la mayoría de los casos el problema de optimización (9.3) se puede resolver más fácilmente en su formulación dual. Además, la formulación dual proporciona la clave para la extensión de las máquinas de vector soporte a funciones no lineales. Por tanto, se usa un método de dualización estándar utilizando multiplicadores de Lagrange [Fletcher 89].

9.1.2. Formulación dual de las SVM

La idea clave [Smola 04] es construir una función de Lagrange de la función objetivo y las restricciones correspondientes, mediante la introducción de un conjunto de variables duales. Se puede demostrar que esta función tiene un punto de silla con respecto a las variables primales y duales en la solución. Para más detalles se puede consultar

[Mangasarian 69, McCormick 83], o [Vanderbei 97]. Se procede de la siguiente manera:

$$\begin{aligned}
 L := & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\ell} (\zeta_i + \zeta_i^*) - \sum_{i=1}^{\ell} (\eta_i \zeta_i + \eta_i^* \zeta_i^*) \\
 & - \sum_{i=1}^{\ell} \alpha_i (\varepsilon + \zeta_i - y_i + \langle w, x_i \rangle + b) \\
 & - \sum_{i=1}^{\ell} \alpha_i^* (\varepsilon + \zeta_i^* + y_i - \langle w, x_i \rangle - b)
 \end{aligned} \tag{9.5}$$

L es la *función lagrangiana* y $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$, son *multiplicadores de Lagrange*. Luego las variables duales en (9.5) satisfacen:

$$\alpha_i^{(*)}, \eta_i^{(*)} \geq 0. \tag{9.6}$$

donde $\alpha_i^{(*)}$ se refiere a α_i y α_i^* .

Se sigue de la condición de punto de silla, es decir, que las derivadas parciales de L con respecto a las variables primales $(w, b, \zeta_i, \zeta_i^*)$ tienen que desaparecer para optimizar.

$$\partial_b L = \sum_{i=1}^{\ell} (\alpha_i^* - \alpha_i) = 0 \tag{9.7}$$

$$\partial_w L = w - \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) x_i = 0 \tag{9.8}$$

$$\partial_{\zeta_i^{(*)}} L = C - \alpha_i^{(*)} - \eta_i^{(*)} = 0 \tag{9.9}$$

Sustituyendo (9.7), (9.8) y (9.9) en (9.5) se obtiene el problema de optimización dual.

$$\begin{aligned}
 \text{Maximizar} \quad & \begin{cases} -\frac{1}{2} \sum_{i,j=1}^{\ell} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \langle x_i, x_j \rangle \\ -\varepsilon \sum_{i=1}^{\ell} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{\ell} y_i (\alpha_i - \alpha_i^*) \end{cases} \\
 \text{sujeto a:} \quad & \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) = 0 \text{ y } \alpha_i, \alpha_i^* \in [0, C]
 \end{aligned} \tag{9.10}$$

En (9.10) ya se han eliminado las variables duales η_i, η_i^* a través de la condición (9.9) que puede ser reformulada como $\eta_i^{(*)} = C - \alpha_i^{(*)}$. La ecuación (9.8) se puede reescribir como:

$$w = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) x_i, \text{ así } f(x) = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b. \tag{9.11}$$

9.1.3. Cálculo de b

El cálculo de b [Smola 04] se puede hacer aprovechando las condiciones de Karush-Kuhn-Tucker (KKT) [Karush 39, Kuhn 51]. En estas situaciones el punto de la solución del producto entre las variables duales y las restricciones se anula:

$$\begin{aligned}
 \alpha_i (\varepsilon + \zeta_i - y_i + \langle w, x_i \rangle + b) &= 0 \\
 \alpha_i^* (\varepsilon + \zeta_i^* + y_i - \langle w, x_i \rangle - b) &= 0
 \end{aligned} \tag{9.12}$$

$$\begin{aligned}(C - \alpha_i) \zeta_i &= 0 \\ (C - \alpha_i^*) \zeta_i^* &= 0\end{aligned}\tag{9.13}$$

Esto permite varias conclusiones útiles:

- Sólo las muestras (x_i, y_i) con $\alpha_i^{(*)} = C$ quedan fuera del tubo ε -insensible.
- $\alpha_i \alpha_i^* = 0$, es decir, nunca puede haber un conjunto de variables duales α_i, α_i^* que ambas simultáneamente sean distintas de cero. Se concluye que:

$$\varepsilon - y_i + \langle w, x_i \rangle + b \geq 0 \quad \text{y} \quad \zeta_i = 0 \quad \text{si} \quad \alpha_i < C \tag{9.14}$$

$$\varepsilon - y_i + \langle w, x_i \rangle + b \leq 0 \quad \text{si} \quad \alpha_i > 0 \tag{9.15}$$

- Se hace un análisis análogo con α_i^* y se tiene:

$$\begin{aligned} \max\{-\varepsilon + y_i - \langle w, x_i \rangle \mid \alpha_i < C \text{ o } \alpha_i^* > 0\} &\leq b \leq \\ \min\{-\varepsilon + y_i - \langle w, x_i \rangle \mid \alpha_i > 0 \text{ o } \alpha_i^* < C\} &\end{aligned}\tag{9.16}$$

- Si algún $\alpha_i^{(*)} \in (0, C)$, las desigualdades se convierten en igualdades. En [Keerthi 01] se pueden ver otras maneras de escoger b .

9.2. SVM con margen máximo en el espacio de características

Se observa que todo lo desarrollado en la sección anterior para la SVM lineal con margen máximo sólo depende de la existencia de un producto escalar en el espacio de entrada: hiperplanos, distancias, derivadas del producto escalar... [Hernández 04].

La SVM (no lineal) con margen máximo en el espacio de características se basa en la idea de hacer una transformación no lineal del espacio de entrada a un espacio dotado de un producto escalar (un espacio de Hilbert [Berberian 61, Hernández 04]):

- Si existe una transformación no lineal del espacio de entrada en un cierto *espacio de características* \mathfrak{S} :

$$\begin{aligned}\Phi : \mathbb{R}^d &\rightarrow \mathfrak{S} \\ x &\rightarrow \Phi(x)\end{aligned}$$

dotado de un producto escalar $\langle \Phi(x), \Phi(x') \rangle$, \mathfrak{S} es un espacio de Hilbert.

- Si el conjunto de datos es linealmente separable en \mathfrak{S} (con los hiperplanos definidos a partir del producto escalar correspondiente), entonces la SVM con margen máximo en el espacio de características se obtiene sustituyendo $\langle x, x' \rangle$ en la SVM lineal con margen máximo por $\langle \Phi(x), \Phi(x') \rangle$.

Para algunos espacios de características y algunas transformaciones existe una forma de calcular el producto escalar usando las *funciones núcleo* [Hernández 04]:

Tabla 9.1: *Funciones núcleo comúnmente utilizadas en \mathbb{R}^d .*

	Funciones núcleo	Referencias
Polinómica	$k(x, x') = (\langle x, x' \rangle + c)^p$	$p \in \mathbb{N}, c \in \mathbb{R}$ [Boser 92, Vapnik 95]
Sigmoidal	$k(x, x') = \tanh(s \langle x, x' \rangle + r)$	$s, r \in \mathbb{R}$ [Hernández 04]
Gaussiana	$k(x, x') = e^{-\frac{\ x-x'\ ^2}{2\sigma^2}}$	$\sigma \in \mathbb{R}$ [Aizerman 64, Micchelli 86]

- Una *función núcleo* es una función $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ tal que $k(x, x') = \langle \Phi(x), \Phi(x') \rangle$, donde Φ es una transformación de \mathcal{X} en un espacio de Hilbert \mathfrak{H} , ver Tabla 9.1.

El siguiente teorema caracteriza las funciones $k(x, x')$ que corresponden a un producto escalar en algún espacio de características \mathfrak{H} (definidas sobre \mathcal{X}) [Smola 04].

Teorema 9.2.0.1. [Mercer 09]. Sea $k \in L_\infty(\mathcal{X}^2)$ tal que el operador integral $T_k : L_2(\mathcal{X}) \rightarrow L_2(\mathcal{X})$,

$$T_k f(\cdot) := \int_{\mathcal{X}} k(\cdot, x) f(x) d\mu(x) \quad (9.17)$$

es positivo (μ denota una medida sobre \mathcal{X} con $\mu(\mathcal{X})$ finita y $\text{supp}(\mu) = \mathcal{X}$). Sea $\phi_j \in L_2(\mathcal{X})$ una función propia de T_k asociada con el valor propio $\lambda_j \neq 0$ y normalizada tal que $\|\phi_j\|_{L_2} = 1$ y sea $\bar{\phi}_j$ su compleja conjugada. Entonces

1. $(\lambda_j(T))_j \in L_1$.
2. $\phi_j \in L_\infty(\mathcal{X})$ y $\text{supp}_j \|\phi_j\|_{L_\infty} < \infty$.
3. $k(x, x') = \sum_{j \in \mathbb{N}} \lambda_j \bar{\phi}_j(x) \phi_j(x')$ para casi todo (x, x') , donde la serie converge absoluta y uniformemente para casi todo (x, x') .

Este teorema significa que si

$$\int_{\mathcal{X} \times \mathcal{X}} k(x, x') f(x) f(x') dx dx' \geq 0 \quad \forall \quad f \in L_2(\mathcal{X}) \quad (9.18)$$

se puede escribir $k(x, x')$ como un producto escalar en algún espacio de características.

Como consecuencia del teorema de Mercer, dada una función k que cumpla las condiciones del teorema de Mercer, existe un espacio de Hilbert \mathfrak{H} y una transformación $\Phi : \mathbb{R}^d \rightarrow \mathfrak{H}$ tal que $k(x, x') = \langle \Phi(x), \Phi(x') \rangle$. La transformación del espacio de entrada en el espacio de características es $\Phi(x) = (\phi_1(x), \phi_2(x), \dots)$, y el producto escalar k en \mathfrak{H} es el expresado en el teorema [Hernández 04].

Lo interesante de usar funciones núcleo con SVM es que el producto escalar se puede calcular implícitamente, sin conocer de manera explícita \mathfrak{H} ni la transformación Φ y evitando el coste computacional derivado de la posible alta dimensión de \mathfrak{H} [Hernández 04].

Capítulo 10

Sistemas basados en lógica borrosa

Los *sistemas de inferencia borrosa* combinan unas variables de entrada (definidas en términos de conjuntos borrosos) por medio de grupos de reglas que producen uno o varios valores de salida [Martín 06].

10.1. Lógica borrosa

La *lógica borrosa* permite tratar información imprecisa, como *temperatura baja*, en términos de conjuntos borrosos o difusos (imprecisos) [Martín 06].

En los siguientes apartados de este capítulo se exponen los principales conceptos que aparecen tanto en lógica borrosa como en los conjuntos borrosos, se explican las funciones de inclusión, las reglas y su utilización, y se introducen los conceptos de partición borrosa y algunas medidas borrosas [Wang 94, Jang 97, Martín 06].

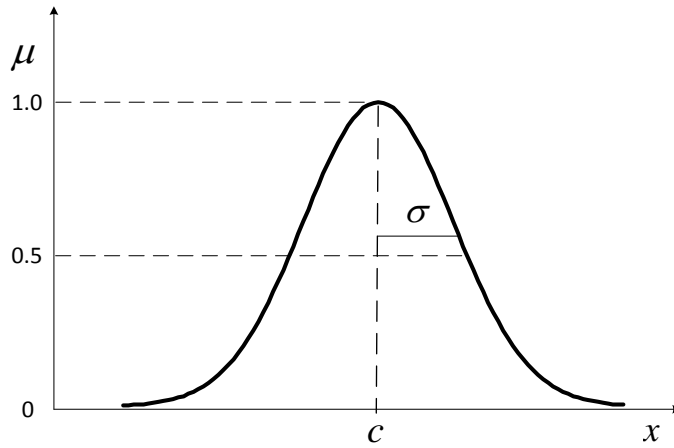
Conjuntos borrosos

- El *Universo del discurso* es un conjunto de objetos X , por ejemplo $X = \mathbb{R}^m$.
- Un *conjunto borroso* F en X queda caracterizado por una función de inclusión μ_F que toma valores en el rango $[0,1]$, es decir, $\mu_F : X \rightarrow [0,1]$; donde $\mu_F(x)$ representa el grado en el que $x \in X$ pertenece al conjunto borroso F .

Funciones de inclusión de conjuntos borrosos

La *función de inclusión o pertenencia* de un conjunto borroso consiste en un conjunto de pares ordenados $F = \{(x, \mu_F(x)) \mid x \in X\}$ si la variable es discreta, o una función continua si no lo es.

Para la definición de estas funciones de pertenencia se utilizan ciertas familias, las más frecuentes son la función de tipo trapezoidal, triangular, S, exponencial y de tipo π . A continuación, se describen algunas (todas toman valores en $[0,1]$):

Figura 10.1: *Función de pertenencia gaussiana.*

- La *función gaussiana* G se define por

$$G(x; c, \sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \quad -\infty < x < \infty \quad (10.1)$$

donde c representa el centro y σ determina la anchura de la función de pertenencia (ver Figura 10.1).

- La *función de tipo S*, se caracteriza por tener un valor de inclusión distinto de 0 para un rango de valores por encima de un cierto punto a , siendo 0 por debajo de a y 1 para valores mayores de c . Su punto de cruce es $b = (a + c)/2$; y entre los puntos a y c es de tipo cuadrático (suave):

$$S(x; a, b, c) = \begin{cases} 0 & x < a \\ 2\left(\frac{x-a}{c-a}\right)^2 & a \leq x \leq b \\ 1 - 2\left(\frac{x-a}{c-a}\right)^2 & b \leq x \leq c \\ 1 & x > c \end{cases} \quad (10.2)$$

También se utilizan funciones exponenciales para definir funciones de tipo S , como

$$S(x; k, c) = \frac{1}{1 + e^{-k(x-b)}} \quad (10.3)$$

- La *función de tipo π* tiene forma de campana, y es adecuada para los conjuntos definidos en torno a un valor c , como: *medio*, *cero*...

$$\pi(x; b, c) = \begin{cases} S(x; c - b, c - b/2, c) & x \leq c \\ 1 - S(x; c - b, c - b/2, c) & x \geq c \end{cases} \quad (10.4)$$

Variable lingüística

Una *variable lingüística* se define por una tupla $(A, T(A), X, G, R)$, donde:

- A es el nombre de la variable.
- $T(A)$ es el conjunto de términos que nombran los valores que puede tomar A , valores que son conjuntos borrosos en X .
- X es el universo del discurso de la variable x , el conjunto de valores numéricos que puede tomar para una variable discreta, o el rango de valores posibles para una continua.
- G es una regla sintáctica para la generación de los nombres de los valores de x .
- R es una regla semántica para asociar un significado a cada valor.

Ejemplo:

- *Temperatura* puede considerarse como una variable lingüística, $A = \text{Temperatura}$.
- $T(\text{temperatura})$ es el conjunto de todos términos (*muy fría, fría, normal, alta, muy alta*).
- El universo del discurso X , se puede restringir al rango 0 y 40 °C.

Particiones borrosas

Una *partición* de A es uno de los subconjuntos que pueden formarse con los elementos (términos) de $T(A)$. Una *partición es completa* si los subconjuntos definidos cubren todo X . Dos subconjuntos están *solapados* si su intersección es no nula.

Medidas borrosas

Las *medidas borrosas* son magnitudes medibles del *conjunto borroso* A . La *distancia* entre dos *conjuntos borrosos* A y C se puede definir utilizando diversas medidas. Una de las más frecuentes es la *euclídea*:

$$f(A) = \left(\sum (\mu_A(x) - \mu_C(x))^2 \right)^{1/2} \quad (10.5)$$

Operaciones borrosas

Los subconjuntos borrosos A y B , asociados a una variable lingüística x , pueden realizar operaciones entre ellos (ver Tabla 10.1). Estas operaciones básicas se expresan en términos de las funciones de pertenencia de los conjuntos borrosos A y B .

Reglas borrosas

Las *reglas borrosas* combinan uno o más conjuntos borrosos de entrada, llamados *antecedentes*, y les asocian un conjunto borroso de salida, llamado *consecuente*. Permiten expresar el conocimiento que se dispone sobre la relación entre *antecedentes* y *consecuentes*. Para expresar este conocimiento de forma completa se precisa de varias reglas, que se agrupan formando una *base de reglas*, que se puede representar como una tabla de las reglas que la forman. El formato de una *base de reglas* puede ser:

Tabla 10.1: Operaciones entre conjuntos borrosos [Martín 06].

	Operación	Rango
Igualdad	$\mu_A(x) = \mu_B(x)$	$x \in X$
Unión	$\mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)]$	$\forall x \in X$
Intersección	$\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)]$	$\forall x \in X$
Complemento	$\bar{\mu}_A(x) = 1 - \mu_A(x)$	$x \in X$
Norma	$\mu_{Norm(A)}(x) = \frac{\mu_A(x)}{\max[\mu_A(x)]}$	$x \in X$
Concentración	$\mu_{Conc(A)}(x) = (\mu_A(x))^2$	$x \in X$
Dilatación	$\mu_{Dilat(A)}(x) = (\mu_A(x))^{0.5}$	$x \in X$

- Tipo *Mamdani*: una colección de reglas $R^{(l)}$ con el formato

$$R^{(l)} : IF \ x_1 \text{ is } F_1^l \text{ and } \dots \text{ and } x_n \text{ is } F_n^l \text{ THEN } y \text{ is } G^l \quad (10.6)$$

donde F_i^l y G^l son conjuntos borrosos en $X_i \subset \mathbb{R}$ y $Y \subset \mathbb{R}$, respectivamente, y $\vec{x} = (x_1, \dots, x_n) \in X_1 \times \dots \times X_n$ e $y \in Y$ son variables lingüísticas.

- Tipo *Sugeno* [Sugeno 85a, Sugeno 85b]: una colección de reglas $R^{(l)}$ con el formato

$$R^{(l)} : IF \ x_1 \text{ is } F_1^l \text{ and } \dots \text{ and } x_n \text{ is } F_n^l \text{ THEN } y^l = f^l(\vec{x}) \quad (10.7)$$

donde la parte consecuyente de la regla corresponde a la función de salida que es una combinación lineal de las variables de entrada, o en su caso más general, una función genérica de las variables de entrada [Takagi 83].

Si se llama L al número de reglas IF-THEN de la base de reglas, entonces $l = 1, \dots, L$ en las ecuaciones (10.6) y (10.7). El vector \vec{x} representa el conjunto de las entradas, mientras que y es la salida del sistema borroso.

Dispositivos de inferencia borrosa

Los *dispositivos de inferencia borrosa* son los sistemas que interpretan las reglas de tipo IF-THEN de una base de reglas, con el fin de obtener los valores de salida a partir de los actuales valores de las variables lingüísticas de entrada al sistema.

El término $\mu_A(\vec{x}) = \mu_{F_1^l \times \dots \times F_n^l}(\vec{x})$, puede ser definido por:

- La *regla del mínimo*

$$\mu_{F_1^l \times \dots \times F_n^l}(\vec{x}) = \min \{ \mu_{F_1^l}(x_1), \dots, \mu_{F_n^l}(x_n) \} \quad (10.8)$$

- La *regla del producto*

$$\mu_{F_1^l \times \dots \times F_n^l}(\vec{x}) = \mu_{F_1^l}(x_1) \cdot \dots \cdot \mu_{F_n^l}(x_n) \quad (10.9)$$

La *salida final* de un dispositivo de inferencia borrosa consiste en L escalares y^l , con $l = 1, 2, \dots, L$. Si las reglas son de tipo Sugeno, según la ecuación (10.7), cada uno de los escalares es el resultado de aplicar la entrada A' a cada una de las L reglas de la base de reglas.

Borrosificador

El *borrosificador* establece una relación entre puntos de entrada no borrosos al sistema $\vec{x} = (x_1, \dots, x_n)^T$, y sus correspondientes conjuntos borrosos A en X .

El método *borrosificador no singleton* utiliza una función exponencial del tipo

$$\mu_{A'}(x') = a e^{-\left(\frac{x'-x}{\sigma}\right)^2} \quad (10.10)$$

con forma de campana, centrada en el valor x , de anchura σ y amplitud a .

Desborrosificador

El *desborrosificador* es la función que transforma un conjunto borroso, en un valor no borroso $y \in Y$.

Si las reglas utilizadas son de tipo Sugeno (10.7), el valor de salida no borrosa se obtiene como media ponderada de las salidas de cada regla de la base de reglas según

$$y = \frac{\sum_{l=1}^L (\mu_{A'}(\vec{x})) y^l}{\sum_{l=1}^L (\mu_{A'}(\vec{x}))} \quad (10.11)$$

donde y^l es la salida de la regla l , y el término $\mu_{A'}(\vec{x})$ se calcula utilizando (10.8) o (10.9). Este valor y^l de la salida de una regla de tipo Sugeno (10.7) se calcula como una combinación lineal de las entradas:

$$y^l = a_{l,0} + \sum_{i=1}^n a_{l,i} x_i \quad (10.12)$$

Agrupamiento borroso

El *agrupamiento borroso o clustering* divide un conjunto de datos en varios subconjuntos de tal manera que la similitud dentro de un subconjunto sea mayor. Lograr tal partición requiere una métrica que tome dos vectores de entrada y devuelva un valor que refleje su similitud [Jang 97]. Dado que la mayoría de las métricas son sensibles a los rangos de elementos en los vectores de entrada, cada una de las variables de entrada se normaliza para estar dentro del hipercubo unidad.

Las técnicas de agrupamiento se utilizan con el modelado difuso, principalmente para determinar reglas difusas de si-entonces. Con este fin, las técnicas de agrupamiento se validan sobre la base de los siguientes supuestos [Jang 97]:

- Entradas similares al sistema objetivo a ser modelado deben producir resultados similares.
- Estos pares de entrada-salida similares se agrupan en clústeres en el conjunto de datos de entrenamiento.

Los algoritmos de agrupación que más se utilizan con modelado difuso son el k -medias, el c -medias difuso, el método de montaña y el substractivo [Jang 97]. A continuación se describe el método del *agrupamiento substractivo* propuesto en [Chiu 94, Jang 97].

Agrupamiento substractivo

Todos los puntos de datos (no los puntos de la red) se tienen en cuenta como candidatos para los centros de agrupación, y el cálculo es simplemente proporcional al número de puntos de datos e independiente de la dimensión del problema considerado.

Se considera una colección de p puntos de datos $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$ en un espacio k -dimensional. Sin pérdida de generalidad, se supone que los puntos de datos han sido normalizados dentro de un hipercubo. Puesto que cada punto de datos es un candidato para centro de agrupación, una medida de densidad en el punto de datos x_i se define como

$$D_i = \sum_{j=1}^p e^{-\frac{\|\vec{x}_i - \vec{x}_j\|^2}{(r_a/2)^2}}$$

donde r_a es una constante positiva. Por lo tanto, un punto de datos tendrá un valor de alta densidad si tiene muchos puntos de datos vecinos. El radio r_a define un vecindario; los puntos de datos fuera de este radio contribuyen sólo ligeramente a la medida de densidad.

Después de calcular la medida de densidad de cada punto de datos, se selecciona el punto de datos con la medida de densidad más alta como el primer centro de agrupación. Se denomina al punto seleccionado \vec{x}_{c_1} y a su medida de densidad D_{c_1} . A continuación, la medida de densidad para cada punto de datos \vec{x}_i se revisa mediante la fórmula

$$D_i = D_i - D_{c_1} e^{-\frac{\|\vec{x}_i - \vec{x}_{c_1}\|^2}{(r_b/2)^2}}$$

donde r_b es una constante positiva. Por lo tanto, los puntos de datos cerca del primer centro de clúster \vec{x}_{c_1} habrán reducido significativamente sus medidas de densidad, lo que hace que estos puntos sean poco probables de ser seleccionados como el centro de cluster siguiente. La constante r_b define un vecindario que tiene reducciones apreciables en la medida de densidad. La constante r_b es normalmente mayor que r_a para evitar centros de agrupación estrechamente espaciados; generalmente r_b es igual a $1.5 r_a$, como se sugiere en [Chiu 94].

Después de revisar la medida de densidad para cada punto de datos, se selecciona el siguiente centro de agrupamiento \vec{x}_{c_2} y se vuelven a revisar todas las medidas de densidad para los puntos de datos. Este proceso se repite hasta que se genera un número suficiente de centros de clúster. Un criterio de detención más sofisticado para determinar automáticamente el número de clusters se puede encontrar en [Chiu 94].

Cuando se aplica el agrupamiento substractivo a un conjunto de datos de entrada-salida, cada uno de los centros de agrupación representa un prototipo que presenta ciertas características del sistema a modelar. Estos centros de agrupación se utilizarían razonablemente como los centros para las premisas de las reglas difusas en un modelo difuso de Sugeno de orden cero. Por ejemplo, se supone que el centro para el i -ésimo grupo es \vec{c}_i con una dimensión k . El \vec{c}_i se puede descomponer en dos componentes, los vectores \vec{v}_i y \vec{w}_i , donde \vec{v}_i es la parte de entrada y contiene los primeros n elementos de \vec{c}_i , y \vec{w}_i es la parte de salida y contiene los últimos $k - n$ elementos de \vec{c}_i . Entonces, dado un vector de entrada \vec{x} , el grado en el que la regla difusa i se cumple es definido por

$$\mu_i = e^{-\frac{\|\vec{x} - \vec{v}_i\|^2}{(r_a/2)^2}}$$

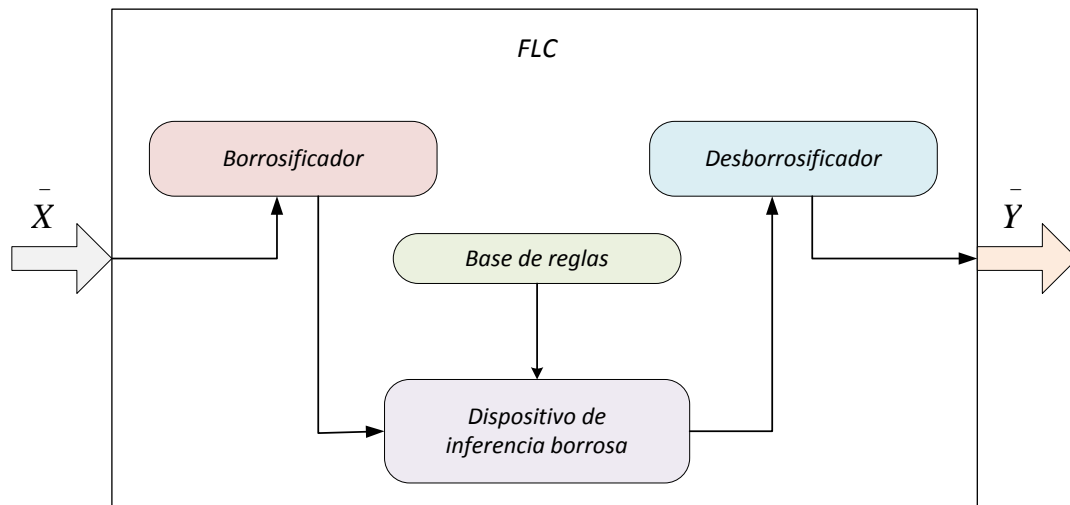


Figura 10.2: Estructura de un controlador borroso o FLC (adaptado de [Martín 06]).

Una vez que se ha determinado la parte de la premisa, la parte consiguiente se puede estimar por el método de los mínimos cuadrados. Una vez completados estos procedimientos, se puede obtener mayor precisión usando el gradiente descendente u otros algoritmos de optimización derivados.

10.2. Controlador borroso

Los sistemas expertos de control borroso basados en reglas, conocidos como *controladores borrosos* o FLC (*Fuzzy Logic Controllers*), o también, sistemas de *inferencia borrosa* o FIS (*Fuzzy Inference Systems*), son la aplicación más extendida de la lógica borrosa [Martín 06].

La estructura interna de un *controlador borroso* se muestra en la Figura 10.2. El *borrosificador* convierte los valores discretos a términos borrosos. Su salida se utiliza por el *dispositivo de inferencia borrosa* para aplicarla a cada una de las reglas de la *base de reglas*, siguiendo el método de inferencia seleccionado. La salida de este bloque son L escalares y^l , con $l = 1, 2, \dots, L$, si las reglas son del tipo Sugeno (10.7). El *desborrosificador* los transforma en un valor no borroso $y \in V$ [Martín 06].

10.3. 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 [Fernández 08]. En esta sección se describe un sistema de inferencia difusa basado en una red adaptativa [Jang 97], la arquitectura se denomina ANFIS (*Adaptive Neuro-Fuzzy Inference Systems*). También, se pueden observar estructuras de redes similares en [Lin 91] y [Wang 92].

Arquitectura ANFIS

Por simplicidad para describir la arquitectura ANFIS, el sistema de inferencia difusa que se considera es con dos entradas x_1 y x_2 , y una salida y . Para un modelo difuso Sugeno de primer orden [Sugeno 88, Takagi 83, Takagi 85], un conjunto de reglas comunes establecidas con dos reglas difusas si-entonces es el siguiente:

$$R^{(1)}: \text{Si } x_1 \text{ es } F_1^1 \text{ y } x_2 \text{ es } F_2^1, \text{ entonces } y^1 = a_{1,0} + a_{1,1} x_1 + a_{1,2} x_2$$

$$R^{(2)}: \text{Si } x_1 \text{ es } F_1^2 \text{ y } x_2 \text{ es } F_2^2, \text{ entonces } y^2 = a_{2,0} + a_{2,1} x_1 + a_{2,2} x_2$$

En la Figura 10.3(a) se ilustra el mecanismo de razonamiento para este modelo Sugeno, y en la Figura 10.3(b) su arquitectura correspondiente, donde los nodos de la misma capa tienen funciones similares, como se describe a continuación (se denota la salida del l -ésimo nodo en la capa num como $O_{num,l}$).

- En la *capa 1* cada nodo l es un nodo adaptativo con una función nodo

$$\begin{aligned} O_{1,l} &= \mu_{F_1^l}(x_1) \quad l = 1, 2 \\ O_{1,l} &= \mu_{F_2^l}(x_2) \quad l = 1, 2 \end{aligned} \quad (10.13)$$

donde x_1 (o x_2) es la entrada al nodo l y F_1^l (o F_2^l) es una etiqueta lingüística (como “pequeña” o “grande”) asociada a este nodo. En otras palabras, $O_{1,l}$ es el grado de pertenencia de un conjunto difuso A ($= F_1^1, F_1^2, F_2^1$ o F_2^2) y especifica el grado en que la entrada dada x_i ($i = 1, 2$) satisface con el cuantificador F_i^l . Aquí la función de pertenencia para F_i^l puede ser cualquier función de pertenencia parametrizada apropiada. En los trabajos de investigación de esta tesis se eligen para $\mu_{F_i^l}(x_i)$ las funciones gaussianas con máximo igual a 1 y mínimo igual a 0 (ver Figura 10.1), tal como

$$\mu_{F_i^l}(x_i) = e^{-\left(\frac{x_i - c_{(l,i)}}{\sigma_{(l,i)}}\right)^2} \quad -\infty < x < \infty; \quad l = 1, 2; \quad i = 1, 2 \quad (10.14)$$

donde $\{c_{(l,i)}, \sigma_{(l,i)}\}$ es el conjunto de parámetros. A medida que cambian los valores de estos parámetros, las funciones gaussianas varían en consecuencia, presentando así diversas formas de funciones de pertenencia para los conjuntos difusos F_i^l . Los parámetros de esta capa se conocen como *parámetros premisa*.

- En la *capa 2* cada nodo es un nodo fijo etiquetado con Π , cuya salida es el producto de todas las señales entrantes:

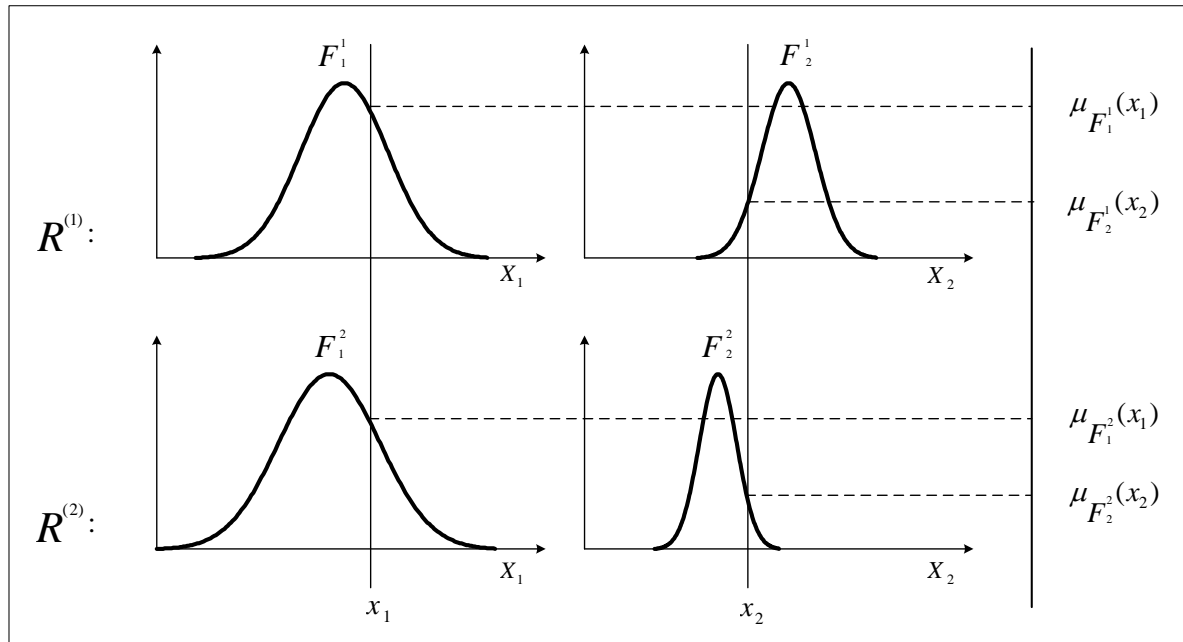
$$O_{2,l} = w_l = \mu_{F_1^l}(x_1) \mu_{F_2^l}(x_2), \quad l = 1, 2 \quad (10.15)$$

Cada salida de nodo representa la fuerza de relación de una regla.

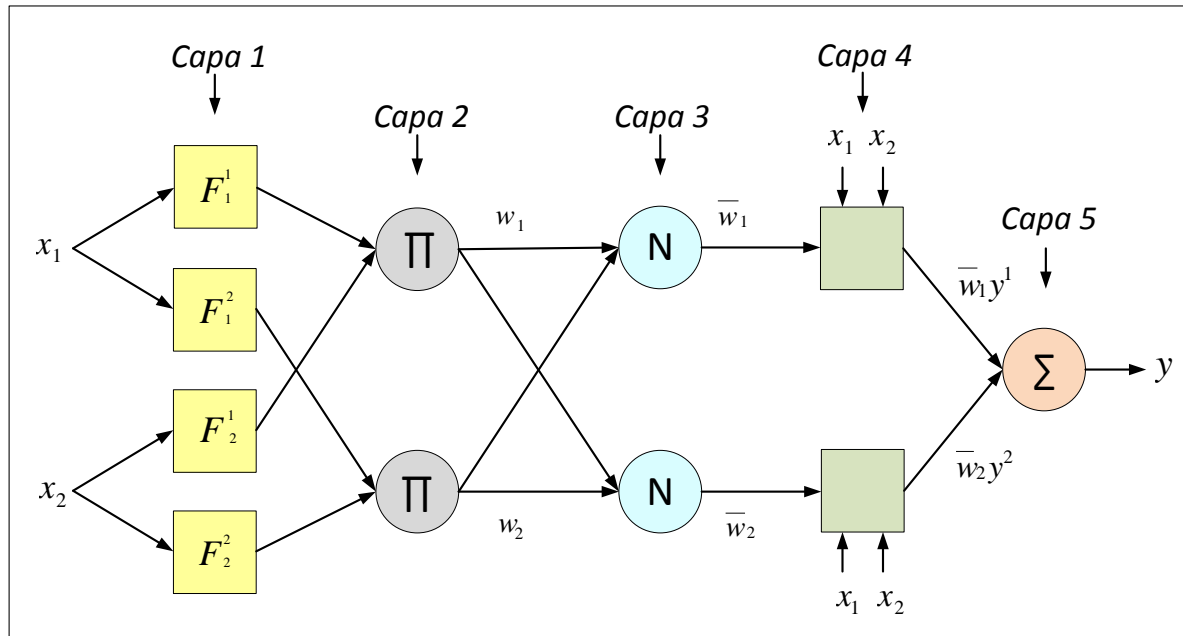
- En la *capa 3* cada nodo es un nodo fijo etiquetado con N . El l -ésimo nodo calcula el radio de la fuerza de relación de la l -ésima regla como la suma de las fuerzas de relación de todas las reglas:

$$O_{3,l} = \bar{w}_l = \frac{w_l}{w_1 + w_2}, \quad l = 1, 2 \quad (10.16)$$

Las salidas de esta capa se denominan *fuerzas de relación normalizadas*.



(a) Sistema de inferencia difusa con dos entradas y dos reglas.



(b) Red equivalente.

Figura 10.3: (a) Un modelo difuso Sugeno de primer orden con dos entradas y dos reglas; (b) Arquitectura ANFIS equivalente (adaptado de [Jang 97]).

- En la *capa 4* cada nodo l es un nodo adaptativo con una función nodo

$$O_{4,l} = \bar{w}_l y^l = \bar{w}_l (a_{l,0} + a_{l,1} x_1 + a_{l,2} x_2) \quad l = 1, 2 \quad (10.17)$$

donde \bar{w}_l es una fuerza de relación normalizada de la capa 3 y $\{a_{l,0}, a_{l,1}, a_{l,2}\}$ es el conjunto de parámetros de este nodo. Los parámetros de esta capa se denominan *parámetros de consecuencia*.

- En la *capa 5* el único nodo es un nodo fijo etiquetado con Σ , que calcula el total de la salida como el sumatorio de todas las señales de entrada:

$$O_{5,l} = \sum_{l=1}^2 \bar{w}_l y^l = \frac{\sum_{l=1}^2 w_l y^l}{\sum_{l=1}^2 w_l} \quad (10.18)$$

Regla de aprendizaje híbrida

Aunque se puede aplicar el algoritmo de retropropagación de errores o BP (*back-propagation*) para identificar los parámetros en una red adaptativa, este método de optimización simple normalmente tarda mucho tiempo antes de converger [Jang 97]. Se puede observar, sin embargo, que la salida (en este caso sólo hay una) de la red adaptativa es lineal en algunos de los parámetros de la red; por lo que se puede identificar esos parámetros lineales por el método de los mínimos cuadrados lineales. Este enfoque conduce a una regla de aprendizaje híbrida [Jang 91, Jang 93], que combina el método de descenso por el gradiente y el estimador por mínimos cuadrados para la identificación rápida de los parámetros [Jang 97].

Se representa la salida de la red adaptativa por

$$O = G(\vec{x}, S) \quad (10.19)$$

donde:

- \vec{x} es el vector de las variables de entrada.
- S es el conjunto de parámetros.
- G es la función global implementada por la red adaptativa.

Si existe una función H tal que la función compuesta $H \circ G$ es lineal en algunos de los elementos de S , entonces esos elementos se pueden identificar por el método de los mínimos cuadrados. Más formalmente, si el conjunto de parámetros S se puede dividir en dos conjuntos

$$S = S_1 \oplus S_2 \quad (10.20)$$

(donde \oplus representa la suma directa) tal que $H \circ G$ es lineal en los elementos de S_2 , entonces se aplica H a la ecuación (10.19), y se tiene

$$H(\circ) = H \circ G(\vec{x}, S) \quad (10.21)$$

que es lineal en los elementos de S_2 . Ahora dados los valores de los elementos de S_1 , se puede trabajar con D datos de entrenamiento en la ecuación (10.21) y obtener una ecuación matriz

$$B\theta = z \quad (10.22)$$

donde θ es un vector desconocido cuyos elementos son los parámetros en S_2 . Este es un problema lineal estándar de mínimos cuadrados, y la mejor solución para θ , que minimiza $\|B\theta - z\|^2$, es el estimador de mínimos cuadrados θ^* :

$$\theta^* = (B^T B)^{-1} B^T z \quad (10.23)$$

donde B^T es la transpuesta de B y $(B^T B)^{-1} B^T$ es la pseudo-inversa de B si $B^T B$ es no singular.

Ahora se puede combinar el descenso por el gradiente y el estimador por mínimos cuadrados para actualizar los parámetros en la red adaptativa. El aprendizaje híbrido se aplica de una vez, cada iteración se compone de un paso hacia adelante y un paso hacia atrás. En el paso hacia adelante, después que se presenta un vector de entrada, se calcula las salidas de los nodos en la red, capa por capa. Este proceso se repite para todos los pares de datos de entrenamiento hasta completar B y z (ecuación (10.22)); entonces los parámetros en S_2 se identifican por la fórmula de la pseudo-inversa de la ecuación (10.23). Después de identificar los parámetros en S_2 , se calcula la medida de error para cada par de datos de entrenamiento. En el paso hacia atrás, las señales de error (la derivada de la medida de error con respecto a cada salida del nodo) se propagan desde la salida hasta el extremo de entrada; el vector de gradiente se acumula para cada entrada de datos de entrenamiento. Al final del paso hacia atrás, para todos los datos de entrenamiento, los parámetros en S_1 se actualizan por el descenso por el gradiente.

Para los valores fijos dados de los parámetros en S_1 , se garantiza que los parámetros hallados en S_2 son el punto óptimo global en el espacio de parámetros S_2 debido a la elección de la medida del error cuadrático. Esta regla de aprendizaje híbrida no sólo puede disminuir la dimensión del espacio de búsqueda, sino que, en general, también reduce sustancialmente el tiempo necesario para alcanzar la convergencia.

Aprendizaje ANFIS

De la arquitectura ANFIS mostrada en la Figura 10.3(b), se observa que cuando los valores de los parámetros premisa son fijos, la salida total se puede expresar como una combinación lineal de los parámetros consecuentes. En símbolos, la salida y en la Figura 10.3(b) se puede reescribir como

$$\begin{aligned} y &= \frac{w_1}{w_1 + w_2} y^1 + \frac{w_2}{w_1 + w_2} y^2 \\ &= \bar{w}_1 (a_{1,0} + a_{1,1} x_1 + a_{1,2} x_2) + \bar{w}_2 (a_{2,0} + a_{2,1} x_1 + a_{2,2} x_2) \\ &= (\bar{w}_1) a_{1,0} + (\bar{w}_1 x_1) a_{1,1} + (\bar{w}_1 x_2) a_{1,2} + (\bar{w}_2) a_{2,0} + (\bar{w}_2 x_1) a_{2,1} + (\bar{w}_2 x_2) a_{2,2} \end{aligned} \quad (10.24)$$

que es lineal en los parámetros consecuentes $a_{1,0}$, $a_{1,1}$, $a_{1,2}$, $a_{2,0}$, $a_{2,1}$ y $a_{2,2}$. De esta observación, se tiene:

S = Conjunto de parámetros totales

S_1 = Conjunto de parámetros premisa (no lineales)

S_2 = Conjunto de parámetros consecuentes (lineales)

Tabla 10.2: *Dos pasos en el proceso de aprendizaje híbrido para ANFIS [Jang 97].*

	Paso hacia adelante	Paso hacia atrás
Parámetros premisa	Fijos	Descenso por el gradiente
Parámetros consecuentes	Estimador de mínimos cuadrados	Fijos
Señales	Salidas de los nodos	Señales de error

en la ecuación 10.20; y $H(\cdot)$ y $G(\cdot, \cdot)$ son la función de identidad y la función del sistema de inferencia difusa, respectivamente, en la ecuación (10.21). Por lo tanto, el algoritmo de aprendizaje híbrido se puede aplicar directamente. Más específicamente, en el paso hacia adelante del algoritmo de aprendizaje híbrido, las salidas de los nodos salen hacia adelante hasta la capa 4 y los parámetros consecuentes se identifican por el método de los mínimos cuadrados. En el paso hacia atrás, las señales de error se propagan hacia atrás y los parámetros premisa se actualizan por descenso por el gradiente. En la Tabla 10.2 se resumen las actividades en cada paso.

Con este proceso de entrenamiento se ajustan los coeficientes de los polinomios de la parte consecuente de las reglas, y los parámetros de las premisas (funciones de pertenencia para los conjuntos definidos en las entradas) [Fernández 08].

En los trabajos de investigación de esta tesis, para evitar el sobre-entrenamiento de los modelos se utiliza un subconjunto de validación cruzada, interrumpiéndose el proceso de entrenamiento cuando el error con los datos del subconjunto de validación cruzada no se reduce tras un número de iteraciones.

Capítulo 11

Algoritmos evolutivos

Los *Algoritmos Evolutivos* (AEs) son una técnica de resolución de problemas de búsqueda y optimización [Araujo 09]. Se aplican como herramientas para la optimización de los parámetros de otros algoritmos [Freitas 02].

11.1. Esquema general

Los diferentes algoritmos evolutivos responden a un esquema común [Araujo 09]:

- Procesan simultáneamente un conjunto de soluciones, que se denominan *individuos*. El conjunto de todos ellos forma la *población* con la que trabaja el algoritmo.
- La composición de la población se modifica a lo largo de las iteraciones del algoritmo que se denominan generaciones (ver Figura 11.1).
- Cada generación incluye un proceso de selección, que da mayor probabilidad de permanecer en la población y participar en las operaciones de reproducción a los mejores individuos.

Se han desarrollado distintas variantes de algoritmos evolutivos [Araujo 09]:

- *Algoritmos Genéticos* [Holland 75, Goldberg 89] que utilizan una representación binaria o entera.
- *Programas de Evolución* [Michalewicz 94] en los que los individuos son cualquier estructura de datos de tamaño fijo.
- *Algoritmos Genéticos de Codificación Real* [Bäck 91, Herrera 98] en los que se utiliza una población de números reales.
- *Programación Evolutiva* [Koza 92, Koza 94] en los que se trabaja con una población de “programas”.

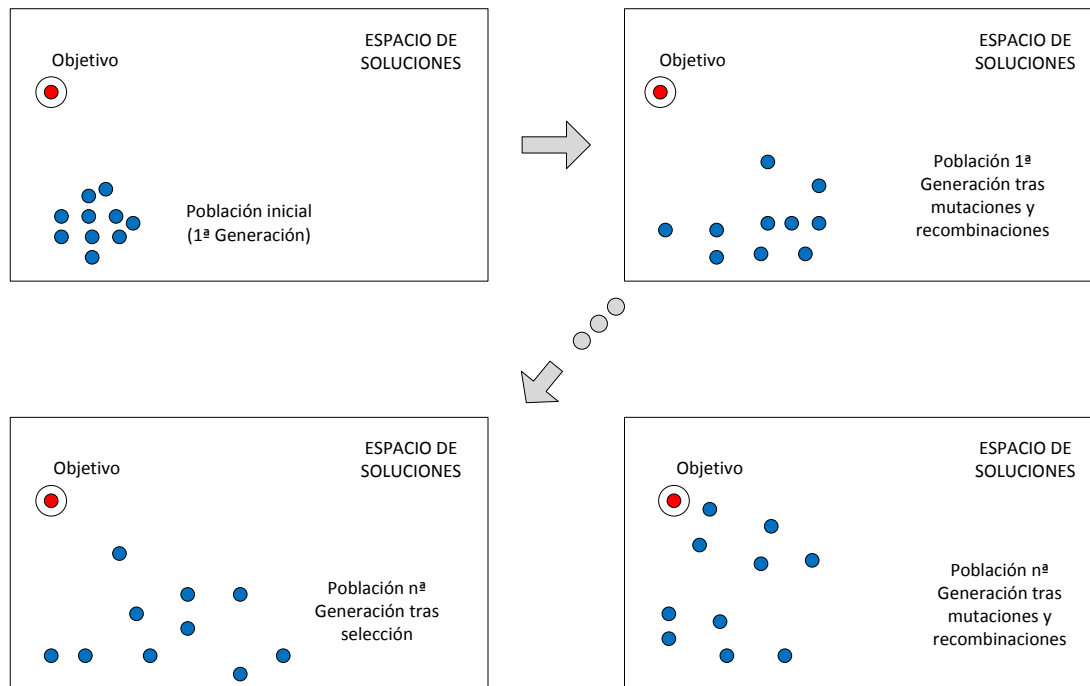


Figura 11.1: *Comportamiento de la búsqueda (adaptado de [Hernández 04]).*

11.2. Algoritmos genéticos

El funcionamiento básico de un algoritmo genético (AG) [Hernández 04] se muestra en la Figura 11.2 y se configura en los siguientes pasos:

- Se parte de una población inicial de individuos que codifican soluciones candidatas al problema propuesto.
- Esta población evoluciona a través de un proceso de competición.
- Cada individuo tiene asociada una medida de adaptación, para determinar cuáles se seleccionan para formar parte de la nueva población. Ésta se creará usando operadores genéticos de cruce y mutación.
- Este ciclo continúa hasta que se verifica una determinada condición de parada.

En los siguientes apartados extraídos de [Araujo 09] se detallan los principales elementos.

Representación de los individuos

En un AG los individuos son cadenas binarias b (ver Tabla 11.1) que representan a puntos x del espacio de búsqueda del problema:

- b se denomina *genotipo* del individuo.

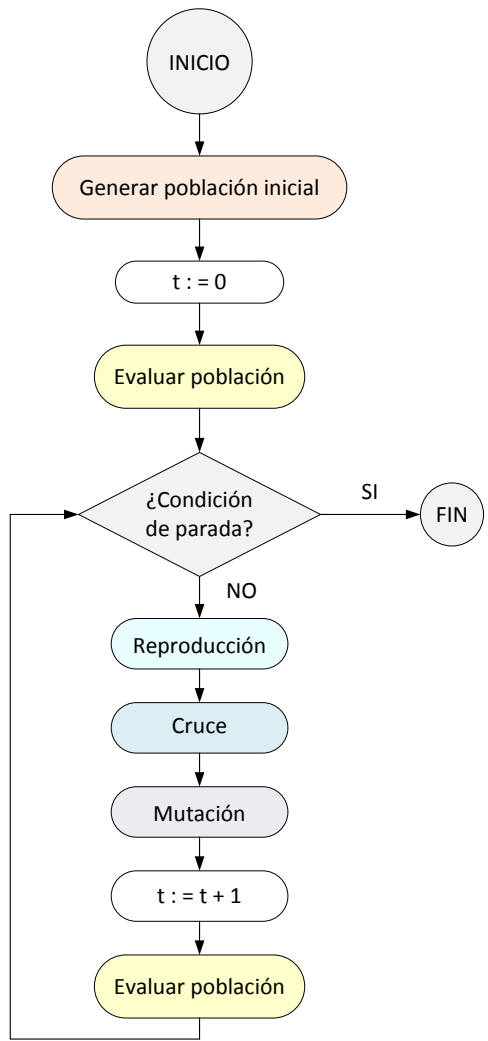


Figura 11.2: *Algoritmo genético básico (adaptado de [Hernández 04]).*

Tabla 11.1: *Individuo genético binario.*

Cromosoma			
Gen	Gen	Gen	
10010	011100	1011101010101000	1010101000101010

- x se denomina *fenotipo*.
- *Gen* se refiere a la codificación de una determinada característica del individuo.
- *Alelo* son los distintos valores que puede tomar un gen.

Generación de la población inicial

Los individuos de la población inicial de un AG suelen ser cadenas de ceros y unos generadas de forma completamente aleatoria.

Para el buen funcionamiento del AG es necesario dotar a la población de suficiente variedad para poder explorar todas las zonas del espacio de soluciones.

Grado de adaptación de los individuos

La evolución de la población depende de la calidad de los individuos que compiten por participar en las operaciones de reproducción. En un problema de búsqueda u optimización, dicha calidad se mide por la *adaptación* de cada individuo a ser solución al problema. Los problemas se presentan como la optimización de una función matemática.

Condiciones de terminación

La condición de terminación más sencilla es alcanzar un número máximo de generaciones.

El proceso de selección

La población del algoritmo genético se somete a un proceso de selección que debe tender a favorecer la cantidad de copias de los individuos más adaptados.

Este proceso se puede realizar de formas diferentes, por ejemplo selección proporcional o por ruleta: la probabilidad de selección p_i de un individuo i con este método es proporcional a su adaptación.

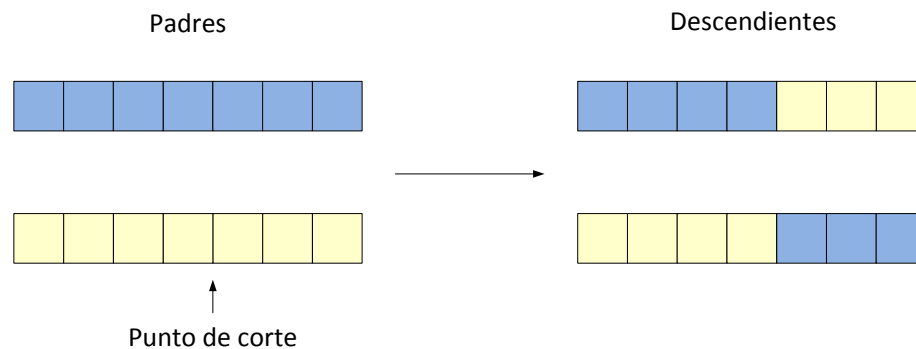
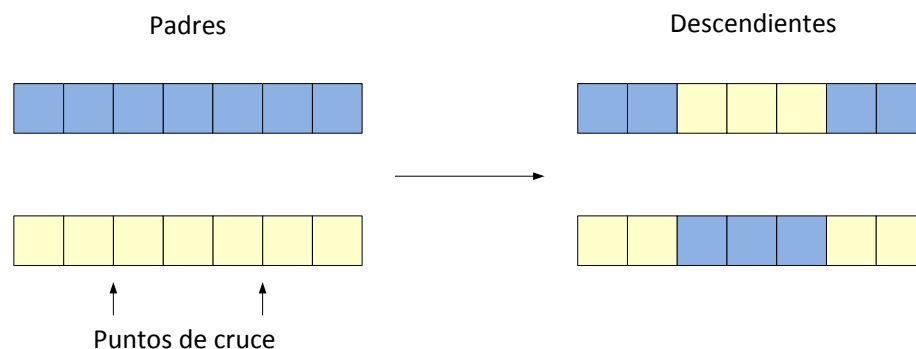
Elitismo

Consiste en asegurar la supervivencia de los mejores individuos de la población. El porcentaje de la población que puede formar parte de la élite debe ser pequeño, no mayor de un 1 o un 2 % del tamaño de la población.

El proceso de reproducción: operadores genéticos

Los nuevos individuos se crean aplicando ciertos *operadores genéticos*:

- El operador de *cruce* combina propiedades de dos individuos de la población anterior para crear nuevos individuos:
 - Cruce *monopunto*: consiste en seleccionar al azar una única posición en la cadena de ambos padres e intercambiar las partes de los padres divididas por dicha posición (ver Figura 11.3).

Figura 11.3: *Cruce monopunto*.Figura 11.4: *Cruce de dos puntos*.

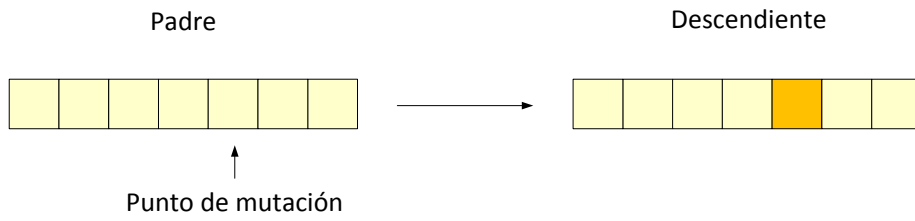
- Cruce de *dos puntos*: se eligen dos puntos de cruce y se intercambia entre los dos padres el segmento definido por dichos puntos de cruce (ver Figura 11.4).
 - Cruce *multipunto* [Eshelman 89]: se tienen múltiples puntos de cruce que definen un conjunto de segmentos que se intercambian entre los padres.
- El operador de *mutación* crea un nuevo individuo realizando alguna alteración pequeña, en un individuo de la población anterior. La forma más sencilla consiste en cambiar el valor de una de las posiciones de la cadena (ver Figura 11.5).

Para cada operador genético se establece una tasa, de manera que el operador sólo se aplica si un valor generado aleatoriamente está por encima de la tasa especificada.

El proceso de reemplazo

Para mantener el tamaño de la población constante, los nuevos individuos creados mediante los operadores genéticos deben reemplazar a otros de la población anterior. Se consideran distintos tipos de AGs:

- *AGs generacionales*: la población se renueva por completo de una generación a otra.

Figura 11.5: *Mutación.*

- *AGs con estado estacionario:* se conserva parte de la población de generación en generación.

Capítulo 12

Probabilidad

En este capítulo se recuerdan algunos conceptos de probabilidad básica, se define la esperanza matemática, la varianza y la desviación estándar, se describe la distribución normal y se muestra la función de probabilidad conjunta. Todos estos conceptos se han extraído de [Spiegel 01] y [Cuadras 88].

12.1. Probabilidad básica

En esta sección se recuerda el concepto, los axiomas y algunos teoremas importantes sobre probabilidad definidos en [Spiegel 01].

Espacios de la muestra

Un conjunto S que consta de todos los resultados posibles de un experimento aleatorio se llama *espacio muestral*.

Evento

Un *evento* es un subconjunto A del espacio muestral S , es decir, un conjunto de resultados posibles.

El concepto de probabilidad

En cualquier experimento aleatorio hay siempre incertidumbre sobre si ocurrirá un evento en particular. Como una medida de la *oportunidad*, o *probabilidad*, con que se espera que ocurra cierto evento, es conveniente asignar un número entre 0 y 1. Si se está seguro de que tal evento ocurrirá, se dice que tiene 100 % de probabilidad o 1, pero si se está seguro de que el evento no ocurrirá, se dice que su probabilidad es cero. Si, por ejemplo, la probabilidad es $\frac{1}{4}$, se debe decir que la oportunidad de que ocurra es 25 %, y 75 % de que no ocurra.

Los axiomas de la probabilidad

Para cada evento A en la clase C de eventos, se asocia un número real $P(A)$. Entonces P se denomina la *función de probabilidad*, y $P(A)$ la probabilidad del evento A , si se cumplen los siguientes axiomas.

Axioma 12.1.1. Para cada evento A en la clase C ,

$$P(A) \geq 0$$

Axioma 12.1.2. Para cada evento cierto o seguro S en la clase C ,

$$P(S) = 1$$

Axioma 12.1.3. Para cualquier número de eventos mutuamente excluyentes A_1, A_2, \dots , en la clase C ,

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

Teoremas

A continuación se enuncian algunos teoremas importantes sobre probabilidad.

Teorema 12.1.1. Si $A_1 \subset A_2$, entonces $P(A_1) \leq P(A_2)$ y $P(A_2 - A_1) = P(A_2) - P(A_1)$.

Teorema 12.1.2. Para todo evento A ,

$$0 \leq P(A) \leq 1,$$

o sea, la probabilidad está entre 0 y 1.

Teorema 12.1.3.

$$P(\emptyset) = 0$$

es decir, el evento imposible tiene probabilidad cero.

Teorema 12.1.4. Si A' es el complemento de A , entonces

$$P(A') = 1 - P(A)$$

Teorema 12.1.5. Si $A = A_1 \cup A_2 \cup \dots \cup A_n$, donde A_1, A_2, \dots, A_n son eventos mutuamente excluyentes, entonces

$$P(A) = P(A_1) + P(A_2) + \dots + P(A_n)$$

En particular, si $A = S$, el espacio muestral, entonces

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1$$

Teorema 12.1.6. Si A y B son dos eventos cualesquiera, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

También se pueden hacer generalizaciones para n eventos.

12.2. Esperanza matemática, varianza y desviación estándar

En esta sección se definen los conceptos de esperanza matemática, varianza y desviación estándar extraídos de [Spiegel 01].

Esperanza matemática

Para una variable aleatoria discreta X , que tiene los posibles valores x_1, \dots, x_n , la *esperanza* de X se define como

$$E(X) = x_1 P(X = x_1) + \dots + x_n P(X = x_n) = \sum_{j=1}^n x_j P(X = x_j)$$

o de manera equivalente, si $P(X = x_j) = f(x_j)$,

$$E(X) = x_1 f(x_1) + \dots + x_n f(x_n) = \sum_{j=1}^n x_j f(x_j) = \sum x f(x) \quad (12.1)$$

donde la última sumatoria toma todos los valores apropiados de x . Como un caso especial de (12.1), donde todas las probabilidades son iguales, se tiene

$$E(X) = \frac{x_1 + x_2 + \dots + x_n}{n}$$

la cual recibe el nombre de *media aritmética*, o simplemente *media* de x_1, \dots, x_n .

La esperanza de X se denomina con frecuencia la *media* de X y se denota por μ_x , o simplemente μ , cuando se sobreentiende de cuál variable aleatoria se trata.

La media, o esperanza de X , da un valor que es representativo o promedio de los valores de X , y por esta razón se denomina la *medida de tendencia central*.

Varianza y desviación estándar

La *varianza* es un número positivo, que se define así

$$Var(X) = E[(X - \mu)^2]$$

La raíz cuadrada positiva de la varianza se llama *desviación estándar*, y está dada por

$$\sigma_X = \sqrt{Var(X)} = \sqrt{E[(X - \mu)^2]}$$

La desviación estándar con frecuencia se denomina σ , en lugar de σ_X , y en tal caso la varianza es σ^2 .

Si X es una variable aleatoria discreta que toma los valores x_1, x_2, \dots, x_n , y tiene función de probabilidad $f(x)$, entonces la *varianza* está dada por

$$\sigma_X^2 = E[(X - \mu)^2] = \sum_{j=1}^n (x_j - \mu)^2 f(x_j) = \sum (x - \mu)^2 f(x) \quad (12.2)$$

En el caso especial de (12.2) donde las probabilidades son iguales, se tiene que

$$\sigma^2 = [(x_1 - \mu)^2 + (x_2 - \mu)^2 + \cdots + (x_n - \mu)^2]/n$$

la cual es la varianza para un conjunto de n números x_1, \dots, x_n .

La varianza (o la desviación estándar) es una medida de la *dispersión*, o de la *variación*, de los valores de la variable aleatoria alrededor de la media μ . Si los valores tienden a concentrarse cerca de la media, la varianza es pequeña, y si los valores tienden a distribuirse lejos de la media, la varianza es grande.

12.3. La distribución normal

Entre todas las distribuciones de probabilidad de variables aleatorias absolutamente continuas ésta es la de mayor importancia [Cuadras 88].

Función de densidad

Una variable aleatoria X con distribución absolutamente continua se dice que tiene una *distribución normal o de Laplace-Gauss* si tiene como función de densidad

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty \quad (12.3)$$

donde $\pi = 3,14159\dots$ y $e = 2,71828\dots$. Esta distribución depende de dos parámetros μ y σ tales que

$$-\infty < \mu < \infty \quad y \quad \sigma > 0$$

Si X tiene una distribución normal de parámetros μ y σ se indica abreviadamente por $X = N(\mu, \sigma)$.

Para que la función $f(x)$ dada por (12.3) sea una *función de densidad* de probabilidad se han de verificar las dos condiciones siguientes:

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x) dx = 1$

A veces la función de densidad de la distribución normal se escribe así:

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

indicando con $f(x; \mu, \sigma)$ que la distribución normal depende de los parámetros μ y σ .

La fórmula (12.3) da realmente una *clase o familia de distribuciones normales*, puesto que para cada valor diferente de μ o σ y también para cada pareja de valores diferentes de μ y de σ , se obtiene una distribución normal diferente de la familia.

Una vez especificadas μ y σ la curva normal está completamente determinada.

En la Figura 12.1 se muestra el resultado de trazar dos curvas normales que tienen diferentes medias y desviaciones típicas. Observemos que están centradas en posiciones diferentes sobre el eje horizontal y que sus formas reflejan los dos valores diferentes de σ .

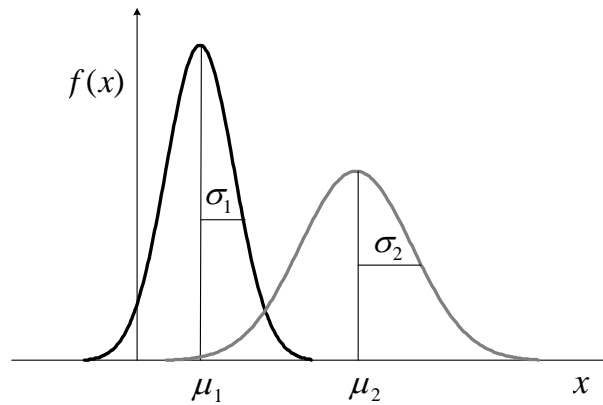


Figura 12.1: Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 < \sigma_2$.

La gráfica de la función de densidad

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

verifica:

- $f(x)$ posee un *máximo absoluto* en el punto $(\mu, \frac{1}{\sigma\sqrt{2\pi}})$.
- La función $f(x)$ tiene una asíntota horizontal que es el eje x ; es decir, la curva normal se aproxima asintóticamente al eje horizontal a medida que nos alejamos de ella en cualquier dirección, ya que

$$\lim_{x \rightarrow \infty} f(x) = 0$$

- Por ser

$$f(\mu - \epsilon) = f(\mu + \epsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\epsilon^2}{2\sigma^2}}$$

resulta que la gráfica de $f(x)$ es simétrica respecto de la recta $x = \mu$.

- Calculando la segunda derivada se ve que la curva presenta dos puntos de inflexión para los valores de las abscisas: $x = \mu - \sigma$ y $x = \mu + \sigma$. La curva es cóncava hacia abajo si $\mu - \sigma < X < \mu + \sigma$, y cóncava hacia arriba en caso contrario.
- El valor del área encerrada bajo la curva y arriba del eje x es igual a 1.

Propiedades

1. La distribución normal depende de los parámetros μ y σ^2 , precisamente iguales a su esperanza y a su varianza

$$E(x) = \mu$$

$$\text{var}(X) = \sigma^2$$

2. Si X es una variable aleatoria normal $N(\mu, \sigma)$ y $a \neq 0$ es una constante, la variable aleatoria $Y = aX + b$ es $N(a\mu + b, |a|\sigma)$.
3. Si una variable X_1 es $N(\mu_1, \sigma_1)$ y otra variable X_2 es $N(\mu_2, \sigma_2)$ y ambas son independientes entre sí, entonces la nueva variable $X = X_1 \pm X_2$ sigue también una distribución normal con media $\mu = \mu_1 \pm \mu_2$ y desviación típica $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$ es decir, X es $N(\mu_1 \pm \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$

Esta propiedad se puede generalizar a n variables aleatorias independientes.

12.4. Variables aleatorias y distribución de probabilidad

En esta sección se definen los conceptos de variable aleatoria, distribución de probabilidad discreta y la función de distribución conjunta [Spiegel 01].

Variables aleatorias

Se supone que a cada punto del espacio muestral se le asigna un número. Se tiene entonces una *función* definida en el espacio muestral. Esta función recibe el nombre de *variable aleatoria* (o *variable estocástica*). Usualmente se denota con una letra mayúscula, como X o Y .

Una variable aleatoria que toma un número finito o contable infinito de valores se llama una *variable aleatoria discreta*, mientras que una que toma un número de valores infinito no contable se llama *variable aleatoria no discreta*.

Distribuciones de probabilidad discreta

Sea X una variable aleatoria discreta, y se supone que los valores posibles que ésta puede asumir están dados por x_1, x_2, x_3, \dots , con determinado orden. Se supone también que estos valores se asumen con probabilidades dadas por

$$P(X = x_k) = f(x_k) \quad k = 1, 2, \dots \quad (12.4)$$

Es conveniente presentar la *función de probabilidad*, también denominada *distribución de probabilidad*, dada por

$$P(X = x) = f(x) \quad (12.5)$$

Para $x = x_k$, se reduce a (12.4), mientras que para otros valores de x , $f(x) = 0$.

En general, $f(x)$ es una función de probabilidad si

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

donde la suma en (12.5) toma todos los valores posibles de x .

Tabla 12.1: *Función de probabilidad conjunta para X y Y .*

	y_1	y_2	\cdots	y_n	Totales \downarrow
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	\cdots	$f(x_1, y_n)$	$f_1(x_1)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	\cdots	$f(x_2, y_n)$	$f_1(x_2)$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	\cdots	$f(x_m, y_n)$	$f_1(x_m)$
Totales \rightarrow	$f_2(y_1)$	$f_2(y_2)$	\cdots	$f_2(y_n)$	1

Distribuciones conjuntas. Caso discreto

Si X y Y son dos variables aleatorias discretas, se define la función de probabilidad conjunta de X y Y por

$$P(X = x, Y = y) = f(x, y)$$

donde

1. $f(x, y) \geq 0$
2. $\sum_x \sum_y f(x, y) = 1$

es decir, la suma de todos los valores de x y y es igual a uno.

Se supone que X puede asumir cualquiera de los m valores x_1, \dots, x_m y Y puede asumir cualquiera de los n valores y_1, \dots, y_n . Entonces la probabilidad del evento $X = x_j$ y $Y = y_k$ está dada por

$$P(X = x_j, Y = y_k) = f(x_j, y_k)$$

La función de probabilidad conjunta para X y Y se puede representar en una tabla de probabilidad conjunta, como es el ejemplo de la Tabla 12.1. La probabilidad de $X = x_j$ se obtiene sumando todas las entradas en la fila correspondiente a x_j , y está dada por

$$P(X = x_j) = f_1(x_j) = \sum_{k=1}^n f(x_j, y_k) \quad (12.6)$$

Para $j = 1, 2, \dots, m$, éstas se indican por la entrada de totales en la columna de la extrema derecha o margen de la Tabla 12.1. La probabilidad de $Y = y_k$ se obtiene sumando todas las entradas en la columna correspondiente a y_k y está dada por

$$P(Y = y_k) = f_2(y_k) = \sum_{j=1}^m f(x_j, y_k) \quad (12.7)$$

Para $k = 1, 2, \dots, n$, éstas se indican por la entrada de totales en la hilera inferior o margen de la Tabla 12.1.

Debido a que las probabilidades (12.6) y (12.7) se obtienen a partir de los márgenes de la Tabla, con frecuencia se refiere a $f_1(x_j)$ y $f_2(y_k)$ como funciones de probabilidad marginal de X y Y , respectivamente. También se debe notar que

$$\sum_{j=1}^m f_1(x_j) = 1 \quad \sum_{k=1}^n f_2(y_k) = 1$$

que puede escribirse como

$$\sum_{j=1}^m \sum_{k=1}^n f(x_j, y_k) = 1$$

Esto es tan sólo la proposición de que la probabilidad total de todas las entradas es 1.

La *función de distribución conjunta* de X y Y se define como

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} f(u, v)$$

En la Tabla 12.1, $F(x, y)$ es la suma de todas las entradas para las cuales $x_j \leq x$ y $y_k \leq y$.

Variables aleatorias independientes

Se supone que X y Y son variables aleatorias discretas. Si los eventos $X = x$ y $Y = y$ son independientes para todo x y y , entonces se dice que X y Y son *variables aleatorias independientes*. En tal caso,

$$P(X = x, Y = y) = P(X = x) P(Y = y)$$

o de manera equivalente

$$f(x, y) = f_1(x) f_2(y)$$

Inversamente, si para todo x y y la función de probabilidad conjunta $f(x, y)$ se puede expresar sólo como el producto de una función de x sola y como una función de y sola (las cuales son entonces funciones de probabilidad marginal de X y Y), X y Y son independientes. Sin embargo, si $f(x, y)$ no se puede expresar de tal manera, entonces X y Y son *dependientes*.

Parte III

EVALUACIÓN DE LOS MODELOS

Capítulo:

13. Técnicas de evaluación

Capítulo 13

Técnicas de evaluación

Los métodos de aprendizaje permiten construir modelos a partir de un conjunto de datos D y es importante conocer con exactitud el nivel de precisión de estos modelos aprendidos [Hernández 04]. En este capítulo se presentan las técnicas que se utilizan en esta tesis para la evaluación de los modelos, y que se van a aplicar con los datos disponibles, que son los correspondientes a los años 2010, 2011 y 2012.

13.1. Diferentes técnicas de evaluación

Se abordan diferentes técnicas de evaluación basadas en la partición de los datos. En las siguientes páginas me referiré a E como el subconjunto de datos correspondientes a los años 2010 y 2011, y a T como el subconjunto de datos del año 2012.

13.1.1. Evaluación mediante validación cruzada

En este método los datos del subconjunto E se dividen aleatoriamente en k subconjuntos disjuntos de similar tamaño [Hernández 04] (ver Figura 13.1):

- Se aprende una hipótesis utilizando el subconjunto formado por la unión de $k-1$ subconjuntos y el subconjunto restante se emplea para calcular un error de muestra parcial (este último subconjunto hace de validación).
- Este procedimiento se repite k veces, utilizando siempre un subconjunto diferente para estimar el error de muestra parcial.
- El error de muestra final se calcula como la media aritmética de los k errores de muestras parciales.

13.1.2. Evaluación con entrenamiento, validación y test

Un mecanismo para calcular el error de muestra de una hipótesis con respecto a una evidencia [Hernández 04], consiste en separar el conjunto de datos D en dos subconjuntos disjuntos (ver Figura 13.2):

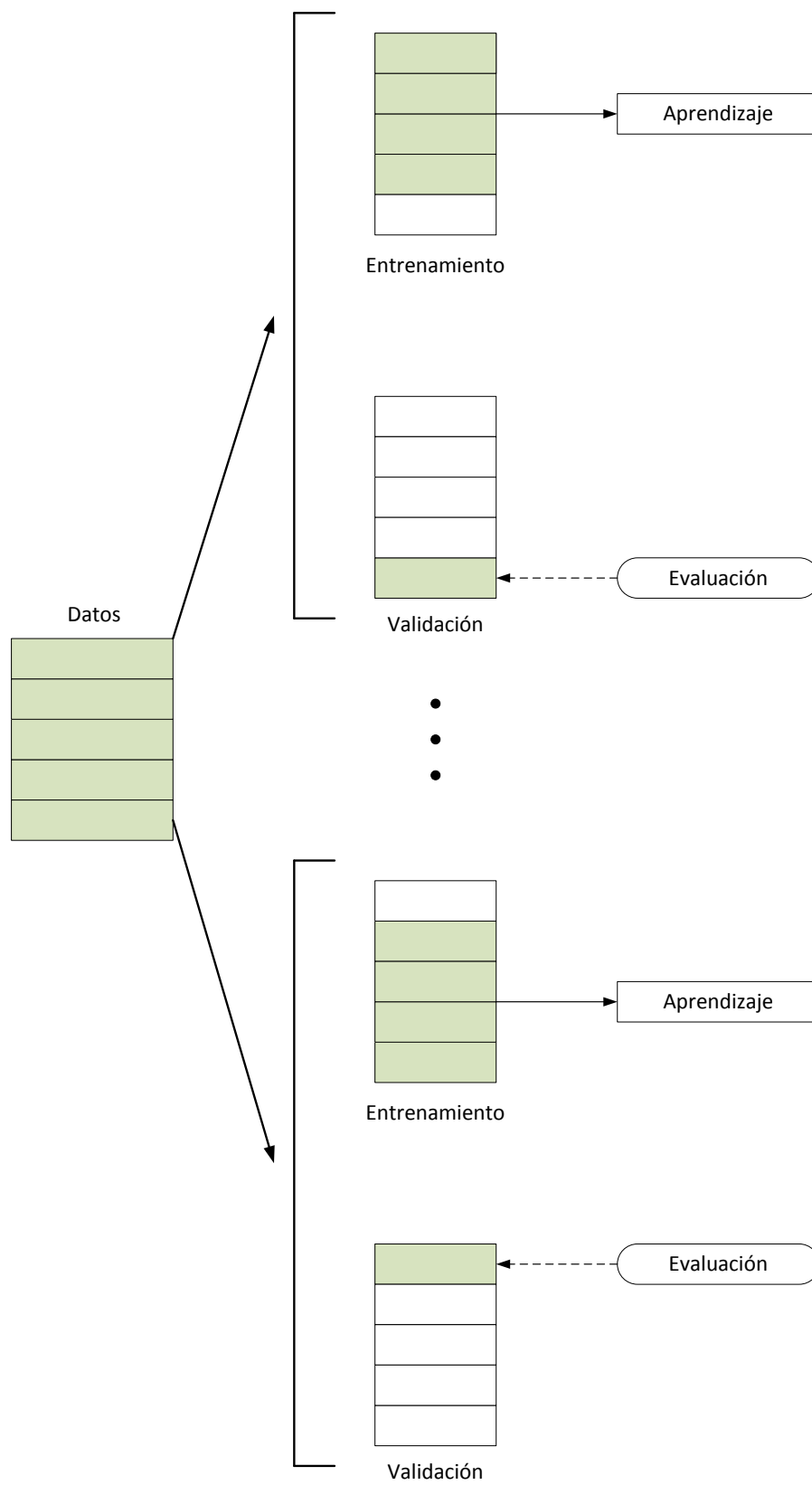


Figura 13.1: *Evaluación mediante validación cruzada (adaptado de [Hernández 04]).*

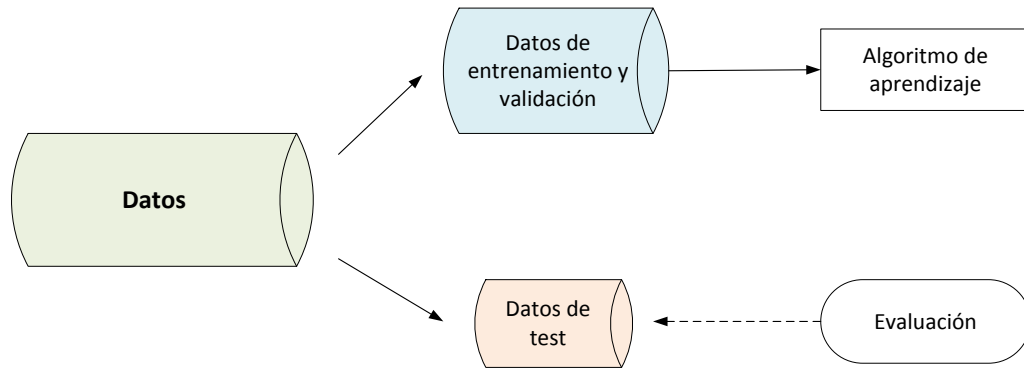


Figura 13.2: *Evaluación de una hipótesis utilizando dos subconjuntos de datos (adaptado de [Hernández 04]).*

- El subconjunto de datos E que se divide en dos subconjuntos. El primero es el subconjunto de datos de entrenamiento E_n , se utiliza para el aprendizaje de las hipótesis, y el segundo es el subconjunto de datos de validación V , que algunos modelos utilizan para elegir entre posibles modelos antes de la salida final del algoritmo.
- El subconjunto de test T sólo se emplea para calcular el error de la muestra de las hipótesis construidas con los datos entrenamiento.

13.1.3. Evaluación de los modelos en los trabajos de esta tesis

Para la evaluación de los modelos en los trabajos desarrollados en esta tesis se utilizan datos horarios desde el 1 de enero de 2010 hasta el 28 de noviembre de 2012, y se reparten los datos en diferentes subconjuntos (ver Tabla 13.1). Durante este período solo se han eliminado unos pocos días enteros, correspondientes a los días en que los datos de la serie de potencia eléctrica media horaria generada en la planta fotovoltaica eran cero. Estos datos nulos se pueden deber a que el aparato de medición de la potencia no funcionó correctamente o la instalación no estaba operativa por tareas de mantenimiento.

Modelos de días con cielo claro

Los días con cielo claro se eligen representando las gráficas con los valores de la potencia eléctrica media horaria de cada día agrupándolos por meses. En total se seleccionan 289 días con cielo claro comprendidos entre el 1 de enero de 2010 y el 31 de diciembre de 2011, y 133 días con cielo claro comprendidos entre el 1 de enero de 2012 y el 28 de noviembre de 2012.

En los modelos desarrollados basados en técnicas con reglas, árboles o redes neuronales, para los datos horarios de los días con cielo claro comprendidos entre el 1 de enero de 2010 y el 31 de diciembre de 2011, se utiliza la técnica de validación cruzada de 5-iteraciones (también conocida como de 5 k-folds). Es decir, se divide aleatoriamente el subconjunto de datos de entrenamiento $Claros_E$ en $k = 5$ subconjuntos disjuntos de similar tamaño.

Tabla 13.1: Descripción de los subconjuntos de datos.

Modelos de	Subconjuntos	Datos horarios comprendidos entre	(%)
Días	<i>Claros_E</i>	01-01-2010 hasta el 31-12-2011	(100 %)
con	<i>Claros_En</i>	01-01-2010 hasta el 31-12-2011	(75 %)
cielo	<i>Claros_V</i>	01-01-2010 hasta el 31-12-2011	(25 %)
claro	<i>Claros_T</i>	01-01-2012 hasta el 28-11-2012	(100 %)
Cualquier	<i>E</i>	01-01-2010 hasta el 31-12-2011	(100 %)
tipo	<i>En</i>	01-01-2010 hasta el 31-12-2011	(75 %)
de	<i>V</i>	01-01-2010 hasta el 31-12-2011	(25 %)
día	<i>T</i>	01-01-2012 hasta el 28-11-2012	(100 %)

Entonces, se aprende una hipótesis utilizando el subconjunto formado por la unión de 4 subconjuntos y el otro subconjunto se emplea para calcular el error de cada elemento de la muestra parcial. Este procedimiento se repite 5 veces, utilizando siempre un subconjunto distinto para estimar el error de cada elemento de la muestra parcial (ver Figura 13.1). El error de la muestra final se calcula como la raíz cuadrada del promedio de la suma de los errores (la potencia real menos la potencia pronosticada) al cuadrado, de cada uno de los elementos de los 5 subconjuntos. Es decir, se calcula la raíz cuadrada del error cuadrático medio (RMSE). Para cada una de estas técnicas, se elige el modelo de días con cielo claro que proporciona menor valor del *RMSE* con los datos del subconjunto de entrenamiento *Claros_E*, y éste se aplica para calcular el valor del RMSE con los datos del subconjunto de test, *Claros_T*.

En los tipos de modelos basados en ANFIS, los datos horarios de los días con cielo claro comprendidos entre el 1 de enero de 2010 y el 31 de diciembre de 2011 se dividen y se crean dos subconjuntos, entrenamiento *Claros_En* y validación *Claros_V* (ver Figura 13.2). Para ello, estos datos se separan de la siguiente manera, los primeros 3 días que sean claros para el subconjunto *Claros_En* y el siguiente día que sea claro para el subconjunto *Claros_V*, y así hasta el último día que sea claro con estos datos. Se calcula la ponderación, *Claros_P*, del *RMSE* de los subconjuntos *Claros_En* y *Claros_V*, de la siguiente forma:

$$Claros_P = \sqrt{\frac{3(RMSE_{Claros_En})^2 + (RMSE_{Claros_V})^2}{4}} \quad (13.1)$$

Se elige el modelo que tiene el menor valor *Claros_P*, y éste se aplica para calcular el valor del RMSE con los datos del subconjunto de test, *Claros_T*.

Modelos de cualquier tipo de día

En los modelos desarrollados basados en técnicas con reglas, árboles, redes neuronales, máquinas de vectores soporte o HISIMI, para los datos horarios comprendidos entre el 1 de enero de 2010 y el 31 de diciembre de 2011 se utiliza la técnica de validación cruzada de 5-iteraciones (también conocida como de 5 k-folds). Es decir, se divide aleatoriamente

el subconjunto de datos de entrenamiento E en $k = 5$ subconjuntos disjuntos de similar tamaño. Entonces, se aprende una hipótesis utilizando el subconjunto formado por la unión de 4 subconjuntos y el otro subconjunto se emplea para calcular el error de cada elemento de la muestra parcial. Este procedimiento se repite 5 veces, utilizando siempre un subconjunto distinto para estimar el error de cada elemento de la muestra parcial (ver Figura 13.1). El error de la muestra final se calcula como la raíz cuadrada del promedio de la suma de los errores (la potencia real menos la potencia pronosticada) al cuadrado, de cada uno de los elementos de los 5 subconjuntos. Es decir, se calcula la raíz cuadrada del error cuadrático medio (RMSE). Para cada una de estas técnicas se elige el modelo que proporciona menor valor del $RMSE$ con los datos del subconjunto de entrenamiento E , y éste se aplica para calcular el RMSE con los datos del subconjunto de test T .

En los tipos de modelos basados en ANFIS o redes neuronales artificiales optimizadas con un algoritmo genético, los datos horarios comprendidos entre el 1 de enero de 2010 y el 31 de diciembre de 2011 se dividen y se crean dos subconjuntos, entrenamiento En y validación V (ver Figura 13.2). Para ello, estos datos se separan de la siguiente manera, 3 días seguidos para el subconjunto En y el siguiente día para el subconjunto V , y así hasta el último día de estos datos. Se calcula la ponderación, P , del $RMSE$ de los subconjuntos En y V , de la siguiente forma:

$$P = \sqrt{\frac{3(RMSE_{En})^2 + (RMSE_V)^2}{4}} \quad (13.2)$$

Para cada una de estas técnicas, se elige el modelo con menor valor P , y éste se aplica para calcular el RMSE con los datos del subconjunto de test T .

13.1.4. Elección de los modelos en los trabajos de esta tesis

A continuación, se describe el criterio que se sigue en los trabajos de investigación de esta tesis para elegir el mejor modelo de predicción de la potencia eléctrica media horaria, para cada uno de los diferentes horizontes de predicción estudiados. Para esto se tiene en cuenta, que de los datos que se disponen, se utilizan los de los años 2010 y 2011 para el subconjunto de entrenamiento y los del año 2012 para el subconjunto de test, es decir, aproximadamente el doble para entrenamiento que para test.

Modelos de días con cielo claro

Para cada horizonte de predicción estudiado, de todos los modelos de días con cielo claro seleccionados (uno de cada técnica estudiada) se elige como mejor modelo el que proporciona el menor valor del RMSE medio ponderado con los datos de los subconjuntos de entrenamiento $Claros_E$ y de test $Claros_T$ calculado de la siguiente forma:

$$Claros_Pond = \sqrt{\frac{2(RMSE_{Claros_E})^2 + (RMSE_{Claros_T})^2}{3}} \quad (13.3)$$

Modelos de cualquier tipo de día

Para cada horizonte de predicción estudiado, de todos los modelos de cualquier tipo de día seleccionados (uno de cada técnica estudiada) se elige como mejor modelo el que

Tabla 13.2: Notaciones (adaptado de [Madsen 04]).

Notación	Descripción
P_{inst}	Capacidad o potencia eléctrica instalada en la planta fotovoltaica
$m = 1, 2, \dots, m_{max}$	Horizonte de predicción
m_{max}	Horizonte de predicción máximo
N	Número de datos utilizados para la evaluación del modelo
$P(t + m)$	Medida de la potencia eléctrica media horaria real, generada en la planta fotovoltaica en el tiempo $t + m$
$\hat{P}(t + m t)$	Potencia eléctrica media horaria pronosticada por un modelo para el tiempo $t + m$, hecha en el tiempo t
$e(t + m t)$	Error predicción correspondiente en el tiempo $t + m$, para la predicción hecha en el tiempo t
$\epsilon(t + m t)$	Error de predicción normalizado correspondiente en el tiempo $t + m$, para la predicción hecha en el tiempo t

proporciona el menor valor del RMSE medio ponderado con los datos de los subconjuntos de entrenamiento E y de test T calculado de la siguiente forma:

$$Pond = \sqrt{\frac{2(RMSE_E)^2 + (RMSE_T)^2}{3}} \quad (13.4)$$

13.2. Definiciones de error de predicción

En la Tabla 13.2, se introducen las notaciones que se utilizan para la definición del error en los modelos de predicción de la potencia eléctrica media horaria, generada en una planta fotovoltaica.

En general, en el campo de predicción de las series temporales, el error de predicción se define como la diferencia entre el valor real y el valor de predicción [Madsen 04]:

- El *error de predicción* correspondiente al tiempo $t + m$, para la predicción hecha en el tiempo t

$$e(t + m|t) = P(t + m) - \hat{P}(t + m|t) \quad (13.5)$$

- El *error de predicción normalizado* correspondiente al tiempo $t + m$, para la predicción hecha en el tiempo t

$$\epsilon(t + m|t) = \frac{1}{P_{inst}}(P(t + m) - \hat{P}(t + m|t)) \quad (13.6)$$

Cualquier error de predicción puede ser descompuesto en el error sistemático μ_e y en el error aleatorio ξ_e [Madsen 04]:

$$e = \mu_e + \xi_e, \quad (13.7)$$

donde μ_e es una constante y ξ_e es una variable aleatoria con media cero.

13.3. Definiciones de medidas de error

A continuación, se definen las medidas de error más comúnmente utilizadas (se denota por p al número de parámetros estimados usando los datos, por m el horizonte de predicción, y por N el total de datos en el grupo de evaluación) [Madsen 04]:

- El error medio

$$ME(m) = \hat{\mu}_e(m) = \bar{e}(m) = \frac{1}{N} \sum_{t=1}^N e(t + m|t) \quad (13.8)$$

- El error absoluto medio

$$MAE(m) = \frac{1}{N} \sum_{t=1}^N |e(t + m|t)| \quad (13.9)$$

- El error cuadrático medio

$$MSE(m) = \frac{\sum_{t=1}^N (e(t + m|t))^2}{N - p} \quad (13.10)$$

- La raíz cuadrada del error cuadrático medio

$$RMSE(m) = \sqrt{MSE(m)} = \sqrt{\frac{\sum_{t=1}^N (e(t + m|t))^2}{N - p}} \quad (13.11)$$

- La desviación estándar de los errores

$$SDE(m) = \left(\frac{\sum_{t=1}^N (e(t + m|t) - \bar{e}(m))^2}{N - (p + 1)} \right)^{\frac{1}{2}} \quad (13.12)$$

El error cuadrático medio da más peso a las divergencias mayores que a las menores. Utilizar la raíz cuadrada permite reducir el valor final a la misma magnitud de los errores [Hernández 04]. En [Wei 90, De Lurgio 98, Makridakis 98], o [Armstrong 01] se pueden encontrar otras definiciones de medidas de error.

Todas las medidas de error introducidas anteriormente se pueden calcular usando el error de predicción $e(t + m|t)$ o el error de predicción normalizado $\epsilon(t + m|t)$. El interés del uso de las medidas del error normalizado es producir resultados independientes del tamaño de la planta [Madsen 04].

En los trabajos desarrollados en esta tesis se utiliza como medida de evaluación de los modelos la raíz del error cuadrático medio (RMSE).

13.4. Comparación entre modelos de predicción

En [Madsen 04] se indica que para efectuar una comparación entre dos modelos de predicción, esa comparación debe ser realizada a través de la mejora con respecto a un modelo de referencia. En muchos casos el modelo de referencia escogido es el llamado modelo persistente o de persistencia: $\hat{P}(t + m|t) = P(t)$, donde $\hat{P}(t + m|t)$ es la potencia

pronosticada en el tiempo t , para el tiempo $t+m$ y $P(t)$ la potencia en el momento actual t .

Para la determinación de la mejora se utiliza la ecuación:

$$Imp_{ref, EC}(m) = \frac{EC_{ref}(m) - EC(m)}{EC_{ref}(m)}, \quad (13.13)$$

donde EC es el criterio de evaluación considerado, que en el caso de los trabajos desarrollados en esta tesis ha sido el RMSE.

Otro camino para ilustrar la destreza de los métodos de predicción avanzados, es calcular el coeficiente de determinación R^2 , para cada horizonte de predicción [Madsen 04]:

$$R^2(m) = \frac{MSE_0(m) - MSE(m)}{MSE_0(m)}, \quad (13.14)$$

donde $MSE_0(m)$ es el error cuadrático medio por el modelo promedio global ($\hat{P}_0(t+m|t) = \bar{P}(t)$, donde $\bar{P}(t)$ es el promedio de todas las observaciones disponibles de la potencia eléctrica media horaria en el tiempo t).

Por consiguiente, R^2 se puede interpretar como:

$$R^2 = \frac{\text{variación explicada}}{\text{variación total}} \quad (13.15)$$

que se explica por la recta de regresión de mínimos cuadrados. R mide *como de bien* se ajusta la recta de regresión de mínimos cuadrados a los datos muestrales. Si la variación total se explica *toda* por la recta de regresión, es decir, si $R^2 = 1$ o $R = \pm 1$, se dice que hay una *correlación lineal perfecta* (y en tal caso *regresión lineal perfecta*). Por otro lado, si toda la variación total es no explicada, entonces la variación explicada es cero y así $R = 0$. En la práctica la cantidad R^2 , se ubica entre 0 y 1 [Spiegel 01].

Modelos de persistencia

En los trabajos de investigación desarrollados en esta tesis se consideran distintos modelos de persistencia según sea el horizonte de predicción:

- Para horizontes de predicción con $m \in \{1, \dots, 8\}$ horas

$$\hat{P}(t+m|t) = \begin{cases} \frac{P(t)}{\widehat{PC}(t)} \widehat{PC}(t+m) & \text{si } \widehat{PC}(t) \neq 0 \\ \frac{1}{2} \widehat{PC}(t+m) & \text{si } \widehat{PC}(t) = 0 \end{cases} \quad (13.16)$$

donde:

- $\widehat{PC}(t+m)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora $t+m$ si el día fuese claro.
- $\widehat{PC}(t)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora t si el día fuese claro.
- $P(t)$ el valor real de la potencia eléctrica media horaria en la hora t .

- Para horizontes de predicción con m entre 0 y 24 horas

$$\hat{P}(t + m|t) = P(t + m - 24) \quad (13.17)$$

donde $P(t + m - 24)$ es el valor real de la potencia eléctrica media horaria en la hora $t + m - 24$.

- Para horizontes de predicción con m entre 25 y 48 horas

$$\hat{P}(t + m|t) = P(t + m - 48) \quad (13.18)$$

donde $P(t + m - 48)$ es el valor real de la potencia eléctrica media horaria en la hora $t + m - 48$.

- Para horizontes de predicción con m entre 49 y 72 horas

$$\hat{P}(t + m|t) = P(t + m - 72) \quad (13.19)$$

donde $P(t + m - 72)$ es el valor real de la potencia eléctrica media horaria en la hora $t + m - 72$.

- Para horizontes de predicción con m mayor a 72 horas, si los días son claros, la potencia eléctrica media horaria pronosticada para el próximo día claro es la que corresponde para la misma hora del día pero del último día claro anterior.

Parte IV

NUEVOS MODELOS DE PREDICCIÓN, RESULTADOS COMPUTACIONALES Y CONCLUSIONES

Capítulos:

- 14. Modelos con técnicas de minería de datos
- 15. Modelo con evaluación de incertidumbre
- 16. Resultados computacionales
- 17. Conclusiones

Capítulo 14

Nuevos modelos para la predicción de la potencia eléctrica media horaria con técnicas de minería de datos

En este capítulo se indican las variables o atributos disponibles para los trabajos de investigación de esta tesis. Se describen los métodos utilizados para su selección y normalización. También, se presentan los nuevos modelos desarrollados con distintas técnicas de minería de datos, para obtener las mejores predicciones de la potencia eléctrica media horaria generada en una planta fotovoltaica, con horizontes de predicción a corto plazo. Los modelos se desarrollan para días con cielo claro y para cualquier tipo de día.

14.1. Variables o atributos disponibles

Los diferentes tipos de variables o atributos disponibles son datos horarios comprendidos entre el 1 de enero de 2010 hasta el 28 de noviembre de 2012.

14.1.1. Potencia eléctrica

Serie temporal de la potencia eléctrica media horaria generada en la planta fotovoltaica con una capacidad de 7000 kWp, situada en España, en la provincia de Almería (no se detalla la situación exacta por razones de confidencialidad).

14.1.2. Atributos astronómicos

Valores medios horarios de atributos astronómicos (ver Tabla 14.1).

Tabla 14.1: Atributos con valores astronómicos.

Horizonte	Atributo	Descripción
Horas entre: 0 y años	$\text{sen } \delta$	Seno de la <i>declinación</i>
	$\text{cos } \delta$	Coseno de la <i>declinación</i>
	$\text{sen } \omega$	Seno del <i>ángulo</i> horario
	$\text{cos } \omega$	Coseno del <i>ángulo</i> horario
	$\text{sen } \gamma$	Seno de la <i>altura o elevación</i>
	$\text{cos } \gamma$	Coseno de la <i>altura o elevación</i>
	$\text{sen } \psi$	Seno del <i>azimut</i>
	$\text{cos } \psi$	Coseno del <i>azimut</i>
	$\text{cos } i$	Coseno del ángulo de incidencia
	$hluzsol$	Número de horas de luz solar diarias
	K_d	Factor corrector de la distancia Tierra-Sol

14.1.3. Atributos con predicciones meteorológicas

Valores medios horarios de atributos con predicciones meteorológicas (ver Tabla 14.2) obtenidas con el modelo NWP de mesoscala WRF-NMM.

Tabla 14.2: Atributos con predicciones meteorológicas para los horizontes de predicción entre 0 y 24 horas ($m = 24$), 25 y 48 horas ($m = 48$), 49 y 72 horas ($m = 72$).

Horizonte	Atributo	Descripción
Horas entre: 0 y 24, 25 y 48, 49 y 72	THS_m	Temperatura en superficie (°K)
	TSHLTR_m	Temperatura potencial en superficie (°K)
	TH10_m	Temperatura a 10 m (°K)
	QS_m	Humedad relativa en superficie (entre 0 y 1)
	QSHLTR_m	Humedad potencial relativa en superficie (entre 0 y 1)
	Q10_m	Humedad relativa a 10 m (entre 0 y 1)
	CFRACL_m	Cobertura de nubes a nivel bajo (entre 0 y 1)
	CFRACM_m	Cobertura de nubes a nivel medio (entre 0 y 1)
	CFRACH_m	Cobertura de nubes a nivel alto (entre 0 y 1)
	RSWIN_m	Radiación de onda corta (W/m^2)
	RSWINC_m	Radiación de onda corta de día claro (W/m^2)
	U10_m	Componente u del viento a 10 m de altura
	V10_m	Componente v del viento a 10 m de altura
	VEL10_m	Velocidad del viento a 10 m de altura
	PSHLTR_m	Presión atmosférica a 2 m de altura (Pa)

14.1.4. Atributos calculados

En esta sección se describen las fórmulas que se utilizan para calcular los valores medios horarios y diarios de diferentes atributos que también están disponibles para ser usados como entradas en los modelos de predicción de la potencia eléctrica media horaria (ver Tabla 14.3). Se denota por $\max\{x\} = \max\{x_k\}_{k=1,\dots,J}$ donde x_k es un valor de un atributo x con un total de J valores.

Tabla 14.3: *Atributos calculados.*

Horizonte	Atributo	Descripción	Valor
Horas	$RSWCal_m$	= ecuación (14.1)	Horario
de:	$RSWCalMed_m$	$= \frac{(\sum_{hora=0}^{23} RSWCal_m)}{hluzsol}$	Diario
$m \in \{1, \dots, 8\}$	$RSWCalCuad_m$	$= \left[\frac{(I_e - RSWCal_m)}{\max\{I_e\}} \right]^2$	Horario
	$RSWCalRaiz_m$	$= \sqrt{\frac{\sum_{hora=0}^{23} RSWCalCuad_m}{hluzsol}}$	Diario
	$GBCal_m$	= ecuación (14.2)	Horario
	$GBCalMed_m$	$= \frac{(\sum_{hora=0}^{23} GBCal_m)}{hluzsol}$	Diario
	$GBCalCuad_m$	$= \left[\frac{(G_{ef,\beta} - GBCal_m)}{\max\{G_{ef,\beta}\}} \right]^2$	Horario
	$GBCalRaiz_m$	$= \sqrt{\frac{\sum_{hora=0}^{23} GBCalCuad_m}{hluzsol}}$	Diario
	$POTCalC_m$	= ecuación (14.3)	Horario
	$POTCalC24_m$	= ecuación (14.4)	Horario
Horas	$RSWINMed_24$	$= \frac{(\sum_{hora=0}^{23} RSWIN_24)}{hluzsol}$	Diario
entre:	$RSWINCuad_24$	$= \left[\frac{(RSWINC_24 - RSWIN_24)}{\max\{RSWINC_24\}} \right]^2$	Horario
0 y 24	$RSWINRaiz_24$	$= \sqrt{\frac{\sum_{hora=0}^{23} RSWINCuad_24}{hluzsol}}$	Diario
	$POTCalC24_24$	= ecuación (14.4)	Horario
Horas	$RSWINMed_48$	$= \frac{(\sum_{hora=0}^{23} RSWIN_48)}{hluzsol}$	Diario
entre:	$RSWINCuad_48$	$= \left[\frac{(RSWINC_48 - RSWIN_48)}{\max\{RSWINC_48\}} \right]^2$	Horario
25 y 48	$RSWINRaiz_48$	$= \sqrt{\frac{\sum_{hora=0}^{23} RSWINCuad_48}{hluzsol}}$	Diario
Horas	$RSWINMed_72$	$= \frac{(\sum_{hora=0}^{23} RSWIN_72)}{hluzsol}$	Diario
entre:	$RSWINCuad_72$	$= \left[\frac{(RSWINC_72 - RSWIN_72)}{\max\{RSWINC_72\}} \right]^2$	Horario
49 y 72	$RSWINRaiz_72$	$= \sqrt{\frac{\sum_{hora=0}^{23} RSWINCuad_72}{hluzsol}}$	Diario

Se describen dos modelos de predicción de irradiancia, uno sobre un plano horizontal y otro sobre un plano de inclinación β . Estos modelos se utilizan para calcular las predicciones medias horarias de las irradiancias $RSWCal_m$ (sobre un plano horizontal) y $GBCal_m$ (sobre un plano de inclinación β), para los horizontes de predicción de $m \in \{1, \dots, 8\}$ horas. Estas irradiancias se utilizan como atributos de entrada en mode-

los de predicción de la potencia eléctrica media horaria desarrollados en los trabajos de investigación de esta tesis.

Modelo de predicción de irradiancia sobre un plano horizontal

En los trabajos de investigación desarrollados en esta tesis, para modelos de predicción de potencia eléctrica media horaria que no utilizan como atributos de entradas valores medios horarios de predicciones meteorológicas pronosticadas por el modelo atmosférico, se utiliza un atributo $RSWCal_m(t+m|t)$. Este atributo es el valor medio horario de la irradiancia global total sobre un plano horizontal, pronosticada para el tiempo $t+m$, hecha en el tiempo t , y expresada en (W/m^2) . Para horizontes de predicción con $m \in \{1, \dots, 8\}$ horas se calcula de la siguiente forma:

$$RSWCal_m(t+m|t) = \begin{cases} \frac{P(t)}{\widehat{PC}(t)} I_e(t+m) & \text{si } \widehat{PC}(t) \neq 0 \\ \frac{1}{2} I_e(t+m) & \text{si } \widehat{PC}(t) = 0 \end{cases} \quad (14.1)$$

donde:

- $P(t)$ el valor real de la potencia eléctrica media horaria en la hora t .
- $\widehat{PC}(t)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora t si el día fuese claro.
- $I_e(t+m)$ es el valor pronosticado de la irradiancia media horaria sobre un plano horizontal para la hora $t+m$ si el día fuese claro (ver ecuación 6.1).

Modelo de predicción de irradiancia sobre un plano de inclinación β

Para horizontes de predicción de $m \in \{1, \dots, 8\}$ horas, en los trabajos de investigación desarrollados en esta tesis, se desarrolla un nuevo modelo de predicción de irradiancia media horaria sobre un plano de inclinación β . El valor medio horario de esta irradiancia $GBCal_m(t+m|t)$ pronosticada para el tiempo $t+m$, hecha en el tiempo t y expresada en (W/m^2) , se calcula de la siguiente manera:

$$GBCal_m(t+m|t) = \begin{cases} \frac{P(t)}{\widehat{PC}(t)} G_{ef,\beta}(t+m) & \text{si } \widehat{PC}(t) \neq 0 \\ \frac{1}{2} G_{ef,\beta}(t+m) & \text{si } \widehat{PC}(t) = 0 \end{cases} \quad (14.2)$$

donde:

- $P(t)$ el valor real de la potencia eléctrica media horaria en la hora t .
- $\widehat{PC}(t)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora t si el día fuese claro.
- $G_{ef,\beta}(t+m)$ es el valor pronosticado de la irradiancia media horaria sobre un plano de inclinación β y orientación Sur para la hora $t+m$ si el día fuese claro (ver ecuación 6.28)

Modelos de predicción de potencia

A continuación, se describen dos tipos de modelos muy simples para poder calcular la predicción de la potencia eléctrica media horaria para horizontes de predicción de m horas:

- Los que no utilizan valores medios horarios de predicciones meteorológicas

$$POTCalc_m(t+m|t) = \begin{cases} \frac{P(t)}{\widehat{PC}(t)} \widehat{PC}(t+m) & \text{si } \widehat{PC}(t) \neq 0 \\ \frac{1}{2} \widehat{PC}(t+m) & \text{si } \widehat{PC}(t) = 0 \end{cases} \quad (14.3)$$

donde:

- $\widehat{PC}(t+m)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora $t+m$ si el día fuese claro.
 - $\widehat{PC}(t)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora t si el día fuese claro.
 - $P(t)$ el valor real de la potencia eléctrica media horaria en la hora t .
- Los que utilizan valores medios horarios de predicciones meteorológicas con horizontes de predicción entre 0 y 24 horas

$$POTCalc24_m(t+m|t) = \begin{cases} \frac{P(t)}{\widehat{PC24}(t)} \widehat{PC24}(t+m) & \text{si } \widehat{PC24}(t) \neq 0 \\ \frac{1}{2} \widehat{PC24}(t+m) & \text{si } \widehat{PC24}(t) = 0 \end{cases} \quad (14.4)$$

donde:

- $\widehat{PC24}(t+m)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora $t+m$ si el día fuese claro.
- $\widehat{PC24}(t)$ es el valor pronosticado de la potencia eléctrica media horaria para la hora t si el día fuese claro.
- $P(t)$ el valor real de la potencia eléctrica media horaria en la hora t .

14.2. Selección de las variables o atributos

Una buena selección de las variables o atributos tiene como objetivos principales [Hernández 04]:

- Reducir el tamaño de los datos, al eliminar los atributos irrelevantes o redundantes.
- Mejorar la calidad del modelo, al permitir al método de minería de datos centrarse en las características relevantes.

Existen muchos métodos para seleccionar los atributos [Liu 98]. En los trabajos de investigación de esta tesis, se han utilizado dos tipos de selección [Fernández 08]:

- Primero, se ha creado una relación ordenada de los atributos disponibles (ver Tablas: 14.1, 14.2, 14.3), considerando los atributos uno a uno por modelo. El atributo que ha dado lugar al modelo con menor RMSE se retiene, formándose con él y otro más todos los modelos posibles con dos atributos. El atributo correspondiente al modelo con menor RMSE, se retiene también. Este proceso continúa escogiendo los atributos de acuerdo al criterio seguido con los dos primeros. El proceso de selección se finaliza, cuando no se obtiene una mejora en los resultados del RMSE con los nuevos modelos de predicción. Esta selección se hace para cada uno de los horizontes de predicción a corto plazo.
- Además, para los modelos de predicción basados en redes neuronales artificiales, también se han hecho pruebas seleccionando los atributos disponibles, con algoritmos genéticos.

14.3. Normalización de las variables o atributos

Los rangos de las variables o atributos de entrada disponibles para los modelos de predicción de la potencia eléctrica media horaria son muy diferentes, y para dar un trato similar a todos atributos se pueden normalizar sus valores. Existen varias técnicas de normalización [Fernández 08] para una muestra x_j de un atributo x con un total de J muestras. A continuación, se describen las que se utilizan en los trabajos de investigación de esta tesis:

- Para cada atributo x a normalizar se puede ajustar su valor mínimo al 0 y su valor máximo al 1. Para ello se utiliza

$$x_{j_normalizado} = \frac{x_j - \min\{x_k\}_{k=1,\dots,J}}{\max\{x_k\}_{k=1,\dots,J} - \min\{x_k\}_{k=1,\dots,J}} \quad (14.5)$$

- Se pueden “escalar” los valores de cada atributo x entre -1 y 1. Esta técnica se emplea con la función de activación tangente hiperbólica en una red neuronal. Viene dada por

$$x_{j_normalizado} = 2 \frac{x_j - \min\{x_k\}_{k=1,\dots,J}}{\max\{x_k\}_{k=1,\dots,J} - \min\{x_k\}_{k=1,\dots,J}} - 1 \quad (14.6)$$

- Si se considera que el rango de los datos del subconjunto de entrenamiento no tiene porque incluir todos los valores posibles para cada atributo x , sus valores se pueden normalizar entre -0.9 y 0.9, teniendo en cuenta así un margen tanto por encima como por debajo de los valores del subconjunto de entrenamiento. Esta técnica se utiliza en los modelos con redes neuronales optimizadas con un algoritmo genético. Se calcula de la siguiente forma

$$x_{j_normalizado} = 1.8 \frac{x_j - \min\{x_k\}_{k=1,\dots,J}}{\max\{x_k\}_{k=1,\dots,J} - \min\{x_k\}_{k=1,\dots,J}} - 0.9 \quad (14.7)$$

14.4. Modelos con técnicas de minería de datos

A continuación, se describen los nuevos modelos desarrollados con diferentes técnicas de minería de datos.

14.4.1. Árboles de decisión y sistemas de reglas

Para la elaboración de estos modelos se utiliza WEKA, una herramienta visual de libre distribución (licencia GNU) elaborada y mantenida por un grupo de investigadores de la universidad de Waikato, Nueva Zelanda.

En el desarrollo de los modelos con árboles se usa el algoritmo M5P (ver sección 7.4). Se hacen pruebas considerando diferentes combinaciones de atributos, de entre los disponibles, para cada horizonte de predicción (ver Tablas 14.1, 14.2 y 14.3). Se siguen los siguientes criterios:

- La división con el algoritmo M5P se finaliza porque hay pocas instancias que llegan a un nodo. En las pruebas realizadas, el mínimo número de instancias que se permite en un nodo hoja varía entre 4 y 30.
- Los modelos lineales para un nodo R particular del árbol con k atributos que se dividen por debajo de ese nodo, se calculan como una combinación lineal de esos atributos y se incluye un término constante. Primero, se simplifican los modelos lineales de los nodos sucesores de un determinado nodo y después, se realiza una búsqueda en escalada de forma que en cada iteración se elimina aquel atributo tal que, si se elimina, el modelo lineal restante minimiza el error [Riaño 16].
- El árbol se poda, es decir, se elimina el subárbol de regresión por debajo de cada nodo del árbol de regresión siempre y cuando el error que produce ese subárbol sea mayor o igual que el producido por el del modelo lineal asociado al nodo [Riaño 16].
- En la etapa final, se utiliza un procedimiento de suavizado descrito en [Quinlan 92]. Primero se utiliza el modelo hoja para calcular el valor pronosticado, y luego se filtra ese valor a lo largo del camino de regreso a la raíz, suavizándose en cada nodo mediante la combinación con el valor pronosticado por el modelo lineal para ese nodo. El valor que se emplea para la constante de suavizado es $s = 15$ [Wang 97].

En el desarrollo de los modelos con sistemas de reglas, se utiliza el método M5Rules (ver sección 7.5). En este método se aplica un árbol que utiliza el algoritmo M5P (ver sección 7.4) con el subconjunto completo de datos de entrenamiento y se genera un árbol podado. Se hacen pruebas considerando diferentes combinaciones de atributos, de entre los disponibles para cada horizonte de predicción (ver Tablas 14.1, 14.2 y 14.3) variando el mínimo número de instancias que se permite en un nodo hoja entre 4 y 30. A continuación, la mejor hoja conforme a una heurística, se convierte en una regla y se descarta el árbol. Todas las instancias cubiertas por esa regla se eliminan del conjunto de datos. El proceso se aplica de forma recursiva a las instancias restantes y se termina cuando todas las instancias son cubiertas por una o más reglas. Se siguen los siguientes criterios:

- Se construyen árboles completos y no árboles parcialmente explorados [Holmes 99].

- Se centra en la generalización de reglas que utilizan modelos lineales sin suavizar, ya que el árbol del cual se genera una regla, se desecha en cada etapa [Holmes 99].
- Para determinar en cada etapa, que hoja del árbol es la candidata mejor para añadir al conjunto de reglas, se elige la hoja que cubre la mayoría de los ejemplos [Frank 98].

Los modelos de días con cielo claro se entrenan con los datos del subconjunto de entrenamiento *Claros_E*, y los modelos de cualquier tipo de día con los datos del subconjunto de entrenamiento *E*, se utiliza la técnica de validación cruzada de 5-iteraciones. En la mayoría de las pruebas realizadas con estos modelos de predicción de la potencia eléctrica media horaria para diferentes horizontes de predicción, se obtienen menores valores del RMSE con el mínimo número de instancias que llegan al nodo hoja comprendidas entre 10 y 21, que son los resultados que se muestran en el capítulo 16.

14.4.2. Redes neuronales artificiales

Para la creación de los modelos con redes neuronales artificiales (ver capítulo 8) se utiliza también la herramienta WEKA. Se desarrollan diferentes modelos con redes neuronales perceptrón multicapa (ver sección 8.4), con una y dos capas ocultas, variando el número de neuronas de cada una de ellas de 2 a 30. Las funciones de activación de los nodos de las capas ocultas son de tipo sigmoideo, y en el nodo de la capa de salida se usa la función identidad. Estas redes se entrenan con el algoritmo de retropropagación de errores (BP). Las pruebas se realizan tomando diferentes valores comprendidos entre 0.0 y 0.6 (en intervalos de 0.05) para el término de inercia (*momentum*), β . También se consideran distintos valores comprendidos entre 0.1 y 0.5 (en intervalos de 0.05) para el tamaño del paso α tomado en cada iteración. Además, si el proceso de entrenamiento no converge se permite al sistema modificar la razón de aprendizaje automáticamente. Se normalizan los atributos numéricos de entrada para que estén comprendidos entre -1 y 1. Con este tipo de redes neuronales se hacen pruebas utilizando un porcentaje del 20 % de los datos del subconjunto de entrenamiento *E* como subconjunto de validación. De esta forma, tras cada entrenamiento se valida el sistema, y se detiene el aprendizaje si el proceso con los datos del subconjunto de validación da unos resultados del RMSE que empeoran durante 20 veces consecutivas, o si se supera el número de iteraciones configurado. También, se hacen otras pruebas en las que no se utiliza ningún subconjunto de validación y en su lugar la red se entrena hasta un número especificado de iteraciones. Las pruebas se realizan configurando distintos números de iteraciones (500, 1000, y 2500).

También se desarrollan modelos con otras redes neuronales perceptrón multicapa, con una capa oculta variando su número de neuronas de 2 a 30, y utilizando una versión aproximada de la función de tipo sigmoideo como su función de activación. Estas redes se entrenan con dos métodos distintos (ver subsección 8.4.3) el BFGS y el gradiente conjugado descendente (CGD). En las pruebas se toman diferentes valores comprendidos entre 0.01 y 0.09 para el tamaño del paso α en cada iteración, y diferentes valores comprendidos entre 10^{-3} y 10^{-8} para la tolerancia de la convergencia ϵ . Se normalizan los atributos incluyendo el objetivo.

Otros modelos se desarrollan con redes neuronales de función de base radial (ver sección 8.5), con una capa oculta variando su número de neuronas de 2 a 30. Se calculan las salidas de las neuronas de la capa oculta a partir de una función gaussiana, los centros

iniciales de las funciones de base radial con forma gaussiana se encuentran aplicando el algoritmo de las k -medias [Tou 74], los valores σ_j iniciales se ajustan a la máxima distancia entre su centro y su vecino más cercano en el conjunto de centros, y se utiliza un valor σ_j por atributo. Estas redes se entrenan con dos métodos distintos (ver subsección 8.4.3) el BFGS y el gradiente conjugado descendente (CGD). En las pruebas que se realizan se toman diferentes tamaños para el paso α en cada iteración y distintos valores para la tolerancia de la convergencia ϵ . Todos los atributos incluyendo el objetivo se normalizan entre los valores 0 y 1.

Estos tipos de redes neuronales se entrenan con los datos del subconjunto *Claros_E* si el modelo es de días con cielo claro, y con los datos del subconjunto *E* si el modelo es de cualquier tipo de día, se utiliza la técnica de validación cruzada de 5-iteraciones. Se han probado con diferentes combinaciones de atributos de entradas, entre los disponibles (ver Tablas 14.1, 14.2, 14.3), para seleccionar los mejores. Los menores valores del RMSE en los modelos de predicción de la potencia eléctrica media horaria para diferentes horizontes de predicción, con los días de cielo claro se han obtenido con los atributos de entrada descritos en la Tabla 16.1, con los días de cualquier tipo para horizontes de predicción entre 1 y 8 horas con los atributos descritos en la Tabla 16.2, y para los horizontes de predicción entre 0 y 24, 25 y 48, 49 y 72 horas con los atributos descritos en la Tabla 16.3. Son redes neuronales de una sola capa oculta. En el capítulo 16 se muestran estos valores del RMSE, con el número de neuronas de la capa oculta comprendido entre 10 y 21. Los resultados obtenidos con las redes neuronales perceptrón multicapa (MLP) entrenadas con el algoritmo de retropropagación de errores (BP) se muestran con dos valores diferentes para el término de inercia (*momentum*) $\beta = 0.0$ y $\beta = 0.2$, con el tamaño del paso $\alpha = 0.3$ tomado en cada iteración y entrenadas con 500 iteraciones. Además, también se muestran los resultados obtenidos con las redes neuronales perceptrón multicapa (MLP) y con las funciones de base radial (RBF) entrenadas con los métodos CGD y BFGS, tomando el tamaño del paso en cada iteración $\alpha = 0.01$ y el valor de la tolerancia de la convergencia $\epsilon = 10^{-6}$.

14.4.3. Máquinas de vectores soporte

Para el desarrollo de de estas técnicas de aprendizaje se utiliza de nuevo el *software* WEKA. Se hacen pruebas considerando diferentes combinaciones de atributos, de entre los disponibles, para cada horizonte de predicción (ver Tablas 14.1, 14.2 y 14.3). Se normalizan los valores de estos atributos escalándolos a un intervalo $[0, 1]$.

La aplicación de las máquinas de vectores soporte (ver capítulo 9) en los modelos de predicción de la potencia eléctrica media horaria, para diferentes horizontes de predicción, requiere resolver dos cuestiones: el tipo de función núcleo que se va a utilizar en el modelo, y el margen blando controlado por un parámetro C .

En las pruebas realizadas en los trabajos de investigación de esta tesis, se usan tres tipos de funciones núcleo (ver Tabla 9.1): polinómica, sigmoidal y gaussiana. Para cada una de estas funciones se trata de encontrar la mejor configuración de sus parámetros, junto con:

- El parámetro ε , que mide la finura con que se va a tolerar el error con los datos del subconjunto de entrenamiento.

- Los valores de las variables de holgura, ζ_i , $1 \leq i \leq \ell$, con ℓ el número total de datos en el subconjunto de entrenamiento, que permiten que las restricciones no se cumplan de manera rigurosa. El valor mínimo para estas variables es cero.
- Una constante estrictamente positiva C , que determina la holgura del margen blando.

Para refinar la búsqueda, en los trabajos de investigación de esta tesis, se hace una exploración con muchas combinaciones de valores para estos parámetros. Se observa que para valores del parámetro C muy grandes, a veces el algoritmo de optimización puede no converger. Estos parámetros se optimizan con los datos del subconjunto de entrenamiento E , utilizando la técnica de validación cruzada de 5-iteraciones.

Para las funciones núcleo polinómicas, el parámetro que se optimiza es el grado p . Se hacen pruebas combinando diferentes valores de los parámetros: ε , ζ_i , C , y variando los valores del grado p , pero se consiguen aproximaciones malas.

Para las funciones núcleo sigmoidales, los parámetros que se optimizan son la pendiente de la sigmoide s y su desplazamiento r . Se hacen pruebas combinando diferentes valores de los parámetros: ε , ζ_i , C , y variando los valores de la pendiente de la sigmoide s y su desplazamiento r , pero no se obtienen modelos con buenos resultados.

Por último, para los núcleos gaussianos, el parámetro que se optimiza es, $\gamma = 1/(2\sigma^2)$, con σ la desviación estándar. Se hacen pruebas considerando diferentes combinaciones de valores de los parámetros: ε , ζ_i , C , y variando los valores de γ . Finalmente, se obtienen modelos con buenos resultados. En el capítulo 16 se muestran los valores del RMSE que se obtienen con los valores, $\varepsilon = 0.001$, $\zeta_i = 0.001$, $C \in \{5, 10, 15, 20\}$, y utilizando las funciones núcleo gaussianas con los valores del parámetro $\gamma \in \{0.2, 0.5, 1\}$.

14.4.4. Sistemas de inferencia neuro-difusos

Los modelos de inferencia borrosa (FIS) se han desarrollado con la herramienta informática “fuzzy toolbox” de MATLAB. Su diseño se realiza del tipo TSK (Takagi-Sugeno-Kang) de primer orden [Takagi 85]. En estos sistemas la formulación de la parte consecuente de las reglas son combinaciones lineales (polinomios de primer grado) de los atributos de entradas del sistema. Se hacen pruebas considerando diferentes combinaciones de atributos, de entre los disponibles para cada horizonte de predicción (ver Tablas 14.1, 14.2 y 14.3). Todos los atributos de entrada se normalizan a valores entre 0 y 1.

La partición del espacio de entrada se realiza con el método de agrupamiento subtractivo [Chiu 94, Jang 97] (se calcula el número de *clusters* durante el proceso de agrupamiento). Cada vector de valores de atributos de entrada al sistema de inferencia borrosa, se toma como un punto de un espacio multi-dimensional. Como primer centro de agrupación se selecciona el punto de datos con la medida de densidad más alta (con mayor número de puntos vecinos cercanos). Después se revisa la medida de densidad para todos los puntos de datos. El algoritmo busca otro nuevo centro, el siguiente con mayor densidad. Este proceso se repite hasta que se genera un número suficiente de centros de *clusters*.

Para cada atributo de entrada, se toma el mismo número de conjuntos borrosos que el número de *clusters*. Los centros calculados por el algoritmo subtractivo se toman como centros de las funciones de pertenencia gaussianas de la parte antecedente de las reglas.

La desviación estándar, σ , de las funciones de pertenencia gaussianas se toman iguales para todos los conjuntos difusos de un mismo atributo de acuerdo al valor [Fernández 08]:

$$\sigma_{ij} = \text{Ra-i} \frac{\max(x_j) - \min(x_j)}{\sqrt{8}} \quad (14.8)$$

donde:

- σ_{ij} es la desviación estándar de la función de pertenencia del conjunto difuso i del atributo j .
- x_j es el atributo de entrada j .
- Ra-i es el rango de influencia que se define para el agrupamiento substractivo.

Se combina en una misma estructura la representación del conocimiento de un sistema de inferencia borrosa con la capacidad de aprendizaje supervisado de una red neuronal, lo que conduce a un sistema neuro-difuso [Fernández 08]. En [Jang 97] estos sistemas se denominan *Adaptive Neuro-Fuzzy Inference Systems*, ANFIS.

En los modelos de días con cielo claro, el sistema de inferencia borrosa TSK en forma de red adaptativa, se entrena con los datos del subconjunto *Claros_En*, con el número de iteraciones necesarias para alcanzar el mínimo con los datos del subconjunto de validación *Claros_V*. En los modelos de cualquier tipo de día, el sistema de inferencia borrosa TSK en forma de red adaptativa, se entrena con los datos del subconjunto *En*, con el número de iteraciones necesarias para alcanzar el mínimo con los datos del subconjunto de validación *V*. Para ello se utiliza una técnica híbrida [Jang 97] que combina el método de los mínimos cuadrados con el descenso por el gradiente.

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 funciones de pertenencia para los conjuntos definidos en los atributos de entrada [Fernández 08]. Los parámetros de los modelos de predicción corresponden a los que presentan los modelos cuando el RMSE con los datos del subconjunto de validación *Claros_V* o *V* es menor.

Para el desarrollo de estos modelos se han hecho pruebas variando en los parámetros del agrupamiento substractivo el rango de influencia $\text{Ra-i} \in \{0.50, 0.55, \dots, \sqrt{n_atrib}\}$, donde n_atrib es el número de atributos de entrada usados por el modelo, y el resto de los parámetros se mantienen constantes: 1.25 para el factor de reducción de densidad, 0.5 para el factor de aceptación, y 0.15 para el factor de rechazo. Se observa que con un valor pequeño del Ra-i se obtiene un excesivo número de centros, y esto puede ocasionar un ajuste excesivo a los datos del subconjunto de entrenamiento. Se comprueba que para valores del rango de influencia, $\text{Ra-i} > 2.1$, los valores que se obtienen del RMSE son generalmente mayores. En el capítulo 16 se muestran los valores del RMSE que se obtienen con estos modelos para el $\text{Ra-i} \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$.

14.4.5. Algoritmos genéticos y redes neuronales artificiales

En los trabajos de investigación de esta tesis también se realizan pruebas desarrollando modelos de predicción de la potencia eléctrica media horaria utilizando diferentes tipos de redes neuronales optimizadas con un algoritmo genético. Para el desarrollo de

los modelos se utiliza la herramienta NeuroSolutions, se puede ver una descripción detallada en [Principe 00]. Con esta herramienta que posee una interfaz gráfica se crean diferentes tipos de redes neuronales artificiales como: perceptron multicapa (MLP) (ver sección 8.4), multicapa generalizada (FFBP) [Fernández 08], modular [Fernández 08], Elman [Elman 90], con análisis de componentes principales (PCA) [Jolliffe 02], de funciones de base radial (RBF) (ver sección 8.5), de mapas de rasgos autoorganizados (SOFM) [Kohonen 82a, Kohonen 82b, Kohonen 89, Ritter 91], con retardo en el tiempo (TDNN) [Fernández 08] y recurrentes [Fernández 08].

En las pruebas realizadas al comienzo se fijan el número de capas ocultas (una y dos), los tipos de funciones de activación que se utilizan en ellas son: sigmoideal y tangente hiperbólica. Se utiliza la función lineal en la capa de salida, y los pesos iniciales son aleatorios.

Para la optimización de las redes neuronales se establece un total de 50 generaciones con 50 individuos. Todos los individuos de la primera generación se crean aleatoriamente. De una generación a otra se aplica elitismo así se conserva el mejor individuo de la generación anterior. Se seleccionan de forma aleatoria los individuos de la población actual, pero para otorgar mayor probabilidad de ser seleccionados los mejores individuos se utiliza el método de la ruleta. El cruce que se emplea es por un punto. Las tasas de cruce y de mutación que se usan son: 90 % por cruce y un 1 % por mutación.

A través del algoritmo se determina:

- La selección óptima de los atributos de entrada, de entre los disponibles (ver Tablas 14.1, 14.2, 14.3), que aportan mejor información para la predicción.
- El número óptimo de neuronas en las capas ocultas (entre 2 y 50).
- El valor óptimo de los parámetros de los algoritmos de entrenamiento utilizados. En el caso de utilizar la técnica de retropropagación de error los valores de factor de aprendizaje y el momento de cada capa.
- El número de centros en las redes neuronales RBF.
- El número de valores previos de cada entrada que se toman como entrada adicional y en la capa oculta de las redes neuronales con retardo en el tiempo.
- El número de componentes principales en las redes neuronales PCA.

Estos modelos se entrenan con los datos del subconjunto de entrenamiento En , y el proceso de entrenamiento se interrumpe si transcurre 100 epochs sin que se reduzca el valor del RMSE con los datos del subconjunto de validación V . Los valores de los parámetros optimizados con un algoritmo genético corresponden a los que proporcionan el mínimo valor del RMSE con los datos del subconjunto de validación V .

Una vez optimizadas las redes neuronales se determinan los valores de los RMSE con los datos del subconjunto de test T . En las pruebas que se han realizado con diferentes tipos de redes neuronales optimizadas con un algoritmo genético, el menor valor RMSE se obtiene con una red neuronal MLP.

Aun así, el valor ponderado, P , con los datos de los subconjuntos de entrenamiento En y de validación V , es mayor que el que se obtiene con otros modelos que se han desarrollado en la subsección 14.4.2.

Capítulo 15

Modelo con evaluación de incertidumbre

En este capítulo se describe el modelo de predicción desarrollado como trabajo de investigación en el periodo de formación del doctorado. En el estado del arte se ha visto que hasta el año 2013 no se había hecho nada en el campo de evaluar la incertidumbre asociada a la predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica. Precisamente uno de los primeros trabajos de investigación fue el desarrollo de un modelo que pudiera evaluar la incertidumbre asociada a los valores previstos. Nace así el modelo denominado *HISIMIM* (*Historical Similar Mining*), basado en el modelo desarrollado para la predicción de la producción eléctrica a corto plazo para un parque eólico [Santos 09]. En [Santos 09] se desarrolla un modelo original de predicción con evaluación de la incertidumbre, basado en la división de la potencia generada en intervalos y el estudio de la probabilidad de la potencia generada que estaría en cada uno de los intervalos en que se ha dividido esa potencia.

En los trabajos desarrollados en esta tesis, se modificó ese modelo para adaptarlo a la predicción de la planta fotovoltaica y se optimizó el modelo con el uso de un algoritmo genético, que permite la selección de las variables dentro de todas las disponibles y selecciona el número óptimo de intervalos de potencia. Los primeros resultados de este modelo fueron publicados en [Monteiro 13a].

15.1. Base de conocimiento

En el modelo [Monteiro 13a] se utiliza una base con datos históricos de una planta fotovoltaica en estudio, la cual contiene los registros durante el periodo de un año con los valores de:

- Producción histórica real de la potencia eléctrica media horaria generada en una planta solar fotovoltaica de 2.8 MWp de capacidad, ubicada en España, y adquiridos por el sistema SCADA.
- Predicciones medias horarias de la radiación de onda corta y de la temperatura

Tabla 15.1: Representación de una r -tupla de transición.

Transiciones NWP					Transiciones de Potencia (SCADA)	
$V_{j-1,1}$	$V_{j,1}$	\dots	$V_{j-1,q}$	$V_{j,q}$	P_{j-1}	P_j

en la superficie, proporcionadas por el modelo mesoescala *Weather Research and Forecasting* (WRF).

- *Seno y coseno* de la hora solar en la localización de la planta en el instante correspondiente.

En el capítulo de resultados se pueden encontrar los obtenidos con el modelo *HISIMI* utilizando, entre las variables de entrada, las predicciones meteorológicas proporcionadas por el modelo WRF correspondientes al lugar donde se encuentra la planta fotovoltaica en que se han basado todos los resultados de esta tesis.

Dado que este modelo de predicción se basa en la probabilidad de transición de un valor de potencia generada a otro, se utiliza un sistema especial de registros que permite trabajar con dichas transiciones. Un registro, que constituye un caso histórico en un instante j contiene dos valores previstos para cada una de las variables de entrada y también dos valores reales correspondientes a la variable salida de potencia. Estos pares de valores son los correspondientes a dos instantes consecutivos, $j - 1$ y j . Así, un registro puede estar formado por los siguientes valores:

- $V_{j-1,1}$ valor de la variable 1, pronosticado para el instante $j - 1$.
- $V_{j,1}$ valor de la variable 1, pronosticado para el instante j .
- \dots
- $V_{j-1,q}$ valor de la variable q , pronosticado para el instante $j - 1$.
- $V_{j,q}$ valor de la variable q , pronosticado para el instante j .
- P_{j-1} valor de generación de potencia eléctrica para el instante $j - 1$.
- P_j valor de generación de potencia eléctrica para el instante j .

Es decir, a una transición de variables NWP se encuentra asociada una transición de potencia, formando así una r -tupla de transición con $r = 2q + 2$, siendo q el número total de variables de entrada al modelo, tal y como se muestra en la Tabla 15.1.

Así, se puede concluir que la base de conocimiento del modelo es constituida por un conjunto de r -tuplas de transición. Con la utilización de este tipo de registros se establece el modelo de predicción probabilístico optimizado con algoritmo genético.

15.2. Algoritmo genético

El modelo desarrollado se basa en un algoritmo genético donde los individuos (posibles soluciones) se codifican por medio de cadenas numéricas que representan sus cromosomas y mediante un proceso iterativo en el que intervienen los operadores genéticos de selección, cruce y mutación, se van obteniendo soluciones cada vez mejores hasta dar con una solución satisfactoria.

En la Figura 15.1 se presenta un diagrama del procedimiento general de este algoritmo genético. Se puede observar que se parte de una población inicial de individuos, típicamente aleatorios, donde un individuo se entiende como una solución posible (no necesariamente buena). Cada individuo se representa entonces con un cromosoma compuesto por genes donde cada gen hace referencia a una parte o secuencia de tal solución. Luego, dichos individuos son evaluados mediante una determinada función de aptitud y se aplican unos operadores genéticos para obtener una nueva población (siguiente generación). El objetivo de estos operadores es preservar los cromosomas, o parte de ellos, que representen mejores soluciones siguiendo para esto los principios de selección natural, de manera que aquellos individuos más aptos tienen más posibilidades de preservar sus genes sobre aquellos que no.

15.2.1. Representación de los individuos

Cada individuo en el algoritmo genético del modelo se entiende como una solución factible del problema analizado y se va a codificar en forma de una cadena numérica, con una representación en formato binario, así la cadena estará compuesta por un conjunto de bits. A su vez, esta cadena se considera como un cromosoma compuesto por 3 genes, tal y como se representa en la siguiente expresión:

$$S_1 S_2 S_3$$

donde S_i con $i \in \{1, 2, 3\}$ representan los diferentes parámetros o características que definen la solución de la cadena. En este modelo un individuo binario, ver Figura 15.2, codifica los siguientes 3 parámetros:

Entradas

Este parámetro tiene tantos bits aleatorios como el número de entradas disponibles para el modelo. Si el bit i -ésimo del gen es un 1, significa que la entrada i es usada por el modelo, mientras que si es un 0 no lo será. Por lo menos debe haber un 1 en los bits correspondientes a las entradas, ya que el modelo debe usar al menos una entrada.

Intervalos

Se ha comentado anteriormente que el modelo se basa en transiciones entre valores de la potencia generada en la planta fotovoltaica. Ahora bien, pueden existir infinitas transiciones posibles entre dos valores dentro del rango establecido por la capacidad de la planta. Para poder acotar el problema, definimos un conjunto de intervalos de igual anchura para el rango de generación de la planta, así no existirán infinitas transiciones,

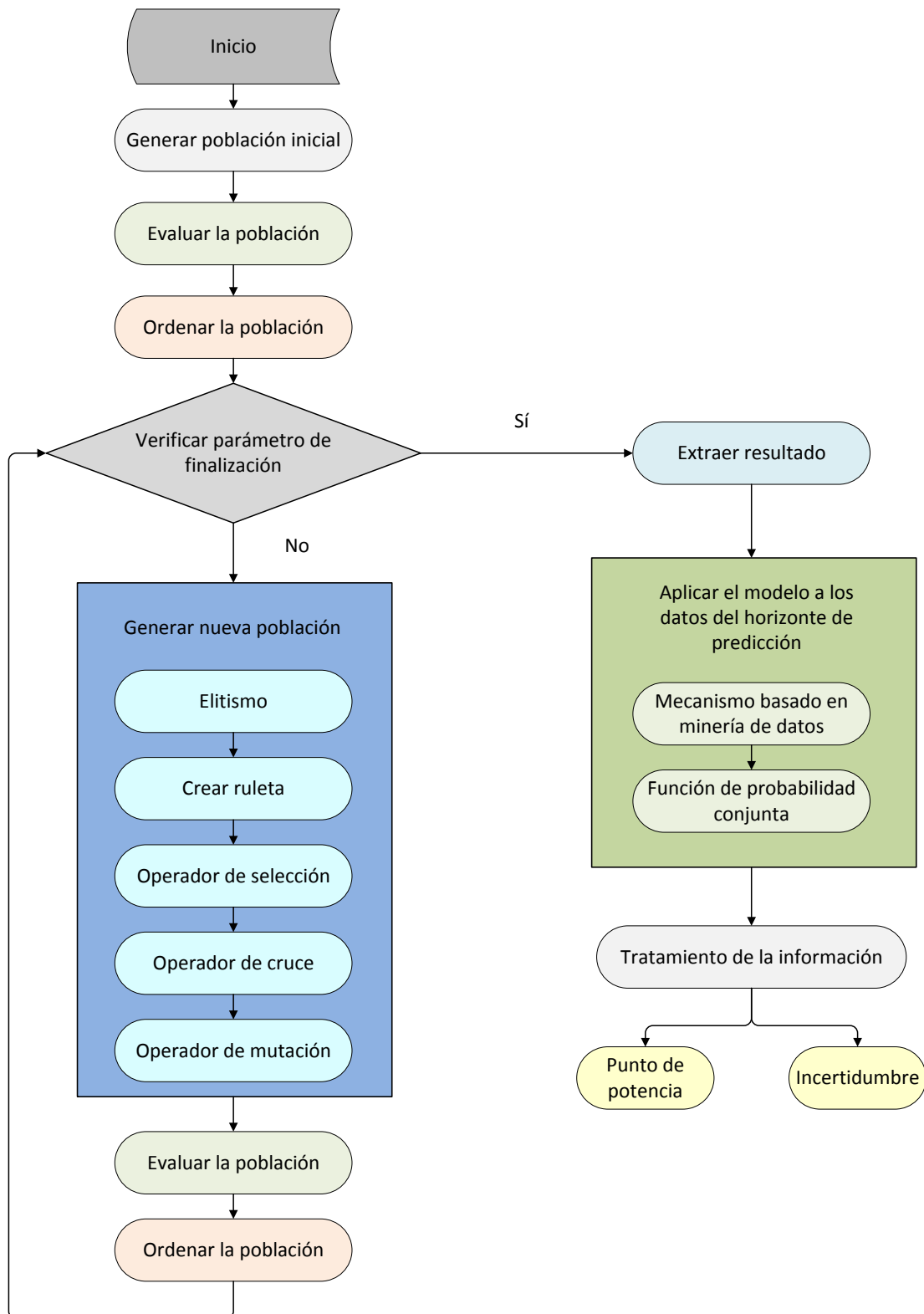


Figura 15.1: Diagrama del algoritmo genético del modelo.

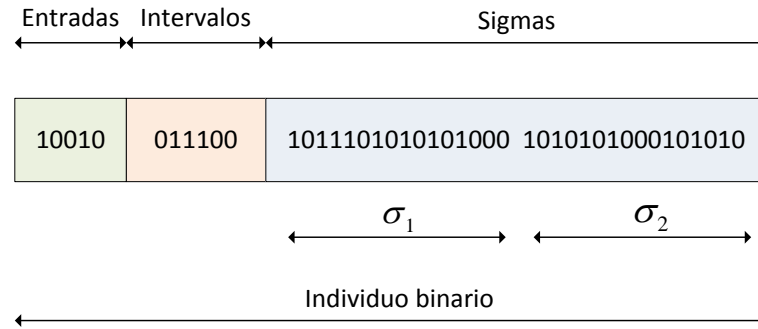


Figura 15.2: Ejemplo de un individuo binario, con dos entradas utilizadas de cinco disponibles.

Tabla 15.2: Para calcular número de intervalos - 1.

Intervalos - 1					
2^5	2^4	2^3	2^2	2^1	2^0
a_1	a_2	a_3	a_4	a_5	a_6

sino un conjunto finito de las mismas (transición de un intervalo a otro en el valor de la potencia generada).

Para establecer el número de intervalos que se van a considerar en el modelo, se establece la variable con su mismo nombre (intervalos). Esta variable está codificada en el segundo gen, que está formado por 6 bits. Aquí se permite que los 6 bits puedan ser 0, porque siempre se añade uno (ya que por lo menos tiene que haber un intervalo).

El valor del número de *intervalos* - 1 se calcula con el apoyo de la Tabla 15.2, de la siguiente forma:

$$\text{número de intervalos} - 1 = \sum_{i=1}^6 a_i 2^{(6-i)} \quad a_i \in \{0, 1\}$$

Sigmas

El modelo de predicción se basa en una búsqueda de transiciones entre intervalos de potencia en el pasado. Esta búsqueda se modeliza con una función gaussiana, distinta para cada una de las entradas usadas por el modelo y para la que tenemos que aportar el valor de la desviación estándar, σ . El valor de cada desviación estándar se codifica en 16 bits, por lo que el número total de bits necesarios para el gen correspondiente será:

$$\text{número total de entradas usadas en el modelo} \cdot 16 \text{ bits.}$$

A cada grupo de estos 16 bits, se denomina σ_i con $i \in \{1, \dots, l\}$ siendo $l = \text{número total de entradas utilizadas en el modelo}$, y se representa de la siguiente forma:

$$\text{Sigmas} = \sigma_1 \sigma_2 \dots \sigma_l$$

Tabla 15.3: Para calcular el valor de σ_i .

σ_i						
2^0	2^{-1}	2^{-3}	\dots	2^{-13}	2^{-14}	2^{-15}
b_1	b_2	b_3	\dots	b_{14}	b_{15}	b_{16}

correspondiendo cada σ_i a la entrada usada i .

El valor de cada $\sigma_i \in [0, 2]$ se calcula con el apoyo de la Tabla 15.3, de la siguiente forma:

$$\sigma_i = \sum_{i=1}^{16} b_i 2^{(1-i)} \quad b_i \in \{0, 1\}$$

15.2.2. Creación de una población inicial y evolución

El algoritmo genético se inicia estableciendo una población inicial de individuos, posibles soluciones al problema, que se crean de manera aleatoria. La representación utilizada es un vector de cadenas (no necesariamente de la misma longitud, ya que el tercer gen tiene una longitud proporcional al número de entradas usadas por el modelo), y cumpliendo la restricción que cada individuo debe estar en una y solo una de las posiciones del vector, es decir, no se puede repetir ningún individuo en la población inicial.

A partir de estas soluciones, se evalúa una función de aptitud que cuantifica el desempeño de las soluciones respecto como se aproximan a la solución que requiere el problema (ver el capítulo de técnicas de evaluación). Después de evaluar la función de aptitud, se ordena la población, en este modelo los individuos de menor a mayor error.

Si no se verifica el parámetro de finalización con las soluciones ordenadas se procede a llevar a cabo el proceso de evolución en el que se pasa de una generación a otra empleando los operadores genéticos, se evalúan las soluciones y se ordenan. Se procede así hasta que se alcance el total de generaciones.

15.2.3. Operadores genéticos

Se utilizan los operadores genéticos de elitismo (si está habilitado), selección, cruce, y mutación. El primero consiste en seleccionar al mejor individuo y copiarlo en la siguiente generación. El segundo, consiste en seleccionar los individuos más aptos, es decir con mejor función de aptitud (en este modelo con menor error), para que sus cromosomas o parte de ellos se conserven en la siguiente generación (bajo el principio de supervivencia del más fuerte o más adaptado). El tercero consiste en generar individuos “hijos” a partir de padres aptos, lo cual se logra combinando los cromosomas de dichos padres (bajo el principio de conservación de los genes aptos). El último consiste en el cambio de un gen de los individuos de la nueva generación (bajo el principio que dicho cambio a nivel de cromosoma puede ser favorable para la población).

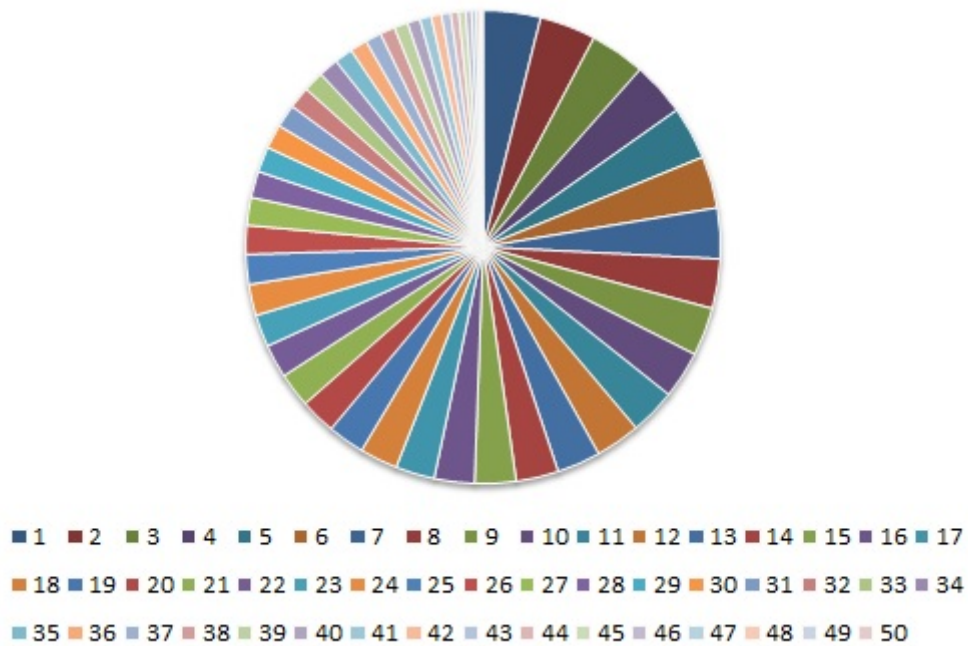


Figura 15.3: Ruleta de Holland con 1275 elementos, para una población de 50 individuos ordenados de menor a mayor error.

Elitismo

Si está habilitado en el programa se aplica elitismo, es decir, se copia en la nueva población el mejor individuo, el que tiene menor error en toda la población.

Selección

Para formar los individuos de una nueva población, se selecciona de forma aleatoria los individuos de la población actual, pero para otorgar mayor probabilidad de ser seleccionados a los mejores individuos, el método que se utiliza es el de la ruleta de Holland. Ésta se forma al repetir en la misma las posiciones de todos los individuos del vector población de una forma proporcional a su valor de aptitud, esto es, la posición del mejor individuo (el que menor error da) se repite un número mayor de veces que la del segundo mejor y así sucesivamente hasta el último.

De esta manera para una población de n individuos, el vector ruleta tiene:

$$\sum_{i=1}^n i = 1 + 2 + 3 + \cdots + n = \frac{n^2 + n}{2}$$

elementos.

Como ejemplo, se representa gráficamente en la Figura 15.3, una ruleta de Holland para una población de $n = 50$ individuos.

Una vez definidas las probabilidades de selección de cada individuo, se selecciona de forma aleatoria dos elementos distintos de la ruleta, con esto se fuerza a que sean dos

posiciones del vector población distintas, es decir dos individuos distintos y sus cromosomas se “cruzan” para formar dos nuevos individuos para la siguiente generación. Este procedimiento se repite hasta que una proporción q de la generación sea seleccionada.

Cruce

Una vez seleccionados dos individuos “padres” P1 y P2, éstos son recombinados para producir de descendencia dos “hijos” H1 y H2, que se insertarán en la siguiente generación.

Su importancia para la transición entre generaciones es elevada puesto que las tasas de cruce con las que se suele trabajar rondan el 90 %.

La idea principal del cruce se basa en que, si se toman dos individuos correctamente adaptados al medio y se obtiene una descendencia que comparte genes de ambos, existe la posibilidad de que los genes heredados sean precisamente los causantes de la bondad de los padres. Al compartir las características buenas de dos individuos, la descendencia, o al menos parte de ella, debería tener una bondad mayor que cada uno de los padres por separado. Si el cruce no agrupa las mejores características en uno de los hijos y la descendencia tiene un peor ajuste que los padres, no significa que se esté dando un paso atrás. Aún con un ajuste peor, se opta por insertar la descendencia, y puesto que los genes de los padres continuarán en la población, aunque dispersos, en posteriores cruces se podrán volver a obtener esos padres, recuperando así la bondad previamente perdida.

El cruce que se emplea en este modelo es por 2 puntos. Dado que el cruce puede afectar a los bits que indican las entradas seleccionadas, estos inicialmente no entran en el cruce, pero si que entran en el cruce los 6 bits que se reservan para intervalos. En el cruce se intercambian el gen que representa el número de intervalos y/o una o varias sigmas.

Mutación

El proceso de mutación se lleva a cabo una vez que se tienen definidos varios individuos de una nueva generación, esto es, una vez se ha realizado el elitismo, la selección y el cruce.

Los individuos a mutar se seleccionan aleatoriamente de la población actual (no de la ruleta), se copian en la nueva población y se utiliza el operador mutación, este proceso se repite hasta completar el tamaño total de la nueva población.

La probabilidad de mutación es muy baja, generalmente menor al 1 %. Esto se debe sobre todo a que los individuos suelen tener un ajuste menor después de mutados. Sin embargo, se realizan mutaciones para garantizar que ningún punto del espacio de búsqueda tenga una probabilidad nula de ser examinado.

En este algoritmo genético no todos los individuos tienen los cromosomas de la misma longitud, ya que dependen del número de entradas usadas, de entre las disponibles, por el modelo representado por el individuo correspondiente. Cuando se utiliza el operador mutación, se cambia aleatoriamente un bit (el “0” por el “1” ó el “1” por el “0”), este puede pertenecer al gen que representa las entradas, los intervalos o las sigmas. Si además, el bit que se cambia pertenece al gen que representa las entradas, será necesario eliminar o añadir grupos de 16 bits en el gen que representa las sigmas, es decir, si el bit seleccionado es un:

- “1”: se elimina el grupo de 16 bits que corresponde a esa entrada en el gen de las sigmas, ver Figura 15.4.

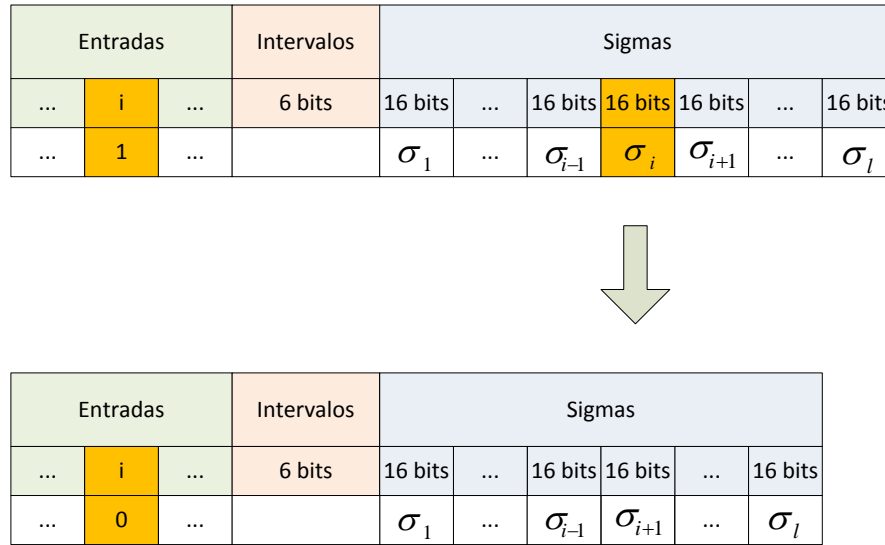


Figura 15.4: Se elimina el grupo de 16 bits, correspondiente a la entrada i .

- “0”: se añade un grupo de 16 bits aleatorios en el gen de las sigmas, justo en el lugar que le corresponde a dicha entrada, ver Figura 15.5.

15.2.4. Resultado

Una vez que se ha alcanzado el total de generaciones, el algoritmo genético devuelve el individuo con menor error, es decir la mejor solución. Ésta se utiliza con los datos de la base de conocimiento y con los valores pronosticados para las variables meteorológicas de entrada del horizonte de predicción, con el fin de proporcionar el valor numérico esperado de potencia y la incertidumbre asociada a él, para un determinado instante de horizonte temporal (ver secciones: formulación y salida del modelo).

15.3. Formulación

Con el objetivo de encontrar casos históricos similares de transiciones de potencia eléctrica, a aquellos que se pretenden prever, el modelo desarrollado realiza un “tipo de búsqueda” en la base de datos históricos recogida de la planta fotovoltaica en estudio.

Por lo tanto, los buenos resultados de predicción de potencia requieren el uso de una base de datos con casos históricos de buena calidad, con un alto volumen de información fiable, lo que implica también la necesidad de encontrar un mecanismo basado en minería de datos, MDM (de Mecanismo de Data Mining), para hacer frente a dicha información, capaz de buscar los casos históricos más relevantes y al mismo tiempo asignar un peso superior a los casos más similares y un peso inferior a los casos más distantes.

Después de la aplicación del MDM, para cada instante futuro $t+m$ con $m \in \{2, \dots, h\}$, surgirá un conjunto $\{FH_2, \dots, FH_t\}$ con los valores de pesos asignados a cada transición

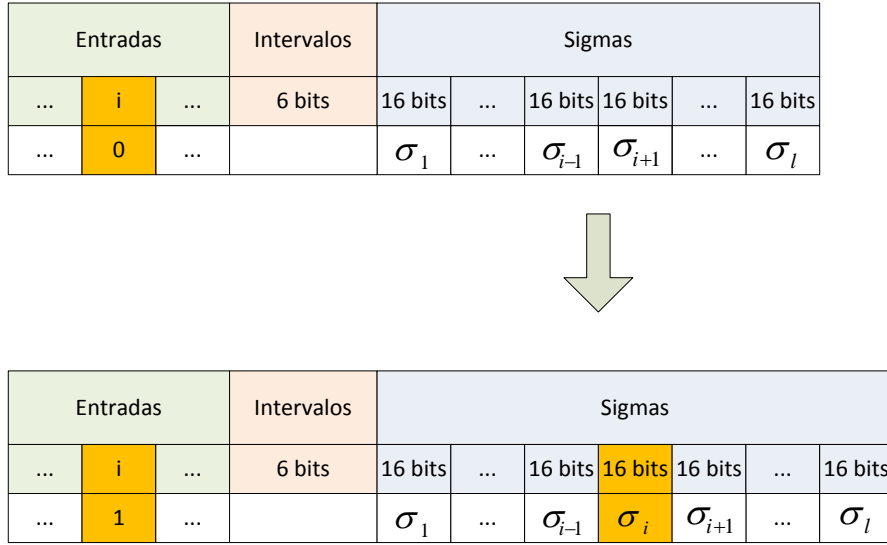


Figura 15.5: Se añade el grupo de 16 bits, correspondiente a la entrada i .

entre P_{j-1} y P_j , con $j \in \{2, \dots, t\}$. Por lo que será posible crear una matriz de probabilidad PM_{t+m} , que contiene las probabilidades de transición de potencia entre los instantes futuros $t + m - 1$ y $t + m$, donde t es el instante cuando se genera la predicción y m es el número de pasos de tiempo en el horizonte de predicción (horas). Esta matriz puede ser representada a través de una función de probabilidad conjunta $f_{XY,t+m}(x, y)$, que proporciona la probabilidad de ocurrencia de una transición de potencia dada.

Los siguientes apartados contienen explicaciones más detalladas sobre la formulación de este modelo.

15.3.1. Mecanismo basado en minería de datos (MDM)

Este mecanismo que se utiliza para procesar la información de las transiciones de potencia eléctrica, tiene como función efectuar una selección de los casos históricos presentes en la base de conocimiento, ver Tabla 15.4, que son similares a los casos que se pretenden prever. De este modo, dado un conjunto de variables de entradas disponibles para el instante futuro $t + m$, el MDM atribuirá diferentes pesos a las variables de entrada disponibles presentes en la base de conocimiento del modelo, de forma que seleccione solo la información relevante.

El MDM propuesto se basa en una modelización gaussiana local, ver Figura 15.6, que es definida por la siguiente ecuación:

$$g(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty \quad (15.1)$$

donde:

x es una variable de búsqueda.

Tabla 15.4: *Representación de una base de datos histórica con las variables meteorológicas de entradas disponibles.*

Instante histórico	Entradas disponibles					Salida P
	x_1	\dots	x_i	\dots	x_l	
1	$x_{1,1}$	\dots	$x_{1,i}$	\dots	$x_{1,l}$	P_1
\dots	\dots	\dots	\dots	\dots	\dots	\dots
$j - 1$	$x_{j-1,1}$	\dots	$x_{j-1,i}$	\dots	$x_{j-1,l}$	P_{j-1}
j	$x_{j,1}$	\dots	$x_{j,i}$	\dots	$x_{j,l}$	P_j
\dots	\dots	\dots	\dots	\dots	\dots	\dots
t (presente)	$x_{t,1}$	\dots	$x_{t,i}$	\dots	$x_{t,l}$	P_t
Instante futuro	Predicción PF					
	μ_1	\dots	μ_i	\dots	μ_l	
$t + 1$	$\mu_{t+1,1}$	\dots	$\mu_{t+1,i}$	\dots	$\mu_{t+1,l}$	
$t + 2$	$\mu_{t+2,1}$	\dots	$\mu_{t+2,i}$	\dots	$\mu_{t+2,l}$	PF_{t+2}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
$t + m - 1$	$\mu_{t+m-1,1}$	\dots	$\mu_{t+m-1,i}$	\dots	$\mu_{t+m-1,l}$	PF_{t+m-1}
$t + m$	$\mu_{t+m,1}$	\dots	$\mu_{t+m,i}$	\dots	$\mu_{t+m,l}$	PF_{t+m}
$t + m + 1$	$\mu_{t+m+1,1}$	\dots	$\mu_{t+m+1,i}$	\dots	$\mu_{t+m+1,l}$	PF_{t+m+1}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
$t + h - 1$	$\mu_{t+h-1,1}$	\dots	$\mu_{t+h-1,i}$	\dots	$\mu_{t+h-1,l}$	PF_{t+h-1}
$t + h$	$\mu_{t+h,1}$	\dots	$\mu_{t+h,i}$	\dots	$\mu_{t+h,l}$	

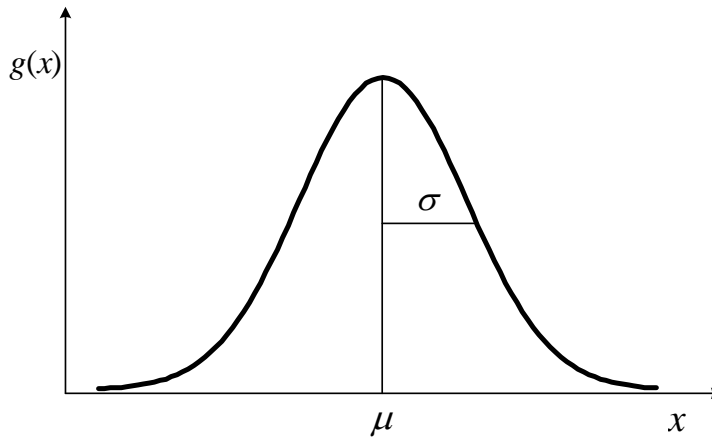


Figura 15.6: Representación gráfica de una función gaussiana local.

μ representa la “media o central” valor asignado a la variable de búsqueda.

σ representa la “desviación estándar” asignada a la variable de búsqueda, define la anchura del rango de exploración.

g representa la función de peso correspondiente al valor de la variable de búsqueda x .

Analizando la Figura 15.6 es conveniente elegir este modelo, ya que la gaussiana local permite la selección de un valor central o media pretendida μ y el valor de la desviación estándar σ . Los valores de la media μ seleccionados para las funciones gaussianas, corresponden a los valores de variables de casos en los instantes futuros del horizonte de predicción y los valores de la desviación estándar σ se fijan inicialmente. El parámetro σ define la anchura de los intervalos de prospección. Se verifica que cuanto mayor es el parámetro σ , mayor es el número de valores que son considerados cuando se efectúa la prospección. Esto es importante, ya que este parámetro define la cantidad de valores con relevancia, que son seleccionados cuando se aplica el MDM. En consecuencia, la σ para una variable de entrada no debe ser muy alta, en el caso que el parámetro σ sea elevado, el modelo no consigue diferenciar los valores que realmente interesan. En este modelo, donde las variables de entradas están normalizadas en el rango de 0 a 1, se opta por elegir la σ con un valor entre 0 y 2.

También es posible ver que el modelo escogido es capaz de distribuir los valores de los pesos teniendo en cuenta la similitud entre los casos históricos y el caso del horizonte de predicción, es decir, de acuerdo con la proximidad a la media μ como se pretende, ya que la función asigna mayor peso cuanto más cerca está del punto central y un menor peso cuando más lejos está del punto central. Así el MDM identifica los casos históricos con mayor semejanza con el caso que está destinado a proporcionar la predicción de potencia eléctrica, es decir, el MDM determina valores de pesos más altos para un conjunto de casos de transiciones de potencia eléctrica, que se utilizarán para proporcionar la información prevista.

Así pues, en cada instante futuro $t + m$, para cada variable de entrada usada en el modelo $\mu_{t+m,i}$ con $i \in \{1, \dots, l\}$ es necesario definir una gaussiana local.

Tabla 15.5: Representación de un registro j con las variables de entradas disponibles.

Transiciones NWP								Transiciones de Potencia (SCADA)	
$x_{j-1,1}$	$x_{j,1}$	\cdots	$x_{j-1,i}$	$x_{j,i}$	\cdots	$x_{j-1,l}$	$x_{j,l}$	P_{j-1}	P_j

Teniendo en cuenta que cada registro histórico j , con $j \in \{2, \dots, t\}$, contiene dos valores previstos: $x_{j-1,i}$ y $x_{j,i}$, para cada una de las variables meteorológicas de entradas usadas x_i con $i \in \{1, \dots, l\}$, ver Tabla 15.5 y con el fin de calcular para un instante futuro $t + m$, el valor de peso asignado a la transición entre P_{j-1} y P_j (el cual define la “similitud” existente entre las transiciones NWP del instante j y las transiciones NWP del instante futuro $t + m$), se definen las funciones gaussianas locales para los instantes futuros $t + m - 1$ y $t + m$. Este valor de peso se calcula de la siguiente forma:

$$\begin{aligned}
 FH_{t+m,j} = & g_{t+m-1,1}(x_{j-1,1}, \mu_{t+m-1,1}, \sigma_1) \cdot g_{t+m,1}(x_{j,1}, \mu_{t+m,1}, \sigma_1) \\
 & \cdot \dots \cdot g_{t+m-1,i}(x_{j-1,i}, \mu_{t+m-1,i}, \sigma_i) \cdot g_{t+m,i}(x_{j,i}, \mu_{t+m,i}, \sigma_i) \\
 & \cdot \dots \cdot g_{t+m-1,l}(x_{j-1,l}, \mu_{t+m-1,l}, \sigma_l) \cdot g_{t+m,l}(x_{j,l}, \mu_{t+m,l}, \sigma_l) \quad (15.2)
 \end{aligned}$$

donde para $i \in \{1, \dots, l\}$:

- $x_{j-1,i}$ es el valor de la variable de entrada seleccionada x_i en el instante histórico $j - 1$.
- $x_{j,i}$ es el valor de la variable de entrada seleccionada x_i en el instante histórico j .
- $\mu_{t+m-1,i}$ representa el valor de la “media o central” de la función gaussiana. Es el valor de la variable de entrada usada i que corresponde con el valor pronosticado para el instante futuro $t + m - 1$.
- $\mu_{t+m,i}$ representa el valor de la “media o central”. Es el valor de la variable de entrada usada i que corresponde con el valor pronosticado para el instante futuro $t + m$.
- σ_i representa el valor de la “desviación estándar”.
- $g_{t+m-1,i}(x_{j-1,i}, \mu_{t+m-1,i}, \sigma_i)$ representa el valor de peso, asociado a la función gaussiana local correspondiente a $x_{j-1,i}$, en el instante futuro $t + m - 1$.
- $g_{t+m,i}(x_{j,i}, \mu_{t+m,i}, \sigma_i)$ representa el valor de peso, asociado a la función gaussiana local correspondiente a $x_{j,i}$, en el instante futuro $t + m$.
- j corresponde al índice del instante histórico en análisis, que puede variar entre 2 y t , siendo t el instante actual.

Después de la aplicación del MDM, para cada instante futuro $t+m$, con $m \in \{2, \dots, h\}$, surge un conjunto $\{FH_{t+m,2}, \dots, FH_{t+m,j}, \dots, FH_{t+m,t}\}$ con los valores de los pesos asignados a cada transición entre P_{j-1} y P_j con $j \in \{2, \dots, t\}$, de acuerdo con la similitud existente entre las transiciones NWP de entrada en el instante futuro $t+m$ (para el cual se pretende efectuar la predicción), y las transiciones NWP que se encuentran en la base del conocimiento del modelo.

15.3.2. Intervalos de potencia

Como incluso con una base de datos reducida, puede existir un amplio conjunto de transiciones de potencia diferente, no se consideran las transiciones entre los valores específicos de potencia eléctrica, pero sí las transiciones entre intervalos de valores de la potencia eléctrica media horaria.

Este aspecto requiere definir un conjunto de intervalos de potencia, lo que permite transformar la variable continua de potencia eléctrica en una variable discreta. Así, se definen, un total de n intervalos de potencia iguales y no superpuestos, cubriendo todos los valores posibles para la generación de potencia eléctrica en la planta fotovoltaica. Cada intervalo de potencia $int \in \{1, \dots, n\}$ se define por sus valores de potencia mínima a_{int} y máxima b_{int} (en kW), que corresponden a los valores mínimo y máximo de ese intervalo. El valor medio del intervalo int se define como:

$$I_{int} = \frac{1}{2} (a_{int} + b_{int}) \quad (15.3)$$

Con los valores promedio de todos los intervalos de potencia, se crea el vector de valores de potencia eléctrica media, definido como:

$$[\mathbf{I}] = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} \quad (15.4)$$

donde $I_1 < I_2 < \dots < I_n$.

Esta técnica permite agrupar los valores de potencia eléctrica en una zona, que va a ser representada por el valor de la potencia eléctrica media del intervalo correspondiente. Es importante señalar que el número de intervalos de valores de potencia eléctrica debe ser razonable.

15.3.3. Matriz de probabilidad (PM)

Se crea la matriz de probabilidades, para ello se define un número n de intervalos para agrupar los valores de potencia eléctrica. Este hecho, es importante, ya que, como existe un elevado conjunto de transiciones de potencia eléctrica diferentes es bastante complicado crear una matriz en la que se presentan todas las transiciones posibles. Con este método es posible agregar las diferentes transiciones de potencia eléctrica en un conjunto de intervalos, haciendo el enfoque del problema más simple.

Se definen dos variables aleatorias discretas: X asociada con el intervalo correspondiente al valor de la potencia eléctrica en el instante futuro $t+m-1$, e Y asociada con

Tabla 15.6: Representación de la matriz PPM_{t+m} de pseudo-probabilidades, para el instante futuro $t + m$, con n intervalos de potencia.

PPM_{t+m}		Intervalos de potencia en $t + m$		
		1	...	n
Intervalos de potencia en $t + m - 1$	1

	n

el intervalo correspondiente al valor de la potencia eléctrica en el instante futuro $t + m$. Así, el intervalo de potencia eléctrica $x \in X$ varía (y también el intervalo $y \in Y$) de 1 a n . Las probabilidades de transición del intervalo de potencia eléctrica desde el instante $t + m - 1$ al siguiente $t + m$, pueden ser expresadas por una matriz cuadrada con n filas y n columnas, aplicando el mecanismo MDM.

Esta matriz PPM_{t+m} , llamada como la matriz pseudo-probabilidades para el instante futuro $t + m$, se muestra en la Tabla 15.6. Cada elemento en esta matriz representa la pseudo-probabilidad de una transición de potencia de un intervalo de potencia en el instante $t + m - 1$ (intervalo x , esto es la fila x en la matriz) a otro intervalo de potencia en el instante $t + m$ (intervalo y , esto es la columna y en la matriz).

El elemento $PPM_{t+m}(x, y)$, correspondiente a la fila x y columna y , es calculado usando la suma de los valores $FH_{t+m,j}$ de (15.2), y se expresa por:

$$PPM_{t+m}(x, y) = PPM_{t+m}(x, y) + FH_{t+m,j} \quad (15.5)$$

donde,

x = intervalo correspondiente al valor de P_{j-1} .

y = intervalo correspondiente al valor de P_j .

Se rellena la matriz, considerando todos los registros históricos de la base de datos.

La “normalización” de los valores de la matriz PPM_{t+m} , a valores entre 0 y 1, crea la matriz PM_{t+m} . Esta matriz representa una función $f_{XY,t+m}(x, y)$ de probabilidad conjunta de transiciones de potencia eléctrica, de un intervalo de potencia eléctrica x (asociado con el instante $t + m - 1$) a otro intervalo de potencia eléctrica y (asociado con el instante $t + m$).

Esta función $f_{XY,t+m}(x, y)$, ver Tabla 15.7, satisface:

1. $f_{XY,t+m}(x, y) \geq 0$
2. $\sum_{x=1}^n \sum_{y=1}^n f_{XY,t+m}(x, y) = 1$

donde $f_{XY,t+m}(x, y)$ representa la probabilidad que el intervalo de la variable de potencia eléctrica es x en un instante dado e y en el siguiente, es decir,

$$f_{XY,t+m}(x, y) = P(X = x, Y = y)$$

Tabla 15.7: Función $f_{XY,t+m}(x, y)$ de probabilidad conjunta de transiciones de potencia.

$f_{XY,t+m}(x, y)$		y		
		1	\dots	n
x	1	$f_{XY,t+m}(1, 1)$	\dots	$f_{XY,t+m}(1, n)$
	\dots	\dots	\dots	\dots
	n	$f_{XY,t+m}(n, 1)$	\dots	$f_{XY,t+m}(n, n)$

se tiene en cuenta que X puede asumir cualquiera de los valores $x \in \{1, 2, \dots, n\}$ e Y puede asumir cualquiera de los valores $y \in \{1, 2, \dots, n\}$.

Notar que para cada instante futuro $t + m$, con $m \in \{2, \dots, h\}$ se crea una matriz de probabilidad, la cual representa una función de probabilidad conjunta de transiciones de potencia eléctrica.

15.4. Herramientas del modelo

La función de probabilidad conjunta puede proporcionar múltiple información al usuario, para ello se definen dos herramientas que se basan en esta función:

15.4.1. Predicción de la incertidumbre

Una de las ventajas de este modelo, es su capacidad para proporcionar la incertidumbre asociada a un punto de predicción de potencia eléctrica esperada. Esto permitirá al usuario tomar decisiones con mayor nivel de confianza.

Con esta herramienta, para un determinado instante de tiempo, el usuario tendrá información sobre los intervalos de potencia eléctrica, donde puede surgir el valor real de la potencia eléctrica generada en dicho instante. Además de los intervalos, también se proporciona la probabilidad correspondiente.

Para el cálculo de la incertidumbre en el instante futuro $t + m$, con $m \in \{2, \dots, h - 1\}$, es necesario tomar en consideración dos funciones de probabilidad conjunta, una que representa la transición para el instante futuro $t + m$, $f_{XY,t+m}(x, y)$ y otra que representa una transición, para el instante siguiente futuro $t + m + 1$, $f_{XY,t+m+1}(x, y)$. Esto es importante ya que, en esta situación también interesa analizar las transiciones de potencia anteriores y posteriores.

Para efectuar la predicción de incertidumbre se utilizan las funciones de probabilidad marginal:

- función de probabilidad marginal de X , asociada a $f_{XY,t+m+1}(x, y)$, ver Tabla 15.8:

$$\begin{aligned}
 f_{m1,t+m+1}(x) = f_{X,t+m+1}(x) = P(X = x) &= P(X = x, Y = 1) + \dots + \\
 &+ P(X = x, Y = n) = \sum_{y=1}^n f_{XY,t+m+1}(x, y) \quad (15.6)
 \end{aligned}$$

Tabla 15.8: Funciones de probabilidad marginal de X , para el instante futuro $t + m + 1$.

$f_{XY,t+m+1}(x,y)$			Intervalos de potencia en $t + m + 1$			Totales
			y			
			1	\dots	n	
Intervalos de potencia en $t + m$	x	1	$f_{XY,t+m+1}(1,1)$	\dots	$f_{XY,t+m+1}(1,n)$	$f_{m1,t+m+1}(1)$
		\dots	\dots	\dots	\dots	\dots
		n	$f_{XY,t+m+1}(n,1)$	\dots	$f_{XY,t+m+1}(n,n)$	$f_{m1,t+m+1}(n)$

Tabla 15.9: Funciones de probabilidad marginal de Y , para el instante futuro $t + m$.

$f_{XY,t+m}(x, y)$			Intervalos de potencia en $t + m$		
			y		
			1	...	n
Intervalos de potencia en $t + m - 1$	x	1	$f_{XY,t+m}(1, 1)$...	$f_{XY,t+m}(1, n)$
	
		n	$f_{XY,t+m}(n, 1)$...	$f_{XY,t+m}(n, n)$
Totales			$f_{m2,t+m}(1)$...	$f_{m2,t+m}(n)$

- función de probabilidad marginal de Y , asociada a $f_{XY,t+m}(x, y)$, ver Tabla 15.9:

$$f_{m2,t+m}(y) = f_{Y,t+m}(y) = P(Y = y) = P(X = 1, Y = y) + \dots + P(X = n, Y = y) = \sum_{x=1}^n f_{XY,t+m}(x, y) \quad (15.7)$$

Dicho esto, la incertidumbre asociada a un punto de predicción de potencia eléctrica para el instante futuro $t + m$ es definida por el siguiente vector:

$$[\mathbf{uu}]_{t+m} = \begin{bmatrix} uu_{t+m}(1) \\ uu_{t+m}(2) \\ \vdots \\ uu_{t+m}(n) \end{bmatrix} = \begin{bmatrix} f_{m1,t+m+1}(1) f_{m2,t+m}(1) \\ f_{m1,t+m+1}(2) f_{m2,t+m}(2) \\ \vdots \\ f_{m1,t+m+1}(n) f_{m2,t+m}(n) \end{bmatrix} \quad (15.8)$$

Debe señalarse que el vector incertidumbre (15.8) devuelve la probabilidad de ocurrencia de cada intervalo. Sin embargo, es necesario normalizar los valores de este vector (a valores entre 0 y 1), dando lugar a un nuevo vector:

$$[\mathbf{u}]_{t+m} = \begin{bmatrix} u_{t+m}(1) \\ \vdots \\ u_{t+m}(n) \end{bmatrix} = \begin{bmatrix} \frac{uu_{t+m}(1)}{\sum_{i=1}^n uu_{t+m}(i)} \\ \vdots \\ \frac{uu_{t+m}(n)}{\sum_{i=1}^n uu_{t+m}(i)} \end{bmatrix} \quad (15.9)$$

donde cada elemento de este vector nuevo correspondiente a un intervalo de potencia eléctrica, está asociado a la probabilidad de que el valor pronosticado de potencia eléctrica pertenezca a dicho intervalo, por lo que dicho vector da una medida de la incertidumbre asociada con las predicciones de potencia eléctrica.

15.4.2. Predicción de un punto de potencia

Esta herramienta permite proporcionar el valor esperado de la potencia eléctrica para un instante futuro $t + m$, con $m \in \{2, \dots, h - 1\}$, tal como es común en los modelos de predicción de potencia existentes.

El punto de predicción PF_{t+m} (en kW), se obtiene calculando el valor esperado de potencia eléctrica para un instante futuro $t + m$, de la siguiente forma:

$$PF_{t+m} = \sum_{int=1}^n u_{t+m}(int) I_{int} \quad (15.10)$$

Capítulo 16

Resultados computacionales

En este capítulo se presentan los mejores resultados que se han obtenido en los trabajos de investigación de esta tesis, con los nuevos modelos desarrollados de predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica, para horizontes de predicción a corto plazo. Se distinguen dos tipos de modelos, los de días de cielo claro y los de cualquier tipo de día.

16.1. Variables o atributos seleccionados

En esta sección se presentan las variables o atributos seleccionados de entrada a los modelos de predicción desarrollados, para diferentes horizontes de predicción. Con estos atributos los modelos han obtenido menores valores del RMSE.

16.1.1. Atributos en los modelos para días con cielo claro

En la Tabla 16.1 figuran las variables o atributos de entrada seleccionados para los modelos desarrollados de la predicción de la potencia eléctrica media horaria en una planta fotovoltaica para la situación de día claro. El conjunto C24 hace referencia a los atributos utilizados en los modelos con horizontes de predicción entre 0 y 24 horas, C48 a los utilizados con modelos de horizontes de predicción entre 25 y 48 horas, C72 a los utilizados con modelos de horizontes de predicción entre 49 y 72 horas, y C a los utilizados con modelos con horizonte de predicción mayores a 72 horas.

Tabla 16.1: *Atributos seleccionados para modelos de días con cielo claro, con horizontes de predicción entre 0 y 24 horas (C24), 25 y 48 horas (C48), 49 y 72 horas (C72), y mayor a 72 horas (C).*

Atributo	Significado	C24	C48	C72	C
$\text{sen } \delta$	Seno de la <i>declinación</i>	•	•	•	•
$\text{cos } \delta$	Coseno de la <i>declinación</i>	•	•	•	•

(Sigue en la página siguiente.)

Continuación de la tabla

Atributo	Significado	C24	C48	C72	C
$\text{sen } \omega$	Seno del <i>ángulo</i> horario	•	•	•	•
$\text{cos } \omega$	Coseno del <i>ángulo</i> horario	•	•	•	•
$\text{sen } \gamma$	Seno de la <i>elevación</i> del Sol	•	•	•	•
$\text{cos } \gamma$	Coseno de la <i>elevación</i> del Sol	•	•	•	•
$\text{sen } \psi$	Seno del <i>azimut</i>	•	•	•	•
$\text{cos } \psi$	Coseno del <i>azimut</i>	•	•	•	•
$\text{cos } i$	Coseno del ángulo de incidencia	•	•	•	•
K_d	Factor corrector de distancia Tierra-Sol	•	•	•	•
$RSWIN_24$	Radiación de onda corta (W/m^2)	•			
$TH10_24$	Temperatura a 10 m ($^{\circ}K$)	•			
$Q10_24$	Humedad relativa a 10 m (entre 0 y 1)	•			
$VEL10_24$	Velocidad del viento (m/s) a 10 m	•			
$RSWIN_48$	Radiación de onda corta (W/m^2)		•		
$TH10_48$	Temperatura a 10 m ($^{\circ}K$)		•		
$Q10_48$	Humedad relativa a 10 m (entre 0 y 1)		•		
$VEL10_48$	Velocidad del viento (m/s) a 10 m		•		
$RSWIN_72$	Radiación de onda corta (W/m^2)			•	
$TH10_72$	Temperatura a 10 m ($^{\circ}K$)			•	
$Q10_72$	Humedad relativa a 10 m (entre 0 y 1)			•	
$VEL10_72$	Velocidad del viento (m/s) a 10 m			•	
I_e	Irradiancia extra-atmosférica				•

16.1.2. Atributos en los modelos para cualquier tipo de día

Para horizontes de predicción de 1 a 8 horas

Una de las diferencias de los modelos de predicción de la potencia eléctrica media horaria desarrollados para cualquier tipo de día, con horizontes de predicción de $m \in \{1, \dots, 8\}$ horas, es que utilizan como atributo de entrada diferentes predicciones de irradiancia. A continuación, se describen las principales características de cada uno de los conjuntos de atributos seleccionados:

- El conjunto H_m contiene los valores medios horarios de predicciones de la irradiancia calculada sobre un plano horizontal sin predicciones meteorológicas, para el horizonte de predicción de m horas.
- El conjunto G_m contiene valores medios horarios de predicciones de la irradiancia sobre un plano con un ángulo de inclinación β , para el horizonte de predicción de m horas. Esta irradiancia se calcula con valores medios horarios de las variables meteorológicas de radiación de onda corta de día claro (W/M^2), temperatura a 10 m ($^{\circ}K$), cobertura de nubes a nivel bajo, medio y alto (entre 0 y 1), presión atmosférica a 2 m de altura (Pa), pronosticadas sobre un plano horizontal con el modelo mesoescalar WRF-NMM, para un horizonte de predicción entre 0 y 24 horas.

- El conjunto Rm contiene valores medios horarios de predicciones de la radiación de onda corta (W/m^2) pronosticada sobre un plano horizontal con el modelo mesoescalar WRF-NMM, para un horizonte de predicción entre 0 y 24 horas.

En la Tabla 16.2 figuran las variables o atributos de entrada seleccionados por los modelos desarrollados para la predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica, para la situación de cualquier tipo de día, con los conjuntos Hm, Gm y Rm; $m \in \{1, \dots, 8\}$, que hacen referencia a atributos utilizados en los modelos con horizontes de predicción entre 1 y 8 horas.

Tabla 16.2: *Atributos seleccionados para modelos de cualquier tipo de día, con horizontes de predicción de $m \in \{1, \dots, 8\}$ horas.*

Atributo	Significado	Hm	Gm	Rm
sen δ	Seno de la <i>declinación</i>	•	•	•
cos δ	Coseno de la <i>declinación</i>	•	•	•
sen ω	Seno del <i>ángulo</i> horario	•	•	•
cos ω	Coseno del <i>ángulo</i> horario	•	•	•
sen γ	Seno de la <i>altura o elevación</i>	•	•	•
cos γ	Coseno de la <i>altura o elevación</i>	•	•	•
sen ψ	Seno del <i>azimut</i>	•	•	•
cos ψ	Coseno del <i>azimut</i>	•	•	•
cos i	Coseno del ángulo de incidencia	•	•	•
hluzsol	Número de horas de luz solar diarias	•	•	•
K_d	Factor corrector de la distancia Tierra-Sol	•	•	•
$RSWCal_m$	Irradiancia sobre un plano horizontal (W/m^2)	•		
$RSWCalMed_m$	$RSWCal_m$ media diaria (W/m^2)	•		
$RSWCalCuad_m$	$\left[\frac{(I_e - RSWCal_m)}{\max\{I_e\}} \right]^2$	•		
$RSWCalRaiz_m$	Raíz [$RSWCalCuad_m$ media diaria]	•		
$POTCalC_m$	Predicción de potencia (sin pred. met.)	•		
$GBCal_m$	Irradiancia sobre plano inclinado (W/m^2)		•	
$GBCalMed_m$	$GBCal_m$ media diaria (W/m^2)		•	
$GBCalCuad_m$	$\left[\frac{(G_{ef,\beta} - GBCal_m)}{\max\{G_{ef,\beta}\}} \right]^2$		•	
$GBCalRaiz_m$	Raíz [$GBCalCuad_m$ media diaria]		•	
$POTCalC24_m$	Predicción de potencia (pred. met. 0–24 h)		•	•
$RSWIN_24$	Radiación de onda corta (W/m^2), hor. 0–24 h			•
$RSWINMed_24$	$RSWIN_24$ media diaria (W/m^2)			•
$RSWINCuad_24$	$\left[\frac{(RSWINC_24 - RSWIN_24)}{\max\{RSWINC_24\}} \right]^2$			•
$RSWINRaiz_24$	Raíz [$RSWINCuad_24$ media diaria]			•

Para horizontes de predicción entre 0 y 24, 25 y 48, 49 y 72 horas

En la Tabla 16.3 figuran las variables o atributos de entrada seleccionados por los modelos desarrollados para la predicción de la generación media horaria en la planta para la situación de cualquier tipo de día. El conjunto R24 hace referencia a los atributos utilizados en los modelos con horizontes de predicción entre 0 y 24 horas, R48 a los utilizados con modelos con horizontes de predicción entre 25 y 48 horas, y R72 a los utilizados con modelos con horizontes de predicción entre 49 y 72 horas.

Tabla 16.3: *Atributos seleccionados para modelos de cualquier tipo de día, con horizontes de predicción entre 0 y 24 horas (R24), 25 y 48 horas (R48), 49 y 72 horas (R72).*

Atributo	Significado	R24	R48	R72
$\text{sen } \delta$	Seno de la <i>declinación</i>	•	•	•
$\text{cos } \delta$	Coseno de la <i>declinación</i>	•	•	•
$\text{sen } \omega$	Seno del <i>ángulo</i> horario	•	•	•
$\text{cos } \omega$	Coseno del <i>ángulo</i> horario	•	•	•
$\text{sen } \gamma$	Seno de la <i>altura o elevación</i>	•	•	•
$\text{cos } \gamma$	Coseno de la <i>altura o elevación</i>	•	•	•
$\text{sen } \psi$	Seno del <i>azimut</i>	•	•	•
$\text{cos } \psi$	Coseno del <i>azimut</i>	•	•	•
$\text{cos } i$	Coseno del ángulo de incidencia	•	•	•
$hluzsol$	Número de horas de luz solar diarias	•	•	•
K_d	Factor corrector de la distancia Tierra-Sol	•	•	•
$RSWIN_24$	Radiación de onda corta (W/m^2), hor. 0–24 h	•		
$RSWINMed_24$	$RSWIN_24$ media diaria (W/m^2)	•		
$RSWINCuad_24$	$\left[\frac{(RSWINC_24 - RSWIN_24)}{\max\{RSWINC_24\}}\right]^2$	•		
$RSWINRaiz_24$	Raíz [$RSWINCuad_24$ media diaria]	•		
$POTCalC24_24$	Predicción de potencia (pred. met. 0 y 24 h)	•		
$RSWIN_48$	Radiación de onda corta (W/m^2), hor. 25–48 h		•	
$RSWINMed_48$	$RSWIN_48$ media diaria (W/m^2)		•	
$RSWINCuad_48$	$\left[\frac{(RSWINC_48 - RSWIN_48)}{\max\{RSWINC_48\}}\right]^2$		•	
$RSWINRaiz_48$	Raíz [$RSWINCuad_48$ media diaria]		•	
$RSWIN_72$	Radiación de onda corta (W/m^2), hor. 49–72 h			•
$RSWINMed_72$	$RSWIN_72$ media diaria (W/m^2)			•
$RSWINCuad_72$	$\left[\frac{(RSWINC_72 - RSWIN_72)}{\max\{RSWINC_72\}}\right]^2$			•
$RSWINRaiz_72$	Raíz [$RSWINCuad_72$ media diaria]			•

16.2. Notación utilizada

En esta sección se presenta la notación que se utiliza en los resultados para los modelos de días con cielo claro y los modelos de cualquier tipo de día. Se recuerda que en la

Tabla 13.1 se describen los subconjuntos de datos: *Claros_E*, *Claros_En*, *Claros_V*, *Claros_T*, *E*, *En*, *V* y *T*, que se utilizan para el desarrollo de estos modelos.

16.2.1. Notación en los modelos para días con cielo claro

En la Tabla 16.4 se muestra el significado de la notación que se utiliza en los títulos de las figuras que exponen los resultados de los valores del RMSE (kW) obtenidos con los modelos de días claros. Se utiliza $m = 24$ para horizontes de predicción entre 0 y 24 horas, $m = 48$ para horizontes de predicción entre 25 y 48 horas, $m = 72$ para horizontes de predicción entre 49 y 72 horas y $m = -$ para horizontes de predicción mayores a 72 horas.

Tabla 16.4: *Significado de los títulos en los resultados de los modelos de días claros, para horizontes de predicción entre 0 y 24 horas ($m = 24$), 25 y 48 horas ($m = 48$), 49 y 72 horas ($m = 72$), y mayor a 72 horas ($m = -$).*

Título	Horizonte $m \in \{24, 48, 72, -\}$	Subconjuntos
Claros m - E	m	Claros_E
Claros m - E - T	m	Claros_E, Claros_T
Claros m - En - V	m	Claros_En, Claros_V
Claros m - P - T	m	Claros_En, Claros_V, Claros_T

En la Tabla 16.5 se indica el significado de la notación que se usa para los modelos de días claros en los resultados.

Tabla 16.5: *Notación en los modelos de días claros, para horizontes de predicción entre 0 y 24 horas ($m = 24$), 25 y 48 horas ($m = 48$), 49 y 72 horas ($m = 72$), y mayor a 72 horas ($m = -$).*

Tipo	Notación	Atributos $m \in \{24, 48, 72, -\}$	Técnica	Subconjunto
Reglas	Cm-M5Rules-E	Cm	M5Rules	Claros_E
	Cm-M5Rules-T	Cm	M5Rules	Claros_T
Árboles	Cm-M5P-E	Cm	M5P	Claros_E
	Cm-M5P-T	Cm	M5P	Claros_T
RNAs	Cm-MLP-GD-E	Cm	MLP-GD	Claros_E
	Cm-MLP-GD-T	Cm	MLP-GD	Claros_T
	Cm-MLP-m(0.2)-E	Cm	MLP-m(0.2)	Claros_E
	Cm-MLP-m(0.2)-T	Cm	MLP-m(0.2)	Claros_T
	Cm-MLP-CGD-E	Cm	MLP-CGD	Claros_E
	Cm-MLP-CGD-T	Cm	MLP-CGD	Claros_T
	Cm-MLP-BFGS-E	Cm	MLP-BFGS	Claros_E

(Sigue en la página siguiente.)

Continuación de la tabla

Tipo	Notación	Atributos $m \in \{24, 48, 72, -\}$	Técnica	Subconjunto
	Cm-MLP-BFGS-T	Cm	MLP-BFGS	Claros_T
	Cm-RBF-CGD-E	Cm	RBF-CGD	Claros_E
	Cm-RBF-CGD-T	Cm	RBF-CGD	Claros_T
	Cm-RBF-BFGS-E	Cm	RBF-BFGS	Claros_E
	Cm-RBF-BFGS-T	Cm	RBF-BFGS	Claros_T
ANFIS	Cm-ANFIS-En	Cm	ANFIS	Claros_En
	Cm-ANFIS-V	Cm	ANFIS	Claros_V
	Cm-ANFIS-T	Cm	ANFIS	Claros_T

16.2.2. Notación en los modelos para cualquier tipo de día

En la Tabla 16.6 se muestra el significado de la notación que se utiliza en los títulos de las figuras que exponen los resultados de los valores del RMSE (kW) obtenidos con los modelos de cualquier tipo de día. En dichas figuras el título “Hora m” indica que el horizonte de predicción es de m horas.

Los modelos con m entre 1 y 8 horas se han desarrollado para predecir la potencia eléctrica media horaria generada en la planta fotovoltaica dentro del propio día, mientras que los modelos con $m = 24$ se han desarrollado para predecirla durante las horas de ese día completo antes de la salida del sol, con $m = 48$ para predecirla durante las horas del siguiente día completo, y con $m = 72$ para predecirla durante las horas del segundo día siguiente completo.

Tabla 16.6: *Significado de los títulos en los resultados de los modelos de cualquier tipo de día, para horizontes de predicción de 1 a 8 horas ($m \in \{1, \dots, 8\}$), entre 0 y 24 horas ($m = 24$), 25 y 48 horas ($m = 48$), 49 y 72 horas ($m = 72$).*

Título	Horizonte $m \in \{1, \dots, 8, 24, 48, 72\}$	Subconjuntos
Hora m - E	m	E
Hora m - En	m	En
Hora m - V	m	V
Hora m - En - V	m	En, V
Hora m - En - V - P	m	En, V
Hora m - P	m	En, V
Hora m - T	m	T

En la Tabla 16.7 se indica el significado de la notación que se usa para los modelos de cualquier tipo de día en los resultados.

Tabla 16.7: *Notación en los modelos de cualquier tipo de día, para horizontes de predicción de 1 a 8 horas ($m \in \{1, \dots, 8\}$), entre 0 y 24 horas ($m = 24$), 25 y 48 horas ($m = 48$), 49 y 72 horas ($m = 72$).*

Tipo	Notación	Atributos $m \in \{1, \dots, 8, 24, 48, 72\}$	Técnica	Subconjunto
Reglas	Hm-M5Rules-E	Gm	M5Rules	E
	Gm-M5Rules-E	Gm	M5Rules	E
	Rm-M5Rules-E	Rm	M5Rules	E
	Hm-M5Rules-T	Gm	M5Rules	T
	Gm-M5Rules-T	Gm	M5Rules	T
	Rm-M5Rules-T	Rm	M5Rules	T
Árboles	Hm-M5P-E	Gm	M5P	E
	Gm-M5P-E	Gm	M5P	E
	Rm-M5P-E	Rm	M5P	E
	Hm-M5P-T	Gm	M5P	T
	Gm-M5P-T	Gm	M5P	T
	Rm-M5P-T	Rm	M5P	T
RNAs	Hm-MLP-GD-E	Hm	MLP-GD	E
	Gm-MLP-GD-E	Gm	MLP-GD	E
	Rm-MLP-GD-E	Rm	MLP-GD	E
	Hm-MLP-GD-T	Hm	MLP-GD	T
	Gm-MLP-GD-T	Gm	MLP-GD	T
	Rm-MLP-GD-T	Rm	MLP-GD	T
	Hm-MLP-m(0.2)-E	Hm	MLP-m(0.2)	E
	Gm-MLP-m(0.2)-E	Gm	MLP-m(0.2)	E
	Rm-MLP-m(0.2)-E	Rm	MLP-m(0.2)	E
	Hm-MLP-m(0.2)-T	Hm	MLP-m(0.2)	T
	Gm-MLP-m(0.2)-T	Gm	MLP-m(0.2)	T
	Rm-MLP-m(0.2)-T	Rm	MLP-m(0.2)	T
	Hm-MLP-CGD-E	Hm	MLP-CGD	E
	Gm-MLP-CGD-E	Gm	MLP-CGD	E
	Rm-MLP-CGD-E	Rm	MLP-CGD	E
	Hm-MLP-CGD-T	Hm	MLP-CGD	T
	Gm-MLP-CGD-T	Gm	MLP-CGD	T
	Rm-MLP-CGD-T	Rm	MLP-CGD	T
	Hm-MLP-BFGS-E	Hm	MLP-BFGS	E
	Gm-MLP-BFGS-E	Gm	MLP-BFGS	E
	Rm-MLP-BFGS-E	Rm	MLP-BFGS	E
	Hm-MLP-BFGS-T	Hm	MLP-BFGS	T
	Gm-MLP-BFGS-T	Gm	MLP-BFGS	T
	Rm-MLP-BFGS-T	Rm	MLP-BFGS	T

(Sigue en la página siguiente.)

Continuación de la tabla

Tipo	Notación	Atributos $m \in \{1, \dots, 8, 24, 48, 72\}$	Técnica	Subconjunto
	Hm-RBF-CGD-E	Hm	RBF-CGD	E
	Gm-RBF-CGD-E	Gm	RBF-CGD	E
	Rm-RBF-CGD-E	Rm	RBF-CGD	E
	Hm-RBF-CGD-T	Hm	RBF-CGD	T
	Gm-RBF-CGD-T	Gm	RBF-CGD	T
	Rm-RBF-CGD-T	Rm	RBF-CGD	T
	Hm-RBF-BFGS-E	Hm	RBF-BFGS	E
	Gm-RBF-BFGS-E	Gm	RBF-BFGS	E
	Rm-RBF-BFGS-E	Rm	RBF-BFGS	E
	Hm-RBF-BFGS-T	Hm	RBF-BFGS	T
	Gm-RBF-BFGS-T	Gm	RBF-BFGS	T
	Rm-RBF-BFGS-T	Rm	RBF-BFGS	T
SVM	Hm-SVM-E	Hm	SVM	E
	Gm-SVM-E	Gm	SVM	E
	Rm-SVM-E	Rm	SVM	E
	Hm-SVM-T	Hm	SVM	T
	Gm-SVM-T	Gm	SVM	T
	Rm-SVM-T	Rm	SVM	T
ANFIS	Hm-ANFIS-En	Gm	ANFIS	En
	Gm-ANFIS-En	Gm	ANFIS	En
	Rm-ANFIS-En	Rm	ANFIS	En
	Hm-ANFIS-V	Gm	ANFIS	V
	Gm-ANFIS-V	Gm	ANFIS	V
	Rm-ANFIS-V	Rm	ANFIS	V
	Hm-ANFIS-T	Gm	ANFIS	T
	Gm-ANFIS-T	Gm	ANFIS	T
	Rm-ANFIS-T	Rm	ANFIS	T
HISIMI	Rm-HISIMI-E	Rm	MDM-AG	E
	Rm-HISIMI-T	Rm	MDM-AG	T

16.3. Resultados de los modelos para días con cielo claro

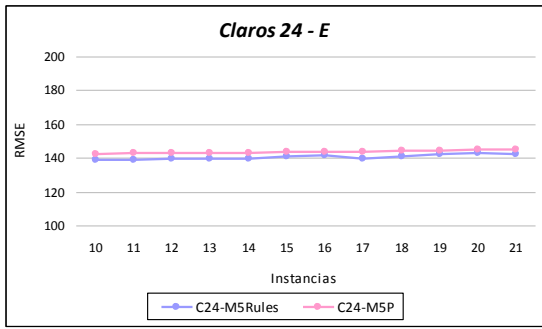
A continuación, se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria generada en la planta fotovoltaica para días claros utilizando diferentes técnicas de minería de datos, para horizontes de predicción entre 0 y 24 horas, entre 25 y 48 horas, entre 49 y 72 horas, y mayores a 72 horas.

16.3.1. Horizonte de predicción de 0 a 24 horas

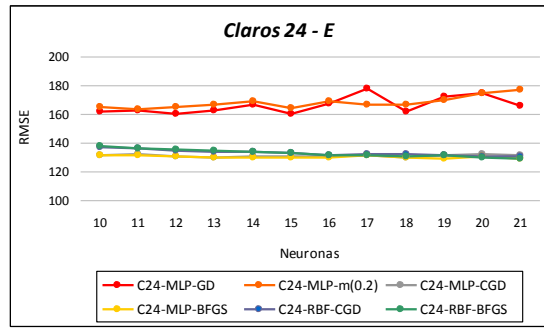
En esta sección se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria para días claros utilizando técnicas de minería de datos como sistemas de reglas, árboles de decisión, redes neuronales artificiales, y sistemas neuro-difusos, utilizando los atributos C24, y para horizontes de predicción entre 0 y 24 horas. La notación que se utiliza para designar estos modelos se recoge en la Tabla 16.5 con $m = 24$.

En la Figura 16.1 se presentan gráficamente los valores del RMSE (kW) que se han obtenido con los modelos de la predicción de la potencia eléctrica media horaria en días claros, para horizontes de predicción entre 0 y 24 horas y utilizando distintos subconjuntos de datos:

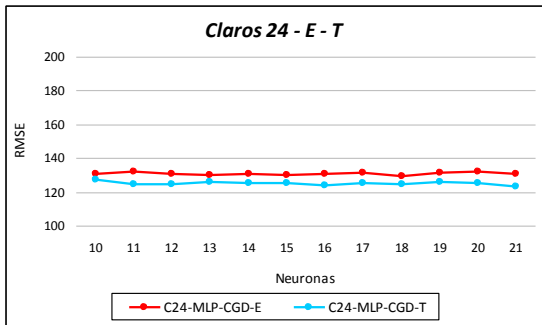
- En la Figura 16.1(a) se comparan los obtenidos con modelos desarrollados con técnicas M5Rules y M5P, variando el número de instancias de 10 a 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 138 kW y 145 kW, siendo los obtenidos con las dos técnicas muy parecidos.
- En la Figura 16.1(b) se comparan los obtenidos con modelos basados en redes neuronales artificiales MLP-GD, MLP-m(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS, todos con una capa oculta cuyo número de neuronas varía entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE con los modelos MLP-GD y MLP-m(0.2) oscilan aproximadamente entre 160 kW y 180 kW, y son mayores que los obtenidos con los modelos MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Los valores del RMSE con estos últimos modelos oscilan aproximadamente entre 130 kW y 140 kW. Los valores menores del RMSE se obtienen, en la mayor parte de los casos analizados, con el modelo MLP-BFGS.
- En la Figura 16.1(c) se muestran los obtenidos con modelos basados en red neuronal artificial MLP-CGD con una capa oculta variando su número de neuronas de 10 a 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilan aproximadamente entre 125 kW y 135 kW.
- En la Figura 16.1(d) se muestran los obtenidos con modelos basados en red neuronal artificial MLP-BFGS con una capa oculta variando su número de neuronas de 10 a 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 125 kW y 135 kW.
- En la Figura 16.1(e) se muestran los obtenidos con modelos basados en red neuronal artificial RBF-CGD con una capa oculta variando su número de neuronas de 10 a 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*.



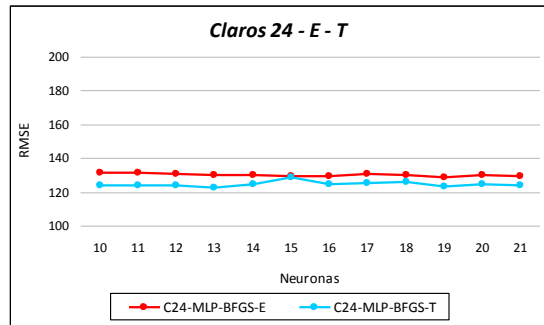
(a) M5Rules y M5P



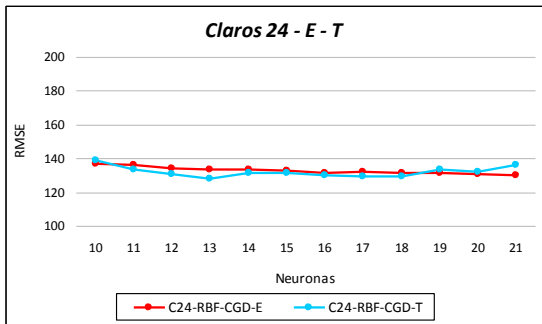
(b) RNAs



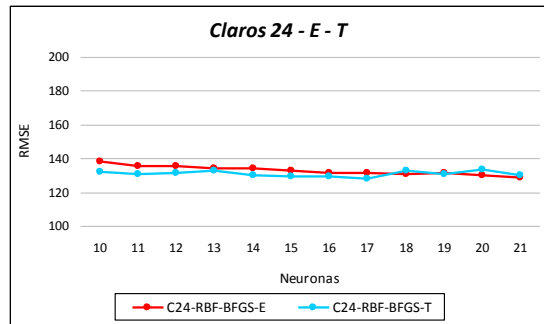
(c) MLP-CGD



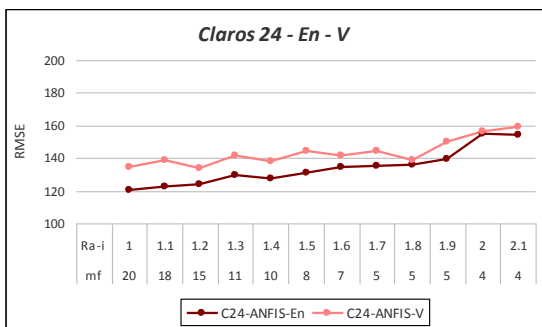
(d) MLP-BFGS



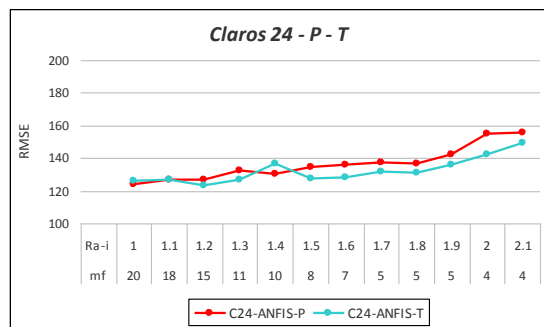
(e) RBF-CGD



(f) RBF-BFGS



(g) ANFIS



(h) ANFIS

Figura 16.1: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C24, para horizontes de predicción entre 0 y 24 horas con los datos de entrenamiento y test.

Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 125 kW y 140 kW.

- En la Figura 16.1(f) se muestran los obtenidos con modelos basados en red neuronal artificial RBF-BFGS con una capa oculta variando su número de neuronas de 10 a 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 125 kW y 140 kW.
- En la Figura 16.1(g) se muestran los obtenidos con modelos ANFIS desarrollados variando su rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de entrenamiento *Claros_En*, y los datos del subconjunto de validación *Claros_V*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 120 kW y 160 kW. El valor del RMSE en la mayoría de los casos disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.
- En la Figura 16.1(h) se muestran los obtenidos con modelos ANFIS desarrollados con la aplicación del algoritmo substractivo variando su rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de test *Claros_T*. También, se muestran los valores del RMSE ponderado, *Claros_P*, obtenidos con los modelos y los datos de los subconjuntos de entrenamiento *Claros_En* y de validación *Claros_V*. Se observa que los valores del RMSE con los datos del subconjunto de test *Claros_T* y el ponderado *Claros_P* son muy parecidos, oscilando aproximadamente entre 120 kW y 160 kW. El valor del RMSE en la mayoría de los casos disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.

En la Tabla 16.8 se muestran los valores del RMSE (kW) más relevantes de cada tipo de modelo (desarrollados con M5Rules, M5P, RNAs y ANFIS) usando los atributos C24, y su mejora con respecto al modelo persistente, para horizontes de predicción entre 0 y 24 horas, con los datos de los subconjuntos de entrenamiento *Claros_E* y de test *Claros_T*.

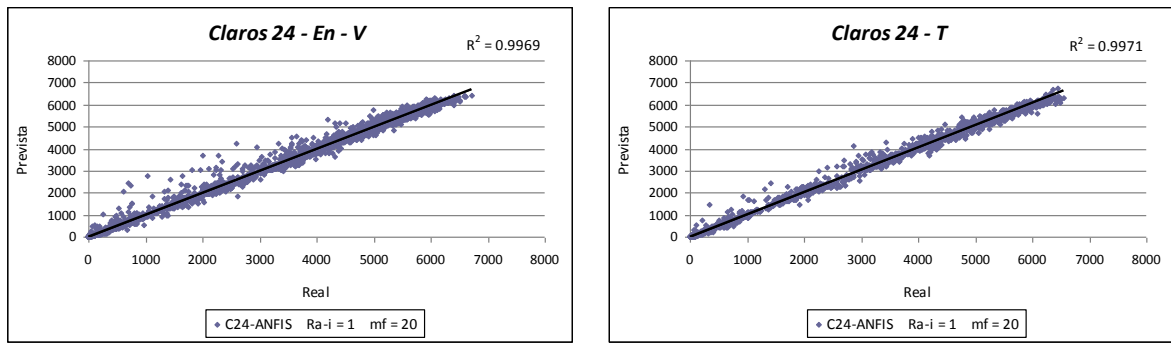
Como se puede apreciar en la Tabla 16.8, los modelos que ofrecen los mejores resultados en la predicción de la potencia eléctrica media horaria de días claros son el basado en la red neuronal MLP-BFGS con 19 neuronas en la capa oculta y el sistema ANFIS con la partición del espacio de entrada por agrupamiento substractivo con el valor del rango de influencia $Ra_i = 1.0$ y 20 reglas. Así el modelo ANFIS obtiene, con los datos de los subconjuntos de entrenamiento y de validación (*Claros_En* y *Claros_V*), un valor del RMSE medio ponderado igual a 123.9652 kW (mejora con respecto al modelo persistente del 31.59 %), y con los datos del subconjunto de test *Claros_T* obtiene un valor del RMSE igual a 126.3366 kW (mejora con respecto al modelo persistente del 19.79 %). El modelo basado en la red neuronal MLP-BFGS obtiene, con los datos del subconjunto de entrenamiento *Claros_E*, un valor del RMSE igual a 128.8561 kW (mejora con respecto al modelo persistente del 28.90 %), y con los datos del subconjunto de test *Claros_T* obtiene un valor del RMSE igual a 123.0392 kW (mejora con respecto al modelo persistente del

Tabla 16.8: Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 0 y 24 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	PC24		181.2343	157.5208	—	—
M5Rules	C24	instancias = 10	138.5881	136.7766	23.53 %	13.16 %
M5P	C24	instancias = 10	142.0306	131.3191	21.63 %	16.63 %
MLP-GD	C24	neuronas = 15	160.2247	150.2805	11.59 %	4.59 %
MLP-m(02)	C24	neuronas = 11	163.0585	141.4332	10.02 %	10.21 %
MLP-CGD	C24	neuronas = 18	129.5293	124.3177	28.52 %	21.07 %
MLP-BFGS	C24	neuronas = 19	128.8561	123.0392	28.90 %	21.89 %
RBF-CGD	C24	neuronas = 21	130.2260	136.0726	28.14 %	13.61 %
RBF-BFGS	C24	neuronas = 21	128.6730	129.8743	29.00 %	17.55 %
ANFIS	C24	mf=20, Ra-i=1.0	123.9652	126.3366	31.59 %	19.79 %

Tabla 16.9: Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción entre 0 y 24 horas.

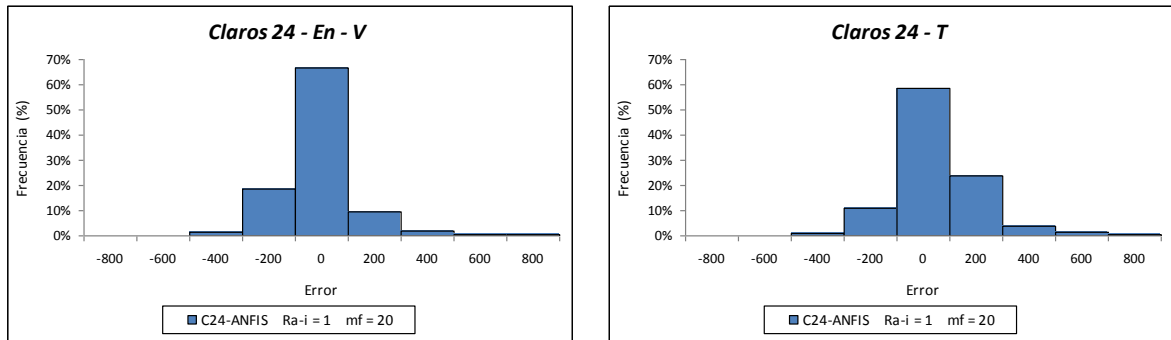
Horizonte	Modelos	Atributos	Características	R^2 Entr.	R^2 Test
De 0 a 24 horas	MLP-CGD	C24	neuronas = 18	0.9967	0.9972
	MLP-BFGS	C24	neuronas = 19	0.9967	0.9973
	RBF-CGD	C24	neuronas = 21	0.9966	0.9967
	RBF-BFGS	C24	neuronas = 21	0.9967	0.9969
	ANFIS	C24	mf=20, Ra-i=1.0	0.9969	0.9971



(a) Entrenamiento y Validación

(b) Test

Figura 16.2: Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 20 reglas, para horizontes de predicción entre 0 y 24 horas.



(a) Entrenamiento y Validación

(b) Test

Figura 16.3: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 20 reglas, para horizontes de predicción entre 0 y 24 horas.

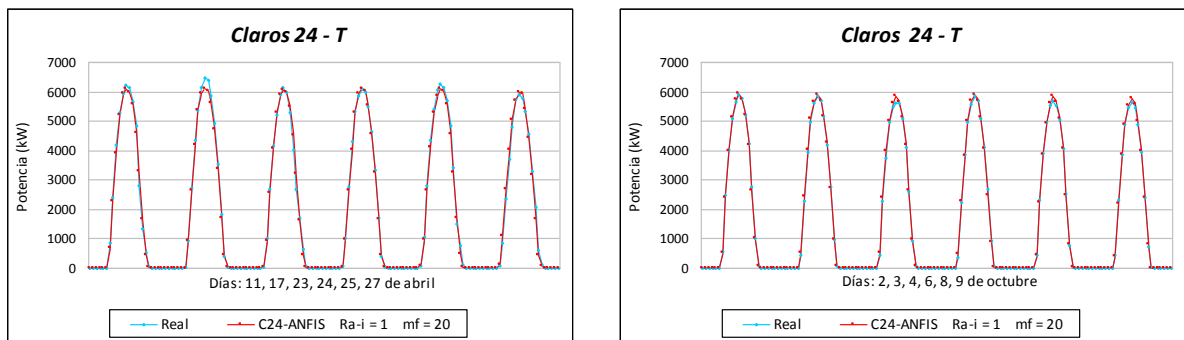


Figura 16.4: Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 20 reglas, para horizontes de predicción entre 0 y 24 horas con datos de test.

21.89 %). El valor del RMSE medio ponderado, considerando los datos de entrenamiento y test, es prácticamente el mismo para los dos modelos, aunque muy ligeramente mejor el del sistema ANFIS.

En la Tabla 16.9 se muestran los coeficientes de determinación R^2 en días claros de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los modelos MLP-CGD con 18 neuronas en la capa oculta, MLP-BFGS con 19 neuronas en la capa oculta, RBF-CGD con 21 neuronas en la capa oculta, RBF-BFGS con 21 neuronas en la capa oculta, y ANFIS con rango de influencia $Ra_i = 1$ y 20 reglas, usando los atributos C24, para las siguientes 24 horas. Se muestran los coeficientes de determinación para los datos de entrenamiento $Claros_E$ y de test $Claros_T$. Se observa que los coeficientes de determinación R^2 oscilan entre 0.9966 y 0.9973, estando así muy próximos al valor 1, por lo que hay una *correlación lineal casi perfecta (una regresión lineal casi perfecta)*.

En la Figura 16.2 se muestran los diagramas de dispersión en días claros de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS desarrollado mediante la partición del espacio de entrada por agrupamiento subtractivo con el valor del rango de influencia $Ra_i = 1$, 20 reglas, y los atributos C24, para horizontes de predicción entre 0 y 24 horas, utilizando los datos de los subconjuntos de entrenamiento $Claros_En$ y de validación $Claros_V$ (ver Figura 16.2(a)), y los datos del subconjunto de test $Claros_T$ (ver Figura 16.2(b)). El coeficiente de determinación tiene un valor de $R^2 = 0.9969$ con los datos de los subconjuntos $Claros_En$ y $Claros_V$, y $R^2 = 0.9971$ con los datos del subconjunto $Claros_T$, ambos muy cercanos a 1, por lo que hay una *correlación lineal casi perfecta (una regresión lineal casi perfecta)*.

En la Figura 16.3 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros, obtenidos con dicho sistema ANFIS con los datos de los subconjuntos de entrenamiento $Claros_En$ y validación $Claros_V$ (ver Figura 16.3(a)), y los datos del subconjunto de test $Claros_T$ (ver Figura 16.3(b)).

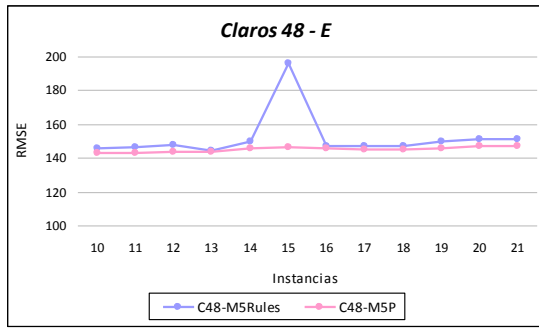
En la Figura 16.4 se representan los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS propuesto para doce días claros pertenecientes al subconjunto de test (días 11, 17, 23, 24, 25 y 27 de abril de 2012 en la gráfica izquierda, y días 2, 3, 4, 6, 8 y 9 de octubre de 2012 en la gráfica de la derecha).

16.3.2. Horizonte de predicción de 25 a 48 horas

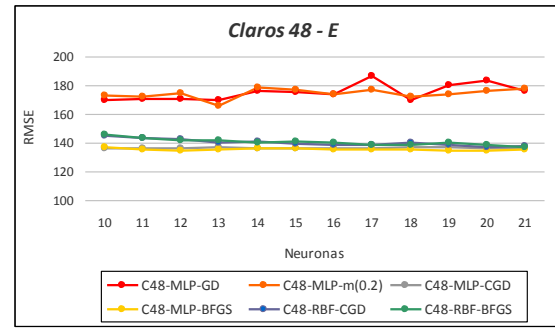
En esta sección se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria en días claros utilizando las técnicas de minería de datos de sistemas de reglas, árboles de decisión, redes neuronales artificiales y sistemas neuro-difusos. Para el desarrollo de estos modelos se han usado los atributos C48, y los horizontes de predicción han estado comprendidos entre 25 y 48 horas. En la Tabla 16.5 con $m = 48$ se presenta el significado de la notación que se utiliza en los resultados.

En la Figura 16.5 se presentan los valores del RMSE (kW) que se han obtenido con los modelos de la predicción de la potencia eléctrica media horaria en días claros, para horizontes de predicción entre 25 y 48 horas, y utilizando distintos subconjuntos de datos:

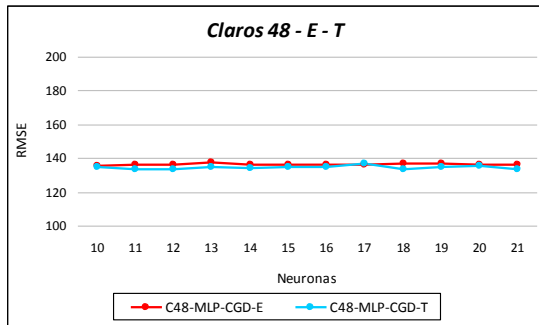
- En la Figura 16.5(a) se comparan los obtenidos con modelos desarrollados con técnicas M5Rules y M5P, variando el número de instancias entre 10 y 21. Se utilizan



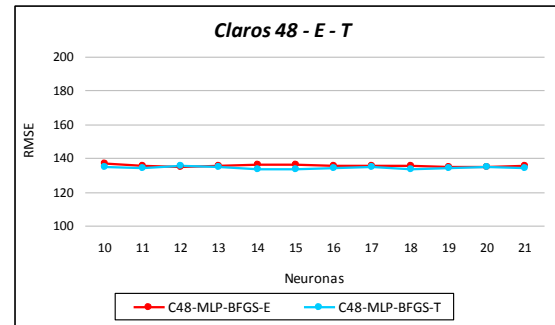
(a) M5Rules y M5P



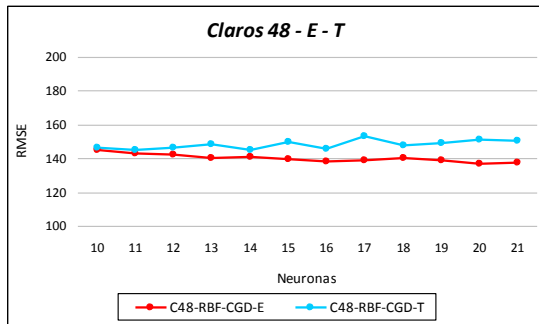
(b) RNAs



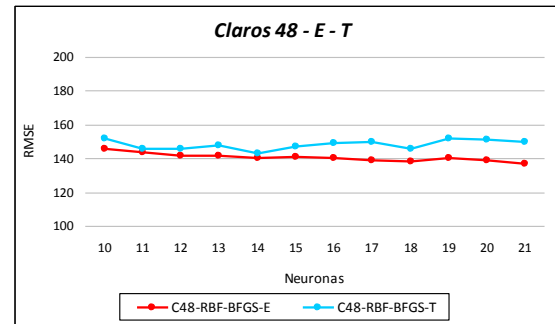
(c) MLP-CGD



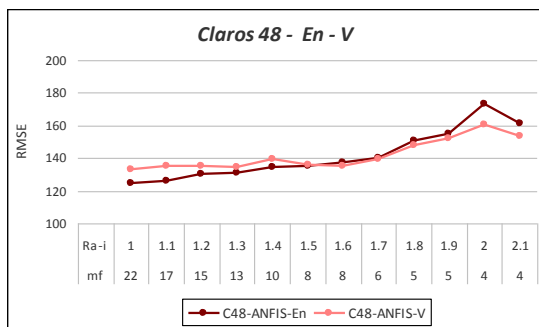
(d) MLP-BFGS



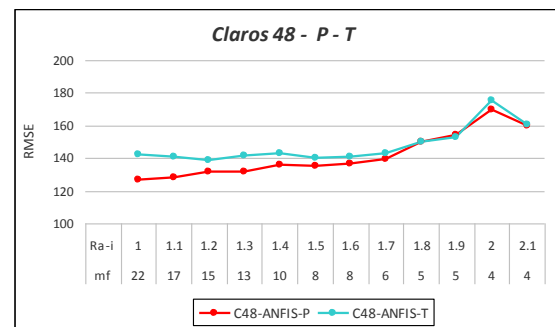
(e) RBF-CGD



(f) RBF-BFGS



(g) ANFIS



(h) ANFIS

Figura 16.5: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C48, para horizontes de predicción entre 25 y 48 horas con los datos de entrenamiento y test.

los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan en la mayoría de los casos entre 140 kW y 155 kW, y que los obtenidos con los modelos M5P son ligeramente mejores que los obtenidos con los modelos basados en la técnica de M5Rules.

- En la Figura 16.5(b) se comparan los obtenidos con modelos basados en redes neuronales artificiales del tipo MLP-GD, MLP-m(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Todas estas redes neuronales presentan una capa oculta con un número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE con los modelos MLP-GD y MLP-m(0.2) oscilan aproximadamente entre 160 kW y 190 kW, y son mayores que los obtenidos con los modelos MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Los valores del RMSE para estos últimos modelos oscilan aproximadamente entre 130 kW y 150 kW. Los valores menores del RMSE, en la mayor parte de los casos, se obtienen con el modelo MLP-BFGS.
- En la Figura 16.5(c) se muestran los obtenidos con modelos basados en red neuronal artificial MLP-CGD con una capa oculta con número de neuronas comprendido entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 130 kW y 140 kW.
- En la Figura 16.5(d) se muestran los obtenidos con modelos basados en red neuronal artificial MLP-BFGS con una capa oculta con número de neuronas comprendido entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 130 kW y 140 kW.
- En la Figura 16.5(e) se muestran los obtenidos con modelos basados en red neuronal artificial RBF-CGD con una capa oculta con número de neuronas comprendido entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son bastante parecidos, oscilando aproximadamente entre 135 kW y 155 kW.
- En la Figura 16.5(f) se muestran los obtenidos con modelos basados en red neuronal artificial RBF-BFGS con una capa oculta con número de neuronas comprendido entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son bastante parecidos, oscilando aproximadamente entre 135 kW y 155 kW.

- En la Figura 16.5(g) se muestran los obtenidos con modelos ANFIS contruidos con la aplicación del algoritmo substractivo con valores del rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de entrenamiento *Claros_En*, y los datos del subconjunto de validación *Claros_V*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son bastante parecidos, oscilando aproximadamente entre 120 kW y 180 kW. El valor del RMSE en la mayoría de los casos disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.
- En la Figura 16.5(h) se muestran los obtenidos con los modelos ANFIS contruidos con la aplicación del algoritmo substractivo con valores del rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de test *Claros_T*. También, se muestran los valores del RMSE ponderado, *Claros_P*, obtenidos con los modelos ANFIS y los datos de los subconjuntos de entrenamiento *Claros_En* y de validación *Claros_V*. Se observa que los valores del RMSE con los datos del subconjunto de test *Claros_T* y el ponderado *Claros_P* son bastante parecidos, oscilando aproximadamente entre 120 kW y 180 kW. El valor del RMSE en la mayoría de los casos disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.

En la Tabla 16.10 se muestran los valores del RMSE (kW) más relevantes de cada tipo de modelo (desarrollados con M5Rules, M5P, RNAs y ANFIS) usando los atributos C48, y su mejora con respecto al modelo persistente, para horizontes de predicción entre 25 y 48 horas, con los datos de los subconjuntos de entrenamiento *Claros_E* y de test *Claros_T*.

Como se puede apreciar en la Tabla 16.10 los modelos que ofrecen los mejores resultados en la predicción de la potencia eléctrica media horaria de días claros son el basado en la red neuronal MLP-BFGS con 12 neuronas en la capa oculta y el sistema ANFIS con la partición del espacio de entrada por agrupamiento substractivo con el valor del rango de influencia $Ra_i = 1.0$ y 22 reglas. Así el modelo ANFIS obtiene, con los datos de los subconjuntos de entrenamiento y de validación (*Claros_En* y *Claros_V*), un valor del RMSE medio ponderado igual a 126.6794 kW (mejora con respecto al modelo persistente del 36.13 %), y con los datos del subconjunto de test *Claros_T*, obtiene un valor del RMSE igual a 142.5325 kW (mejora con respecto al modelo persistente del 18.92 %). El modelo basado en la red neuronal MLP-BFGS obtiene, con los datos del subconjunto de entrenamiento *Claros_E*, un valor del RMSE igual a 134.4542 kW (mejora con respecto al modelo persistente del 32.21 %), y con los datos del subconjunto de test *Claros_T* obtiene un valor del RMSE igual a 135.4430 kW (mejora con respecto al modelo persistente del 22.96 %). El valor del RMSE medio ponderado, considerando los datos de entrenamiento y test, es prácticamente el mismo para los dos modelos, aunque ligeramente mejor el del sistema ANFIS. El modelo con la red neuronal MLP-CGD obtiene un menor valor del RMSE para los datos de test *Claros_T*, pero su valor del RMSE medio ponderado, para todos los datos (entrenamiento y test) es peor que el de los dos modelos antes citados.

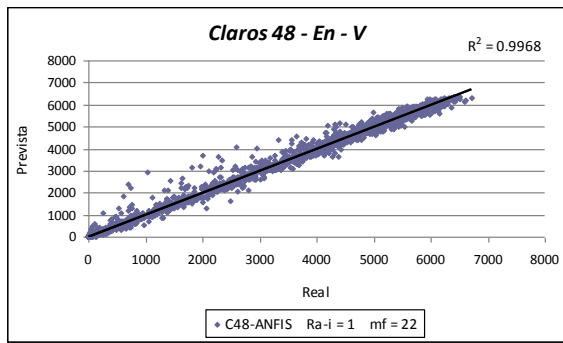
En la Tabla 16.11 se muestran los coeficientes de determinación R^2 en días claros de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los modelos MLP-CGD con 10 neuronas en la capa oculta, MLP-BFGS con 12 neuronas en la capa oculta, RBF-CGD con 20 neuronas en la capa oculta, RBF-BFGS con 21

Tabla 16.10: Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 25 y 48 horas.

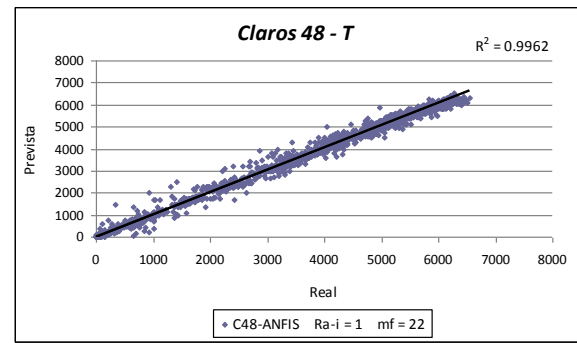
Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	PC48		198.3627	175.8124	—	—
M5Rules	C48	instancias = 10	145.4556	140.0252	26.67 %	20.35 %
M5P	C48	instancias = 10	142.8338	148.6492	27.99 %	15.45 %
MLP-GD	C48	neuronas = 13	169.4973	157.2511	14.55 %	10.55 %
MLP-m(02)	C48	neuronas = 13	165.6247	161.2274	16.50 %	8.29 %
MLP-CGD	C48	neuronas = 10	135.6868	134.8005	31.59 %	23.32 %
MLP-BFGS	C48	neuronas = 12	134.4542	135.4430	32.21 %	22.96 %
RBF-CGD	C48	neuronas = 20	136.4232	150.7344	31.22 %	14.26 %
RBF-BFGS	C48	neuronas = 21	136.9424	149.8280	30.96 %	14.77 %
ANFIS	C48	mf=22, Ra-i=1.0	126.6794	142.5325	36.13 %	18.92 %

Tabla 16.11: Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción entre 25 y 48 horas.

Horizonte	Modelos	Atributos	Características	R^2 Entr.	R^2 Test
De 25 a 48 horas	MLP-CGD	C48	neuronas = 10	0.9963	0.9966
	MLP-BFGS	C48	neuronas = 12	0.9964	0.9966
	RBF-CGD	C48	neuronas = 20	0.9963	0.9960
	RBF-BFGS	C48	neuronas = 21	0.9963	0.9958
	ANFIS	C48	mf=22, Ra-i=1.0	0.9968	0.9962

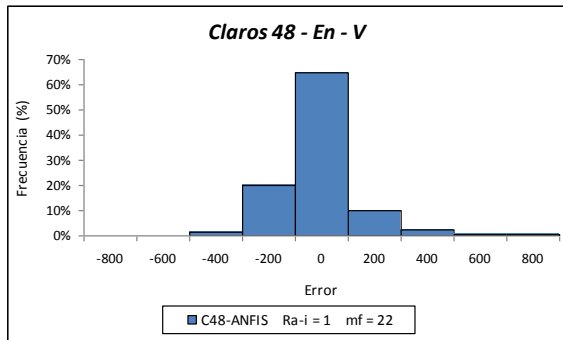


(a) Entrenamiento y Validación

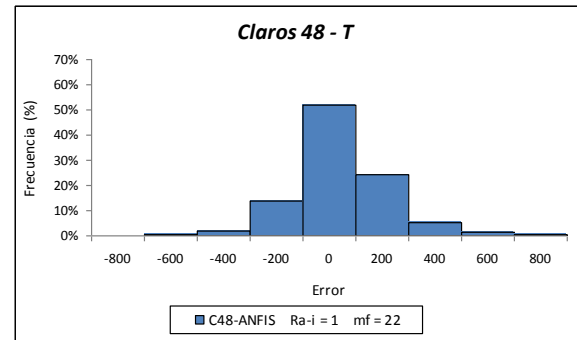


(b) Test

Figura 16.6: Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 22 reglas, para horizontes de predicción entre 25 y 48 horas.



(a) Entrenamiento y Validación



(b) Test

Figura 16.7: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 22 reglas, para horizontes de predicción entre 25 y 48 horas.

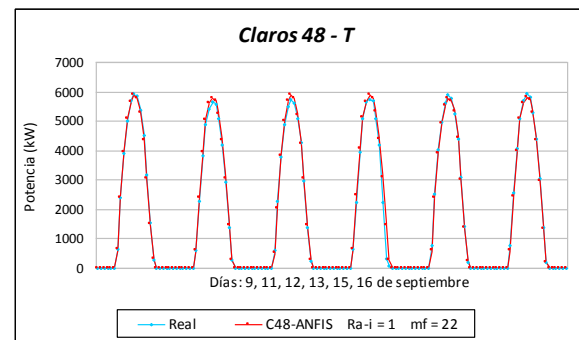
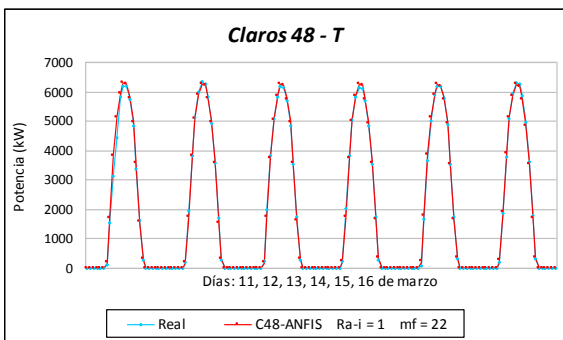


Figura 16.8: Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 22 reglas, para horizontes de predicción entre 25 y 48 horas con datos de test.

neuronas en la capa oculta, y ANFIS con rango de influencia $Ra_i = 1.0$ y 22 reglas, usando los atributos C48, con horizontes entre 25 y 48 horas. Se muestran los coeficientes de determinación para los datos de entrenamiento $Claros_E$ y de test $Claros_T$. Se observa que los coeficientes de determinación R^2 oscilan entre 0.9958 y 0.9968, muy cerca de 1, por lo que hay *una correlación lineal casi perfecta (una regresión lineal casi perfecta)*.

En la Figura 16.6 se muestran los diagramas de dispersión en días claros con valores reales de la potencia eléctrica media horaria (kW), y la predicción obtenida con el modelo ANFIS desarrollado mediante la partición del espacio de entrada por agrupamiento substractivo con el valor del rango de influencia $Ra_i = 1$ y 22 reglas, y los atributos C48, para horizontes de predicción entre 25 y 48 horas, y con los datos de los subconjuntos de entrenamiento $Claros_En$ y validación $Claros_V$ (ver Figura 16.6(a)), y los datos del subconjunto de test $Claros_T$ (ver Figura 16.6(b)). El coeficiente de determinación es $R^2 = 0.9968$ con los datos de los subconjuntos $Claros_En$ y $Claros_V$, y $R^2 = 0.9962$ con los datos del subconjunto $Claros_T$, que al estar muy cerca de 1 indican que hay *una correlación lineal casi perfecta (una regresión lineal casi perfecta)*.

En la Figura 16.7 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de día) en días claros, con el modelo ANFIS propuesto con los datos de los subconjuntos de entrenamiento $Claros_En$ y validación $Claros_V$ (ver Figura 16.7(a)), y los datos del subconjunto de test $Claros_T$ (ver Figura 16.7(b)).

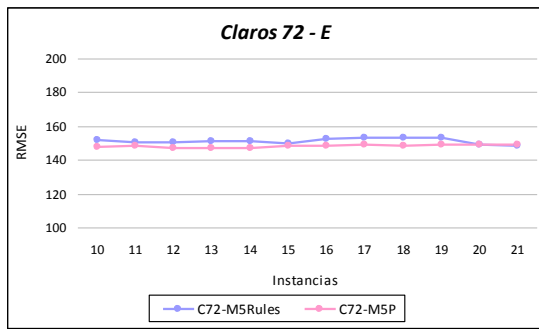
En la Figura 16.8 se representan los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS propuesto para doce días claros del conjunto de datos de test (días 11, 12, 13, 14, 15 y 16 de marzo de 2012 a la izquierda, y días 9, 11, 12, 13, 15 y 16 de septiembre a la derecha).

16.3.3. Horizonte de predicción de 49 a 72 horas

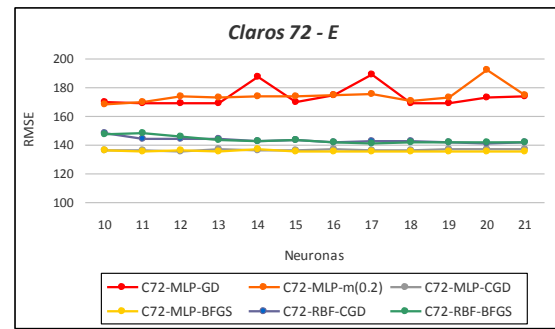
En esta sección se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria en días claros utilizando técnicas de minería de datos como sistemas de reglas, árboles de decisión, redes neuronales artificiales y sistemas neuro-difusos. Los datos utilizados en el desarrollo de los modelos corresponden a los atributos C72, y los horizontes de predicción considerados se encuentran entre las 49 y las 72 horas. En la Tabla 16.5 con $m = 72$ se presenta el significado de la notación que se utiliza en los resultados.

En la Figura 16.9 se presentan los valores del RMSE (kW) que se han obtenido con los modelos desarrollados:

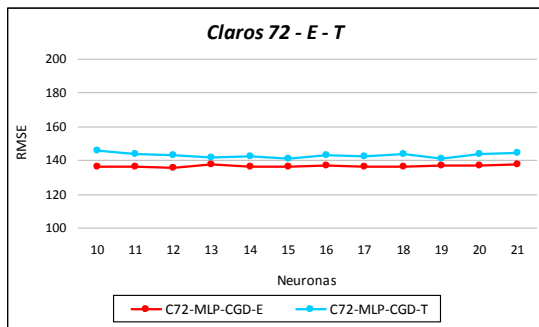
- En la Figura 16.9(a) se comparan los obtenidos con los modelos desarrollados con las técnicas M5Rules y M5P, variando el número de instancias de 10 hasta 21. Se utilizan los datos del subconjunto de entrenamiento $Claros_E$ con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 145 kW y 155 kW; los obtenidos con los modelos desarrollados con la técnica M5P son un poco menores que los obtenidos con los desarrollados con la técnica M5Rules.
- En la Figura 16.9(b) se comparan los obtenidos con modelos basados en redes neuronales artificiales de los tipos MLP-GD, MLP-m(0.2), MLP-CGD, MLP-BFGS,



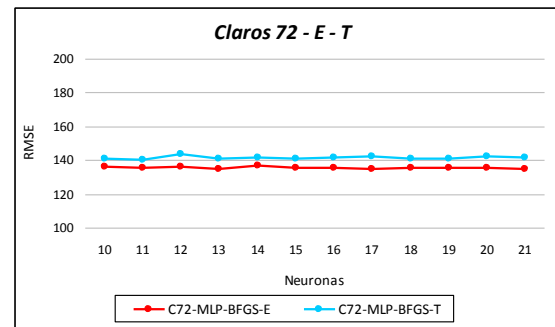
(a) M5Rules y M5P



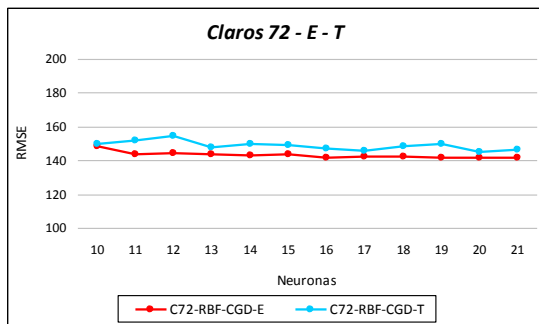
(b) RNAs



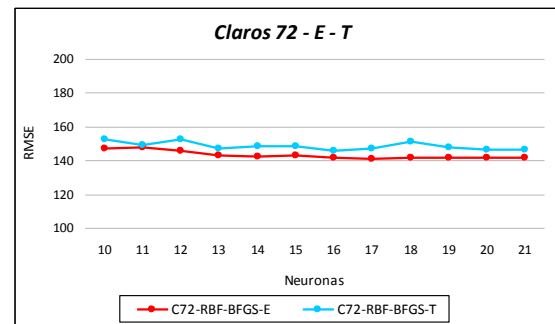
(c) MLP-CGD



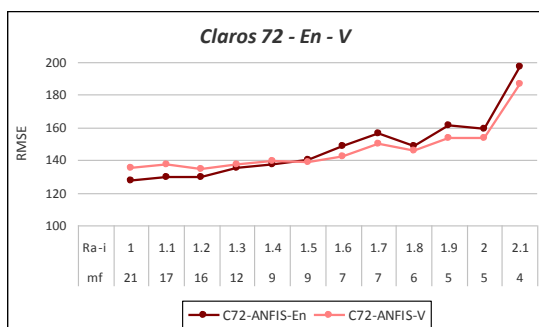
(d) MLP-BFGS



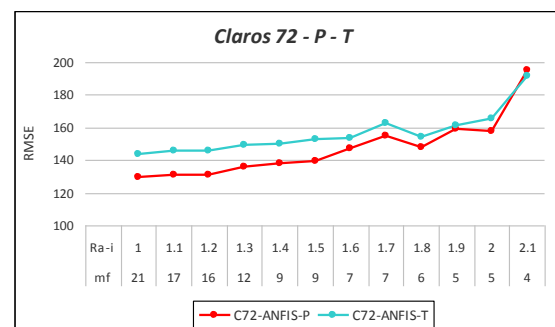
(e) RBF-CGD



(f) RBF-BFGS



(g) ANFIS



(h) ANFIS

Figura 16.9: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C72, para horizontes de predicción entre 49 y 72 horas con los datos de entrenamiento y test.

RBF-CGD y RBF-BFGS. Estos modelos constan de una capa oculta cuyo número de neuronas varía entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE con los modelos de redes neuronales MLP-GD y MLP-m(0.2) oscilan aproximadamente entre 165 kW y 190 kW, y son mayores que los obtenidos con los modelos de redes neuronales MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Los valores obtenidos con los modelos de estas últimas redes oscilan aproximadamente entre 130 kW y 150 kW. Los valores menores del RMSE, en la mayoría de los casos, se obtienen con el modelo MLP-BFGS.

- En la Figura 16.9(c) se muestran los obtenidos con un modelo basado en una red neuronal artificial MLP-CGD con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 135 kW y 150 kW.
- En la Figura 16.9(d) se muestran los obtenidos con un modelo basado en una red neuronal artificial MLP-BFGS con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 135 kW y 145 kW.
- En la Figura 16.9(e) se muestran los obtenidos con un modelo basado en una red neuronal artificial RBF-CGD con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son bastante parecidos, oscilando aproximadamente entre 140 kW y 160 kW.
- En la Figura 16.9(f) se muestran los obtenidos con un modelo basado en una red neuronal artificial RBF-BFGS con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son bastante parecidos, oscilando aproximadamente entre 140 kW y 155 kW.
- En la Figura 16.9(g) se muestran los obtenidos con los modelos ANFIS desarrollados con la aplicación del algoritmo substractivo variando su rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de entrenamiento *Claros_En*, y los datos del subconjunto de validación *Claros_V*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son bastante parecidos, oscilando aproximadamente entre 125 kW y 200 kW. El valor

del RMSE, en la mayoría de los casos, disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.

- En la Figura 16.9(h) se muestran los obtenidos con los modelos ANFIS desarrollados con la aplicación del algoritmo substractivo variando su rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de test *Claros_T*. También, se muestran los valores del RMSE ponderado, *Claros_P*, obtenidos con los modelos ANFIS y los datos de los subconjuntos de entrenamiento *Claros_En* y de validación *Claros_V*. Se observa que los valores del RMSE con los datos del subconjunto de test *Claros_T* y el ponderado *Claros_P* son bastante parecidos, oscilando aproximadamente entre 125 kW y 200 kW. El valor del RMSE, en la mayoría de los casos, disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.

En la Tabla 16.12 se muestran los valores del RMSE (kW) más relevantes de cada tipo de modelo (desarrollados con M5Rules, M5P, RNAs y ANFIS) usando los atributos C72, y su mejora con respecto al modelo persistente y para horizontes de predicción entre 49 y 72 horas, con los datos de los subconjuntos de entrenamiento *Claros_E* y de test *Claros_T*.

Como se puede apreciar en la Tabla 16.12 los modelos que ofrecen los mejores resultados en la predicción de la potencia eléctrica media horaria de días claros son el basado en la red neuronal MLP-BFGS con 13 neuronas en la capa oculta y el sistema ANFIS con la partición del espacio de entrada por agrupamiento substractivo con el valor del rango de influencia $Ra_i = 1.0$ y 21 reglas. Así el modelo ANFIS obtiene, con los datos de los subconjuntos de entrenamiento y validación (*Claros_En* y *Claros_V*), un valor RMSE medio ponderado igual a 129.2890 kW (mejora con respecto al modelo persistente del 38.39%), y con los datos del subconjunto de test *Claros_T* obtiene un valor del RMSE igual a 143.5102 kW (mejora con respecto al modelo persistente del 22.09%). El modelo basado en la red neuronal MLP-BFGS obtiene, con los datos del subconjunto de entrenamiento *Claros_E*, un valor del RMSE igual a 134.9632- kW (mejora con respecto al modelo persistente del 35.69%), y con los datos del subconjunto de test *Claros_T* obtiene un valor del RMSE igual a 140.8739 kW (mejora con respecto al modelo persistente del 23.53%). El valor del RMSE medio ponderado, considerando los datos de entrenamiento y test, es prácticamente el mismo para los dos modelos, aunque ligeramente mejor el del sistema ANFIS.

En la Tabla 16.13 se muestran los coeficientes de determinación R^2 en días claros de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los modelos MLP-CGD con 12 neuronas en la capa oculta, MLP-BFGS con 13 neuronas en la capa oculta, RBF-CGD con 20 neuronas en la capa oculta, RBF-BFGS con 17 neuronas en la capa oculta, y ANFIS con rango de influencia $Ra_i = 1.0$ y 21 reglas, y los atributos C72, para horizontes de predicción entre 49 y 72 horas, con los datos de los subconjuntos de entrenamiento *Claros_E* y de test *Claros_T*. Se observa que los coeficientes de determinación R^2 oscilan entre 0.9959 y 0.9967, estando muy cerca de 1, por lo que hay una *correlación lineal casi perfecta (una regresión lineal casi perfecta)*.

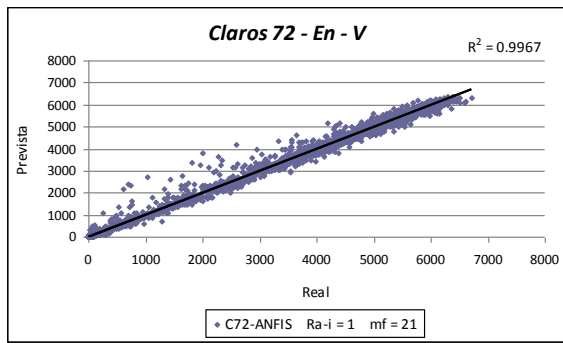
En la Figura 16.10 se muestran los diagramas de dispersión en días claros de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS desarrollado mediante la partición del espacio de entrada por agrupamiento

Tabla 16.12: Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 49 y 72 horas.

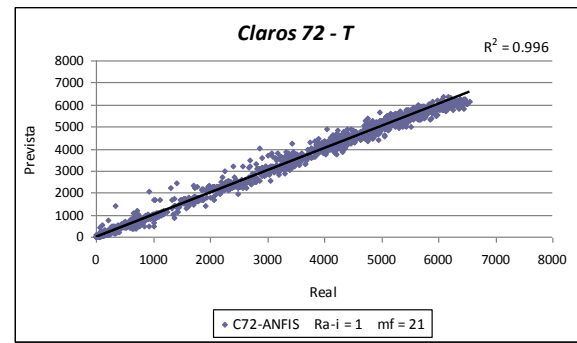
Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	PC72		209.8756	184.2229	—	—
M5Rules	C72	instancias = 21	148.4238	162.6920	29.28 %	11.68 %
M5P	C72	instancias = 13	146.8535	152.2341	30.02 %	17.36 %
MLP-GD	C72	neuronas = 19	168.5502	173.6481	19.69 %	5.74 %
MLP-m(02)	C72	neuronas = 10	168.1849	180.1797	19.86 %	2.19 %
MLP-CGD	C72	neuronas = 12	135.5688	142.7493	35.40 %	22.51 %
MLP-BFGS	C72	neuronas = 13	134.9632	140.8739	35.69 %	23.53 %
RBF-CGD	C72	neuronas = 20	141.1920	145.0327	32.72 %	21.27 %
RBF-BFGS	C72	neuronas = 17	141.1006	146.9714	32.76 %	20.22 %
ANFIS	C72	mf=21, Ra-i=1.0	129.2890	143.5102	38.39 %	22.09 %

Tabla 16.13: Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción entre 49 y 72 horas.

Horizonte	Modelos	Atributos	Características	R^2 Entr.	R^2 Test
De 49 a 72 horas	MLP-CGD	C72	neuronas = 12	0.9963	0.9961
	MLP-BFGS	C72	neuronas = 13	0.9964	0.9962
	RBF-CGD	C72	neuronas = 20	0.9961	0.9960
	RBF-BFGS	C72	neuronas = 17	0.9961	0.9959
	ANFIS	C72	mf=21, Ra-i=1.0	0.9967	0.9960

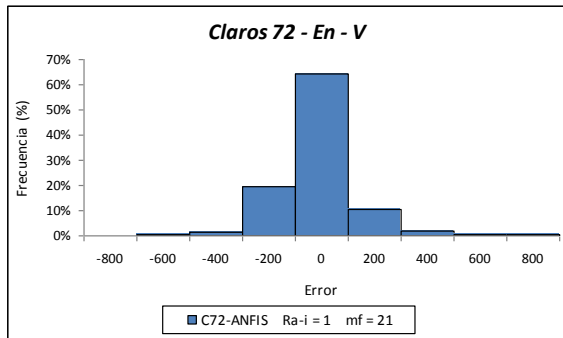


(a) Entrenamiento y Validación

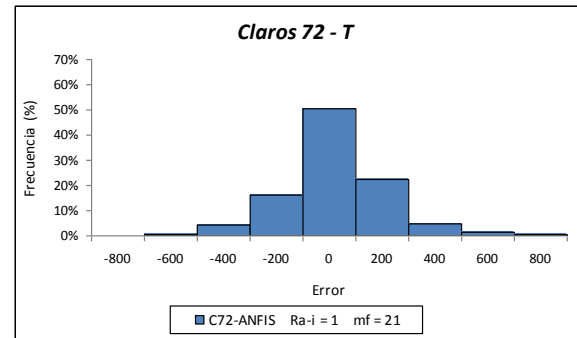


(b) Test

Figura 16.10: Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 21 reglas, para horizontes de predicción entre 49 y 72 horas.



(a) Entrenamiento y Validación



(b) Test

Figura 16.11: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 21 reglas, para horizontes de predicción entre 49 y 72 horas.

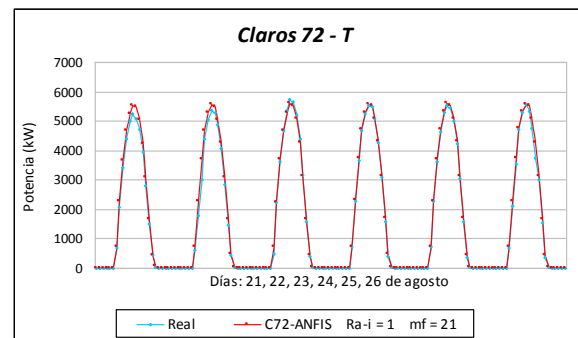
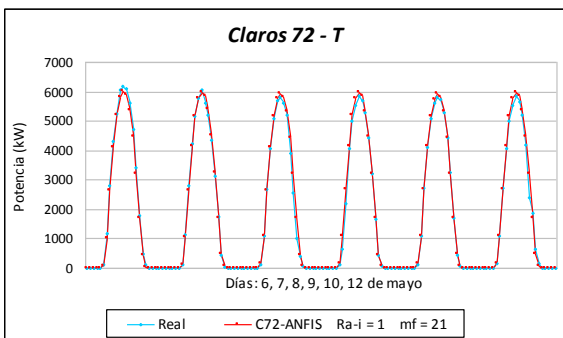


Figura 16.12: Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 21 reglas, para horizontes de predicción entre 49 y 72 horas con datos de test.

substractivo con el valor del rango de influencia $Ra_i = 1$ y 21 reglas, con los datos de los subconjuntos de entrenamiento $Claros_En$ y validación $Claros_V$ (ver Figura 16.10(a)), y los datos del subconjunto de test $Claros_T$ (ver Figura 16.10(b)). El coeficiente de determinación es $R^2 = 0.9967$ para los datos de los subconjuntos $Claros_En$ y $Claros_V$, y $R^2 = 0.9960$ para los datos del subconjunto $Claros_T$, que al estar muy próximos a 1 sugieren que hay una *correlación lineal casi perfecta* (*una regresión lineal casi perfecta*).

En la Figura 16.11 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros, obtenidos con el modelo ANFIS propuesto con los datos de los subconjuntos de entrenamiento $Claros_En$ y validación $Claros_V$ (ver Figura 16.11(a)), y los datos del subconjunto de test $Claros_T$ (ver Figura 16.11(b)).

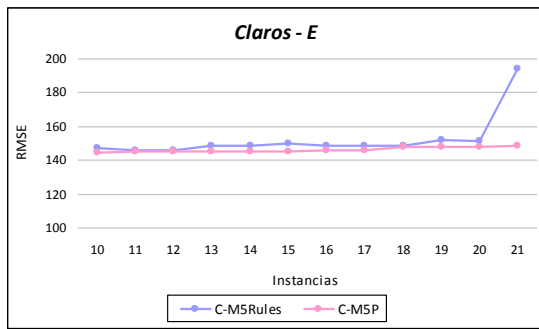
En la Figura 16.12 se representan los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS propuesto para doce días claros del subconjunto de test (días 6, 7, 8, 9, 10 y 12 de mayo de 2012, y días 21, 22, 23, 24, 25 y 26 de agosto de 2012).

16.3.4. Horizonte de predicción mayor a 72 horas

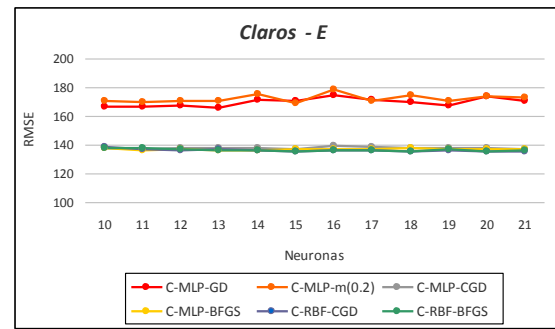
En esta sección se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria en días claros utilizando las técnicas de minería de datos de sistemas de reglas, árboles de decisión, redes neuronales artificiales y sistemas neuro-difusos. Para el desarrollo de estos modelos se han usado los atributos C. El horizonte de predicción es superior a 72 horas (puede ser de años, ya que solo se tiene en cuenta el día del año para el que se hace la predicción). En la Tabla 16.5 con $m = -$ se presenta el significado de la notación que se utiliza en los resultados.

En la Figura 16.13 se presentan los resultados de los valores del RMSE (kW) que se han obtenido con los modelos desarrollados:

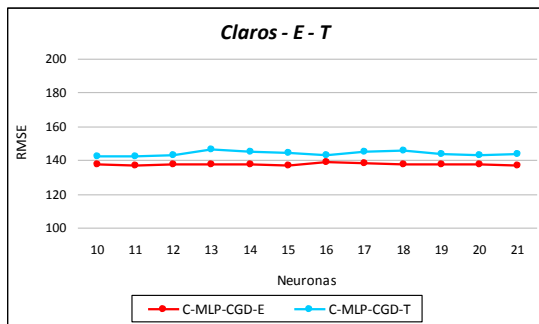
- En la Figura 16.13(a) se comparan los obtenidos con los de los modelos desarrollados con técnicas M5Rules y M5P, variando el número de instancias entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento $Claros_E$ con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan en la mayoría de los casos entre 145 kW y 155 kW, siendo los obtenidos con los modelos basados en la técnica M5P un poco menores que los obtenidos con modelos basados en la técnica M5Rules.
- En la Figura 16.13(b) se comparan los obtenidos con modelos basados en redes neuronales artificiales del tipo MLP-GD, MLP-m(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. En todos los modelos la red neuronal tiene una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento $Claros_E$ con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE con los modelos MLP-GD y MLP-m(0.2) oscilan aproximadamente entre 165 kW y 180 kW, y son mayores que los obtenidos con los modelos MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Para estos últimos modelos, los valores del RMSE oscilan aproximadamente entre 135 kW y 140 kW, siendo, en general, muy parecidos.



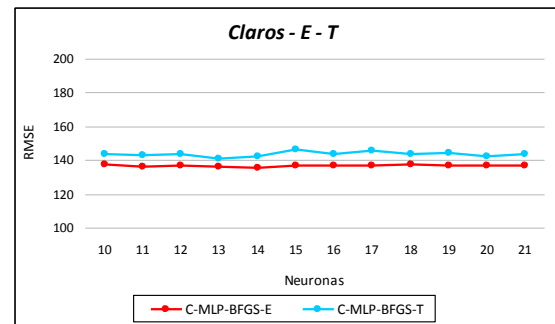
(a) M5Rules y M5P



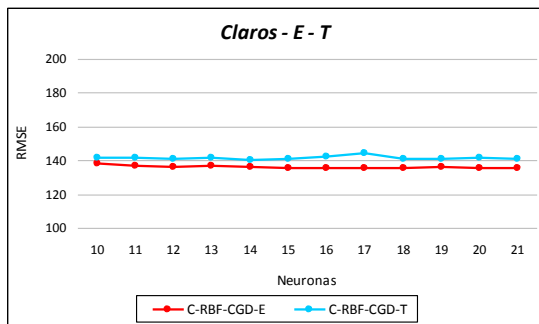
(b) RNAs



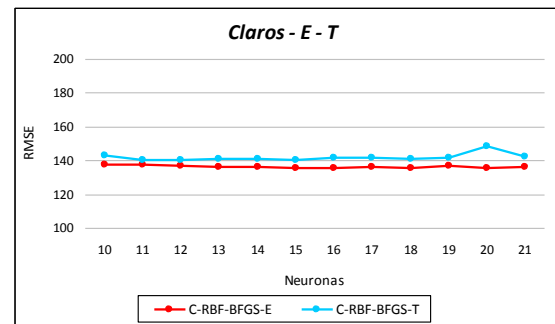
(c) MLP-CGD



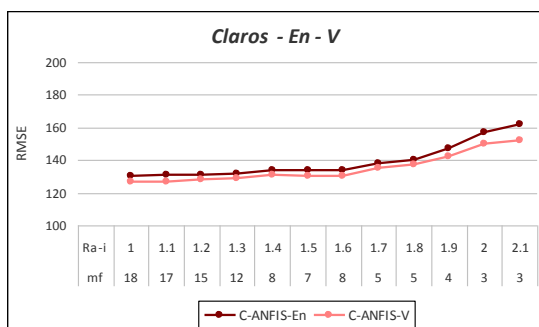
(d) MLP-BFGS



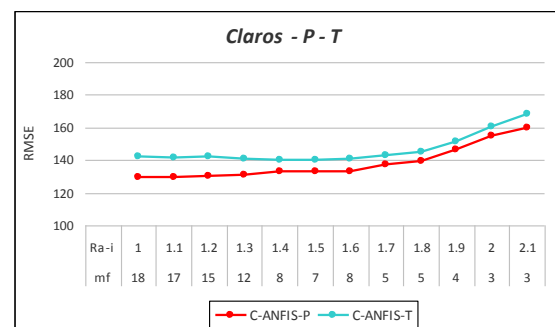
(e) RBF-CGD



(f) RBF-BFGS



(g) ANFIS



(h) ANFIS

Figura 16.13: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con los modelos para días claros desarrollados con los atributos C, para horizontes de predicción mayores a 72 horas con los datos de entrenamiento y test.

- En la Figura 16.13(c) se muestran los obtenidos con los modelos basados en red neuronal MLP-CGD con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos es muy parecido, oscilando aproximadamente entre 135 kW y 150 kW.
- En la Figura 16.13(d) se muestran los obtenidos con los modelos basados en una red neuronal artificial MLP-BFGS con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 135 kW y 150 kW.
- En la Figura 16.13(e) se muestran los obtenidos con los modelos basados en una red neuronal artificial RBF-CGD con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 130 kW y 150 kW.
- En la Figura 16.13(f) se muestran los obtenidos con los modelos basados en una red neuronal artificial RBF-BFGS con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento *Claros_E* con la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test *Claros_T*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 135 kW y 150 kW.
- En la Figura 16.13(g) se muestran los obtenidos con los modelos ANFIS desarrollados con la aplicación del algoritmo substractivo variando su rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de entrenamiento *Claros_En*, y los datos del subconjunto de validación *Claros_V*. Se observa que los valores del RMSE con los datos de los dos subconjuntos son muy parecidos, oscilando aproximadamente entre 125 kW y 165 kW. El valor del RMSE, en la mayoría de los casos, disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.
- En la Figura 16.13(h) se muestran los obtenidos con los modelos ANFIS desarrollados con la aplicación del algoritmo substractivo variando su rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$. Se utilizan los datos del subconjunto de test *Claros_T*. También, se muestran los valores del RMSE ponderado, *Claros_P*, obtenidos con los modelos ANFIS y los datos de los subconjuntos de entrenamiento *Claros_En* y de validación *Claros_V*. Se observa que los valores del RMSE con los datos del subconjunto de test *Claros_T* y los valores del RMSE ponderado *Claros_P* son bastante parecidos, oscilando aproximadamente entre 125 kW

y 170 kW. El valor del RMSE, en la mayoría de los casos, disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.

En la Tabla 16.14 se muestran los valores del RMSE (kW) más relevantes de los modelos obtenidos con la aplicación de las técnicas M5Rules, M5P, RNAs, ANFIS, y los atributos C, en la predicción de la potencia eléctrica media horaria en días claros, y su mejora con respecto al modelo persistente, para horizontes de predicción mayores a 72 horas, con los datos de los subconjuntos de entrenamiento *Claros_E* y de test *Claros_T*.

Los modelos con los que se consiguen los mejores resultados son:

- Sistema ANFIS con partición del espacio de entrada por agrupamiento substractivo con el valor del rango de influencia $Ra_i = 1$ y 18 reglas. El valor del RMSE medio ponderado con los datos de los subconjuntos de entrenamiento y de validación (*Claros_En* y *Claros_V*) es de 129.3417 kW, lo que significa una mejora con respecto al modelo persistente del 28.63 %. En el caso del subconjunto de test *Claros_T*, su valor de RMSE es de 142.2216 kW, lo que significa una mejora con respecto al persistente del 9.71 %. Su comportamiento promedio con los dos subconjuntos de datos es el mejor de entre todos los modelos analizados.
- MLP-BFGS con 14 neuronas en la capa oculta. Con los datos del subconjunto de entrenamiento *Claros_E* se obtiene el valor del RMSE igual a 135.6589 kW, su mejora con respecto al modelo persistente es del 25.14 %, y con los datos del subconjunto de test *Claros_T* se obtiene un valor del RMSE igual a 141.8976 kW, su mejora con respecto al modelo persistente es del 9.91 %.
- RBF-CGD con 21 neuronas en la capa oculta. Con los datos del subconjunto de entrenamiento *Claros_E* se obtiene el valor del RMSE igual a 135.2043 kW, su mejora con respecto al modelo persistente es del 25.39 %, y con los datos del subconjunto de test *Claros_T* se obtiene un valor del RMSE igual a 140.8278 kW, su mejora con respecto al modelo persistente es del 10.59 %.
- RBF-BFGS con 18 neuronas en la capa oculta. Con los datos del subconjunto de entrenamiento *Claros_E* se obtiene el valor del RMSE igual a 135.4201 kW, su mejora con respecto al modelo persistente es del 25.27 %, y con los datos del subconjunto de test *Claros_T* se obtiene un valor del RMSE igual a 141.0653 kW, su mejora con respecto al modelo persistente es del 10.44 %.

En la Tabla 16.15 se muestran los coeficientes de determinación R^2 en días claros de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los modelos desarrollados basados en redes neuronales MLP-CGD con 15 neuronas en la capa oculta, MLP-BFGS con 14 neuronas en la capa oculta, RBF-CGD con 21 neuronas en la capa oculta, RBF-BFGS con 18 neuronas en la capa oculta, y ANFIS con rango de influencia $Ra_i = 1$ y 18 reglas, con los atributos C, para horizontes de predicción de más de 72 horas, con los datos de los subconjuntos de entrenamiento *Claros_E* y de test *Claros_T*. Se observa que los coeficientes de determinación R^2 oscilan entre 0.9961 y 0.9966, estando muy cerca de 1, por lo que hay una *correlación lineal casi perfecta (una regresión lineal casi perfecta)*.

Tabla 16.14: Valores del RMSE (kW) más relevantes de los mejores modelos para días claros desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción mayores a 72 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	PC24		181.2343	157.5208	—	—
M5Rules	C	instancias = 11	145.4365	151.9860	19.75 %	3.51 %
M5P	C	instancias = 10	143.9847	154.1512	20.55 %	2.13 %
MLP-CGD	C	neuronas = 15	136.8300	144.3596	24.50 %	8.35 %
MLP-BFGS	C	neuronas = 14	135.6589	141.8976	25.14 %	9.91 %
RBF-CGD	C	neuronas = 21	135.2043	140.8278	25.39 %	10.59 %
RBF-BFGS	C	neuronas = 18	135.4201	141.0653	25.27 %	10.44 %
ANFIS	C	mf=18, Ra-i=1.0	129.3417	142.2216	28.63 %	9.71 %

Tabla 16.15: Coeficientes de determinación R^2 en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos, para horizontes de predicción mayores a 72 horas.

Horizonte	Modelos	Atributos	Características	R^2 Entr.	R^2 Test
Mayor a 72 horas	MLP-CGD	C	neuronas = 15	0.9963	0.9961
	MLP-BFGS	C	neuronas = 14	0.9963	0.9962
	RBF-CGD	C	neuronas = 21	0.9963	0.9962
	RBF-BFGS	C	neuronas = 18	0.9963	0.9962
	ANFIS	C	mf=18, Ra-i=1.0	0.9966	0.9961

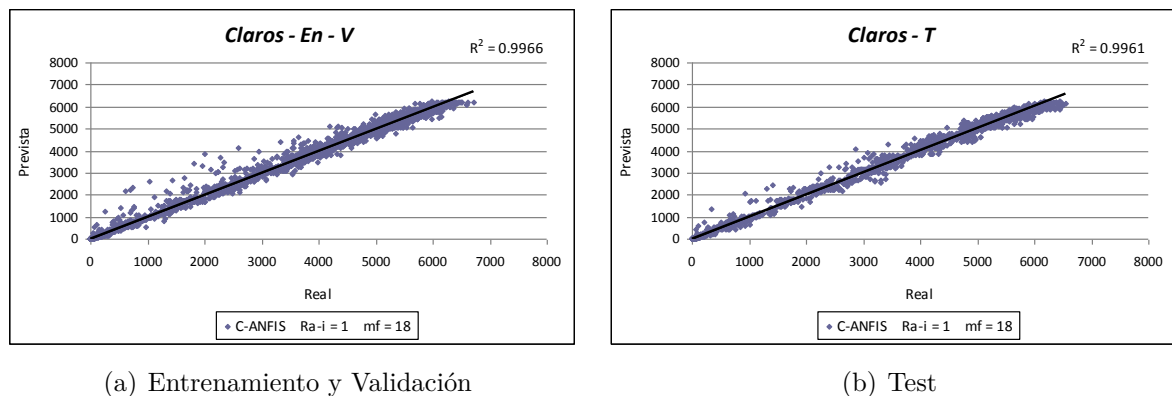


Figura 16.14: Diagramas de dispersión en días claros con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS de 18 reglas, para horizontes de predicción mayores a 72 horas.

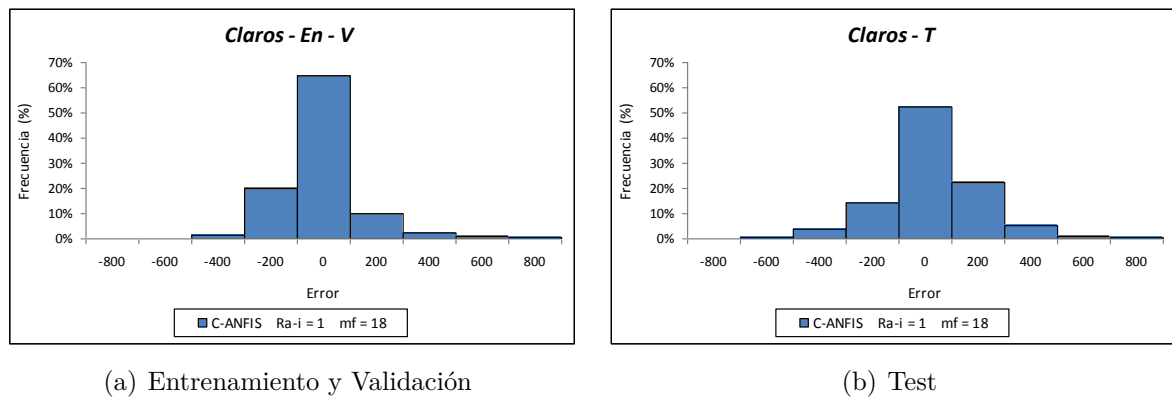


Figura 16.15: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros obtenidos con el modelo ANFIS de 18 reglas, para horizontes de predicción mayores a 72 horas.

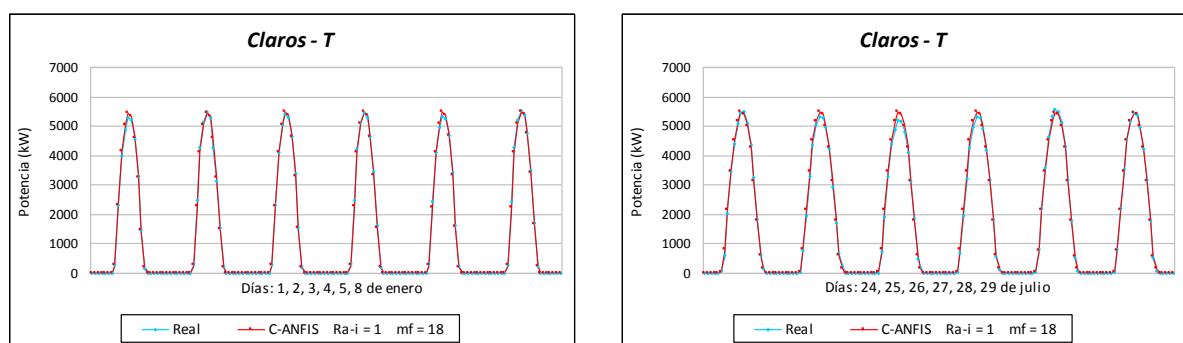


Figura 16.16: Valores reales de la potencia eléctrica media horaria (kW) en días claros y la predicción obtenida con el modelo ANFIS de 18 reglas, para horizontes de predicción mayores a 72 horas con datos de test.

En la Figura 16.14 se muestran los diagramas de dispersión en días claros con valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS obtenido mediante la partición del espacio de entrada por agrupamiento substractivo con el valor del rango de influencia $Ra_i = 1$, 18 reglas, y los atributos C, para horizontes de predicción de más de 72 horas, con los datos de los subconjuntos de entrenamiento *Claros_En* y validación *Claros_V* (ver Figura 16.14(a)), y los datos del subconjunto de test *Claros_T* (ver Figura 16.14(b)). El coeficiente de determinación es $R^2 = 0.9966$ con los datos de los subconjuntos *Claros_En* y *Claros_V*, y $R^2 = 0.9961$ con los datos del subconjunto *Claros_T*, estando muy cerca de 1, por lo que hay una *correlación lineal casi perfecta (una regresión lineal casi perfecta)*.

En la Figura 16.15 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) en días claros, obtenidos con el modelo ANFIS propuesto con los datos de los subconjuntos de entrenamiento *Claros_En* y validación *Claros_V* (ver Figura 16.15(a)), y los datos del subconjunto de test *Claros_T* (ver Figura 16.15(b)).

En la Figura 16.16 se representan los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo ANFIS propuesto para doce días claros pertenecientes al subconjunto de test (días 1, 2, 3, 4, 5 y 8 de enero de 2012 en la gráfica izquierda, y días 24, 25, 26, 27, 28 y 29 de julio de 2012 en la gráfica de la derecha).

16.4. Resultados de los modelos para cualquier tipo de día

A continuación, se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria para cualquier tipo de día utilizando diferentes técnicas de minería de datos, para horizontes de predicción de 1 a 8 horas, entre 0 y 24 horas, entre 25 y 48 horas, y entre 49 y 72 horas.

16.4.1. Sistemas de reglas y árboles de decisión

En esta sección se exponen los resultados de los valores del RMSE que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria, para cualquier tipo de día, utilizando modelos basados en sistemas de reglas y árboles de decisión, para los horizontes de predicción de 1 a 8 horas. En la Tabla 16.7 se presenta el significado de la notación que se utiliza en los resultados.

Se comparan los valores del RMSE (kW) que se han obtenido en los modelos de predicción desarrollados con la técnica M5Rules y los atributos Gm y Rm (donde $m \in \{1, \dots, 8\}$ para los horizontes de predicción de 1 a 8 horas) variando el número instancias entre 10 y 21. En la Figura 16.17 se utilizan los datos del subconjunto de entrenamiento *E* con la técnica de validación cruzada de 5-iteraciones. Se observa que los modelos con los atributos Gm obtienen un valor RMSE menor.

Se comparan los valores del RMSE (kW) que se han obtenido en los modelos de predicción desarrollados con la técnica M5P y los atributos Gm y Rm, variando el número de instancias entre 10 y 21. En la Figura 16.18 se utilizan los datos del subconjunto de

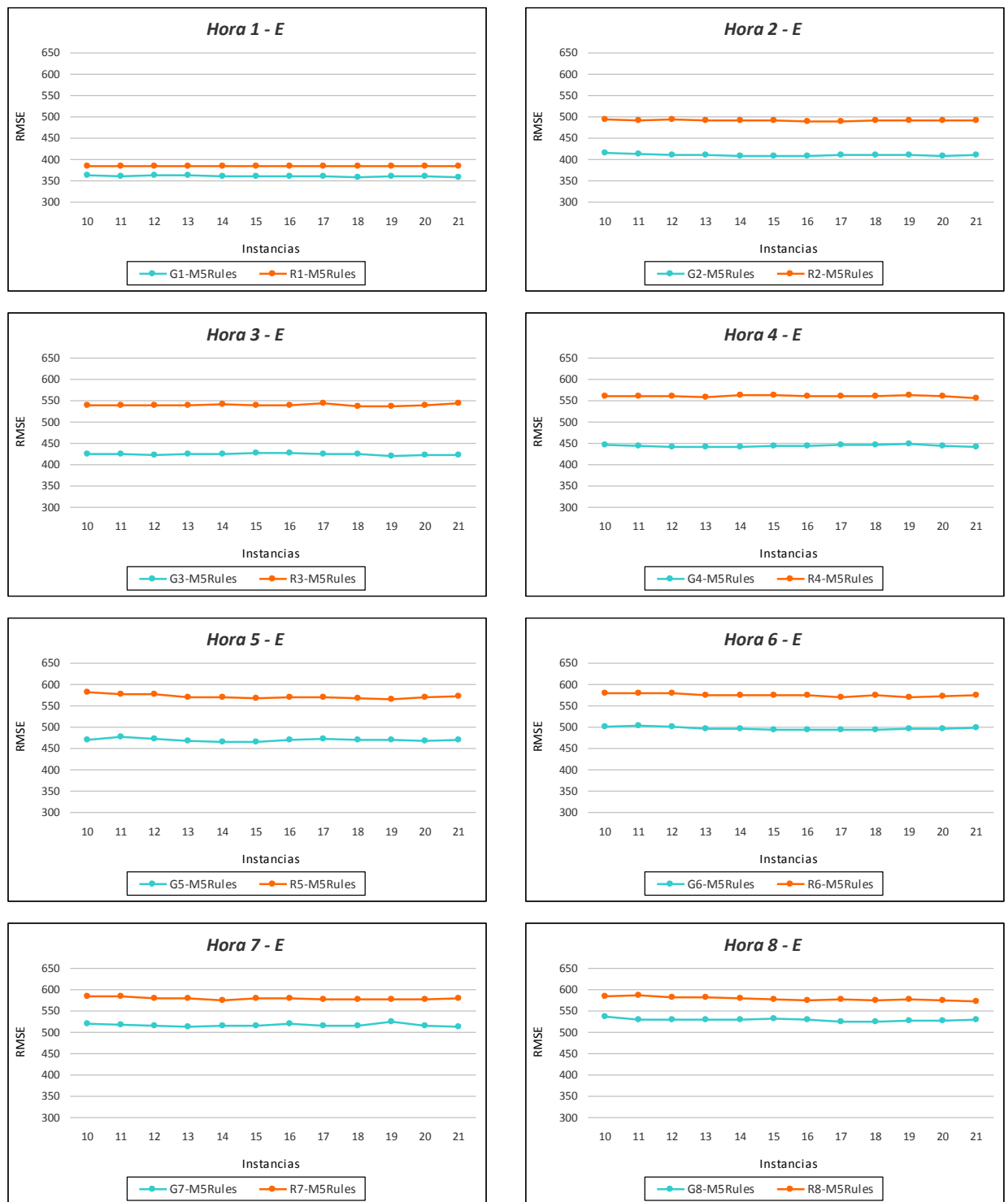


Figura 16.17: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en M5Rules y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

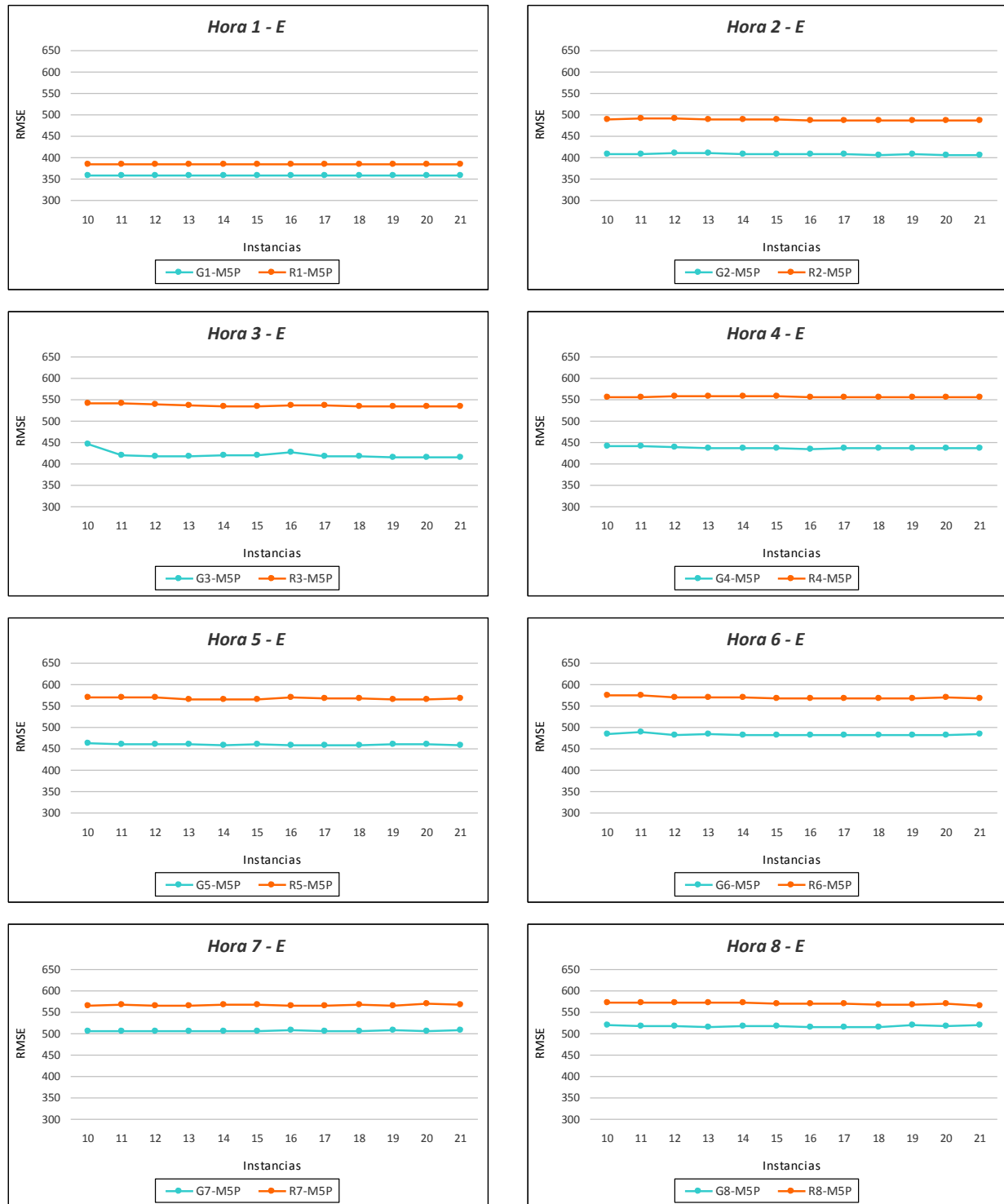


Figura 16.18: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en M5P y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los modelos con los atributos G_m obtienen un valor RMSE menor.

16.4.2. Redes neuronales artificiales

En esta sección se exponen los resultados de los valores del RMSE que se han obtenido con los modelos desarrollados para la predicción de la potencia eléctrica media horaria, para cualquier tipo de día, utilizando los tipos de redes neuronales artificiales MLP-GD, MLP-momentum(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Todos los modelos cuentan con una capa oculta con número de neuronas entre 10 y 21, y se han desarrollado con los atributos H_m , G_m y R_m ($m \in \{1, \dots, 8\}$) para los horizontes de predicción de 1 a 8 horas. Se utilizan los datos del subconjunto de entrenamiento E usando la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test T . En la Tabla 16.7 se presenta el significado de la notación que se utiliza para los resultados. En primer lugar, se comparan los valores del RMSE (kW) que se obtienen utilizando los atributos G_m y R_m , en:

- Los modelos basados en redes MLP-GD con los datos del subconjunto de entrenamiento E (ver Figura 16.19).
- Los modelos basados en redes MLP-momentum(0.2) con los datos del subconjunto de entrenamiento E (ver Figura 16.20).
- Los modelos basados en redes MLP-CGD con los datos del subconjunto de entrenamiento E (ver Figura 16.21), y con los datos del subconjunto de test T (ver Figura 16.22).
- Los modelos basados en redes MLP-BFGS con los datos del subconjunto de entrenamiento E (ver Figura 16.23), y con los datos del subconjunto de test T (ver Figura 16.24).
- Los modelos basados en redes RBF-CGD con los datos del subconjunto de entrenamiento E (ver Figura 16.25), y con los datos del subconjunto de test T (ver Figura 16.26).
- Los modelos basados en redes RBF-BFGS con los datos del subconjunto de entrenamiento E (ver Figura 16.27), y con los datos del subconjunto de test T (ver Figura 16.28).

Se observa que todos los modelos que utilizan como entradas los atributos G_m presentan un valor del RMSE menor que el de los modelos que utilizan como entradas los atributos R_m .

En la Figura 16.29 se comparan los valores del RMSE (kW) que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria, para cualquier tipo de día, con modelos basados en redes neuronales artificiales MLP-GD, MLP-momentum(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Todos estos modelos cuentan con una capa oculta con un número de neuronas entre 10 y 21, y se han desarrollado con los atributos G_m ($m \in \{1, \dots, 8\}$) para los horizontes de predicción de 1 a 8 horas. Se utilizan

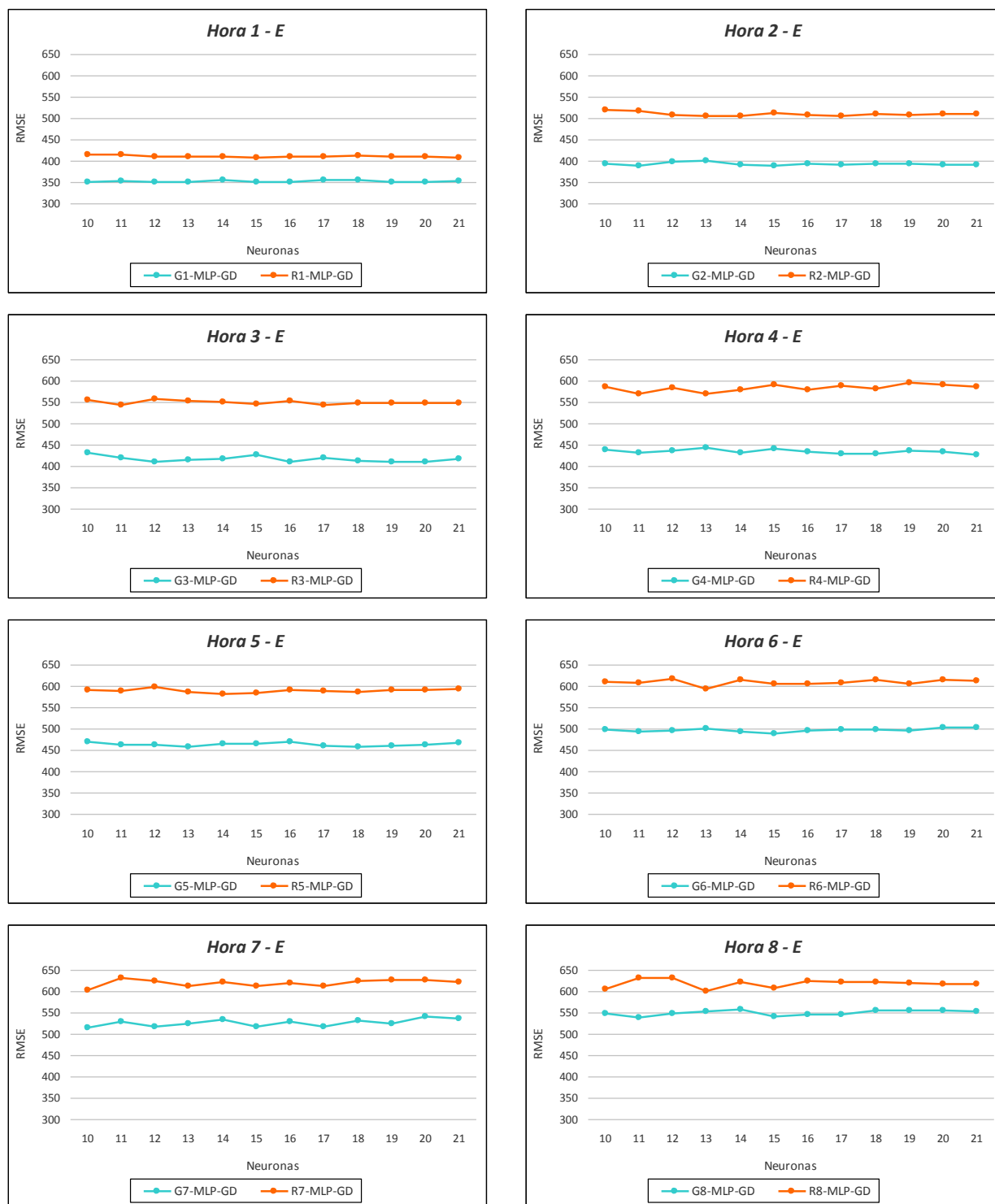


Figura 16.19: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-GD y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

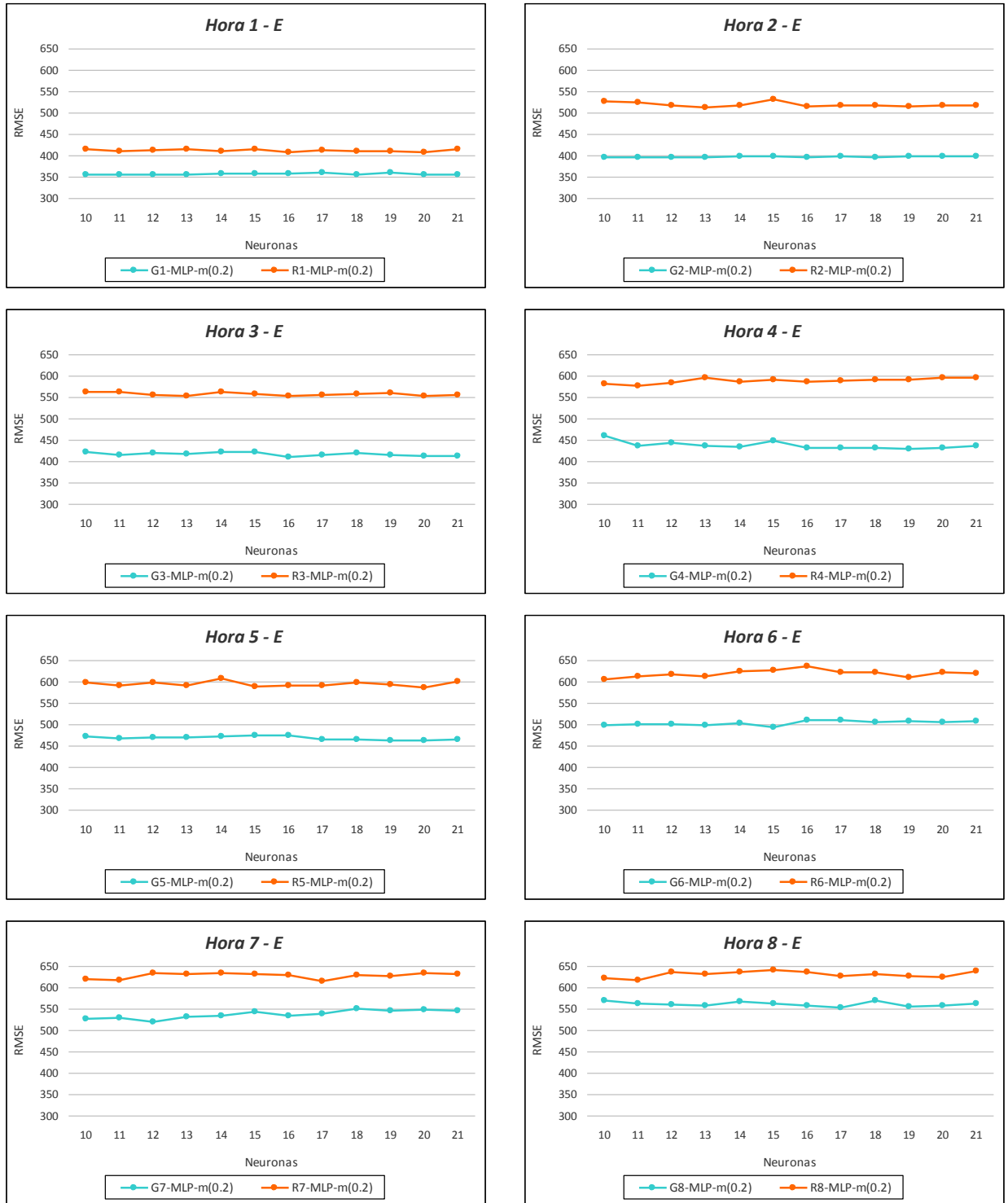


Figura 16.20: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-momentum(0.2) y desarrollados con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

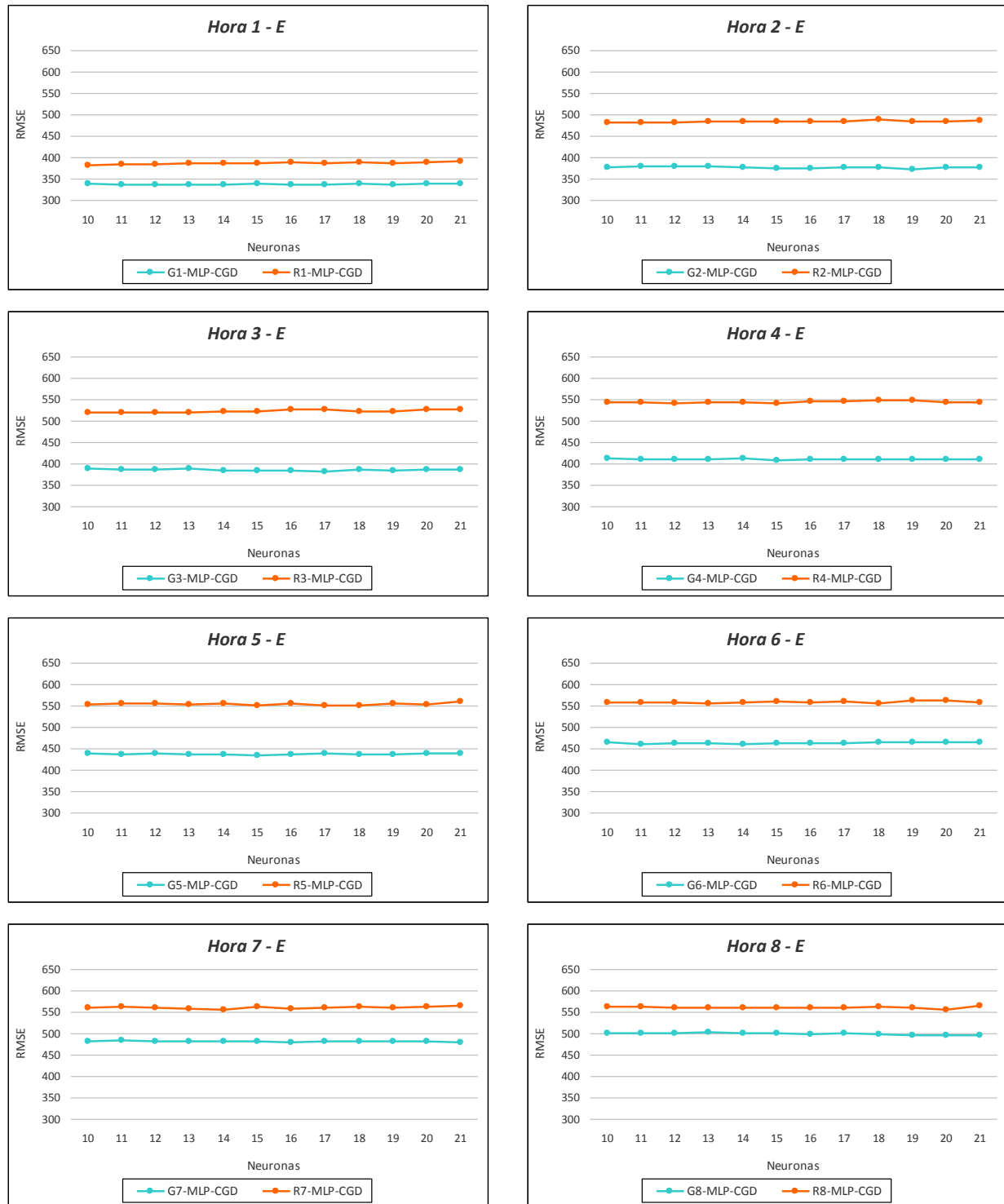


Figura 16.21: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-CGD y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

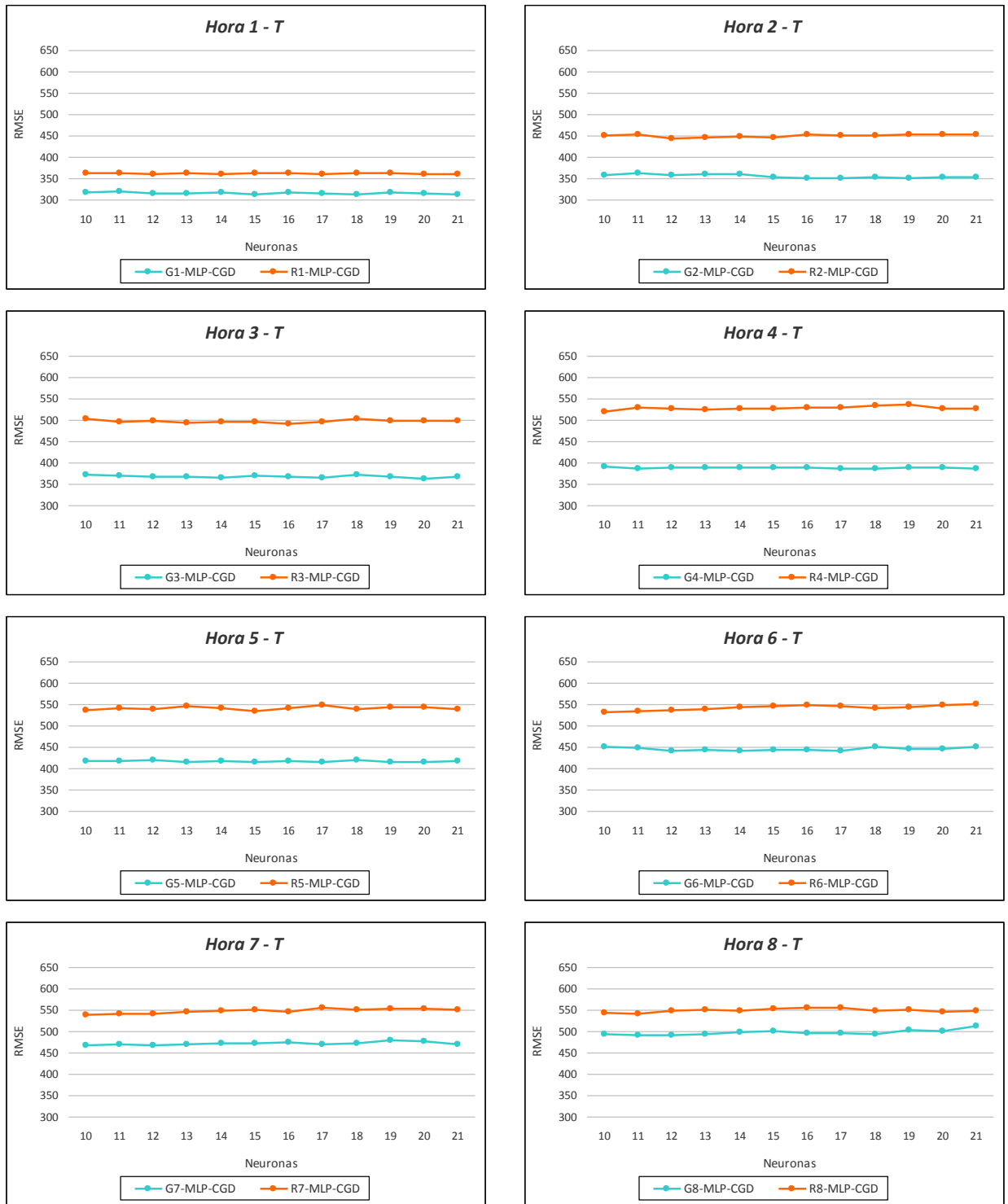


Figura 16.22: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-CGD y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de test.

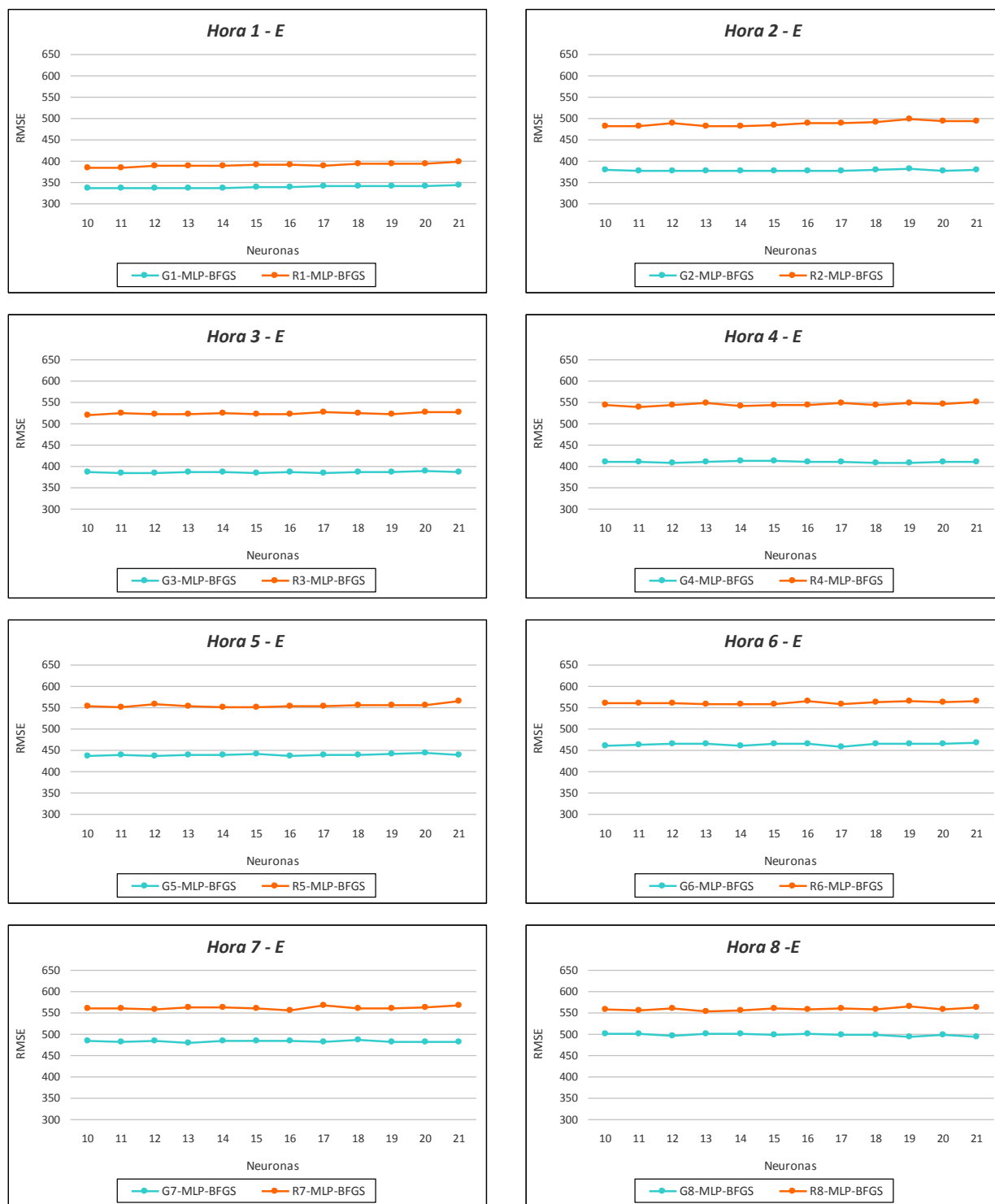


Figura 16.23: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-BFGS y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

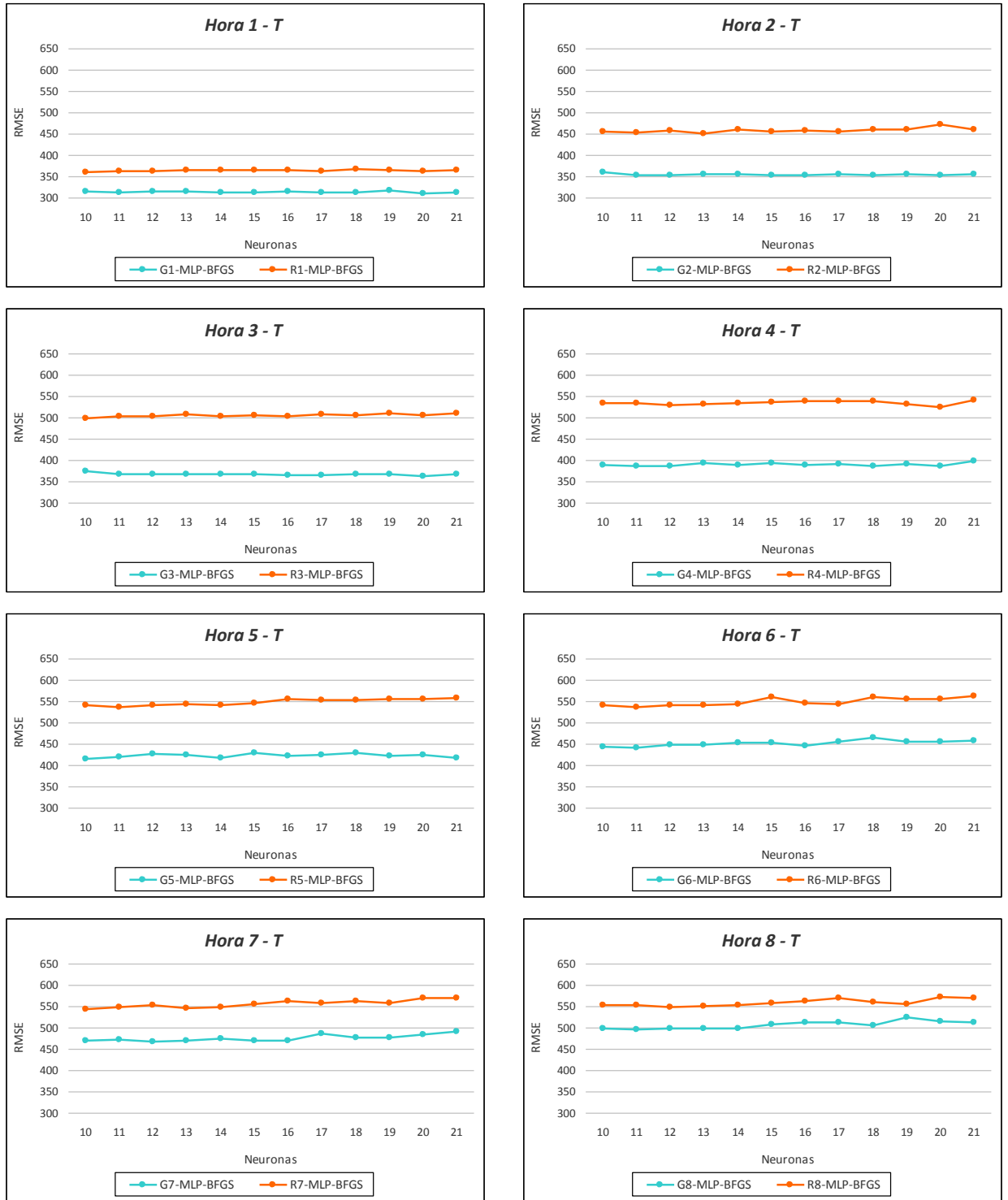


Figura 16.24: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos MLP-BFGS y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de test.

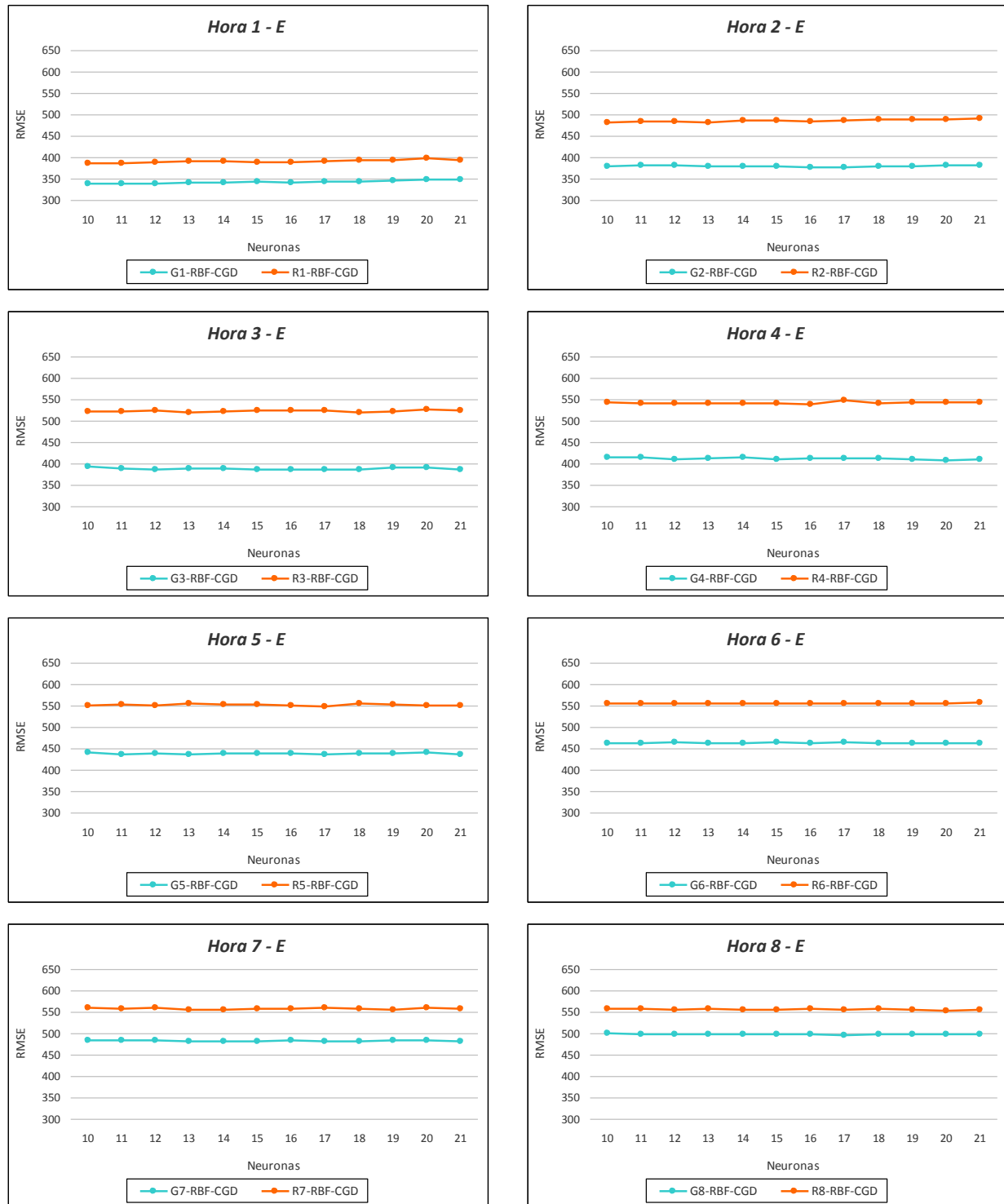


Figura 16.25: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-CGD y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

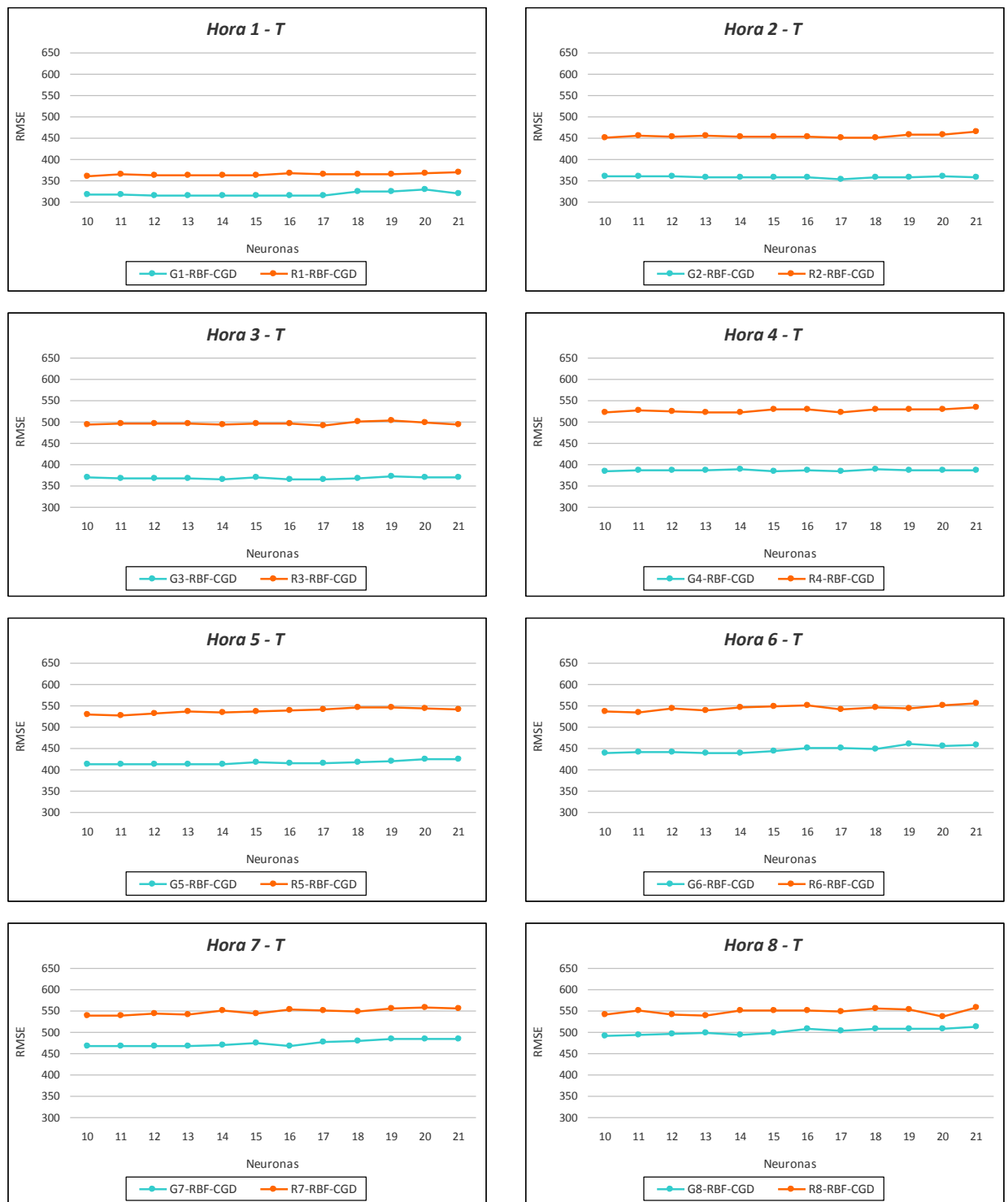


Figura 16.26: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-CGD y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de test.

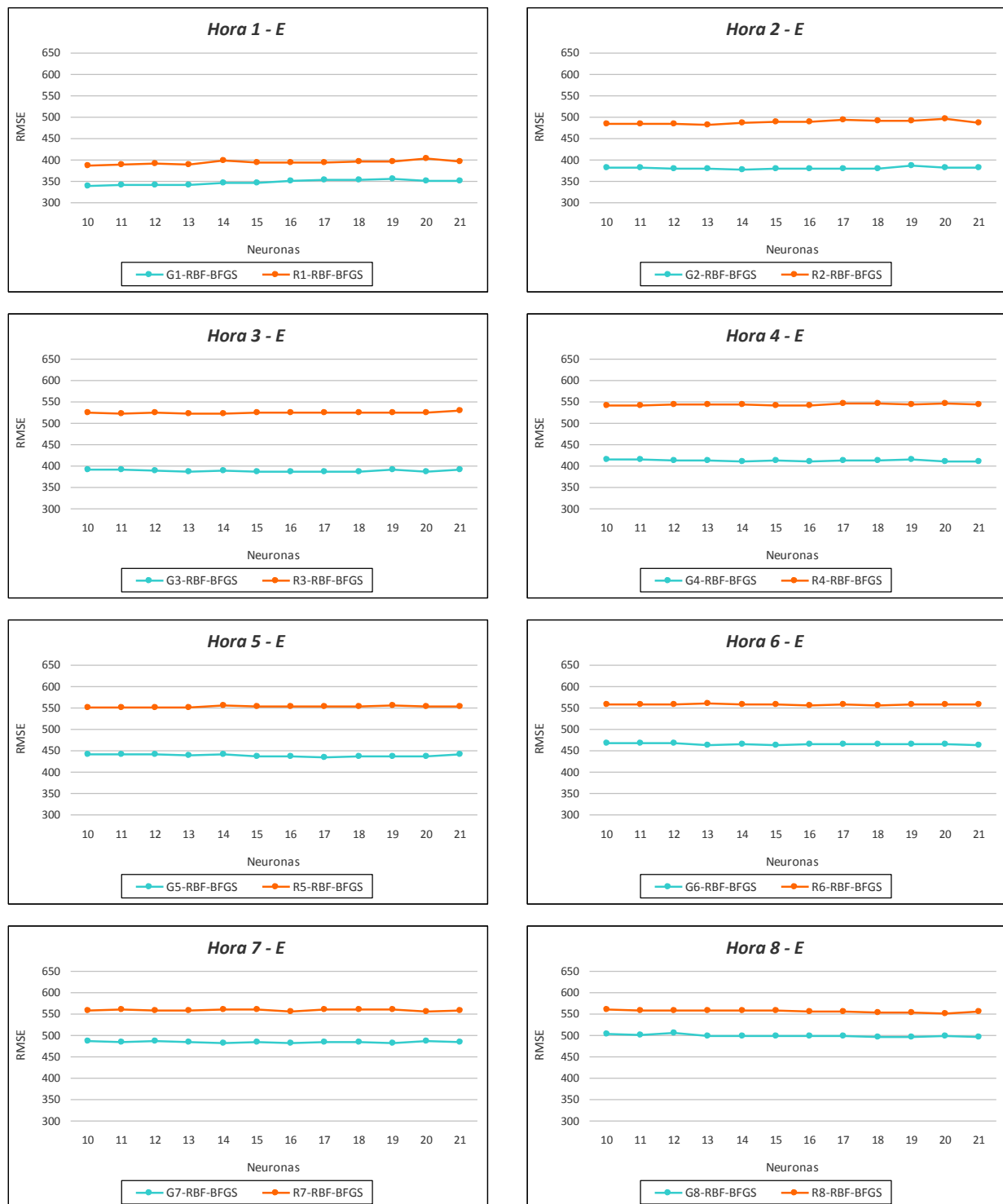


Figura 16.27: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-BFGS y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

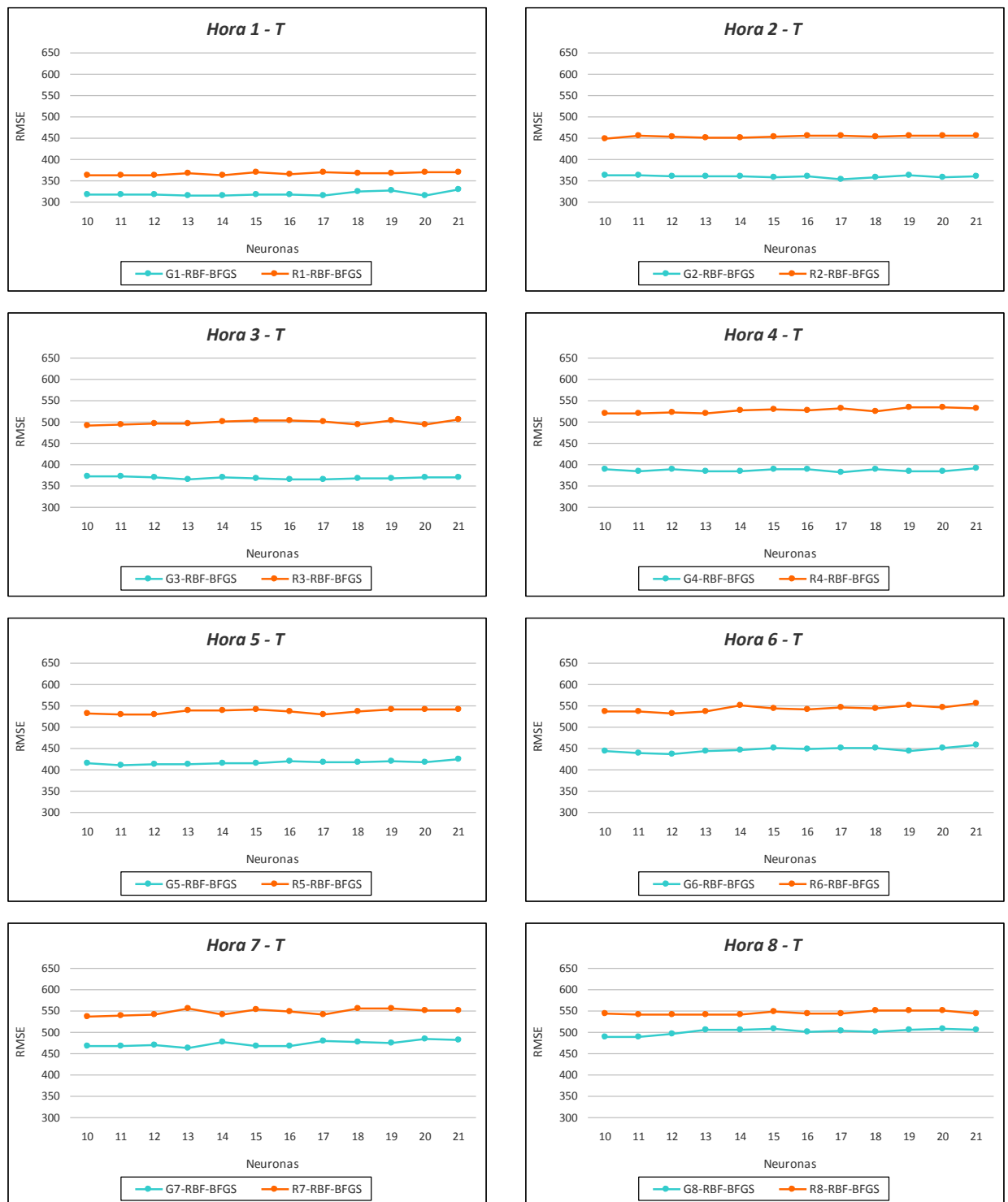


Figura 16.28: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos RBF-BFGS y desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de test.

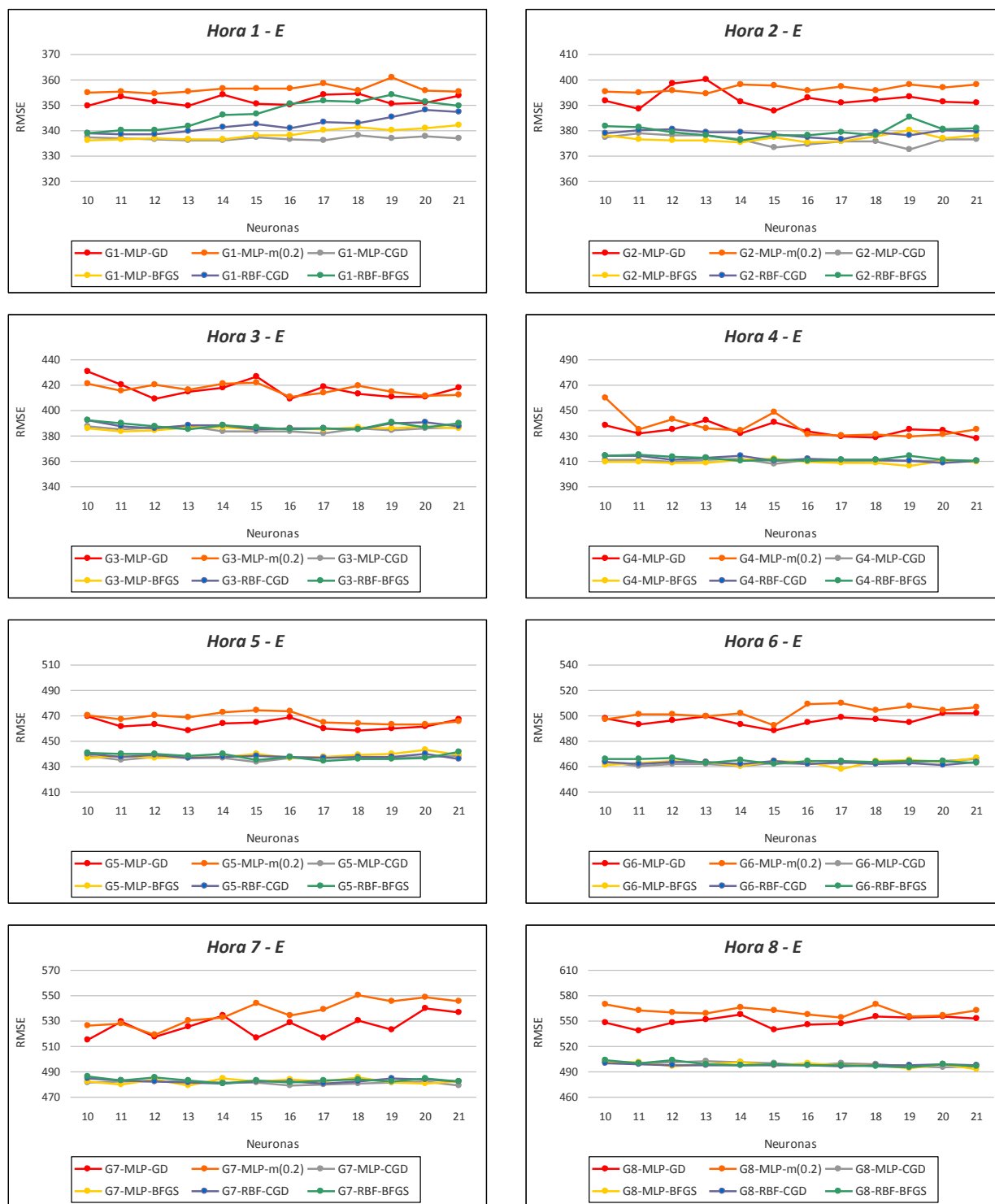


Figura 16.29: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en RNA y desarrollados con los atributos G_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

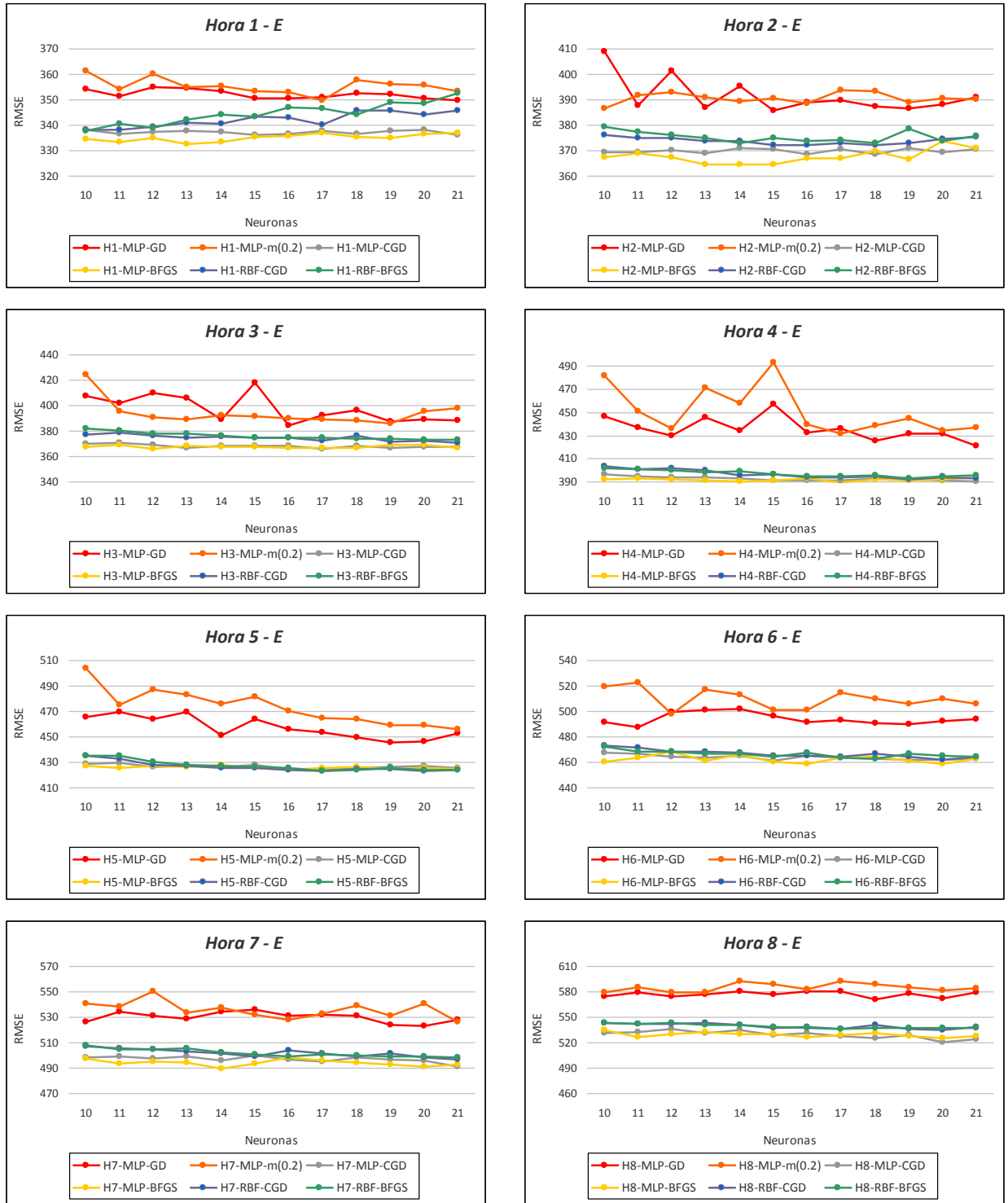


Figura 16.30: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en RNA y desarrollados con los atributos H_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.



Figura 16.31: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en MLP-BFGS y desarrollados con los atributos G_m y H_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

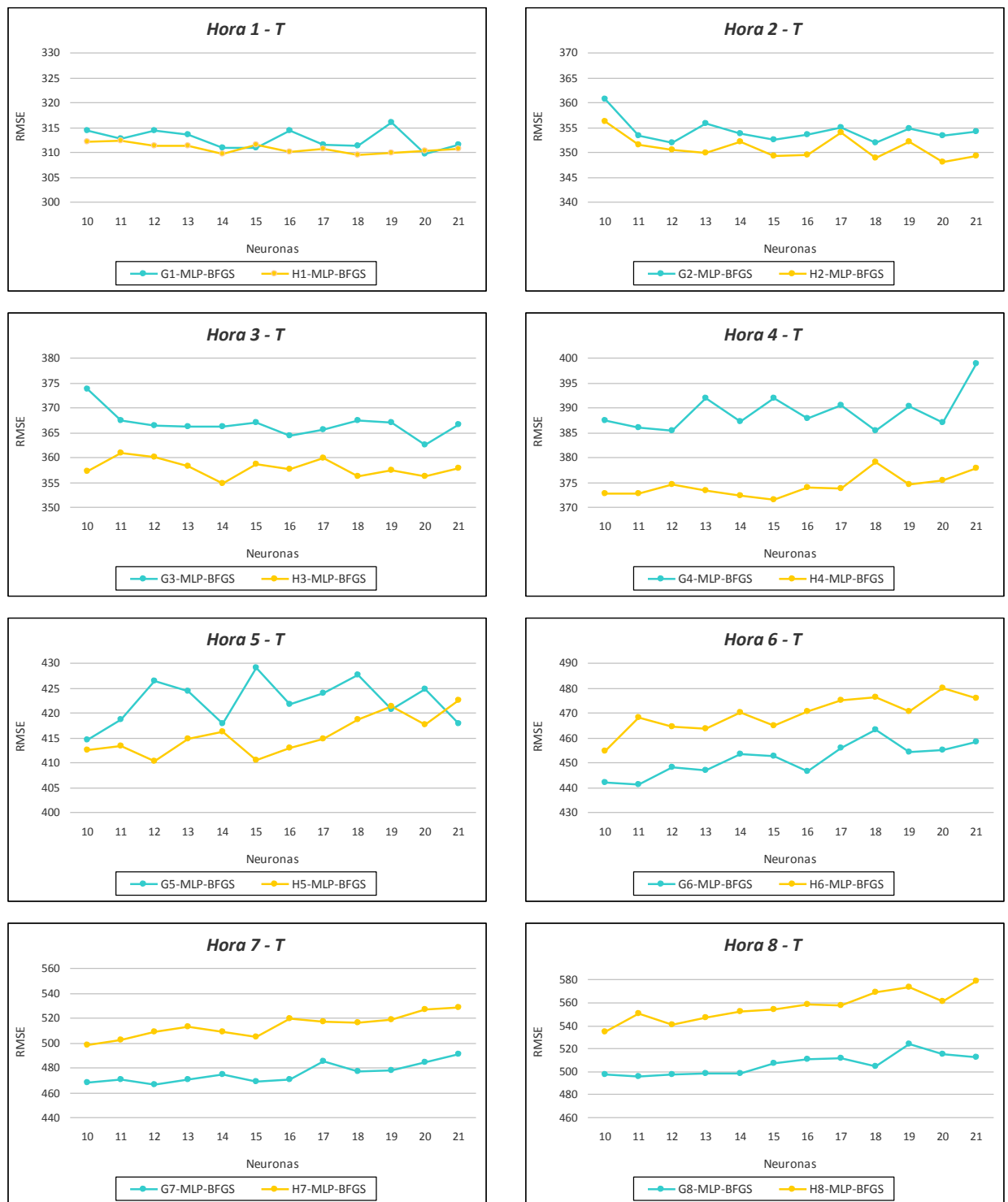


Figura 16.32: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en MLP-BFGS y desarrollados con los atributos Gm y Hm, para horizontes de predicción de 1 a 8 horas con los datos de test.

los datos del subconjunto de entrenamiento E usando la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE (kW) son menores en los modelos MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS.

En la Figura 16.30 se comparan valores del RMSE (kW) que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria, para cualquier tipo de día, con modelos basados en redes neuronales artificiales MLP-GD, MLP-momentum(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Todos estos modelos cuentan con una capa oculta con un número de neuronas entre 10 y 21, y se han desarrollado con los atributos H_m ($m \in \{1, \dots, 8\}$) para los horizontes de predicción de 1 a 8 horas. Se utilizan los datos del subconjunto de entrenamiento E usando la técnica de validación cruzada de 5-iteraciones. Se observa que en la mayoría de los casos el valor del RMSE (kW) es menor en los modelos basados en MLP-BFGS.

Se comparan los valores del RMSE (kW) que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria, para cualquier tipo de día, con modelos basados en la red neuronal MLP-BFGS con una capa oculta con número de neuronas entre 10 y 21. Los modelos se han desarrollado con los atributos G_m y H_m para los horizontes de predicción de 1 a 8 horas. En la Figura 16.31 se representan los valores obtenidos con los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones, y en la Figura 16.32 se representan los valores obtenidos con los datos del subconjunto de test T . Se observa que los menores valores del RMSE (kW) se obtienen si se utilizan los atributos H_m para horizontes entre 1 y 5 horas y los atributos G_m para horizontes entre 6 y 8 horas.

16.4.3. Máquinas de vectores soporte

En esta sección se exponen los resultados de los valores del RMSE que se han obtenido con los modelos desarrollados para la predicción de la potencia eléctrica media horaria, para cualquier tipo de día, utilizando modelos basados en máquinas de vectores soporte (SVM), para los horizontes de predicción de 1 a 8 horas. Con estos modelos se mantienen constantes el parámetro $\varepsilon = 0.001$, que mide la finura con que se va a tolerar el error con los datos del subconjunto de entrenamiento, y los valores de las variables de holgura, $\zeta_i = 0.001$. Se varía los valores de la constante $C \in \{5, 10, 15, 20\}$, que determina la holgura del margen blando. Se usan las funciones núcleo gaussianas con los valores del parámetro $\gamma \in \{0.2, 0.5, 1\}$. Se utilizan los datos del subconjunto de entrenamiento E usando la técnica de validación cruzada de 5-iteraciones, y los datos del subconjunto de test T . En la Tabla 16.7 se presenta el significado de la notación que se utiliza en los resultados. Se comparan los valores del RMSE (kW) que se han obtenido con:

- Los atributos G_m y R_m ($m \in \{1, \dots, 8\}$), para los horizontes de predicción de 1 a 8 horas, con los datos del subconjunto de entrenamiento E en la Figura 16.33. Se observa que los modelos con los atributos G_m obtienen un valor RMSE menor.
- Los atributos H_m y G_m ($m \in \{1, \dots, 8\}$), para los horizontes de predicción de 1 a 8 horas, con los datos del subconjunto de entrenamiento E en la Figura 16.34. Se observa que los menores valores del RMSE (kW) se obtienen si se utilizan los atributos H_m para los horizontes entre 1 y 5 horas, y los atributos G_m para horizontes entre 6 y 8 horas.

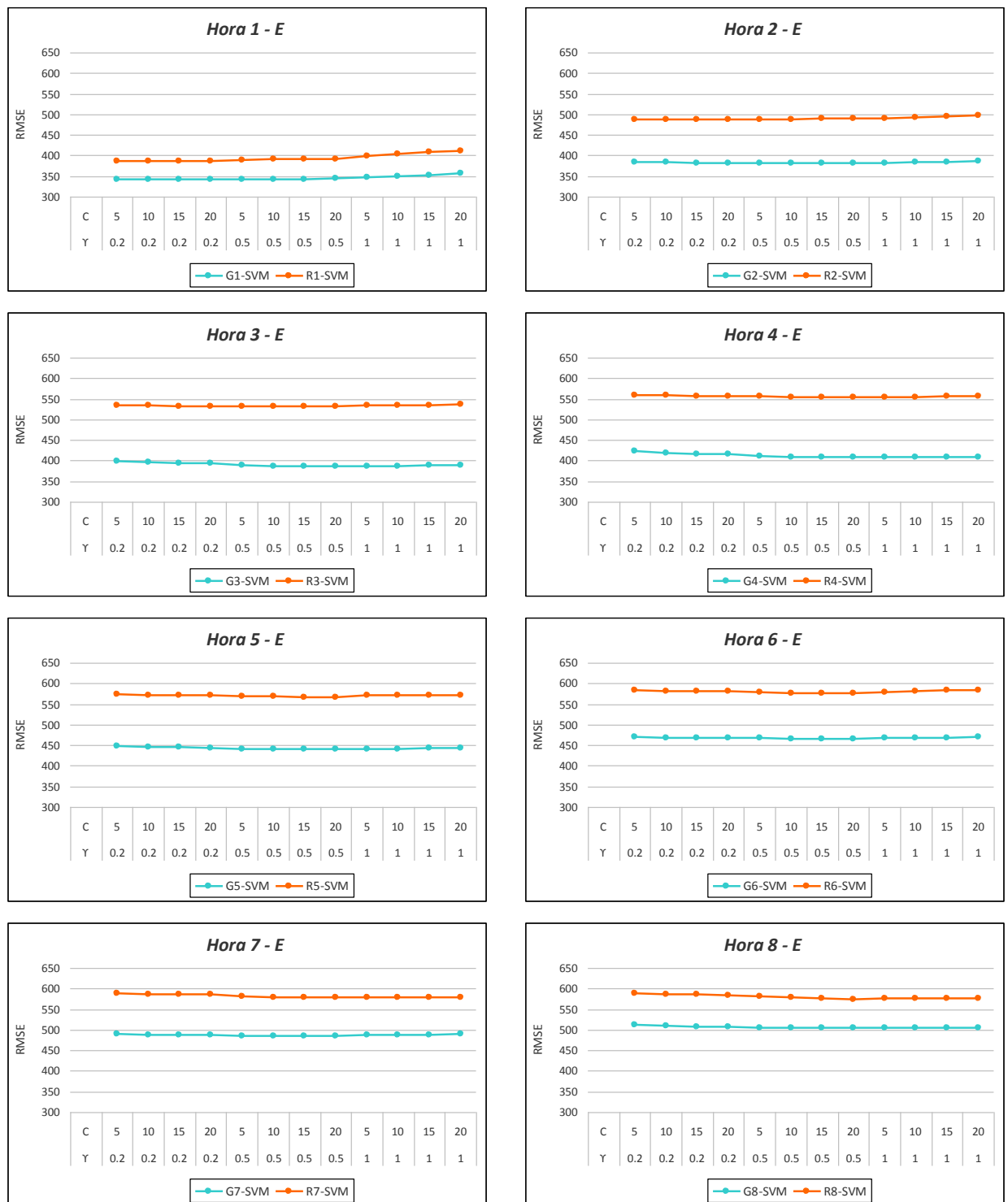


Figura 16.33: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en SVM desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

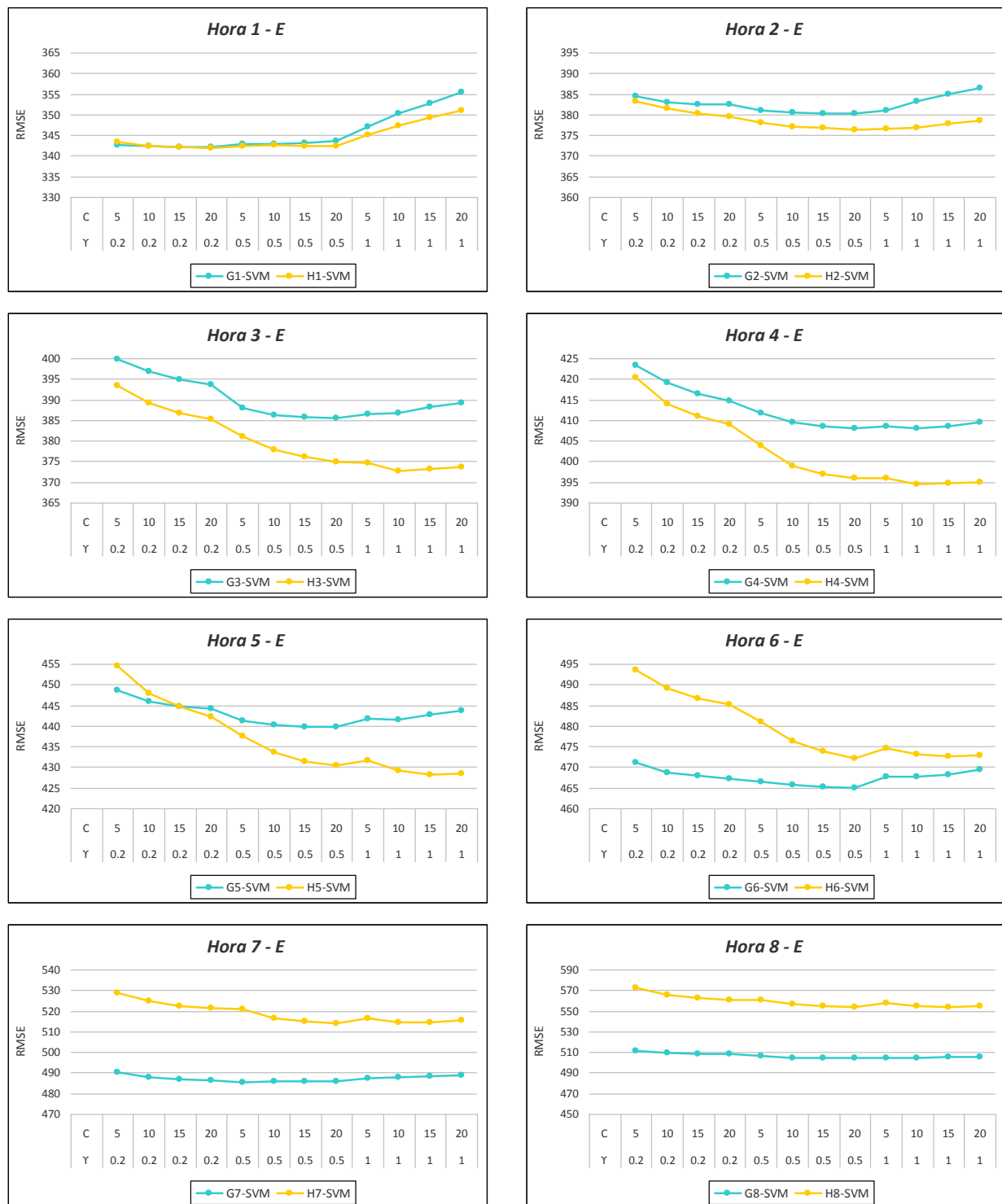


Figura 16.34: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en SVM desarrollados con los atributos G_m y H_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

16.4.4. Sistemas de inferencia neuro-difusos

En esta sección se comparan los valores del RMSE (kW) que se obtienen en la predicción de la potencia eléctrica media horaria, con modelos basados en sistemas ANFIS desarrollados usando los atributos G_m y R_m ($m \in \{1, \dots, 8\}$) para los horizontes de predicción de 1 a 8 horas con diferentes subconjuntos de datos. La partición del espacio de entrada se realiza por medio del agrupamiento substractivo, variando los valores del rango de influencia $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1\}$, y el resto de los parámetros se mantienen constantes: 1.25 para el factor de reducción de densidad, 0.5 para el factor de aceptación, y 0.15 para el factor de rechazo. Además, se muestra el número de las funciones de pertenencia (mf) para cada uno de dichos rangos de influencia. En la Tabla 16.7 se presenta el significado de la notación que se utiliza en los resultados.

En la Figura 16.35 se muestran los valores del RMSE (kW) que se obtienen con los datos del subconjunto de entrenamiento, En , y en la Figura 16.36 con los datos del subconjunto de validación, V . En la Figura 16.37 se muestran los valores del RMSE (kW) medios ponderados, P , calculados en función al porcentaje de datos del subconjunto de entrenamiento, En (75 %) y del subconjunto de validación, V (25 %). En la Figura 16.38 se muestran los valores del RMSE que se obtienen con los datos del subconjunto de test, T . Se observa que para todos los subconjuntos de datos (entrenamiento En , validación V y test T), el valor RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos ANFIS es menor con los atributos G_m que con los atributos R_m .

A modo de ejemplo, para el caso del horizonte de predicción a 1 hora, se presentan parte de los resultados obtenidos con el modelo ANFIS desarrollado con los atributos $H1$ y con la partición del espacio de entrada por medio del agrupamiento substractivo (con 1.2 el valor del rango de influencia, 1.25 el factor de reducción de densidad, 0.5 el factor de aceptación, y 0.15 el factor de rechazo). Así en la Figura 16.39 se muestran los diagramas de dispersión que se obtienen con valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida: en la Figura 16.39(a) con los datos de los subconjuntos de entrenamiento En y de validación V , y en la Figura 16.39(b) con los datos del subconjunto de test T . En la Figura 16.40 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de luz solar): en la Figura 16.40(a) con los datos de los subconjuntos de entrenamiento En y de validación V , y en la Figura 16.40(b) con los datos del subconjunto de test T . Por último, en la Figura 16.41 se representan los valores reales de la potencia eléctrica media horaria (kW) y su valor de predicción obtenido, para los días del 11 al 16 de enero y del 5 al 10 de noviembre de 2012, que pertenecen al subconjunto de test T .

16.4.5. Resultados de los mejores modelos

En esta sección se muestran los mejores valores del RMSE (kW) que se obtienen con cada tipo de modelo basado en las diferentes técnicas de minería de datos utilizadas, en la predicción de la potencia eléctrica media horaria para cualquier tipo de día y su mejora con respecto al modelo persistente, para los horizontes de predicción de 1 a 8 horas, con los datos de los subconjuntos de entrenamiento E y de test T .

- En la Tabla 16.16 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos

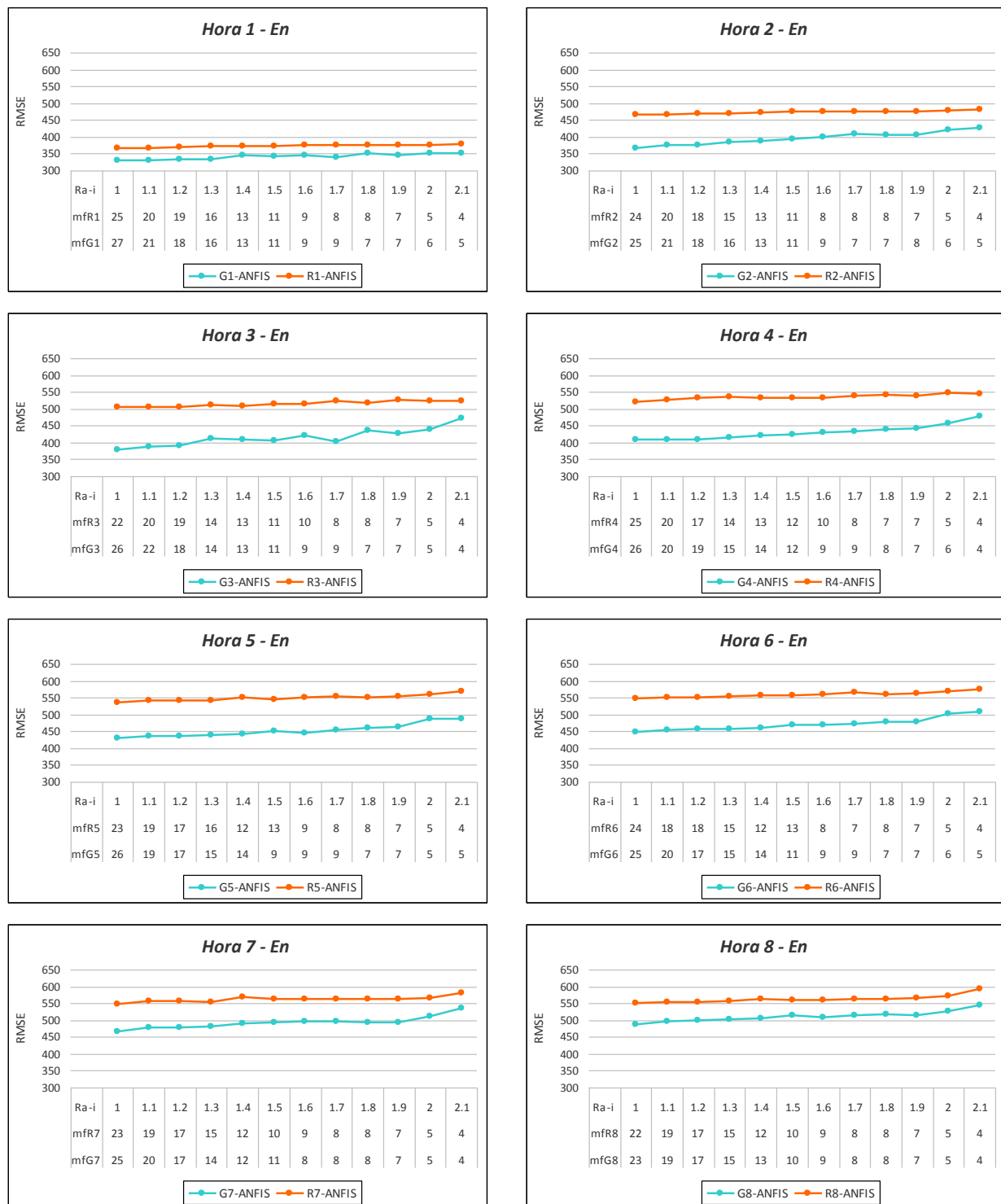


Figura 16.35: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

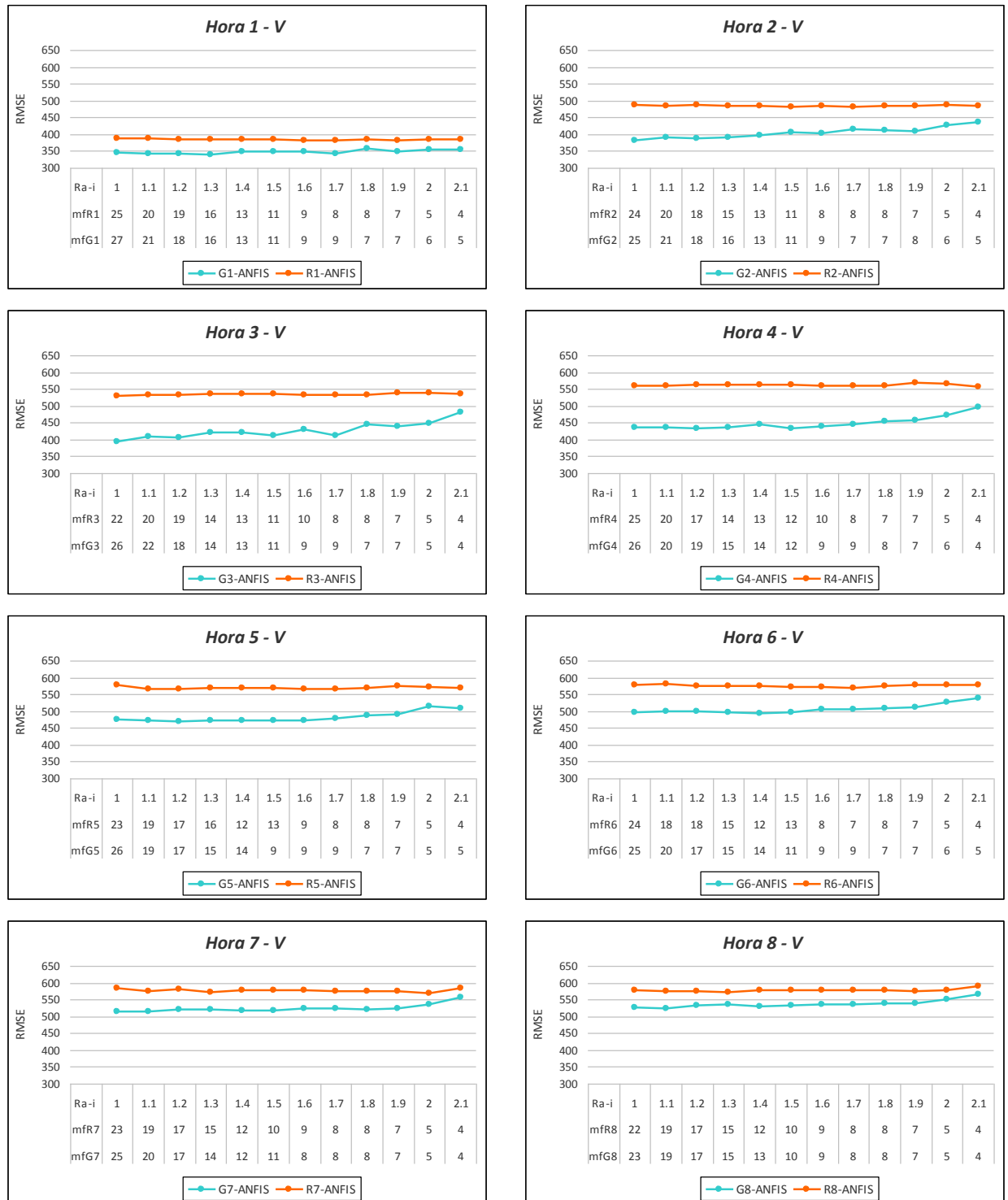


Figura 16.36: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de validación.

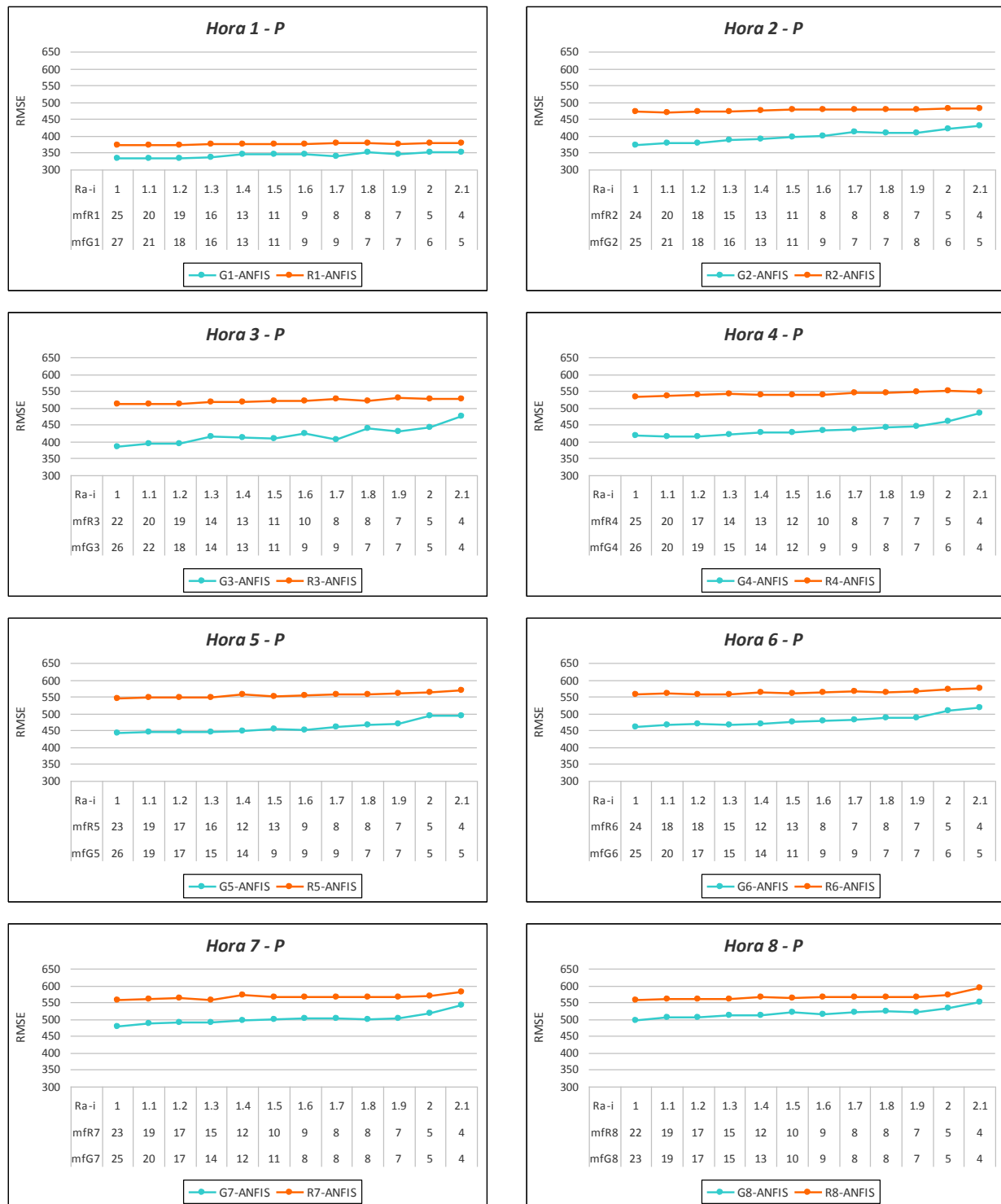


Figura 16.37: Valores ponderados del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento y validación.

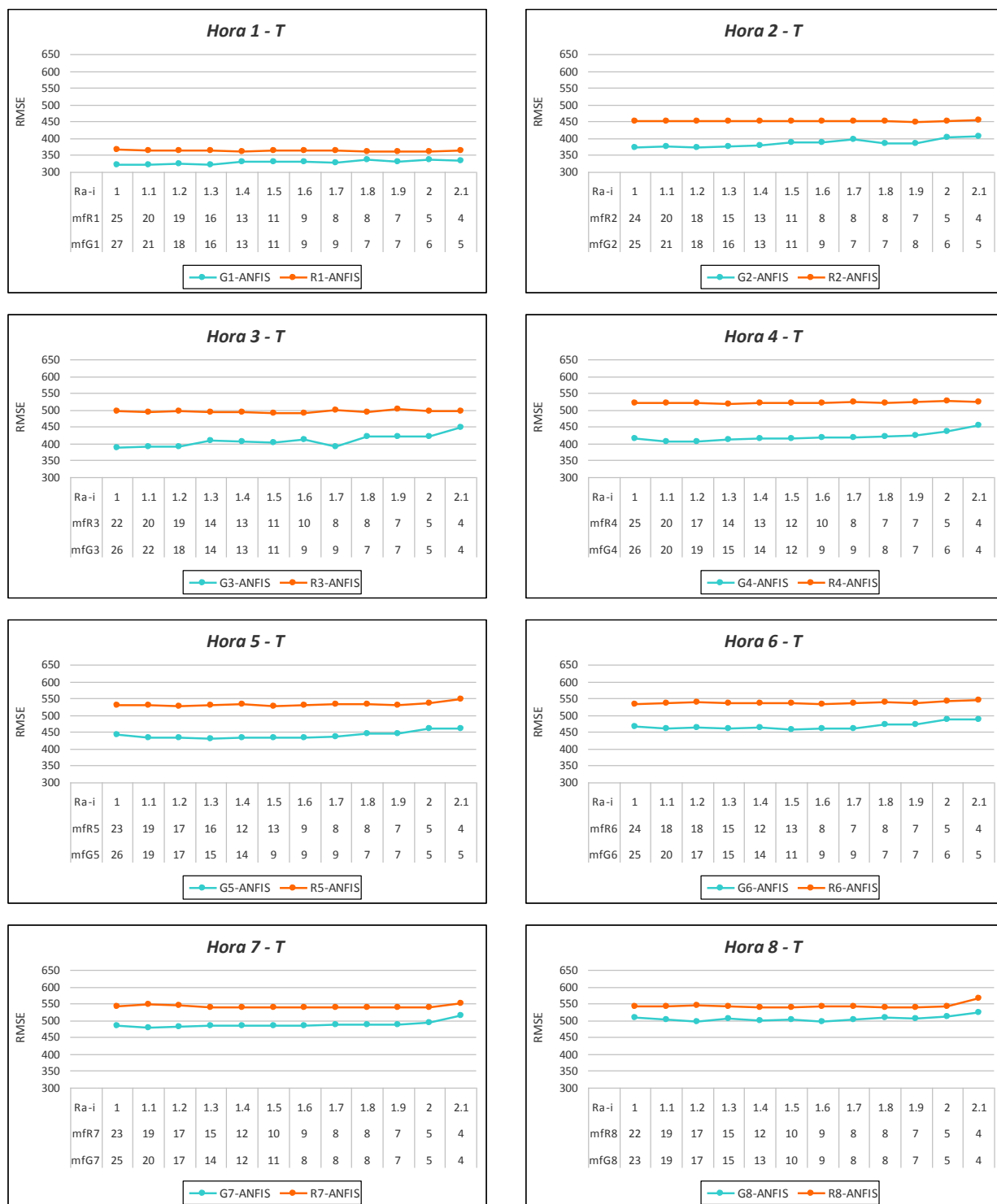
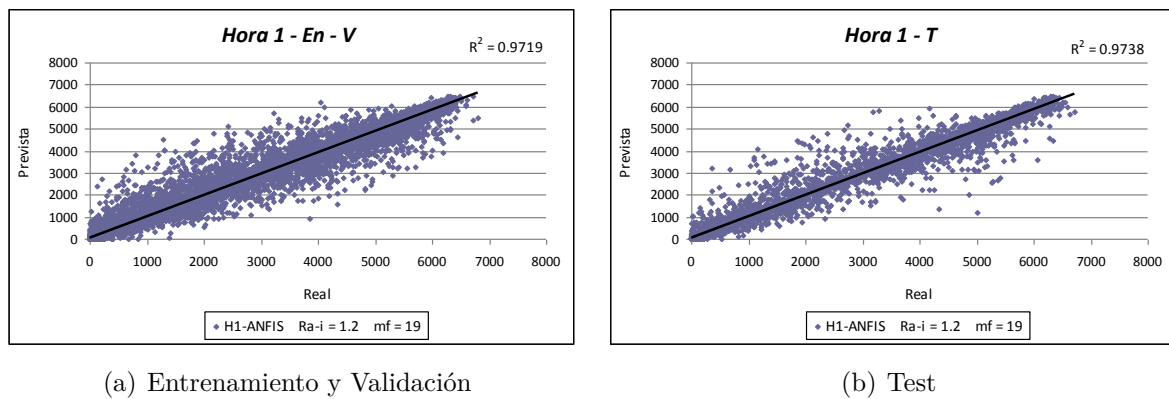


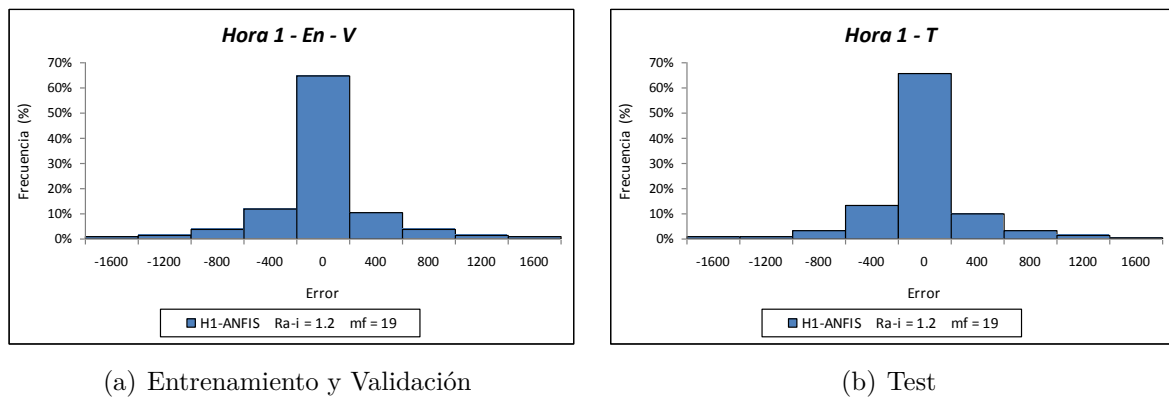
Figura 16.38: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria con modelos basados en sistemas ANFIS desarrollados con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de test.



(a) Entrenamiento y Validación

(b) Test

Figura 16.39: Diagramas de dispersión con valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el sistema ANFIS de 19 reglas usando los atributos H1, para el horizonte de predicción de 1 hora.



(a) Entrenamiento y Validación

(b) Test

Figura 16.40: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con el sistema ANFIS de 19 reglas usando los atributos H1, para el horizonte de predicción de 1 hora.

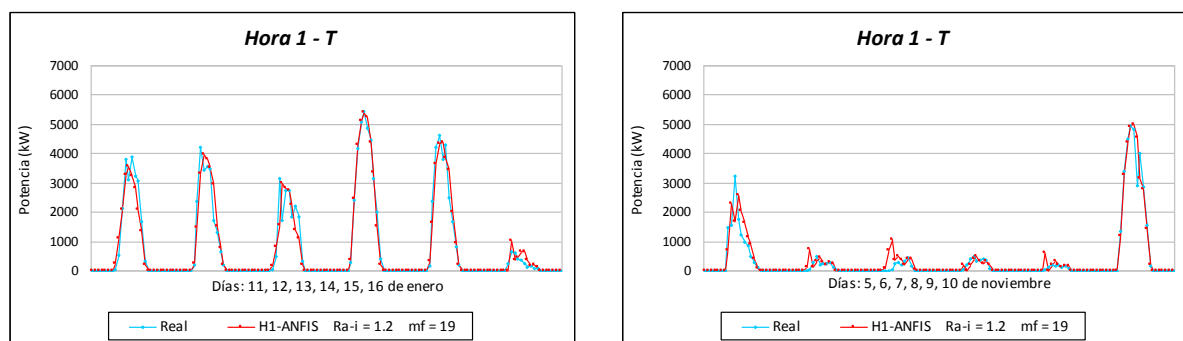


Figura 16.41: Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el sistema ANFIS de 19 reglas usando los atributos H1, para el horizonte de predicción de 1 hora con datos de test.

H1, G1 y R1, para el horizonte de predicción de 1 hora.

Los modelos que mejores resultados ofrecen son: el sistema ANFIS que utiliza los atributos H1, la partición del espacio de entrada por agrupamiento substractivo con el valor del rango de influencia $Ra-i = 1.2$ y 19 reglas, y el modelo basado en la red neuronal artificial MLP-BFGS con 13 neuronas en la capa oculta desarrollado, también, con los atributos H1. Así el modelo ANFIS obtiene, con los datos del subconjunto de entrenamiento En y de validación V , un valor del RMSE medio ponderado igual a 331.7184 kW (mejora con respecto al modelo persistente del 19.76 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 324.1660 kW (mejora con respecto al modelo persistente del 18.34 %). El modelo basado en la red neuronal MLP-BFGS obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 332.5056 kW (mejora con respecto al modelo persistente del 19.57 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 311.1322 kW (mejora con respecto al modelo persistente del 21.63 %). Con estos valores, el comportamiento de la red neuronal MLP-BFGS con 13 neuronas es ligeramente mejor que el del sistema ANFIS (menor valor del RMSE medio ponderado con ambos subconjuntos de datos).

- En la Tabla 16.17 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos H2, G2 y R2, para el horizonte de predicción de 2 horas.

El modelo que mejores resultados ofrece es el basado en la una red neuronal artificial MLP-BFGS con 14 neuronas en la capa oculta y que utiliza los atributos H2. Este modelo, con los datos del subconjunto de entrenamiento E , obtiene un valor del RMSE igual a 364.4054 kW (mejora con respecto al modelo persistente del 38.00 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 352.0382 kW (mejora con respecto al modelo persistente del 36.98 %).

- En la Tabla 16.18 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos H3, G3 y R3, para el horizonte de predicción de 3 horas.

Los modelos que mejores resultados ofrecen son: el basado en la red neuronal artificial MLP-CGD con 17 neuronas en la capa oculta y que utiliza los atributos H3, y el basado en la red neuronal MLP-BFGS con 12 neuronas en la capa oculta y que usa también los atributos H3. Así, el modelo basado en la red MLP-CGD obtiene, con los datos el subconjunto de entrenamiento E , un valor del RMSE igual a 365.7650 kW (mejora con respecto al modelo persistente del 49.42 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 358.1819 kW (mejora con respecto al modelo persistente del 48.61 %). El modelo basado en la red neuronal MLP-BFGS obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 365.9435 kW (mejora con respecto al modelo persistente del 49.40 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 359.9624 kW (mejora con respecto al modelo persistente del 48.36 %). Si se toma el valor del RMSE medio ponderado obtenido con los subconjuntos de entrenamiento E y de test T , el error obtenido con la red MLP-CGD con 17 neuronas es ligeramente menor.

- En la Tabla 16.19 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos H4, G4 y R4, para el horizonte de predicción de 4 horas.

Uno de los modelos que mejores resultados ofrece es el basado en la una red neuronal artificial MLP-BFGS con 17 neuronas en la capa oculta y que utiliza los atributos H4. Este modelo, con los datos del subconjunto de entrenamiento E obtiene un valor del RMSE igual a 389.3066 kW (mejora con respecto al modelo persistente del 53.62 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 373.6758 kW (mejora con respecto al modelo persistente del 54.91 %). La red neuronal MLP-CGD con 21 neuronas en la capa oculta y que utiliza también los atributos H4, consigue unos resultados destacables, con los datos del subconjunto de entrenamiento E obtiene un valor del RMSE igual a 389.6820 kW (mejora con respecto al modelo persistente del 53.58 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 371.4647 kW (mejora del 55.18 % con respecto al modelo persistente). Si se toma el valor del RMSE medio ponderado obtenido con los subconjuntos de entrenamiento E y de test T , el error obtenido con la red MLP-CGD con 21 neuronas es ligeramente menor.

- En la Tabla 16.20 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos H5, G5 y R5, para el horizonte de predicción de 5 horas.

Los modelos que mejores resultados ofrecen son: el basado en una red neuronal artificial MLP-CGD con 17 neuronas en la capa oculta y que utiliza los atributos H5, y el basado en la red neuronal MLP-BFGS con 16 neuronas en la capa oculta y que usa también los atributos H5. Así, el modelo basado en la red MLP-CGD obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 424.2394 kW (mejora con respecto al modelo persistente del 54.86 %) y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 411.4069 kW (mejora con respecto al modelo persistente del 55.93 %). El modelo basado en la red neuronal MLP-BFGS obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 423.7195 kW (mejora con respecto al modelo persistente del 54.92 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 412.9194 kW (mejora con respecto al modelo persistente del 55.77 %). Si se toma el valor del RMSE medio ponderado obtenido con los subconjuntos de entrenamiento E y de test T , el error obtenido con las dos redes es prácticamente igual, siendo con la MLP-CGD con 17 neuronas ligeramente menor.

- En la Tabla 16.21 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos H6, G6 y R6, para el horizonte de predicción de 6 horas.

El modelo que se propone es utilizar los atributos G6 con una red neuronal artificial MLP-CGD con 14 neuronas en la capa oculta. Con los datos del subconjunto de entrenamiento E obtiene el valor del RMSE igual a 460.0089 kW, su mejora con respecto al modelo persistente es del 54.94 %, y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 440.0666 kW, su mejora con respecto al modelo persistente es del 56.69 %.

- En la Tabla 16.22 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos H7, G7 y R7, para el horizonte de predicción de 7 horas.

Los modelos que mejores resultados ofrecen son: el basado en la red neuronal artificial MLP-CGD con 21 neuronas en la capa oculta y que utiliza los atributos G7, y el basado en la red neuronal MLP-BFGS con 13 neuronas en la capa oculta y que usa también los atributos G7. Así, el modelo basado en la red MLP-CGD obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 478.5120 kW (mejora con respecto al modelo persistente del 55.89 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 468.8237 kW (mejora con respecto al modelo persistente del 56.74 %). El modelo basado en la red neuronal MLP-BFGS obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 479.0830 kW (mejora con respecto al modelo persistente del 55.84 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 470.1686 kW (mejora con respecto al modelo persistente del 56.62 %). Si se toma el valor del RMSE medio ponderado entre los obtenidos con los subconjuntos de entrenamiento E y de test T , el error obtenido con la red MLP-CGD con 21 neuronas es ligeramente menor.

- En la Tabla 16.23 se muestran los valores del RMSE (kW) más relevantes de los modelos desarrollados con M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos H8, G8 y R8, para el horizonte de predicción de 8 horas.

El modelo que mejores resultados ofrece es el basado en la una red neuronal artificial MLP-CGD con 20 neuronas en la capa oculta y que utiliza los atributos G8. Este modelo, con los datos del subconjunto de entrenamiento E , obtiene un valor del RMSE igual a 494.9004 kW (mejora con respecto al modelo persistente del 56.29 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 499.6582 kW (mejora con respecto al modelo persistente del 55.73 %). De todos los modelos analizados, para este horizonte de predicción, si se toma el valor del RMSE medio ponderado entre los obtenidos con los subconjuntos de entrenamiento E y de test T , el error obtenido con este modelo es ligeramente menor.

Para comparar más fácilmente los errores de predicción proporcionados por los mejores modelos desarrollados con las diferentes técnicas de minería de datos, se representan los gráficos de barras del RMSE (kW) con cada tipo de modelo, para los horizontes de predicción de 1 a 8 horas:

- Los valores del RMSE que se obtienen usando, en cada uno de estos modelos, los atributos G_m y R_m , con $m \in \{1, \dots, 8\}$, se representan con los datos del subconjunto de entrenamiento E (ver Figura 16.42) y con los datos del subconjunto de test T (ver Figura 16.43). Se observa que los valores del RMSE (kW) son menores cuando se utilizan como entradas de los modelos los atributos G_m .
- Los valores del RMSE que se obtienen usando, en cada uno de estos modelos, los atributos H_m y G_m , con $m \in \{1, \dots, 8\}$, se representan con los datos del subconjunto de entrenamiento E (ver Figura 16.44) y con los datos del subconjunto de test T

Tabla 16.16: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 1 hora.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	$\hat{P}1$		413.3850	396.9875	—	—
M5Rules	H1	instancias = 17	356.9915	339.6672	13.64 %	14.44 %
	G1	instancias = 21	357.9166	333.6961	13.42 %	15.94 %
	R1	instancias = 21	383.6620	369.1024	7.19 %	7.02 %
M5P	H1	instancias = 19	355.1378	335.1888	14.09 %	15.57 %
	G1	instancias = 19	356.0655	333.3967	13.87 %	16.02 %
	R1	instancias = 21	383.7297	369.1024	7.17 %	7.02 %
MLP-GD	H1	neuronas = 21	349.7410	320.1419	15.40 %	19.36 %
	G1	neuronas = 10	349.6609	325.5980	15.42 %	17.98 %
	R1	neuronas = 21	407.0491	368.9376	1.53 %	7.07 %
MLP-m(02)	H1	neuronas = 17	349.7024	322.5958	15.41 %	18.74 %
	G1	neuronas = 12	354.5785	327.3133	14.23 %	17.55 %
	R1	neuronas = 16	406.7691	365.7547	1.60 %	7.87 %
MLP-CGD	H1	neuronas = 15	335.9969	316.3809	18.72 %	20.30 %
	G1	neuronas = 14	335.8249	315.5836	18.76 %	20.51 %
	R1	neuronas = 10	381.1730	362.4036	7.79 %	8.71 %
MLP-BFGS	H1	neuronas = 13	332.5056	311.1322	19.57 %	21.63 %
	G1	neuronas = 10	335.8514	314.3801	18.76 %	20.81 %
	R1	neuronas = 10	382.7316	359.8273	7.42 %	9.36 %
RBF-CGD	H1	neuronas = 10	337.8913	318.6835	18.26 %	19.72 %
	G1	neuronas = 11	338.3642	316.4920	18.15 %	20.28 %
	R1	neuronas = 10	385.2125	359.9418	6.82 %	9.33 %
RBF-BFGS	H1	neuronas = 10	337.5579	319.3699	18.34 %	19.55 %
	G1	neuronas = 10	338.8728	317.5269	18.02 %	20.02 %
	R1	neuronas = 10	386.0720	362.7624	6.61 %	8.62 %
SVM	H1	$C = 20, \gamma = 0.2$	341.9422	320.3764	17.28 %	19.30 %
	G1	$C = 20, \gamma = 0.2$	341.9896	319.5480	17.27 %	19.51 %
	R1	$C = 10, \gamma = 0.2$	385.5261	366.9694	6.74 %	7.56 %
ANFIS	H1	mf=19, Ra-i=1.2	331.7184	324.1660	19.76 %	18.34 %
	G1	mf=21, Ra-i=1.1	332.7743	321.4506	19.50 %	19.03 %
	R1	mf=25, Ra-i=1.0	372.0914	365.4111	9.99 %	7.95 %

Tabla 16.17: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 2 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	$\hat{P}2$		587.7702	558.6576	—	—
M5Rules	H2	instancias = 16	404.2584	393.5680	31.22 %	29.55 %
	G2	instancias = 20	406.9370	389.0998	30.77 %	30.35 %
	R2	instancias = 16	489.2550	462.5395	16.76 %	17.21 %
M5P	H2	instancias = 18	404.8790	379.0834	31.12 %	32.14 %
	G2	instancias = 20	405.2479	381.6231	31.05 %	31.69 %
	R2	instancias = 16	486.2505	462.5522	17.27 %	17.20 %
MLP-GD	H2	neuronas = 15	385.6085	378.4619	34.39 %	32.26 %
	G2	neuronas = 11	388.5650	385.2086	33.89 %	31.05 %
	R2	neuronas = 14	504.3143	462.7234	14.20 %	17.17 %
MLP-m(02)	H2	neuronas = 10	386.5713	393.1939	34.23 %	29.62 %
	G2	neuronas = 13	394.4834	384.7042	32.88 %	31.14 %
	R2	neuronas = 13	512.1067	465.1042	12.87 %	16.75 %
MLP-CGD	H2	neuronas = 16	368.2142	354.4327	37.35 %	36.56 %
	G2	neuronas = 19	372.2870	349.4672	36.66 %	37.45 %
	R2	neuronas = 11	480.1161	452.7848	18.32 %	18.95 %
MLP-BFGS	H2	neuronas = 14	364.4054	352.0382	38.00 %	36.98 %
	G2	neuronas = 16	375.2489	353.4557	36.16 %	36.73 %
	R2	neuronas = 11	481.1105	453.1241	18.15 %	18.89 %
RBF-CGD	H2	neuronas = 18	371.8323	355.3760	36.74 %	36.39 %
	G2	neuronas = 17	376.2085	353.0803	35.99 %	36.80 %
	R2	neuronas = 13	481.4221	454.1518	18.09 %	18.71 %
RBF-BFGS	H2	neuronas = 14	372.7290	356.1148	36.59 %	36.26 %
	G2	neuronas = 14	376.0705	358.7112	36.02 %	35.79 %
	R2	neuronas = 13	481.3759	450.4195	18.10 %	19.37 %
SVM	H2	C = 20, $\gamma = 0.5$	376.3093	359.6091	35.98 %	35.63 %
	G2	C = 20, $\gamma = 0.5$	380.1218	368.7994	35.33 %	33.98 %
	R2	C = 20, $\gamma = 0.2$	486.8926	460.8504	17.16 %	17.51 %
ANFIS	H2	mf=27, Ra-i=1.0	366.0445	372.7897	37.72 %	33.27 %
	G2	mf=25, Ra-i=1.0	371.0171	372.2767	36.88 %	33.36 %
	R2	mf=20, Ra-i=1.1	469.8109	449.6537	20.07 %	19.51 %

Tabla 16.18: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 3 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	$\hat{P}3$		723.1449	697.0362	—	—
M5Rules	H3	instancias = 13	417.9495	418.1894	42.20 %	40.00 %
	G3	instancias = 19	420.1941	398.1977	41.89 %	42.87 %
	R3	instancias = 19	534.8192	508.2875	26.04 %	27.08 %
M5P	H3	instancias = 16	417.7952	408.8696	42.23 %	41.34 %
	G3	instancias = 20	414.3567	392.8292	42.70 %	43.64 %
	R3	instancias = 20	533.0607	506.1821	26.29 %	27.38 %
MLP-GD	H3	neuronas = 16	383.8129	384.2510	46.92 %	44.87 %
	G3	neuronas = 16	408.4154	397.8199	43.52 %	42.93 %
	R3	neuronas = 17	542.9898	499.5514	24.91 %	28.33 %
MLP-m(02)	H3	neuronas = 19	385.9422	403.5852	46.63 %	42.10 %
	G3	neuronas = 16	410.4336	399.5478	43.24 %	42.68 %
	R3	neuronas = 16	551.4088	512.9342	23.75 %	26.41 %
MLP-CGD	H3	neuronas = 17	365.7650	358.1819	49.42 %	48.61 %
	G3	neuronas = 17	381.5392	363.9725	47.24 %	47.78 %
	R3	neuronas = 12	517.8811	498.5748	28.38 %	28.47 %
MLP-BFGS	H3	neuronas = 12	365.9435	359.9624	49.40 %	48.36 %
	G3	neuronas = 11	382.8766	367.4001	47.05 %	47.29 %
	R3	neuronas = 10	519.9480	496.5792	28.10 %	28.76 %
RBF-CGD	H3	neuronas = 19	369.0390	364.4877	48.97 %	47.71 %
	G3	neuronas = 15	384.7243	367.9117	46.80 %	47.22 %
	R3	neuronas = 18	519.7800	499.7347	28.12 %	28.31 %
RBF-BFGS	H3	neuronas = 21	372.6146	367.3817	48.47 %	47.29 %
	G3	neuronas = 16	384.7816	363.5181	46.79 %	47.85 %
	R3	neuronas = 13	520.4603	495.3058	28.03 %	28.94 %
SVM	H3	$C = 10, \gamma = 1.0$	372.7238	368.2176	48.46 %	47.17 %
	G3	$C = 20, \gamma = 0.5$	385.5249	376.6075	46.69 %	45.97 %
	R3	$C = 10, \gamma = 0.5$	531.0376	510.5919	26.57 %	26.75 %
ANFIS	H3	mf=25, Ra-i=1.0	366.2227	382.3321	49.36 %	45.15 %
	G3	mf=26, Ra-i=1.0	383.1074	387.8340	47.02 %	44.36 %
	R3	mf=20, Ra-i=1.1	511.0607	494.4600	29.33 %	29.06 %

Tabla 16.19: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 4 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	$\hat{P}4$		839.4644	828.7346	—	—
M5Rules	H4	instancias = 17	441.8847	442.0674	47.36 %	46.66 %
	G4	instancias = 12	440.2776	444.4783	47.55 %	46.37 %
	R4	instancias = 21	555.7420	534.0228	33.80 %	35.56 %
M5P	H4	instancias = 11	435.3547	422.7892	48.14 %	48.98 %
	G4	instancias = 16	434.2481	433.2631	48.27 %	47.72 %
	R4	instancias = 21	553.8649	526.3128	34.02 %	36.49 %
MLP-GD	H4	neuronas = 21	421.0369	395.8204	49.84 %	52.24 %
	G4	neuronas = 21	427.3362	421.3785	49.09 %	49.15 %
	R4	neuronas = 11	569.5812	552.9464	32.15 %	33.28 %
MLP-m(02)	H4	neuronas = 17	431.6213	410.5563	48.58 %	50.46 %
	G4	neuronas = 19	429.3023	413.6879	48.86 %	50.08 %
	R4	neuronas = 11	577.3658	545.8098	31.22 %	34.14 %
MLP-CGD	H4	neuronas = 21	389.6820	371.4647	53.58 %	55.18 %
	G4	neuronas = 15	407.5181	388.4866	51.45 %	53.12 %
	R4	neuronas = 12	539.3757	526.0197	35.75 %	36.53 %
MLP-BFGS	H4	neuronas = 17	389.3066	373.6758	53.62 %	54.91 %
	G4	neuronas = 19	406.1537	390.1671	51.62 %	52.92 %
	R4	neuronas = 11	538.0843	532.5255	35.90 %	35.74 %
RBF-CGD	H4	neuronas = 19	391.7876	380.5972	53.33 %	54.07 %
	G4	neuronas = 20	408.1017	386.3666	51.39 %	53.38 %
	R4	neuronas = 16	539.1545	528.5775	35.77 %	36.22 %
RBF-BFGS	H4	neuronas = 19	392.7420	382.9516	53.22 %	53.79 %
	G4	neuronas = 14	409.9995	383.0524	51.16 %	53.78 %
	R4	neuronas = 10	539.3690	519.6639	35.75 %	37.29 %
SVM	H4	C = 10, $\gamma = 1.0$	394.4228	387.3159	53.01 %	53.26 %
	G4	C = 20, $\gamma = 0.5$	407.8724	391.7655	51.41 %	52.73 %
	R4	C = 20, $\gamma = 0.5$	554.0033	550.0474	34.01 %	33.63 %
ANFIS	H4	mf=21, Ra-i=1.1	395.0391	400.8711	52.94 %	51.63 %
	G4	mf=19, Ra-i=1.2	414.4315	405.6316	50.63 %	51.05 %
	R4	mf=25, Ra-i=1.0	530.8648	521.2202	36.76 %	37.11 %

Tabla 16.20: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 5 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	$\hat{P}5$		939.8819	933.6220	—	—
M5Rules	H5	instancias = 18	475.4391	472.5758	49.42 %	49.38 %
	G5	instancias = 14	463.4176	451.1175	50.69 %	51.68 %
	R5	instancias = 19	563.5435	542.2776	40.04 %	41.92 %
M5P	H5	instancias = 10	456.6187	466.5829	51.42 %	50.02 %
	G5	instancias = 14	456.8083	457.9578	51.40 %	50.95 %
	R5	instancias = 15	564.6734	545.0905	39.92 %	41.62 %
MLP-GD	H5	neuronas = 19	445.2842	437.2939	52.62 %	53.16 %
	G5	neuronas = 18	457.9221	440.5250	51.28 %	52.82 %
	R5	neuronas = 14	580.3170	548.2385	38.26 %	41.28 %
MLP-m(02)	H5	neuronas = 21	455.7905	448.8840	51.51 %	51.92 %
	G5	neuronas = 20	462.8648	446.8550	50.75 %	52.14 %
	R5	neuronas = 20	585.2509	562.3703	37.73 %	39.76 %
MLP-CGD	H5	neuronas = 17	424.2394	411.4069	54.86 %	55.93 %
	G5	neuronas = 15	433.4590	414.7685	53.88 %	55.57 %
	R5	neuronas = 18	550.0141	537.8645	41.48 %	42.39 %
MLP-BFGS	H5	neuronas = 16	423.7195	412.9194	54.92 %	55.77 %
	G5	neuronas = 12	436.0506	426.2295	53.61 %	54.35 %
	R5	neuronas = 14	549.3822	539.4546	41.55 %	42.22 %
RBF-CGD	H5	neuronas = 20	422.8583	423.8445	55.01 %	54.60 %
	G5	neuronas = 21	435.8227	423.3949	53.63 %	54.65 %
	R5	neuronas = 17	547.8693	540.2831	41.71 %	42.13 %
RBF-BFGS	H5	neuronas = 17	422.9261	420.4956	55.00 %	54.96 %
	G5	neuronas = 17	433.6177	416.9026	53.86 %	55.35 %
	R5	neuronas = 12	550.0238	528.6972	41.48 %	43.37 %
SVM	H5	$C = 15, \gamma = 1.0$	428.1322	434.2348	54.45 %	53.49 %
	G5	$C = 20, \gamma = 0.5$	439.6061	426.3751	53.23 %	54.33 %
	R5	$C = 15, \gamma = 0.5$	566.8986	555.0199	39.68 %	40.55 %
ANFIS	H5	mf=17, Ra-i=1.2	435.8494	435.4229	53.63 %	53.36 %
	G5	mf=26, Ra-i=1.0	441.7337	441.6517	53.00 %	52.69 %
	R5	mf=23, Ra-i=1.0	545.8297	528.9957	41.93 %	43.34 %

Tabla 16.21: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 6 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	\hat{P}_6		1020.8579	1016.1953	—	—
M5Rules	H6	instancias = 17	507.8388	516.4380	50.25 %	49.18 %
	G6	instancias = 16	492.9174	461.1271	51.72 %	54.62 %
	R6	instancias = 19	569.6520	549.9787	44.20 %	45.88 %
M5P	H6	instancias = 12	496.1091	513.4768	51.40 %	49.47 %
	G6	instancias = 17	480.4705	468.0758	52.93 %	53.94 %
	R6	instancias = 15	566.5184	538.6891	44.51 %	46.99 %
MLP-GD	H6	neuronas = 11	487.5473	473.9893	52.24 %	53.36 %
	G6	neuronas = 15	487.8674	467.5270	52.21 %	53.99 %
	R6	neuronas = 13	593.4996	560.4533	41.86 %	44.85 %
MLP-m(02)	H6	neuronas = 15	500.5059	476.4862	50.97 %	53.11 %
	G6	neuronas = 15	491.9224	475.6170	51.81 %	53.20 %
	R6	neuronas = 10	604.2462	563.0322	40.81 %	44.59 %
MLP-CGD	H6	neuronas = 15	460.5901	461.1358	54.88 %	54.62 %
	G6	neuronas = 14	460.0089	440.0666	54.94 %	56.69 %
	R6	neuronas = 13	555.4972	537.7669	45.59 %	47.08 %
MLP-BFGS	H6	neuronas = 20	458.0874	479.5950	55.13 %	52.80 %
	G6	neuronas = 17	457.3747	455.7895	55.20 %	55.15 %
	R6	neuronas = 14	556.3974	543.3794	45.50 %	46.53 %
RBF-CGD	H6	neuronas = 20	461.7680	465.8417	54.77 %	54.16 %
	G6	neuronas = 20	460.8402	454.3933	54.86 %	55.28 %
	R6	neuronas = 19	553.7410	543.3704	45.76 %	46.53 %
RBF-BFGS	H6	neuronas = 18	462.6530	469.2875	54.68 %	53.82 %
	G6	neuronas = 15	461.6162	450.8214	54.78 %	55.64 %
	R6	neuronas = 18	553.6608	543.3018	45.77 %	46.54 %
SVM	H6	C = 20, $\gamma = 0.5$	472.0771	474.7524	53.76 %	53.28 %
	G6	C = 20, $\gamma = 0.5$	464.9581	458.5946	54.45 %	54.87 %
	R6	C = 20, $\gamma = 0.5$	575.5864	563.0327	43.62 %	44.59 %
ANFIS	H6	mf=19, Ra-i=1.1	472.9090	481.7949	53.68 %	52.59 %
	G6	mf=25, Ra-i=1.0	460.3281	466.0041	54.91 %	54.14 %
	R6	mf=24, Ra-i=1.0	555.4471	531.8973	45.59 %	47.66 %

Tabla 16.22: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 7 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	$\hat{P}7$		1084.9314	1083.7357	—	—
M5Rules	H7	instancias = 13	541.9698	547.6123	50.05 %	49.47 %
	G7	instancias = 13	513.0686	497.9693	52.71 %	54.05 %
	R7	instancias = 14	574.5697	584.0512	47.04 %	46.11 %
M5P	H7	instancias = 15	527.9586	548.6621	51.34 %	49.37 %
	G7	instancias = 14	504.0150	510.5974	53.54 %	52.89 %
	R7	instancias = 17	563.4769	564.4370	48.06 %	47.92 %
MLP-GD	H7	neuronas = 20	522.6998	508.5611	51.82 %	53.07 %
	G7	neuronas = 10	515.1899	498.0184	52.51 %	54.05 %
	R7	neuronas = 10	602.7314	573.1516	44.45 %	47.11 %
MLP-m(02)	H7	neuronas = 21	526.3595	516.7485	51.48 %	52.32 %
	G7	neuronas = 12	519.0571	509.2753	52.16 %	53.01 %
	R7	neuronas = 17	613.6093	569.9867	43.44 %	47.41 %
MLP-CGD	H7	neuronas = 21	490.9790	535.7826	54.75 %	50.56 %
	G7	neuronas = 21	478.5120	468.8237	55.89 %	56.74 %
	R7	neuronas = 14	555.4967	546.8525	48.80 %	49.54 %
MLP-BFGS	H7	neuronas = 14	489.2370	508.4016	54.91 %	53.09 %
	G7	neuronas = 13	479.0830	470.1686	55.84 %	56.62 %
	R7	neuronas = 16	553.5731	560.8898	48.98 %	48.24 %
RBF-CGD	H7	neuronas = 21	496.5529	509.2492	54.23 %	53.01 %
	G7	neuronas = 17	480.2763	476.9406	55.73 %	55.99 %
	R7	neuronas = 19	554.0638	555.4699	48.93 %	48.74 %
RBF-BFGS	H7	neuronas = 21	498.0374	506.1689	54.10 %	53.29 %
	G7	neuronas = 14	480.1423	475.4105	55.74 %	56.13 %
	R7	neuronas = 20	553.9841	550.1596	48.94 %	49.23 %
SVM	H7	C = 20, $\gamma = 0.5$	513.9207	517.1991	52.63 %	52.28 %
	G7	C = 5, $\gamma = 0.5$	485.3489	479.8072	55.26 %	55.73 %
	R7	C = 20, $\gamma = 0.5$	577.7328	568.7720	46.75 %	47.52 %
ANFIS	H7	mf=13, Ra-i=1.3	516.0760	517.7655	52.43 %	52.22 %
	G7	mf=25, Ra-i=1.0	478.6691	485.0043	55.88 %	55.25 %
	R7	mf=23, Ra-i=1.0	556.1315	542.8501	48.74 %	49.91 %

Tabla 16.23: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción de 8 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	$\hat{P}8$		1132.2053	1128.6546	—	—
M5Rules	H8	instancias = 18	579.6339	618.0931	48.80 %	45.24 %
	G8	instancias = 18	522.7502	515.0331	53.83 %	54.37 %
	R8	instancias = 21	571.0882	553.0075	49.56 %	51.00 %
M5P	H8	instancias = 14	561.6310	610.9703	50.39 %	45.87 %
	G8	instancias = 17	514.4955	538.4104	54.56 %	52.30 %
	R8	instancias = 21	565.1506	562.2166	50.08 %	50.19 %
MLP-GD	H8	neuronas = 18	570.2826	552.0405	49.63 %	51.09 %
	G8	neuronas = 11	537.8730	522.8593	52.49 %	53.67 %
	R8	neuronas = 13	601.1473	583.1690	46.90 %	48.33 %
MLP-m(02)	H8	neuronas = 10	578.3480	556.1806	48.92 %	50.72 %
	G8	neuronas = 17	553.5463	520.1204	51.11 %	53.92 %
	R8	neuronas = 11	615.5897	577.9709	45.63 %	48.79 %
MLP-CGD	H8	neuronas = 20	520.5991	555.1429	54.02 %	50.81 %
	G8	neuronas = 20	494.9004	499.6582	56.29 %	55.73 %
	R8	neuronas = 20	555.4976	545.7866	50.94 %	51.64 %
MLP-BFGS	H8	neuronas = 20	525.3647	560.5332	53.60 %	50.34 %
	G8	neuronas = 21	492.3227	512.2247	56.52 %	54.62 %
	R8	neuronas = 13	552.9836	550.8289	51.16 %	51.20 %
RBF-CGD	H8	neuronas = 20	534.9008	544.1404	52.76 %	51.79 %
	G8	neuronas = 17	495.6594	503.2436	56.22 %	55.41 %
	R8	neuronas = 20	551.4948	535.9181	51.29 %	52.52 %
RBF-BFGS	H8	neuronas = 21	536.3315	548.7941	52.63 %	51.38 %
	G8	neuronas = 19	495.3685	531.8591	56.25 %	52.88 %
	R8	neuronas = 20	551.1568	548.9616	51.32 %	51.36 %
SVM	H8	C = 15, γ = 1.0	553.8362	564.4373	51.08 %	49.99 %
	G8	C = 15, γ = 0.5	503.9529	511.2834	55.49 %	54.70 %
	R8	C = 20, γ = 0.5	574.3768	569.0372	49.27 %	49.58 %
ANFIS	H8	mf=23, Ra-i=1.0	544.1030	559.3721	51.94 %	50.44 %
	G8	mf=23, Ra-i=1.0	496.8406	508.6344	56.12 %	54.93 %
	R8	mf=22, Ra-i=1.0	557.4063	542.0626	50.77 %	51.97 %

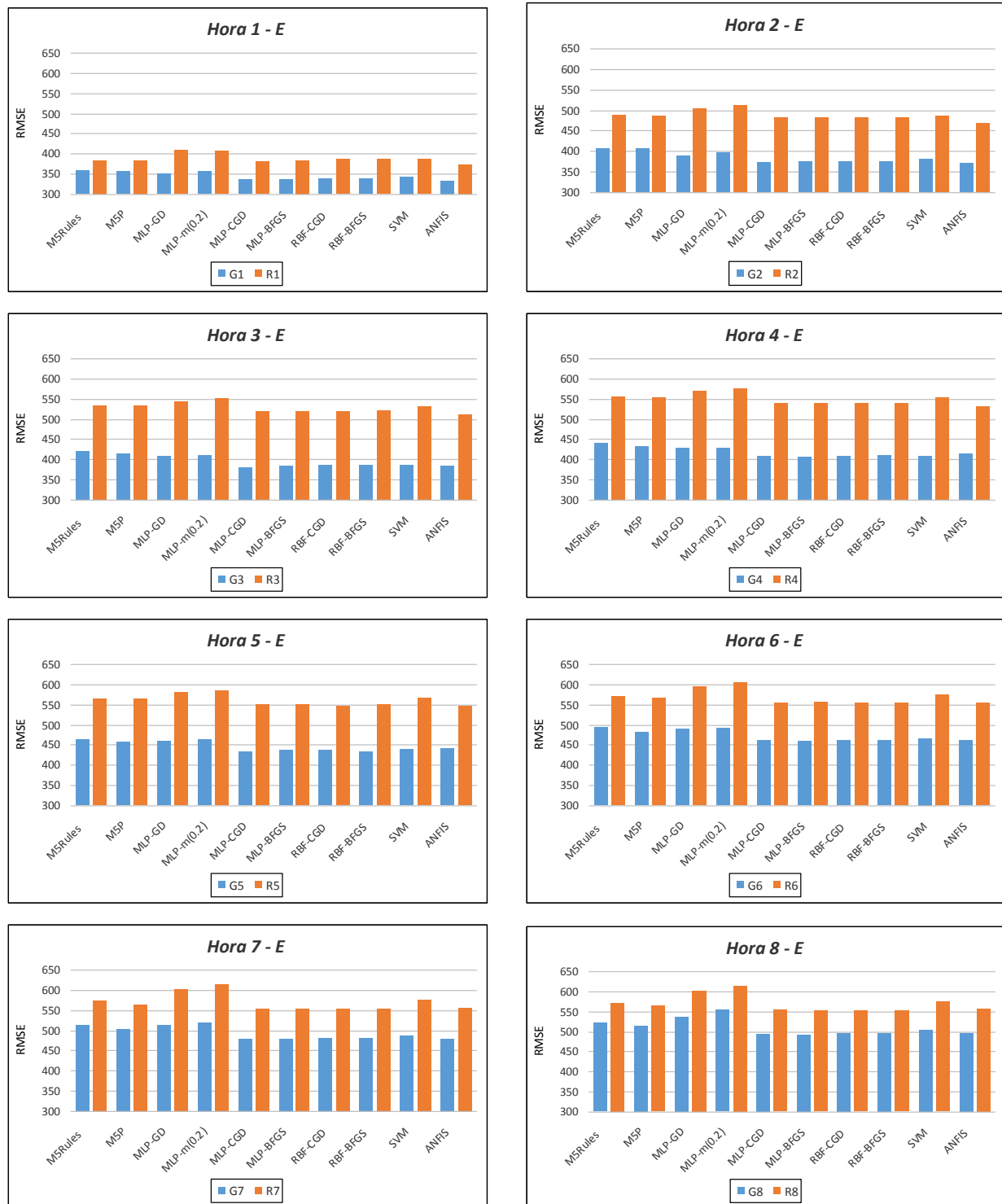


Figura 16.42: Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos G_m y R_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.

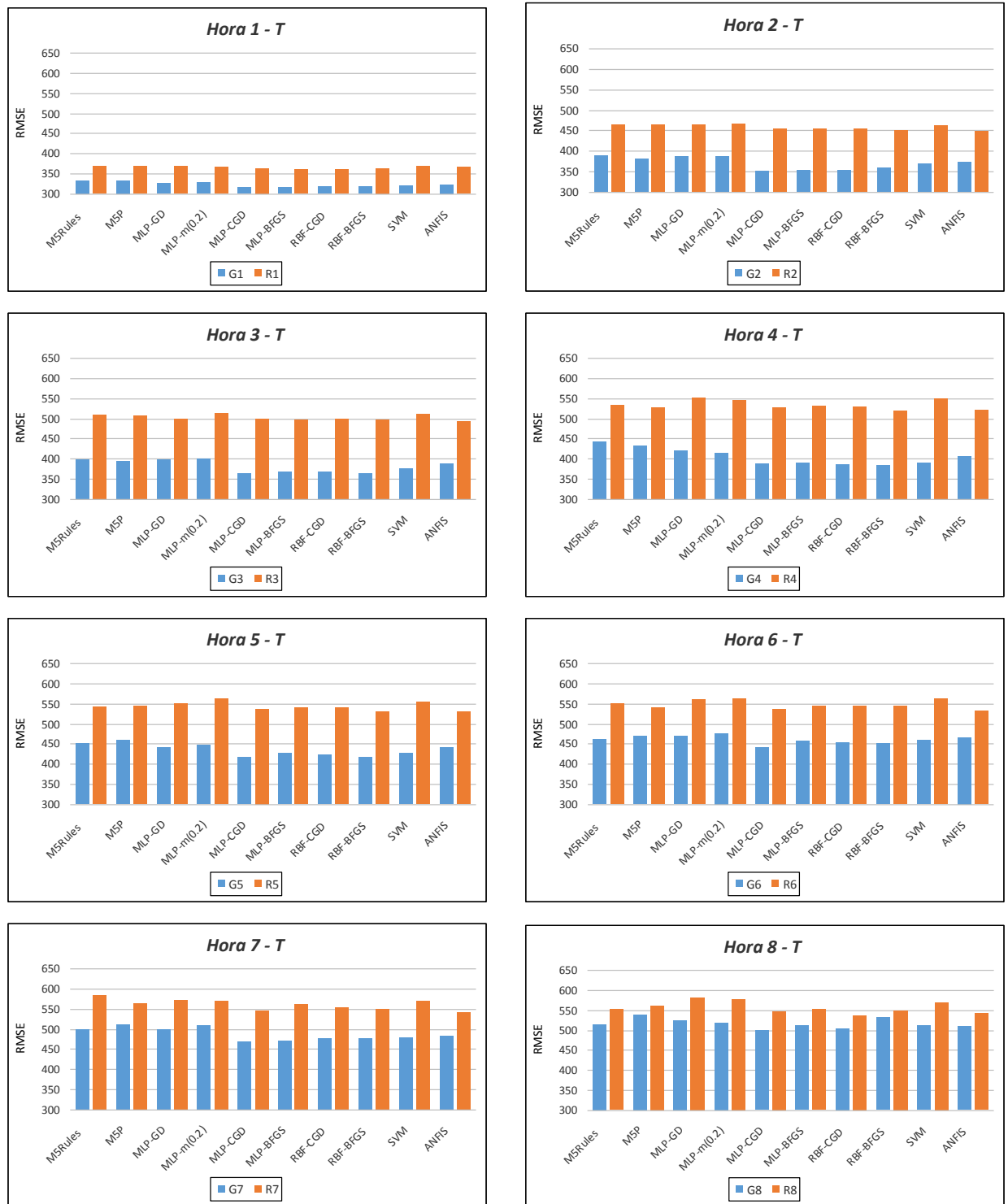


Figura 16.43: Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos Gm y Rm, para horizontes de predicción de 1 a 8 horas con los datos de test.



Figura 16.44: Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos H_m y G_m , para horizontes de predicción de 1 a 8 horas con los datos de entrenamiento.



Figura 16.45: Valores del RMSE (kW) de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria con los atributos H_m y G_m , para horizontes de predicción de 1 a 8 horas con los datos de test.

Tabla 16.24: Coeficientes de determinación R^2 en días de cualquier tipo con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos basados en redes neuronales, para horizontes de predicción de 1 a 8 horas.

Horizonte	Modelos	Atributos	Características	R^2	R^2
				Entr.	Test
1 hora	MLP-CGD	H1	neuronas = 15	0.9712	0.9750
	MLP-BFGS	H1	neuronas = 13	0.9718	0.9758
	RBF-CGD	H1	neuronas = 10	0.9708	0.9746
	RBF-BFGS	H1	neuronas = 10	0.9709	0.9745
2 horas	MLP-CGD	H2	neuronas = 16	0.9655	0.9688
	MLP-BFGS	H2	neuronas = 14	0.9661	0.9691
	RBF-CGD	H2	neuronas = 18	0.9646	0.9684
	RBF-BFGS	H2	neuronas = 14	0.9645	0.9683
3 horas	MLP-CGD	H3	neuronas = 17	0.9659	0.9681
	MLP-BFGS	H3	neuronas = 12	0.9658	0.9676
	RBF-CGD	H3	neuronas = 19	0.9648	0.9668
	RBF-BFGS	H3	neuronas = 21	0.9645	0.9662
4 horas	MLP-CGD	H4	neuronas = 21	0.9612	0.9657
	MLP-BFGS	H4	neuronas = 17	0.9613	0.9651
	RBF-CGD	H4	neuronas = 19	0.9607	0.9638
	RBF-BFGS	H4	neuronas = 19	0.9605	0.9634
5 horas	MLP-CGD	H5	neuronas = 17	0.9540	0.9577
	MLP-BFGS	H5	neuronas = 16	0.9543	0.9574
	RBF-CGD	H5	neuronas = 20	0.9542	0.9550
	RBF-BFGS	H5	neuronas = 17	0.9542	0.9557
6 horas	MLP-CGD	G6	neuronas = 14	0.9459	0.9520
	MLP-BFGS	G6	neuronas = 17	0.9466	0.9488
	RBF-CGD	G6	neuronas = 20	0.9456	0.9488
	RBF-BFGS	G6	neuronas = 15	0.9455	0.9496
7 horas	MLP-CGD	G7	neuronas = 21	0.9417	0.9460
	MLP-BFGS	G7	neuronas = 13	0.9414	0.9455
	RBF-CGD	G7	neuronas = 17	0.9410	0.9437
	RBF-BFGS	G7	neuronas = 14	0.9410	0.9441
8 horas	MLP-CGD	G8	neuronas = 20	0.9376	0.9390
	MLP-BFGS	G8	neuronas = 21	0.9383	0.9358
	RBF-CGD	G8	neuronas = 17	0.9371	0.9376
	RBF-BFGS	G8	neuronas = 19	0.9372	0.9372

Tabla 16.25: Modelos seleccionados para cada horizonte de predicción de 1 a 8 horas.

Horizon.	Modelos	Atr.	Características	RMSE (kW) Entr.	RMSE (kW) Test	R^2 Entr.	R^2 Test
1 hora	MLP-BFGS	H1	neuronas = 13	332.5056	311.1322	0.9718	0.9758
2 horas	MLP-BFGS	H2	neuronas = 14	364.4054	352.0382	0.9661	0.9691
3 horas	MLP-CGD	H3	neuronas = 17	365.7650	358.1819	0.9659	0.9681
4 horas	MLP-CGD	H4	neuronas = 21	389.6820	371.4647	0.9612	0.9657
5 horas	MLP-CGD	H5	neuronas = 17	424.2394	411.4069	0.9540	0.9577
6 horas	MLP-CGD	G6	neuronas = 14	460.0089	440.0666	0.9459	0.9520
7 horas	MLP-CGD	G7	neuronas = 21	478.5120	468.8237	0.9417	0.9460
8 horas	MLP-CGD	G8	neuronas = 20	494.9004	499.6582	0.9376	0.9390

(ver Figura 16.45). Se observa que los menores valores del RMSE (kW) se obtienen si se utilizan los atributos Hm para horizontes de predicción entre 1 y 5 horas, es decir, con $m \in \{1, \dots, 5\}$ y los atributos Gm para horizontes de predicción entre 6 y 8 horas, es decir, con $m \in \{6, \dots, 8\}$.

Para ver la destreza de los mejores modelos desarrollados de predicción de la potencia eléctrica media horaria, en la Tabla 16.24 se muestran los coeficientes de determinación R^2 que se obtienen con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos desarrollados para cualquier tipo de día, para horizontes de predicción de 1 a 8 horas, con los datos de los subconjuntos de entrenamiento E y de test T .

En la Tabla 16.25 se resumen los modelos que se han seleccionado como mejores en la predicción de la potencia eléctrica media horaria para cada horizonte de predicción de 1 a 8 horas, es decir, con los que ha obtenido el menor valor del RMSE medio ponderado con los datos de los subconjuntos de entrenamiento E y de test T . Estos modelos se han desarrollado con los atributos Hm y Gm, y están basados en redes neuronales perceptrón multicapa (MLP) entrenadas con el algoritmo de gradiente conjugado descendente (CGD) o con el método BFGS.

Se representan los diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con estos modelos, con los datos del subconjunto de entrenamiento E (ver Figura 16.46) y con los datos del subconjunto de test T (ver Figura 16.47).

También se representan los histogramas del error en la predicción de la potencia eléctrica media horaria (considerando exclusivamente las horas de luz solar) de esos mismos modelos. Los histogramas que se obtienen con los datos del subconjunto de entrenamiento E se representan en la Figura 16.48, y con los datos del subconjunto de test T en la Figura 16.49.

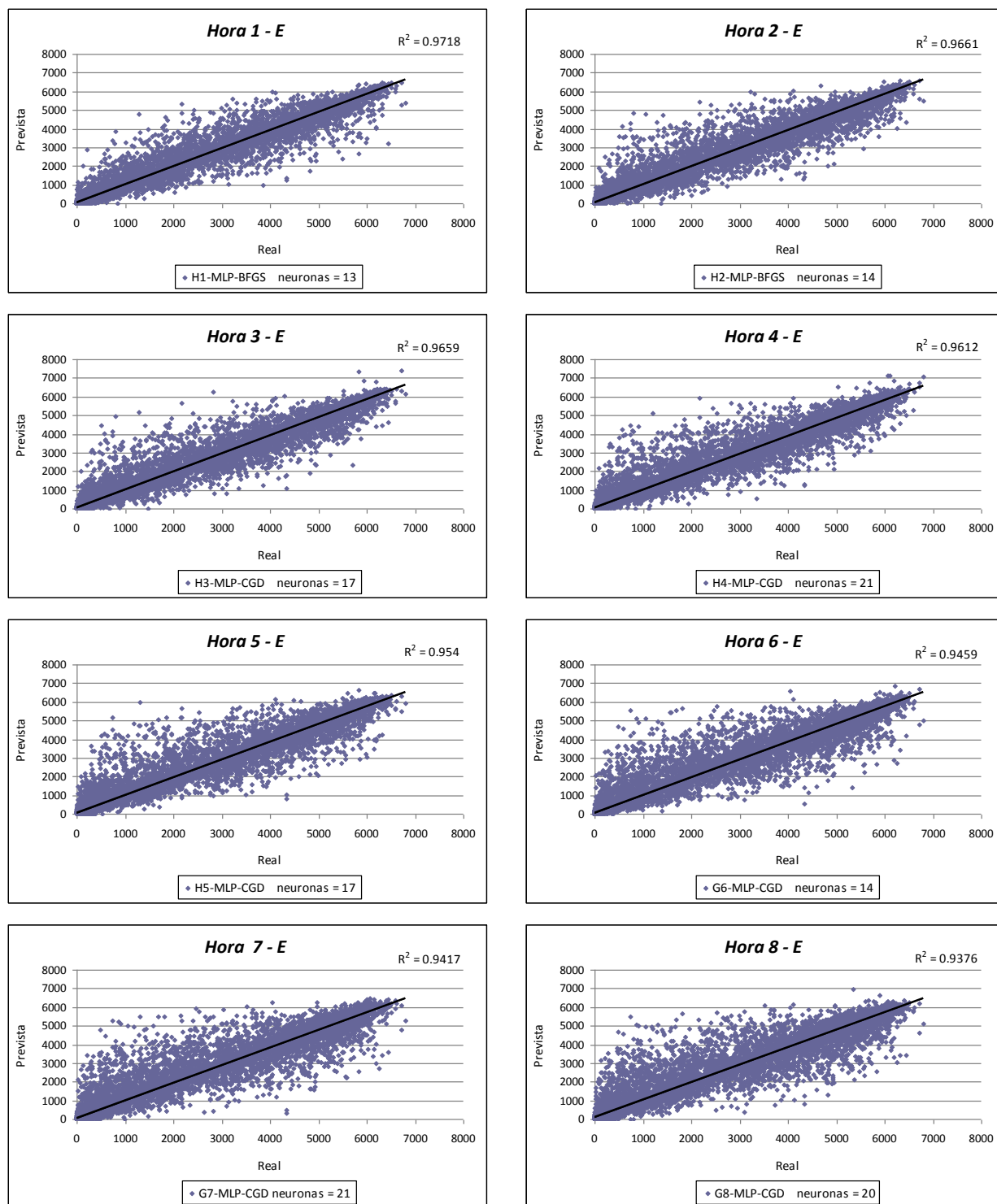


Figura 16.46: Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de entrenamiento.

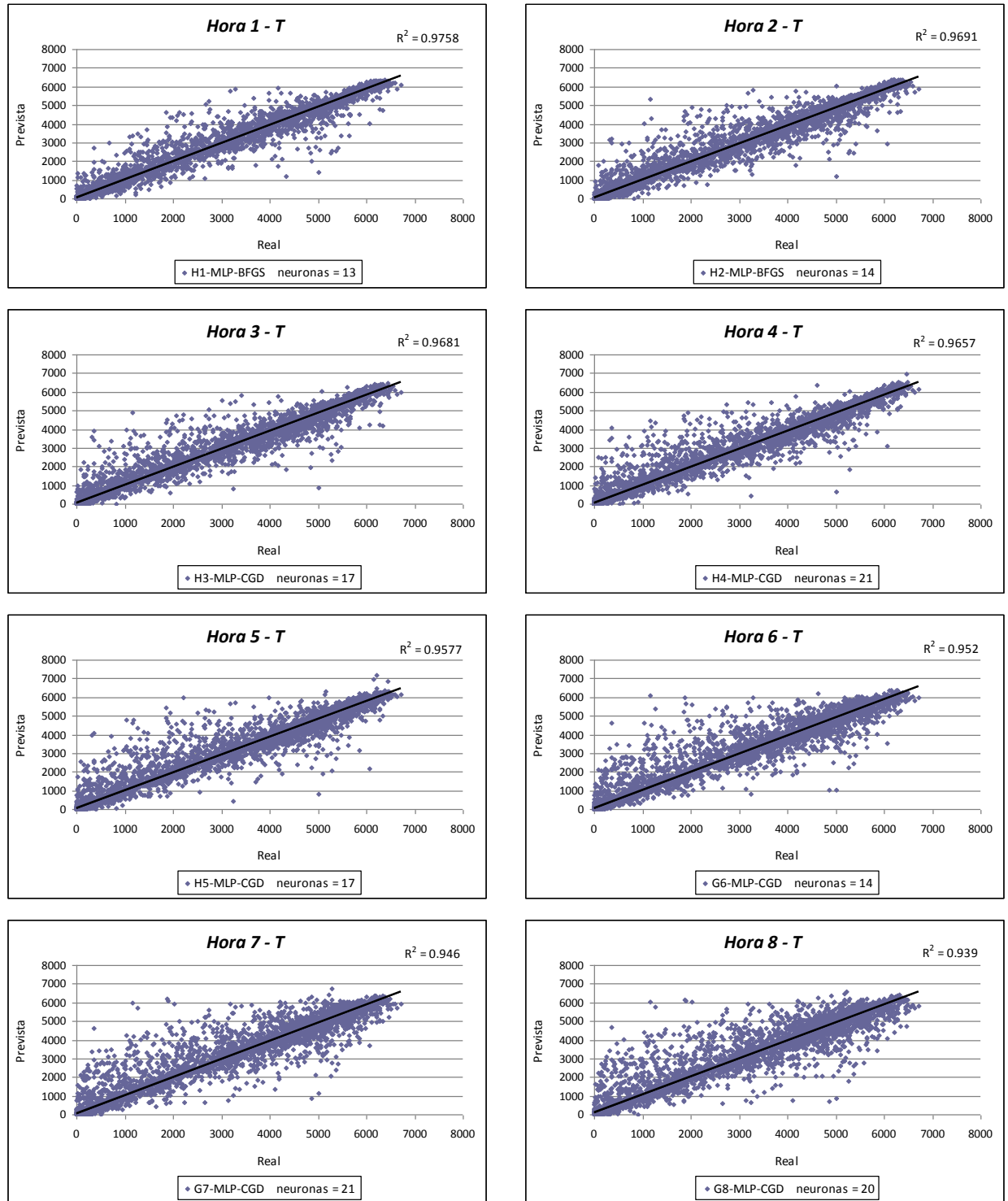


Figura 16.47: Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de test.

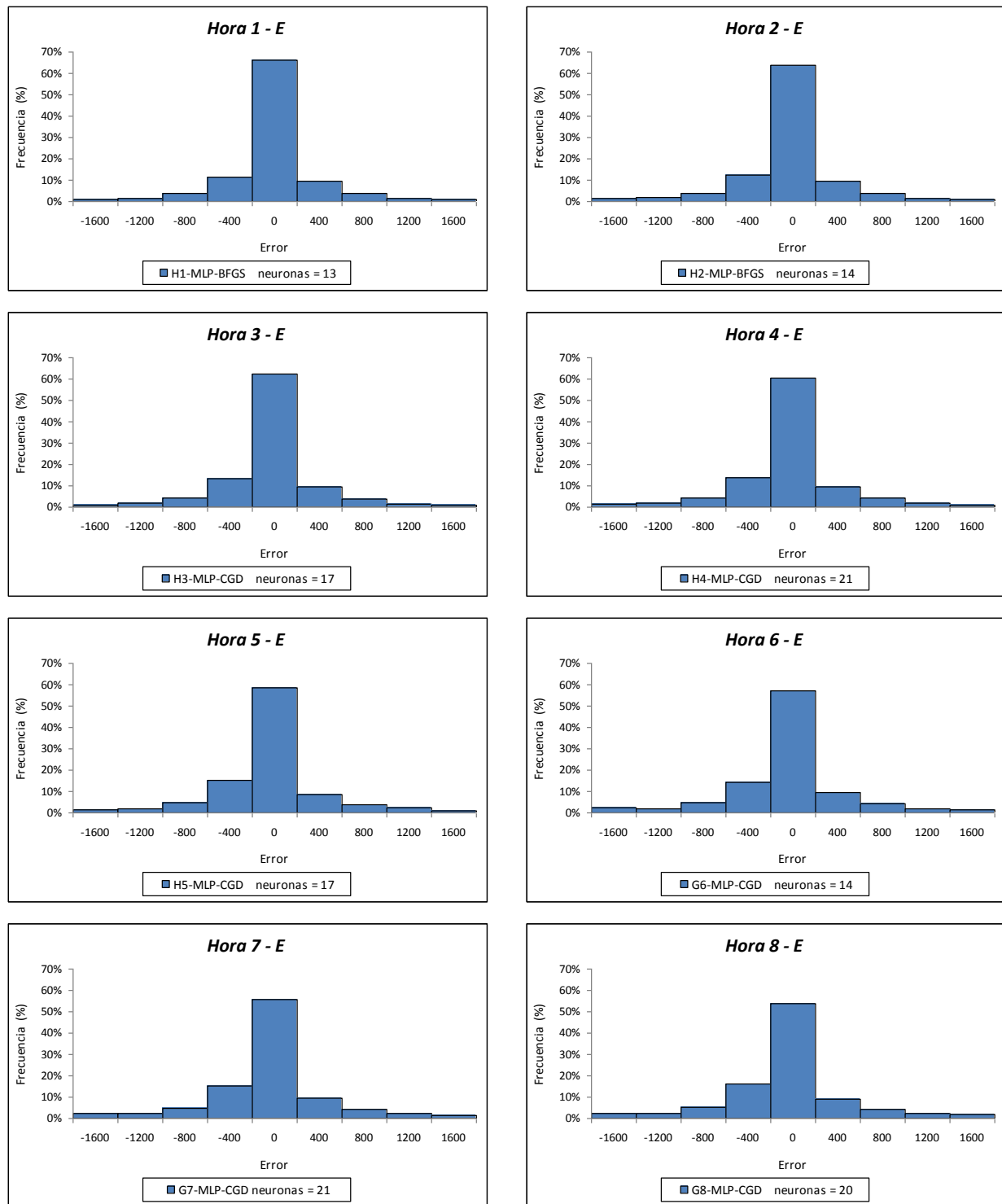


Figura 16.48: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de entrenamiento.

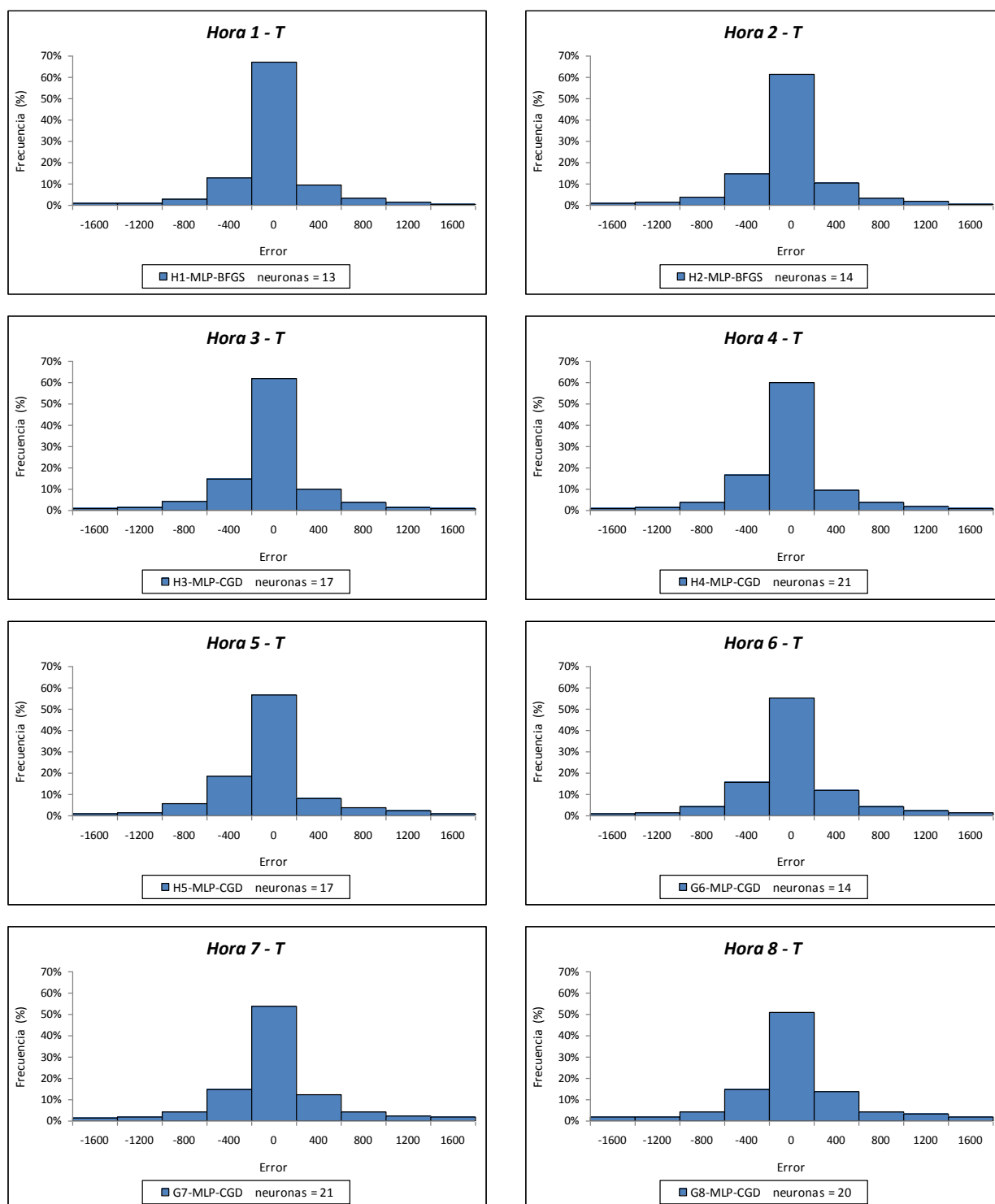


Figura 16.49: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) con el mejor modelo basado en las red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con los datos de test.

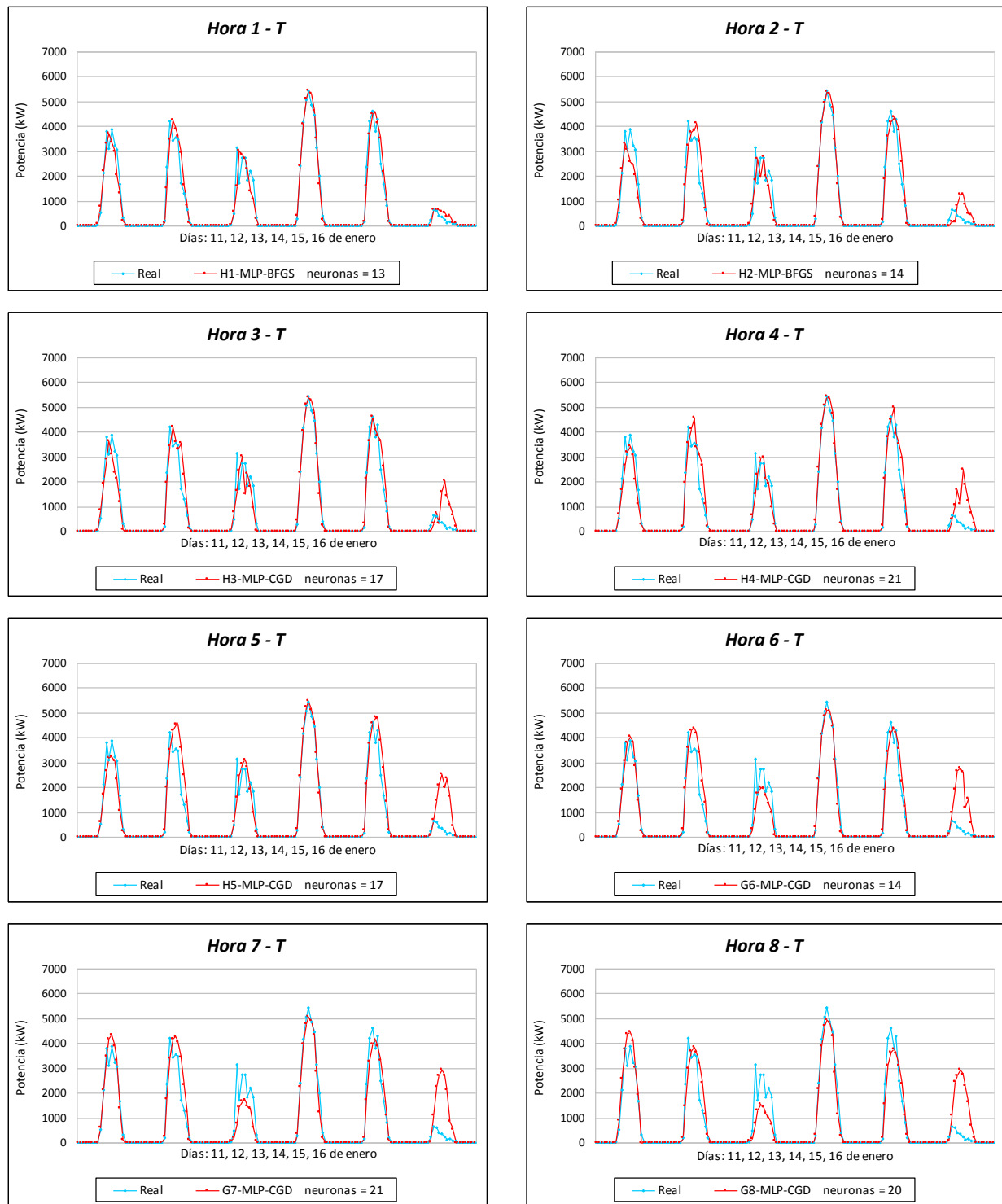


Figura 16.50: Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con datos de test de los días del 11 al 16 de enero de 2012.

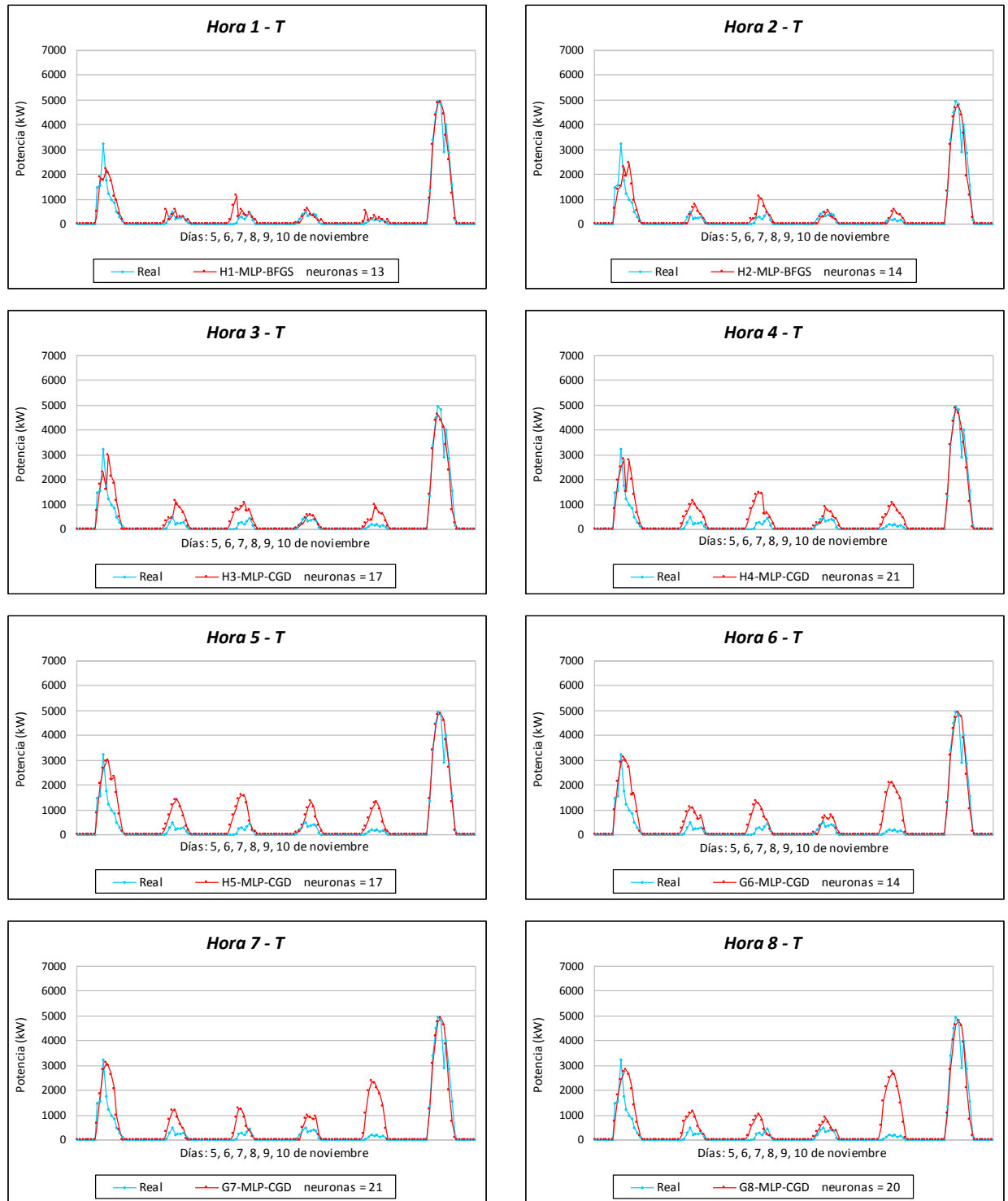


Figura 16.51: Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el mejor modelo basado en la red neuronal MLP-BFGS o MLP-CGD, para cada horizonte de predicción de 1 a 8 horas con datos de test de los días del 5 al 10 de noviembre de 2012.

Se eligen varios días que pertenecen al subconjunto de test T para representar los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con estos modelos (ver Figuras 16.50 y 16.51). Se puede observar como las predicciones son peores conforme aumenta el horizonte de predicción.

16.4.6. Horizonte de predicción de 0 a 24 horas

En esta sección se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria para cualquier tipo de día, desarrollados con las técnicas de minería de datos de sistemas de reglas, árboles de decisión, redes neuronales artificiales, máquinas de vectores soporte y sistemas neuro-difusos, usando los atributos R24, para horizontes de predicción entre 0 y 24 horas y con distintos subconjuntos de datos. En la Tabla 16.7 con $m = 24$ se presenta el significado de la notación que se utiliza en los resultados.

En la Figura 16.52 se presentan los valores del RMSE (kW) que se han obtenido con los modelos desarrollados para los horizontes de predicción entre 0 y 24 horas:

- En la Figura 16.52(a) se comparan los valores obtenidos con los modelos desarrollados con las técnicas M5Rules y M5P, variando el número de instancias entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 565 kW y 595 kW, y que los obtenidos con modelos desarrollados con la técnica M5P son menores a los obtenidos con los desarrollados con la técnica M5Rules.
- En la Figura 16.52(b) se comparan los valores obtenidos con los modelos desarrollados con redes neuronales artificiales del tipo MLP-GD, MLP-m(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS. Todos los modelos cuentan con una capa oculta con un número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE con los modelos MLP-GD y MLP-m(0.2) oscilan aproximadamente entre 600 kW y 640 kW, y son mayores que los obtenidos con los modelos MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS, que oscilan aproximadamente entre 545 kW y 565 kW. Los valores menores del RMSE en la mayoría de los casos se obtienen con modelos MLP-BFGS.
- En la Figura 16.52(c) se muestran los obtenidos con modelos basados en máquinas de vectores soporte. Con estos modelos se mantienen constantes el parámetro $\varepsilon = 0.001$, que mide la finura con que se va a tolerar el error con los datos del subconjunto de entrenamiento, y los valores de las variables de holgura, $\zeta_i = 0.001$. Se varía los valores de la constante $C \in \{5, 10, 15, 20\}$, que determina la holgura del margen blando. Se usan las funciones núcleo gaussianas con los valores del parámetro $\gamma \in \{0.2, 0.5, 1\}$. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 570 kW y 590 kW.
- En la Figura 16.52(d) se muestran los valores obtenidos con los modelos desarrollados con ANFIS, variando el rango de influencia del agrupamiento substractivo

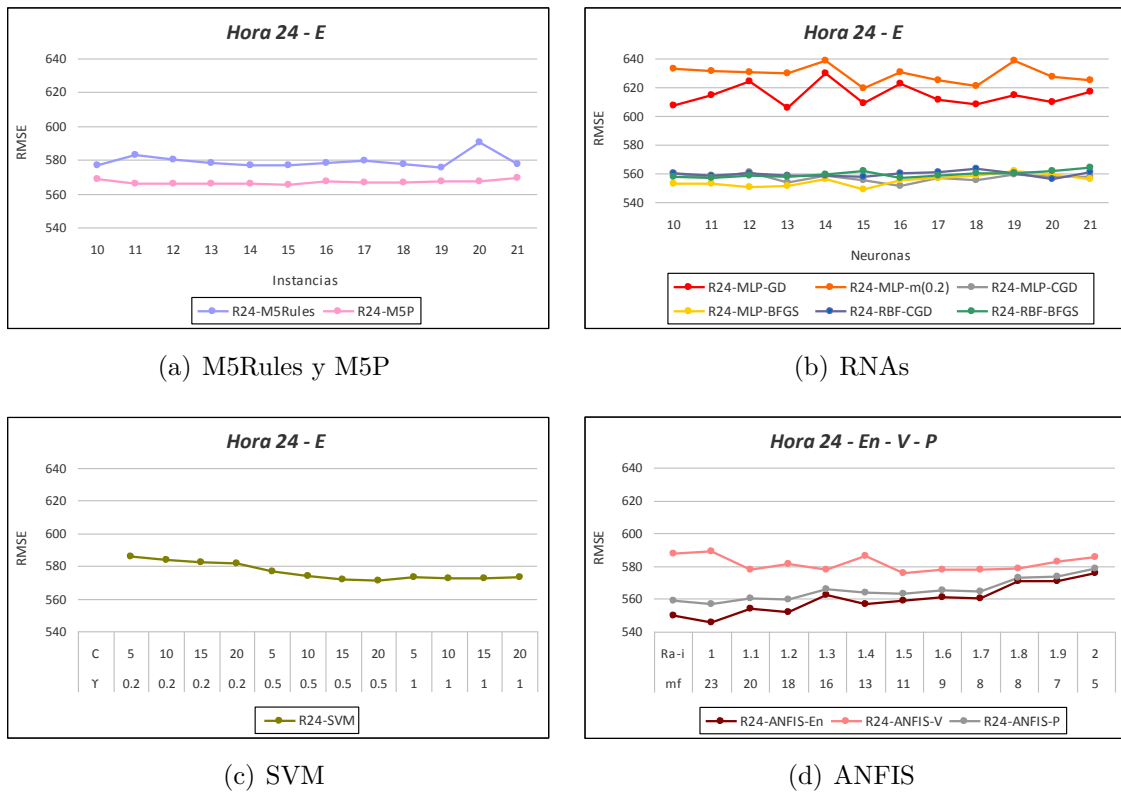


Figura 16.52: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria obtenidos con modelos basados en M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos R24, para horizontes de predicción entre 0 y 24 horas con los datos de entrenamiento.

$Ra-i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2\}$. Se utilizan los datos del subconjunto de entrenamiento En , y los datos del subconjunto de validación V . También, se representan los valores del RMSE ponderado, P , obtenidos con los modelos y los datos de los subconjuntos de entrenamiento En y de validación V . Se observa que los valores del RMSE oscilan aproximadamente entre 545 kW y 590 kW. El valor del RMSE en la mayoría de los casos disminuye conforme menor sea el rango de influencia del agrupamiento substractivo, pero aumenta el número de reglas y de funciones de pertenencia.

En la Tabla 16.26 se muestran los valores del RMSE (kW) más relevantes de cada tipo de modelo desarrollado con técnicas M5Rules, M5P, RNAs, SVM, ANFIS e HISIMI, usando los atributos R24, en la predicción de la potencia eléctrica media horaria para cualquier tipo de día, y su mejora con respecto al modelo persistente, para horizontes de predicción entre 0 y 24 horas, con los datos de los subconjuntos de entrenamiento E y de test T .

Los modelos que mejores resultados ofrecen son: el basado en una red neuronal artificial RBF-CGD con 20 neuronas en la capa oculta y que utiliza los atributos R24, y el basado en la red neuronal RBF-BFGS con 16 neuronas en la capa oculta y que usa también los atributos R24. Así, el modelo basado en la red RBF-CGD obtiene, con los datos

Tabla 16.26: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 0 y 24 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	P24		938.1573	809.7095	—	—
M5Rules	R24	instancias = 19	575.5234	553.7868	38.65 %	31.61 %
M5P	R24	instancias = 15	565.4486	557.5133	39.73 %	31.15 %
MLP-GD	R24	neuronas = 13	605.5818	569.6297	35.45 %	29.65 %
MLP-m(02)	R24	neuronas = 15	619.1985	564.7727	34.00 %	30.25 %
MLP-CGD	R24	neuronas = 16	551.5391	555.4069	41.21 %	31.41 %
MLP-BFGS	R24	neuronas = 15	549.1429	559.2476	41.47 %	30.93 %
RBF-CGD	R24	neuronas = 20	556.1976	543.0104	40.71 %	32.94 %
RBF-BFGS	R24	neuronas = 16	556.9902	535.1618	40.63 %	33.91 %
SVM	R24	C = 20, $\gamma = 0.5$	571.0900	569.2630	39.13 %	29.70 %
ANFIS	R24	mf=20, Ra-i=1.1	556.5407	539.4720	40.68 %	33.37 %
HISIMI	R24	en = 14, in = 12	596.5646	565.0307	36.41 %	30.22 %

Tabla 16.27: Coeficientes de determinación R^2 con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos basados en redes neuronales, para horizontes de predicción entre 0 y 24 horas.

Horizonte	Modelos	Atributos	Características	R^2 Entr.	R^2 Test
De 0 a 24 horas	MLP-CGD	R24	neuronas = 16	0.9221	0.9252
	MLP-BFGS	R24	neuronas = 15	0.9231	0.9242
	RBF-CGD	R24	neuronas = 20	0.9207	0.9278
	RBF-BFGS	R24	neuronas = 16	0.9204	0.9302

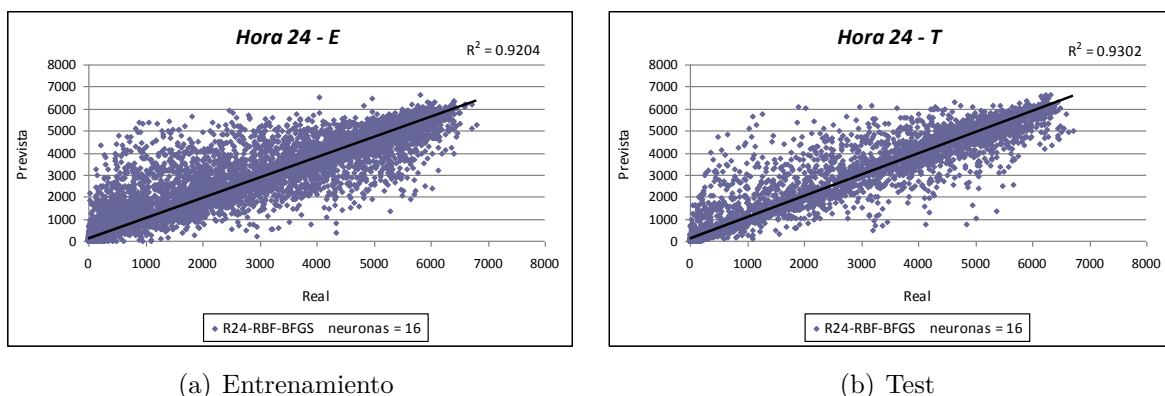


Figura 16.53: Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-BFGS con 16 neuronas, para horizontes de predicción entre 0 y 24 horas.

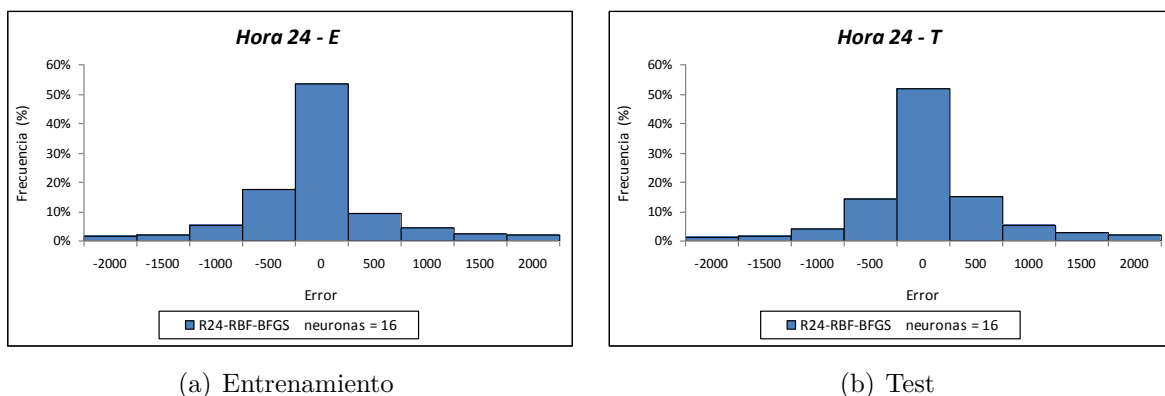


Figura 16.54: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con la red neuronal RBF-BFGS con 16 neuronas, para horizontes de predicción entre 0 y 24 horas.

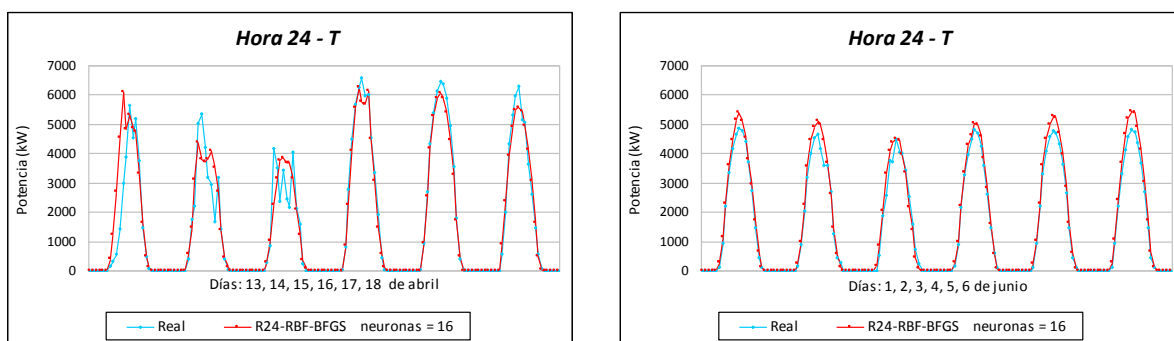


Figura 16.55: Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-BFGS con 16 neuronas, para horizontes de predicción entre 0 y 24 horas con datos de test.

del subconjunto de entrenamiento E , un valor del RMSE igual a 556.1976 kW (mejora con respecto al modelo persistente del 40.71 %) y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 543.0104 kW (mejora con respecto al modelo persistente del 32.94 %). El modelo basado en la red neuronal RBF-BFGS obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 556.9902 kW (mejora con respecto al modelo persistente del 40.63 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 535.1618 kW (mejora con respecto al modelo persistente del 33.91 %). Si se toma el valor del RMSE medio ponderado obtenido con los subconjuntos de entrenamiento E y de test T , el error obtenido con la RBF-BFGS con 16 neuronas es ligeramente menor.

En la Tabla 16.27 se muestran los coeficientes de determinación R^2 de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos basados en las redes neuronales artificiales: MLP-CGD con 16 neuronas en la capa oculta, MLP-BFGS con 15 neuronas en la capa oculta, RBF-CGD con 20 neuronas en la capa oculta, RBF-BFGS con 16 neuronas en la capa oculta. Los modelos se han desarrollado con los datos de los subconjuntos de entrenamiento E y de test T . Se observa que los coeficientes de determinación R^2 oscilan entre 0.9204 y 0.9302.

En la Figura 16.53 se representan los diagramas de dispersión de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo basado en la red neuronal artificial RBF-BFGS con 16 neuronas en la capa oculta y usando los atributos de entrada R24, con los datos del subconjunto de entrenamiento E (ver Figura 16.53(a)) y del subconjunto de test T (ver Figura 16.53(b)).

En la Figura 16.54 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de luz solar), obtenidos con dicho modelo, con los datos del subconjunto de entrenamiento E (ver Figura 16.54(a)) y del subconjunto de test T (ver Figura 16.54(b)).

En la Figura 16.55 se representan los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo que ha obtenido el menor valor del RMSE medio ponderado (RBF-BFGS con 16 neuronas en la capa oculta) con horizontes de predicción entre 0 y 24 horas, para los días 13, 14, 15, 16, 17, 18 de abril de 2012, y 1, 2, 3, 4, 5, 6 de junio de 2012, pertenecientes todos al subconjunto de test.

16.4.7. Horizonte de predicción de 25 a 48 horas

En esta sección se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria para cualquier tipo de día, desarrollados con las técnicas de minería de datos de sistemas de reglas, árboles de decisión, redes neuronales artificiales, máquinas de vectores soporte y sistemas neuro-difusos. El entrenamiento de los modelos se ha hecho con los atributos R48, para horizontes de predicción entre 25 y 48 horas, y con distintos subconjuntos de datos. En la Tabla 16.7 con $m = 48$ se presenta el significado de la notación que se utiliza en los resultados.

En la Figura 16.56 se presentan los valores del RMSE (kW) que se han obtenido con los modelos:

- En la Figura 16.56(a) se comparan los obtenidos con los modelos desarrollados con las técnicas M5Rules y M5P, variando el número de instancias entre 10 y 21. Se

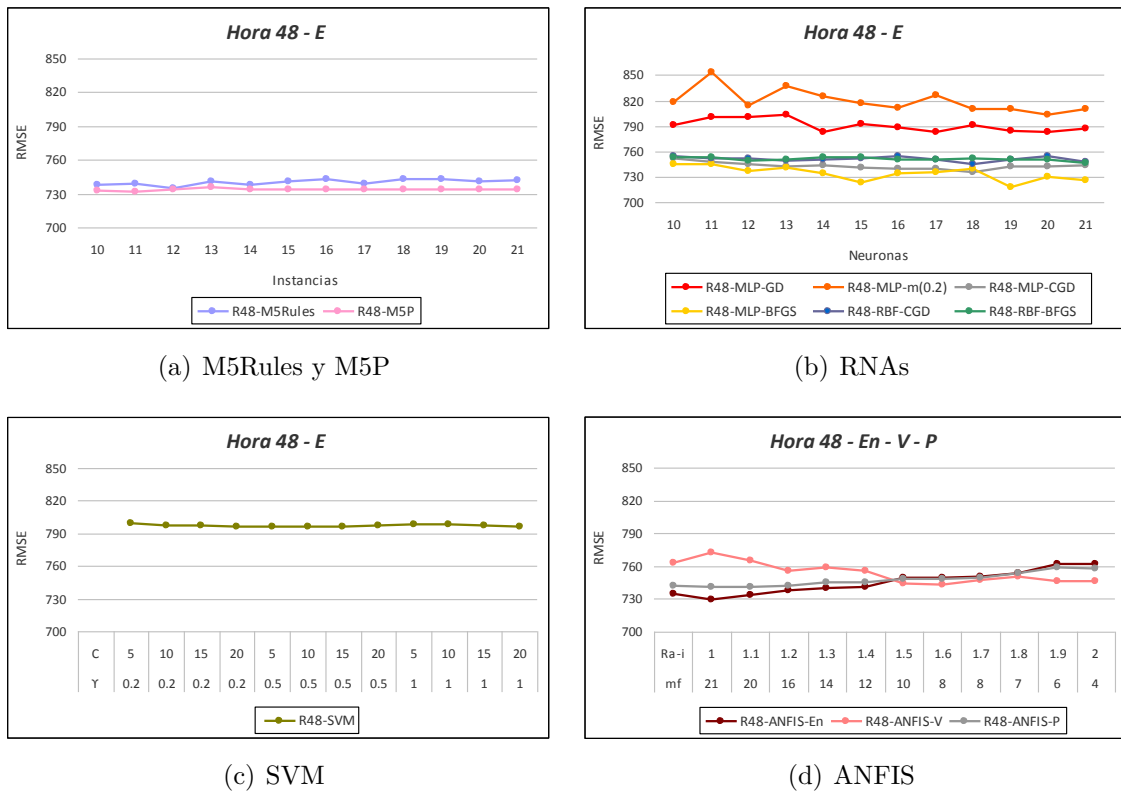


Figura 16.56: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria obtenidos con modelos basados en M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos R48, para horizontes de predicción entre 25 y 48 horas con los datos de entrenamiento.

utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 730 kW y 750 kW, y que los obtenidos con los modelos desarrollados con la técnica M5P son ligeramente menores a los obtenidos con los modelos desarrollados con la técnica M5Rules.

- En la Figura 16.56(b) se comparan los obtenidos con los modelos desarrollados con redes neuronales artificiales MLP-GD, MLP-m(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS, con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE con los modelos basados en MLP-GD y MLP-m(0.2) oscilan aproximadamente entre 785 kW y 850 kW, y son mayores que los obtenidos con los modelos basados en las redes neuronales MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS, que oscilan aproximadamente entre 720 kW y 760 kW. Los valores menores del RMSE se obtienen con modelos basados en la red neuronal MLP-BFGS.
- En la Figura 16.56(c) se muestran los obtenidos con modelos basados en máquinas de vectores soporte. Con estos modelos se mantienen constantes el parámetro $\varepsilon = 0.001$,

que mide la finura con que se va a tolerar el error con los datos del subconjunto de entrenamiento, y los valores de las variables de holgura, $\zeta_i = 0.001$. Se varía los valores de la constante $C \in \{5, 10, 15, 20\}$, que determina la holgura del margen blando. Se usan las funciones núcleo gaussianas con los valores del parámetro $\gamma \in \{0.2, 0.5, 1\}$. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 795 kW y 810 kW.

- En la Figura 16.56(d) se muestran los valores obtenidos con los modelos desarrollados con ANFIS, variando el rango de influencia del agrupamiento substractivo $Ra_i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2\}$. Se utilizan los datos del subconjunto de entrenamiento En , y los datos del subconjunto de validación V . También, se representan los valores del RMSE ponderado, P , obtenidos con los modelos y los datos de los subconjuntos de entrenamiento En y de validación V . Se observa que los valores del RMSE oscilan aproximadamente entre 725 kW y 780 kW.

En la Tabla 16.28 se muestran los valores del RMSE (kW) más relevantes obtenidos con los modelos basados en técnicas de M5Rules, M5P, RNAs, SVM, ANFIS e HISIMI, en la predicción de la potencia eléctrica media horaria para cualquier tipo de día, y su mejora con respecto al modelo persistente, para horizontes de predicción entre 25 y 48 horas, con los datos de los subconjuntos de entrenamiento E y de test T .

Los modelos que mejores resultados ofrecen son: el basado en una red neuronal artificial RBF-CGD con 18 neuronas en la capa oculta y que utiliza los atributos R48, y el basado en la red neuronal RBF-BFGS con 21 neuronas en la capa oculta y que usa también los atributos R48. Así, el modelo basado en la red RBF-CGD obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 745.3241 kW (mejora con respecto al modelo persistente del 27.17 %) y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 712.6638 kW (mejora con respecto al modelo persistente del 22.21 %). El modelo basado en la red neuronal RBF-BFGS obtiene, con los datos del subconjunto de entrenamiento E , un valor del RMSE igual a 746.4945 kW (mejora con respecto al modelo persistente del 27.05 %), y con los datos del subconjunto de test T obtiene un valor del RMSE igual a 713.8387 kW (mejora con respecto al modelo persistente del 22.08 %). Sin embargo, si se considera como mejor modelo aquel con el que se obtiene el menor valor del RMSE medio ponderado de los errores obtenidos con los datos de entrenamiento E y de test T , el basado en la red neuronal RBF-CGD con 18 neuronas en la capa oculta, obtiene el menor valor.

En la Tabla 16.29 se muestran los coeficientes de determinación R^2 de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los modelos MLP-CGD con 18 neuronas en la capa oculta, MLP-BFGS con 19 neuronas en la capa oculta, RBF-CGD con 18 neuronas en la capa oculta y RBF-BFGS con 21 neuronas en la capa oculta. Los modelos se han desarrollado con los datos de los subconjuntos de entrenamiento E y de test T . Se observa que los coeficientes de determinación R^2 oscilan entre 0.8551 y 0.8729.

En la Figura 16.57 se representan los diagramas de dispersión de los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo basado en la red neuronal artificial RBF-CGD con 18 neuronas en la capa oculta y usando los

Tabla 16.28: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 25 y 48 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	P48		1023.3407	916.1243	—	—
M5Rules	R48	instancias = 12	734.8857	748.1017	28.19 %	18.34 %
M5P	R48	instancias = 11	731.6907	744.3047	28.50 %	18.76 %
MLP-GD	R48	neuronas = 14	782.7576	706.6659	23.51 %	22.86 %
MLP-m(02)	R48	neuronas = 20	802.7526	701.2920	21.56 %	23.45 %
MLP-CGD	R48	neuronas = 18	734.8885	757.4332	28.19 %	17.32 %
MLP-BFGS	R48	neuronas = 19	718.2104	769.6914	29.82 %	15.98 %
RBF-CGD	R48	neuronas = 18	745.3241	712.6638	27.17 %	22.21 %
RBF-BFGS	R48	neuronas = 21	746.4945	713.8387	27.05 %	22.08 %
SVM	R48	C = 10, $\gamma = 0.5$	795.9757	758.5141	22.22 %	17.20 %
ANFIS	R48	mf=20, Ra-i=1.1	740.5433	728.2744	27.63 %	20.50 %
HISIMI	R48	en = 8, in = 8	774.3649	692.2753	24.33 %	24.43 %

Tabla 16.29: Coeficientes de determinación R^2 con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con los mejores modelos basados en redes neuronales, para horizontes de predicción entre 25 y 48 horas.

Horizonte	Modelos	Atributos	Características	R^2 Entr.	R^2 Test
De 25 a 48 horas	MLP-CGD	R48	neuronas = 18	0.8628	0.8583
	MLP-BFGS	R48	neuronas = 19	0.8692	0.8551
	RBF-CGD	R48	neuronas = 18	0.8576	0.8729
	RBF-BFGS	R48	neuronas = 21	0.8572	0.8725

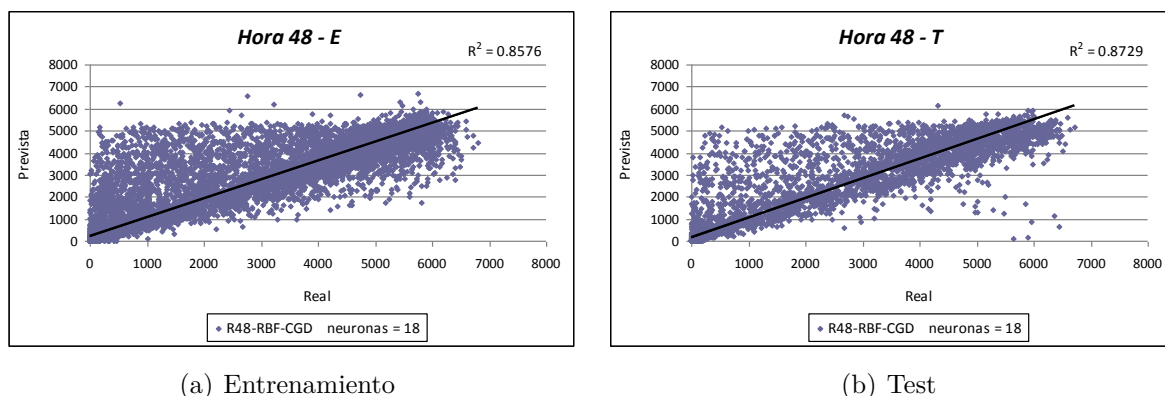


Figura 16.57: Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-CGD con 18 neuronas, para horizontes de predicción entre 25 y 48 horas.

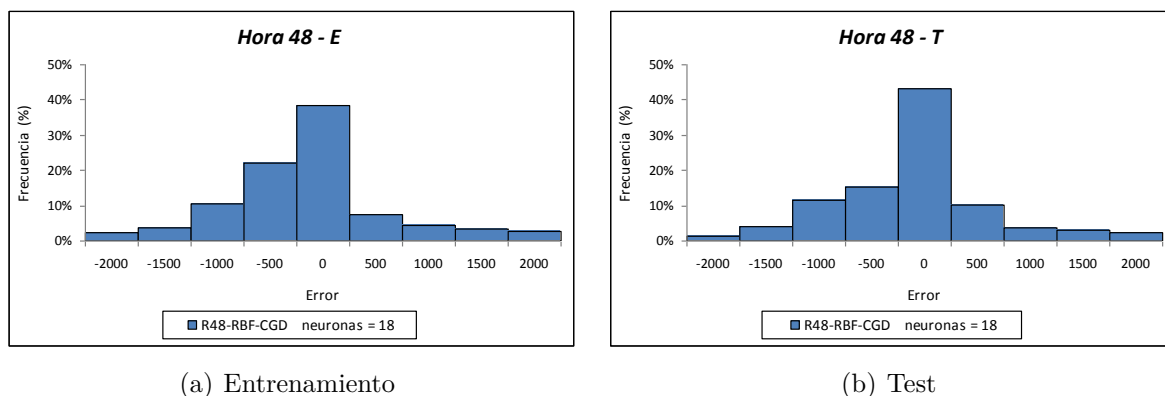


Figura 16.58: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con la red neuronal RBF-CGD con 18 neuronas, para horizontes de predicción entre 25 y 48 horas.

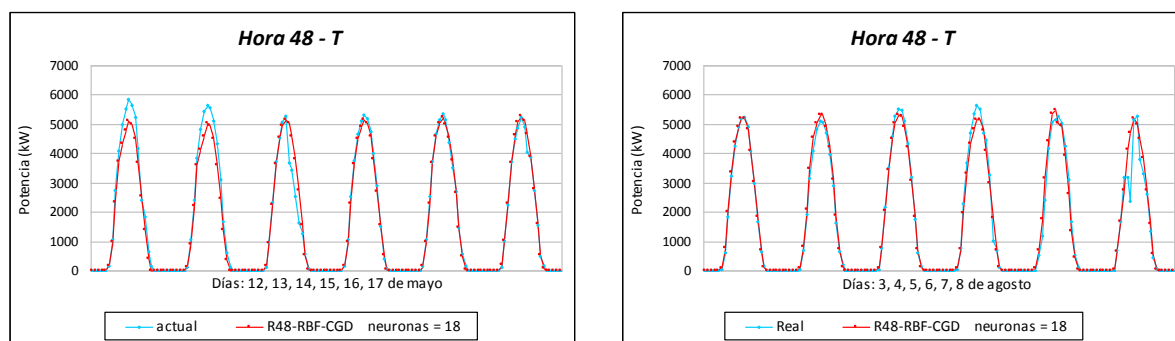


Figura 16.59: Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con la red neuronal RBF-CGD con 18 neuronas, para horizontes de predicción entre 25 y 48 horas con datos de test.

atributos de entrada R48, con los datos del subconjunto de entrenamiento E (ver Figura 16.57(a)) y del subconjunto de test T (ver Figura 16.57(b)).

En la Figura 16.58 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de luz solar), obtenidos con dicho modelo, con los datos del subconjunto de entrenamiento E (ver Figura 16.58(a)) y del subconjunto de test T (ver Figura 16.58(b)).

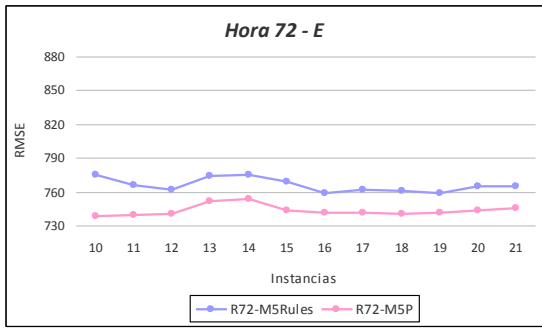
En la Figura 16.59 se representan los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo que ha obtenido el menor valor del RMSE medio ponderado (RBF-CGD con 18 neuronas en la capa oculta) con horizontes de predicción entre 25 y 48 horas, para los días 12, 13, 14, 15, 16, 17 de mayo de 2012, y 3, 4, 5, 6, 7, 8 de agosto de 2012, pertenecientes todos al subconjunto de test.

16.4.8. Horizonte de predicción de 49 a 72 horas

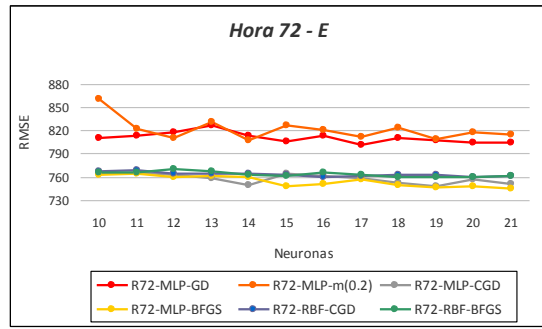
En esta sección se exponen los resultados que se han obtenido con los modelos de predicción de la potencia eléctrica media horaria para cualquier tipo de día, desarrollados con las técnicas de minería de datos de sistemas de reglas, árboles de decisión, redes neuronales artificiales, máquinas de vectores soporte y sistemas neuro-difusos. Todos los modelos se han desarrollado usando los atributos R72, para horizontes de predicción entre 49 y 72 horas, y con distintos subconjuntos de datos. En la Tabla 16.7 con $m = 72$ se presenta el significado de la notación que se utiliza en los resultados.

En la Figura 16.60 se presentan los valores del RMSE (kW) que se han obtenido con los modelos:

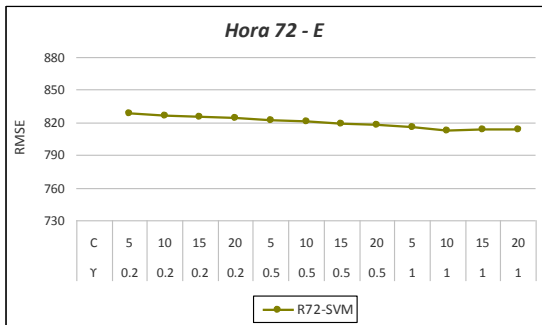
- En la Figura 16.60(a) se comparan los obtenidos con los modelos desarrollados con las técnicas M5Rules y M5P, variando el número de instancias entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 735 kW y 785 kW, y que los obtenidos con los modelos basados en la técnica M5P son menores a los obtenidos con los modelos basados en la técnica M5Rules.
- En la Figura 16.60(b) se comparan los obtenidos con los modelos desarrollados con redes neuronales artificiales MLP-GD, MLP-m(0.2), MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS, con una capa oculta variando su número de neuronas entre 10 y 21. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE con los modelos basados en MLP-GD y MLP-m(0.2) oscilan aproximadamente entre 800 kW y 865 kW, y son mayores que los obtenidos con los modelos basados en MLP-CGD, MLP-BFGS, RBF-CGD y RBF-BFGS, que oscilan aproximadamente entre 745 kW y 780 kW. Los valores menores del RMSE en la mayoría de los casos se obtienen con modelos basados en la red neuronal MLP-BFGS.
- En la Figura 16.60(c) se muestran los obtenidos con modelos basados en máquinas de vectores soporte. Con estos modelos se mantienen constantes el parámetro $\varepsilon = 0.001$, que mide la finura con que se va a tolerar el error con los datos del subconjunto de entrenamiento, y los valores de las variables de holgura, $\zeta_i = 0.001$. Se varía los



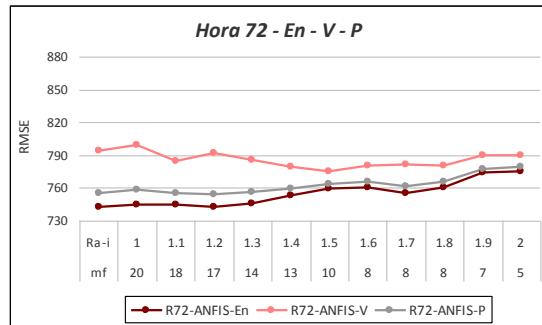
(a) M5Rules y M5P



(b) RNAs



(c) SVM



(d) ANFIS

Figura 16.60: Valores del RMSE (kW) en la predicción de la potencia eléctrica media horaria obtenidos con modelos basados en M5Rules, M5P, RNAs, SVM y ANFIS, usando los atributos R72, para horizontes de predicción entre 49 y 72 horas con los datos de entrenamiento.

valores de la constante $C \in \{5, 10, 15, 20\}$, que determina la holgura del margen blando. Se usan las funciones núcleo gaussianas con los valores del parámetro $\gamma \in \{0.2, 0.5, 1\}$. Se utilizan los datos del subconjunto de entrenamiento E con la técnica de validación cruzada de 5-iteraciones. Se observa que los valores del RMSE oscilan aproximadamente entre 810 kW y 830 kW.

- En la Figura 16.60(d) se muestran los valores obtenidos con los modelos desarrollados con ANFIS, variando el rango de influencia del agrupamiento substractivo $Ra-i \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2\}$. Se utilizan los datos del subconjunto de entrenamiento En , y los datos del subconjunto de validación V . También, se representan los valores del RMSE ponderado, P , obtenidos con los modelos y los datos de los subconjuntos de entrenamiento En y de validación V . Se observa que los valores del RMSE oscilan aproximadamente entre 740 kW y 800 kW. El valor del RMSE en la mayoría de los casos disminuye conforme menor sea el rango de influencia, pero aumentan el número de reglas y de funciones de pertenencia.

En la Tabla 16.30 se muestran los valores del RMSE (kW) más relevantes obtenidos con los modelos basados en técnicas de M5Rules, M5P, RNAs, SVM, ANFIS e HISIMI, en la predicción de la potencia eléctrica media horaria para cualquier tipo de día, y su

Tabla 16.30: Valores del RMSE (kW) más relevantes de los mejores modelos desarrollados con cada técnica en la predicción de la potencia eléctrica media horaria y mejora con respecto al modelo persistente, para horizontes de predicción entre 49 y 72 horas.

Modelos	Atributos	Características	RMSE (kW) Entr.	RMSE (kW) Test	Mejora (%) Entr.	Mejora (%) Test
Persistente	P72		1055.6028	954.9182	—	—
M5Rules	R72	instancias = 16	758.4094	746.7144	28.15 %	21.80 %
M5P	R72	instancias = 10	738.1058	783.1051	30.08 %	17.99 %
MLP-GD	R72	neuronas = 17	801.0454	761.4378	24.11 %	20.26 %
MLP-m(02)	R72	neuronas = 14	807.4104	751.3373	23.51 %	21.32 %
MLP-CGD	R72	neuronas = 19	747.1480	772.6981	29.22 %	19.08 %
MLP-BFGS	R72	neuronas = 21	745.3999	811.2311	29.39 %	15.05 %
RBF-CGD	R72	neuronas = 17	760.3585	744.8399	27.97 %	22.00 %
RBF-BFGS	R72	neuronas = 19	759.6719	772.2987	28.03 %	19.12 %
SVM	R72	$C = 10, \gamma = 0.5$	812.6664	797.2680	23.01 %	16.51 %
ANFIS	R72	mf=14, Ra-i=1.3	754.6276	751.4154	28.51 %	21.31 %
HISIMI	R72	en = 9, in = 6	794.5619	737.7609	24.73 %	22.74 %

mejora con respecto al modelo persistente, para horizontes de predicción entre 49 y 72 horas, con los datos de los subconjuntos de entrenamiento E y de test T .

El modelo que mejores resultados ofrece con los datos del subconjunto de entrenamiento E es el basado en árboles de decisión con la técnica M5P. La división con M5P cesa cuando los valores de todas las instancias que llegan a un nodo varían poco, o solo quedan 10 instancias. Con los datos del subconjunto de entrenamiento E se obtiene el valor del RMSE igual a 738.1058 kW, su mejora con respecto al modelo persistente es del 30.08 %, y con los datos del subconjunto de test T se obtiene un valor del RMSE igual a 783.1051 kW, su mejora con respecto al modelo persistente es del 17.99 %. Si se considera el valor del RMSE medio ponderado con los datos de los subconjuntos de entrenamiento E y de test T , el mejor modelo sigue siendo el basado en árboles de decisión con la técnica M5P, seguido muy de cerca por el sistema ANFIS de 14 reglas y después por el sistema de reglas con la técnica M5Rules con el mínimo de 16 instancias permitidas en un nodo hoja.

En la Figura 16.61 se representan los diagramas de dispersión con valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo desarrollado con la técnica M5P, cesando la división cuando solo quedan 10 instancias en un nodo, y usando los atributos de entrada R72, con los datos del subconjunto de entrenamiento E

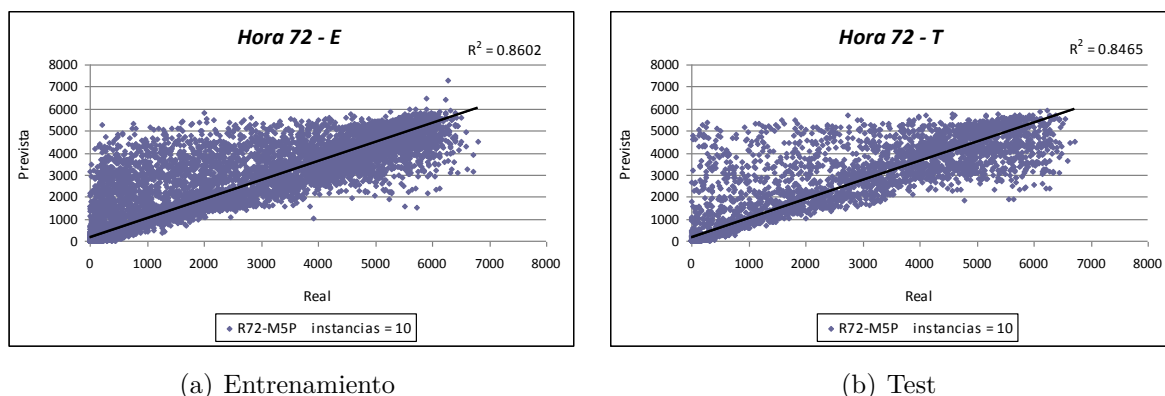


Figura 16.61: Diagramas de dispersión con los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo basado en la técnica M5P con 10 instancias, para horizontes de predicción entre 49 y 72 horas.

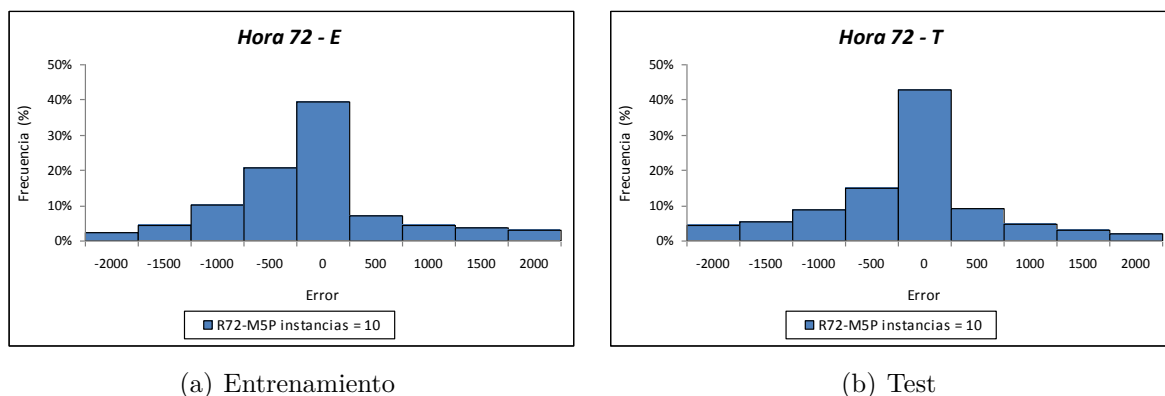


Figura 16.62: Histogramas del error (kW) en la predicción de la potencia eléctrica media horaria (para las horas de luz solar) obtenidos con el modelo basado en la técnica M5P con 10 instancias, para horizontes de predicción entre 49 y 72 horas.

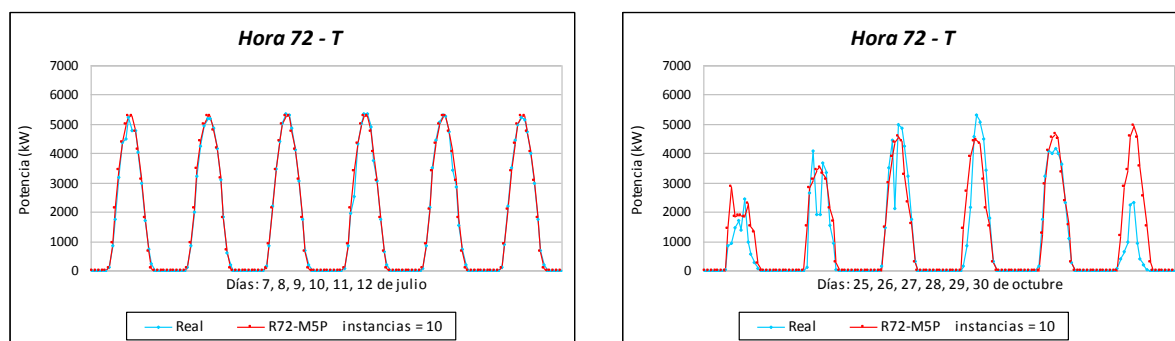


Figura 16.63: Valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo basado en la técnica M5P con 10 instancias, para horizontes de predicción entre 49 y 72 horas con datos de test.

(ver Figura 16.61(a)) y del subconjunto de test T (ver Figura 16.61(b)). El coeficiente de determinación es $R^2 = 0.8602$ con los datos del subconjunto E , y es $R^2 = 0.8465$ con los datos del subconjunto de test T .

En la Figura 16.62 se muestran los histogramas del error en la predicción de la potencia eléctrica media horaria (para las horas de luz solar), obtenidos con dicho modelo, con los datos del subconjunto de entrenamiento E (ver Figura 16.62(a)) y del subconjunto de test T (ver Figura 16.62(b)).

En la Figura 16.63 se representan los valores reales de la potencia eléctrica media horaria (kW) y la predicción obtenida con el modelo que ha obtenido el menor valor del RMSE medio ponderado (M5P con 10 instancias), con horizontes de predicción entre 49 y 72 horas, para los días 7, 8, 9, 10, 11, 12 de julio de 2012, y 25, 26, 27, 28, 29, 30 de octubre de 2012, días que pertenecen al subconjunto de test.

16.5. Resultados con incertidumbre de la predicción

La incertidumbre asociada a cada punto de predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica, se puede calcular fácilmente utilizando el resultado principal del modelo HISIMI (ver capítulo 15), es decir, la distribución de probabilidad discreta asociada con la transición de la potencia eléctrica para cada paso en el horizonte de predicción. Este modelo se optimiza por un algoritmo genético, que permite la selección de los atributos dentro de todos los disponibles y selecciona el número óptimo de intervalos de potencia. En los trabajos de investigación de esta tesis se han hecho pruebas considerando como entradas diferentes atributos disponibles para cada horizonte de predicción (ver Tablas 14.1, 14.2, 14.3).

A continuación, se resumen las principales características de los modelos seleccionados para predicción de la potencia eléctrica media horaria con evaluación de la incertidumbre para cualquier tipo de día, con los horizontes de predicción entre 0 y 24 horas, entre 25 y 48 horas, y entre 49 y 72 horas.

16.5.1. Horizonte de predicción de 0 a 24 horas

Los resultados de los valores del RMSE que se han obtenido con este modelo se han expuesto en la Tabla 16.26, se han conseguido considerado como entradas disponibles los atributos $R24$ (ver Tabla 16.3). El modelo se ha optimizado por un algoritmo genético que ha elegido como entradas en $= 14$ de los atributos disponibles ($RSWIN_24$, $\sin \delta$, $\cos \delta$, $\sin \gamma$, $\cos \gamma$, $\sin \psi$, $\cos \psi$, $\cos i$, K_d , $hluzsol$, $RSWINMed_24$, $RSWINCuad_24$, $RSWINRaiz_24$, $POTCalC24_24$), y el número óptimo de intervalos de potencia que ha seleccionado ha sido $in = 12$.

16.5.2. Horizonte de predicción de 25 a 48 horas

Los resultados de los valores del RMSE que se han obtenido con este modelo se han expuesto en la Tabla 16.28, se han conseguido considerado como entradas disponibles los atributos $R48$ (ver Tabla 16.3). El modelo se ha optimizado por un algoritmo genético que ha elegido como entradas en $= 8$ de los atributos disponibles ($RSWIN_48$, $\sin \delta$,

$\cos \delta$, $\sin \omega$, $\cos \omega$, $\sin \psi$, $\cos \psi$, $hluzsol$), y ha seleccionado $in = 8$ intervalos de potencia como el número óptimo.

En la Figura 16.64 se muestra la distribución de probabilidad, correspondiente a la incertidumbre asociada al punto de predicción de la potencia eléctrica media horaria, durante 8 horas seleccionadas de un día de mayo, con un horizonte de predicción entre 25 y 48 horas. Se ha escogido este horizonte por ser el que tendría una aplicación más directa en la formulación de propuestas de venta de energía en la sesión diaria del mercado eléctrico, aplicación para la que la información de la incertidumbre asociada al valor esperado de producción puede ser fundamental para la toma de decisiones. Cuando solo hay un intervalo significativo la incertidumbre es baja. También lo es cuando solo hay dos intervalos significativos pero contiguos. La incertidumbre es alta cuando hay más de dos significativos, o son dos pero no contiguos. En los ejes horizontales se muestran los intervalos de potencia de la planta fotovoltaica, y en los ejes verticales los valores de la probabilidad correspondiente. Por ejemplo, se puede observar que en:

- La hora 6:00 (tiempo solar), el intervalo 2 de potencia eléctrica media horaria es el que mayor probabilidad tiene, con un valor promedio de 0.469, le sigue el intervalo 1 con 0.323. Hay otros intervalos como el 3 con 0.125 y el 4 con 0.055. La incertidumbre en el valor puntual pronosticado, que es 962.8903 kW (valor real = 1045 kW), es un poco alta, ya que hay varios intervalos de potencia con valores de probabilidad que puede ser significativa.
- La hora 11:00 (tiempo solar), el intervalo 7 de potencia eléctrica media horaria es el que mayor probabilidad tiene, con un valor promedio de 0.596, le sigue el intervalo 6 con 0.333. El resto de intervalos presentan valores de probabilidad no significativos. Dado que son dos los intervalos de potencia con valores de probabilidad significativos y además corresponden a intervalos consecutivos, la incertidumbre en el valor puntual pronosticado, que es 5326.0601 kW (valor real = 5376 kW) es baja.
- La hora 15:00 (tiempo solar), el intervalo 4 de potencia eléctrica media horaria es el que mayor probabilidad tiene, con un valor promedio de 0.347, le sigue el intervalo 3 con 0.298. Hay otros intervalos que su probabilidad no se puede despreciar como el 2 con 0.102, el 5 con 0.108 y el 6 con 0.109. Al haber varios intervalos de potencia con valores de probabilidad significativos, la incertidumbre en el valor puntual pronosticado, que es 2692.9354 (valor real = 2721 kW) es alta.

Este tipo de información sobre la incertidumbre es muy importante para los usuarios que sean sensibles al riesgo de desviación del valor puntual de la predicción.

16.5.3. Horizonte de predicción de 49 a 72 horas

Los resultados de los valores del RMSE que se han obtenido con este modelo se han expuesto en la Tabla 16.30, se han conseguido considerado como entradas disponibles los atributos $R72$ (ver Tabla 16.3). El modelo se ha optimizado por un algoritmo genético que ha elegido como entradas en $= 9$ de los atributos disponibles ($RSWIN_72$, $\cos \delta$, $\sin \omega$, $\cos \omega$, $\cos \gamma$, $\cos \psi$, $\cos i$, K_d , $RSWINMed_72$), y ha seleccionado como el número óptimo de intervalos de potencia igual a $in = 6$.

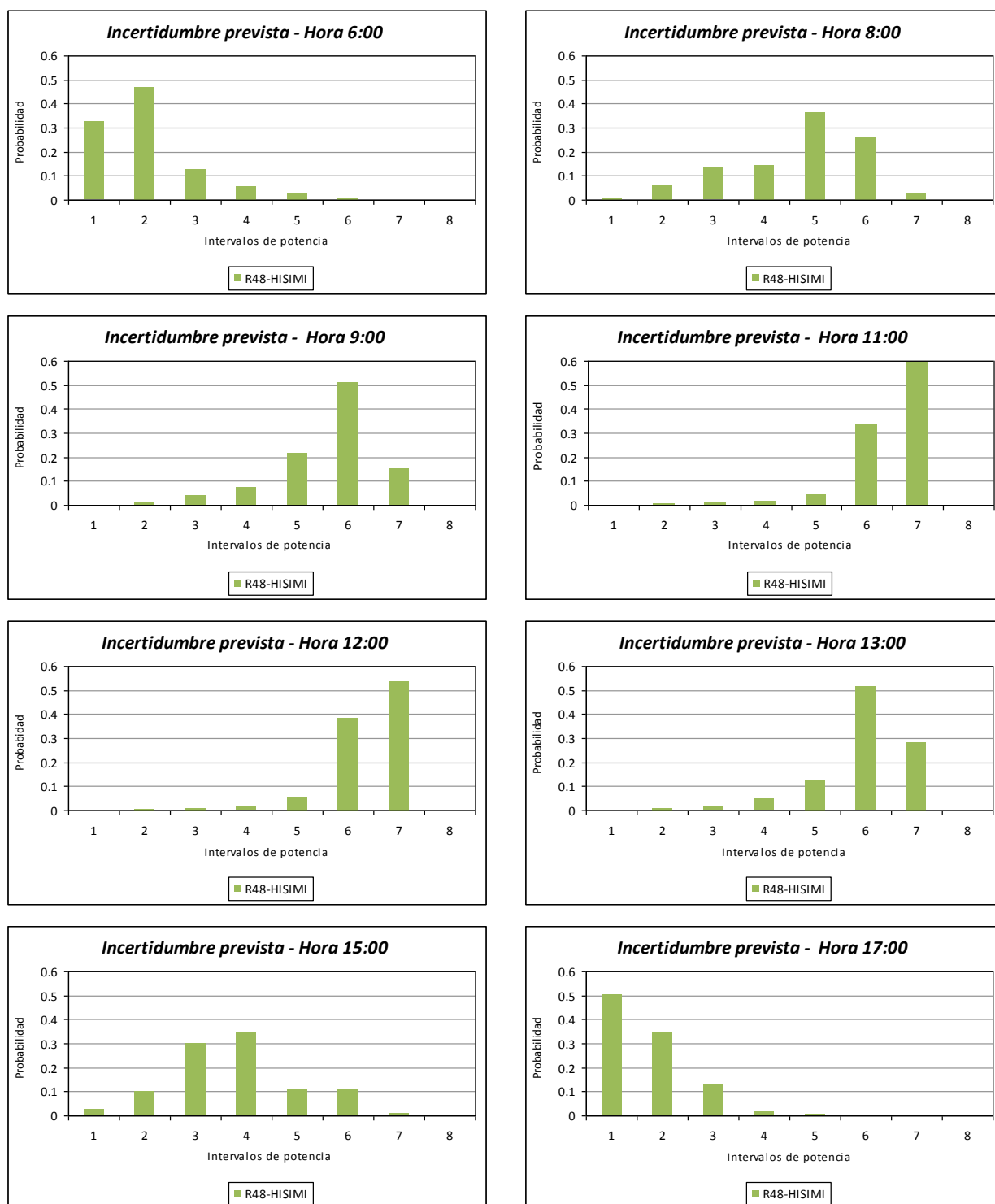


Figura 16.64: Distribución de probabilidad, correspondiente a la incertidumbre asociada al valor puntual de predicción de la potencia eléctrica media horaria, para algunas horas (tiempo solar) del día 16 de mayo de 2012, desarrolladas para un horizonte de predicción entre 25 y 48 horas.

Es de observar que, como cabía esperarse, el número de intervalos seleccionados por el algoritmo genético en la optimización del modelo se reduce conforme aumenta el horizonte de predicción. Esto es consecuencia de la menor precisión de las predicciones meteorológicas al aumentar este horizonte, lo que provoca una mayor dispersión de los datos en la base histórica (datos pasados), que justifica un mayor rango en la anchura de estos intervalos y, por tanto, un menor número de ellos.

Capítulo 17

Conclusiones

En este capítulo se presentan las conclusiones y las aportaciones originales de los trabajos de investigación de esta tesis doctoral, además de las futuras líneas de investigación en las que se puede continuar trabajando.

17.1. Conclusiones

En esta sección se presentan las principales conclusiones de esta tesis doctoral:

1. En muchos modelos de predicción de la potencia eléctrica media horaria publicados en la literatura internacional, se utiliza como criterio de evaluación del error el RMSE, y en la mayoría de ellos se utiliza el modelo persistente como modelo de referencia para cualquier horizonte de predicción.
2. El éxito de los buenos resultados en los modelos de predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica depende en gran medida de la calidad de las predicciones proporcionadas por el modelo meteorológico elegido. En los resultados que se obtienen con los nuevos modelos se observa que, cuanto mayor es el horizonte de predicción, mayores son los valores del RMSE. Esto se debe a que estos modelos usan valores de variables de predicciones meteorológicas cada vez menos precisas conforme aumenta el horizonte de predicción. Es fundamental tener valores con buena precisión para la variable de predicción de la radiación. Así que hay que seguir investigando para desarrollar nuevos modelos que consigan predicciones meteorológicas más precisas.
3. La recopilación de datos debe ir acompañada de la eliminación del mayor número posible de datos incorrectos. Además, una buena elección del método de aprendizaje y del subconjunto de atributos de entrada es clave para obtener buenos resultados en los modelos de predicción de la potencia eléctrica media horaria para horizontes de predicción a corto plazo.
4. El desarrollo de computadores cada vez más potentes permite el desarrollo de estos modelos de predicción con más rapidez, y así se consigue reducir el tiempo que se

tarda en efectuar las predicciones de la potencia eléctrica media horaria, posibilitando su uso en aplicaciones en tiempo real (elaboración de ofertas al mercado eléctrico, ejecución de labores de mantenimiento, etc.).

5. En los resultados que se han obtenido para cualquier tipo de día con horizontes de predicción de 1 a 8 horas, se observa que los modelos que utilizan como entradas los atributos G_m ; $m \in \{1, \dots, 8\}$ consiguen un menor valor del RMSE (kW) que los que utilizan como entradas los atributos R_m ; $m \in \{1, \dots, 8\}$. Y además, se observa que los menores valores del RMSE (kW) se obtienen si se utilizan como entradas los atributos H_m ; $m \in \{1, \dots, 5\}$ para horizontes entre 1 y 5 horas, y los atributos G_m ; $m \in \{6, \dots, 8\}$ para horizontes entre 6 y 8 horas. Como conclusión, el uso de predicciones de variables meteorológicas mejora los resultados en la predicción de la potencia media horaria generada en la planta fotovoltaica para horizontes de 6 o más horas.
6. Los modelos de predicción que aplican como método de aprendizaje las máquinas de vectores soporte con el tipo de función núcleo gaussiano son bastantes aceptables para predecir la potencia eléctrica media horaria de cualquier tipo de día para horizontes a muy corto plazo (entre 1 y 8 horas).
7. Una técnica que se puede aplicar con bastante éxito para conseguir buenas predicciones en los modelos de predicción de la potencia eléctrica media horaria para horizontes de predicción a muy corto plazo, está basada en los sistemas de inferencia neuro-difusos. Además, de todos los modelos desarrollados de días con cielo claro para todos los horizontes de predicción analizados, los modelos basados en sistemas ANFIS obtienen el menor valor RMSE medio ponderado, incluso menor que el obtenido con la mejor de las redes neuronales.
8. Una de las mejores técnicas para conseguir predicciones más precisas en los modelos de predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica para cualquier tipo de día, están basadas en redes neuronales artificiales MLP entrenadas con el algoritmo de gradiente conjugado descendente (CGD) o con el algoritmo de cuasi-Newton más popular, el método BFGS. Con estas técnicas se consiguen muy buenos resultados para horizontes de predicción de 1 a 8 horas. Si se toma el valor del RMSE medio ponderado obtenido con los subconjuntos de entrenamiento y test, el error obtenido con estos dos tipos de redes es prácticamente igual, aunque mejora con el método BFGS a muy corto plazo (para las primeras horas) y después mejora ligeramente con el método CGD para las siguientes horas hasta las 8 horas.
9. Para horizontes de predicción a corto plazo, hasta 48 horas, los modelos de predicción de la potencia eléctrica media horaria para cualquier tipo de día, basados en las redes neuronales de funciones de base radial (RBF) entrenadas con los métodos BFGS o CGD son los que obtienen con los subconjuntos de entrenamiento y test el menor valor del RMSE medio ponderado. El error obtenido con estos dos métodos es prácticamente igual, aunque mejora con el método BFGS hasta el horizonte de predicción de 24 horas y con el método CGD para horizontes entre 25 y 48 horas.

10. Para horizontes de predicción a más largo plazo, entre 49 y 72 horas, los errores en los modelos de predicción de la potencia eléctrica media horaria para cualquier tipo de día son mayores que los obtenidos con los modelos desarrollados para horizontes menores, debido a que éstos utilizan como entradas atributos con variables meteorológicas pronosticadas por un modelo numérico y su precisión empeora conforme aumenta el horizonte de predicción. Para estos horizontes, con los modelos basados en las redes neuronales anteriores no se obtienen buenos resultados. Sin embargo, los modelos basados en árboles entrenados con el algoritmo M5P tienden a aproximar la predicción entre los mejores.
11. Como conclusión final, se puede destacar que para cualquier tipo de día, si se conseguirían predicciones muy precisas de variables meteorológicas, en especial de la radiación, para un horizonte de predicción concreto, los modelos basados en redes neuronales MLP entrenadas con los métodos BFGS o CGD son con los que se obtendrían las mejores predicciones de la potencia eléctrica media horaria generada en una planta fotovoltaica, y además, estas predicciones serían ligeramente más precisas con los modelos que utilizaran el método BFGS.

17.2. Principales aportaciones de la tesis

A continuación, se indican las principales aportaciones originales de los trabajos de investigación de esta tesis:

1. Se ha hecho una comparativa de los resultados que se han obtenido con los mejores modelos de predicción de la potencia eléctrica media horaria generada en una planta fotovoltaica desarrollados con diferentes técnicas de minería de datos, incluyendo las de árboles de decisión, sistemas de reglas, redes neuronales artificiales, máquinas de vectores soporte, y sistemas neuro-difusos. No se ha encontrado en la literatura internacional ningún trabajo en el cual se comparan resultados de modelos desarrollados con tantas técnicas diferentes como se ha realizado en los trabajos de investigación de esta tesis.
2. Todos los datos necesarios para la obtención de los valores de predicción de los nuevos modelos desarrollados son fácilmente accesibles sin coste económico de adquisición. Son valores de medidas de potencia eléctrica media horaria en la misma planta fotovoltaica, o públicos, ya que las predicciones de las variables meteorológicas se han obtenido con el modelo mesoescalar WRF-NMM versión 3 de libre distribución. Así que los modelos de predicción de la potencia eléctrica media horaria se pueden obtener sin coste alguno.
3. Los modelos de predicción desarrollados en esta tesis se pueden aplicar a una planta fotovoltaica ubicada en cualquier lugar del planeta, con solo obtener la serie temporal histórica de la potencia eléctrica media horaria generada en la planta fotovoltaica en estudio, y valores previos horarios históricos y futuros de predicciones meteorológicas obtenidas por un modelo atmosférico para la ubicación del lugar. Cuantos más datos se dispongan en estas series temporales históricas más precisos son los resultados.

En lugares donde predominan los días con cielo claro se obtienen mejores resultados que en lugares donde hay muchos días con cielo cubierto o con lluvia.

4. Se ha creado un nuevo modelo muy simple para pronosticar la irradiancia media horaria sobre la superficie horizontal para un horizonte de predicción a corto plazo. Hasta un horizonte de predicción de 5 horas los modelos de predicción de la potencia media horaria desarrollados que utilizan estos valores en un atributo de entrada obtienen muy buenos resultados, y mejoran el valor del RMSE con respecto al obtenido con los modelos que emplean en un atributo de entrada el valor de la predicción de la radiación media horaria pronosticada por un modelo atmosférico.
5. Se ha conseguido un nuevo modelo para pronosticar la irradiancia media horaria sobre una superficie con un ángulo de inclinación β , para un horizonte de predicción de 1 a 8 horas. Los modelos de predicción de la potencia media horaria desarrollados que utilizan estos valores en un atributo de entrada, mejoran el valor del RMSE con respecto al obtenido con los modelos que emplean en un atributo de entrada el valor de la predicción de la radiación media horaria pronosticada por un modelo atmosférico.
6. Se han desarrollado un conjunto de nuevos modelos capaces de realizar predicciones de la potencia eléctrica media horaria generada en una planta fotovoltaica que puede estar ubicada en cualquier lugar del planeta, para días con cielo claro y para días de cualquier tipo, con horizontes de predicción a corto plazo. Estos modelos son adaptables a los datos que se dispongan de atributos astronómicos, de valores previos de predicciones meteorológicas proporcionadas por un modelo atmosférico para la misma ubicación de la planta, y a la serie temporal histórica de la potencia eléctrica media horaria obtenida en la planta fotovoltaica en estudio. Además, estos modelos presentan unas características diferentes a otros modelos desarrollados con anterioridad. Se han utilizado como entradas atributos astronómicos, y se han usado técnicas de minería de datos, algunas de las cuales no se han aplicado hasta ahora en este tipo de modelos de predicción. Los resultados obtenidos con los mejores modelos mejoran a los obtenidos con otros modelos publicados en la literatura internacional. Esto hace que los propietarios de las plantas fotovoltaicas puedan presentar mejores ofertas de venta de energía eléctrica en el mercado diario o en las sesiones del mercado intradiario y así poder incrementar el beneficio económico en la planta fotovoltaica.
7. Hasta este momento son raros los modelos que predicen hasta el tercer día. En los trabajos de investigación de esta tesis se han desarrollado modelos que predicen entre 49 y 72 horas, superando así los horizontes de la mayoría de los modelos publicados en la literatura internacional. Las predicciones que se obtienen con estos modelos serían más precisas si se consiguiera realizar modelos atmosféricos que pudieran pronosticar variables meteorológicas con mayor fiabilidad para horizontes de predicción hasta 72 horas.
8. Para días con cielo claro se han desarrollado nuevos modelos originales de predicción de la potencia eléctrica media horaria para horizontes de predicción superiores a 72 horas (podrían ser hasta de años) ya que solo se tiene en cuenta el día del año para

el que se hace la predicción. Éstos utilizan como entradas atributos astronómicos, y no utilizan variables de predicciones meteorológicas pronosticadas por un modelo atmosférico. Sus predicciones son muy buenas a pesar de no tener en cuenta la posible temperatura de forma directa.

9. Para cada horizonte de predicción a corto plazo de 1 a 8 horas, entre 0 y 24 horas, entre 25 y 48 horas, entre 49 y 72 horas, se han probado cuáles son los mejores atributos de entradas para los modelos de predicción de la potencia eléctrica media horaria.

Se ha comprobado que la inclusión como entradas en estos modelos de los atributos astronómicos (el *azimut*, la *altura* o *elevación*, el *ángulo horario*, la *declinación* y el *ángulo de incidencia* que forman los rayos solares directos con el plano que les recibe) calculados en función del día del año para el que se realiza la predicción, consiguen mejorar la calidad de la precisión de las predicciones, reduciendo el valor del RMSE con respecto a otros modelos de predicción que nos los utilizan.

10. Se ha hecho un estudio comparativo con los datos disponibles de una misma planta fotovoltaica, con el fin de saber cuáles son las técnicas de minería de datos (árboles de decisión, sistemas de reglas, redes neuronales artificiales, máquinas de vectores soporte, y sistemas neuro-difusos) más apropiadas para utilizar en el desarrollo de los modelos de predicción de la potencia eléctrica media horaria generada en la misma. Para modelos de cualquier tipo de día, se ha logrado determinar en función del horizonte de predicción de 1 a 8 horas, entre 0 y 24 horas, entre 25 y 48 horas, y entre 49 y 72 horas, el modelo que puede lograr las mejores predicciones, y por tanto el más conveniente a utilizar.
11. Se ha determinado que en los modelos a muy corto plazo (unas pocas horas) no se necesita utilizar como atributos predicciones meteorológicas, sin embargo en los modelos con mayor horizonte de predicción si que son necesarias para reducir los errores de predicción. Hasta el momento no hay ningún estudio publicado en la literatura internacional que determine a partir de que horizonte de predicción mejoran los modelos con el uso de las predicciones meteorológicas. En los trabajos de investigación de esta tesis se ha hecho un estudio para determinar a partir de que horizonte de predicción mejora el modelo con el uso de las predicciones meteorológicas.

En los resultados que se han obtenido para cualquier tipo de día con horizontes de predicción de 1 a 8 horas, se observa que los modelos seleccionados de predicción de la potencia eléctrica media horaria obtienen un menor valor del RMSE si no utilizan predicciones de variables meteorológicas hasta el horizonte de predicción de 5 horas. Por lo que se concluye que a partir del horizonte de predicción de 6 horas es mejor utilizar modelos que tengan como atributos de entrada predicciones meteorológicas obtenidas con modelos atmosféricos.

12. Se ha desarrollado una herramienta de predicción con capacidad de proporcionar la incertidumbre asociada a un punto de predicción de potencia eléctrica media horaria esperada, para los horizontes de tiempo entre: 0 y 24, 25 y 48, 49 y 72 horas. Con esta herramienta se dispone de información sobre los intervalos de potencia eléctrica, donde puede surgir el valor real de la potencia eléctrica generada para

dichos horizontes, así como de su probabilidad correspondiente. Esto permite al usuario tomar decisiones con mayor nivel de confianza. Este modelo, presenta un conjunto de características innovadoras, tales como:

- El modelo permite una modelización estocástica, basada en la similitud entre los valores de entrada, en una base de datos de casos históricos. Dicha similitud se centra en las variables meteorológicas previstas por herramientas NWP.
- Se usa una base de datos con un número significativo de casos históricos, para modelar la predicción estocástica, mediante la creación de funciones de probabilidad conjunta.
- La modelización estocástica de transiciones de potencia eléctrica permite la estimación de la incertidumbre en los puntos de predicción. Los resultados de incertidumbre obtenidos del modelo darán las distribuciones de probabilidad asociadas con los valores de los puntos de predicción de potencia eléctrica.
- Los puntos de predicción son calculados a partir de funciones de probabilidad conjunta. También proporciona el mismo tipo de información (valores de potencia de un punto) que otros modelos de pronóstico no estocásticos.
- Debido a que el proceso de minería de datos utilizado por el modelo busca directamente en la base de datos, el modelo no necesita procesos de reentrenamiento.
- Aparentemente parece un modelo pesado a nivel de cálculo, sin embargo este hecho no se verifica, debido al bajo número de variables que toma en consideración, consiguiendo efectuar predicciones en tiempos muy inferiores al que es necesario para funcionar en tiempo real.
- Otra de las ventajas del modelo es su capacidad para actualizar fácilmente la base de datos utilizada por dicho modelo. Tan pronto como los valores nuevos estén disponibles, pueden ser incluidos en la base de datos. Este factor es importante, ya que, cuanto mejor es la base del conocimiento, mejor será el comportamiento del modelo.
- Este modelo nuevo presentado en esta tesis abre perspectivas para dimensiones útiles de información de predicción de potencia eléctrica con respecto a los resultados habituales obtenidos, por lo común, en los modelos de predicción a corto plazo. Además del valor del punto de predicción, proporcionados por los modelos de predicción convencionales, el modelo propuesto logra valores de incertidumbre de los puntos de predicción de potencia. Estos nuevos resultados de predicción permiten la integración del modelo en aplicaciones donde existe un riesgo asociado con el error de predicción. Ese riesgo es evidente en el mercado de la electricidad. En previsión de errores y las consiguientes desviaciones entre lo real y lo ofrecido (pronóstico) en valores de generación de potencia, que pueden dar lugar a sanciones económicas, la información sobre la incertidumbre es muy útil para avanzar en un valor esperado de la penalización, para que el agente de mercado pueda planificar cubrir el riesgo.
- Los resultados de las pruebas realizadas son esperanzadores, ofreciendo resultados de predicción comparables a los obtenidos por otros modelos, descritos

en artículos publicados en revistas internacionales, pero superando a estos, ya que ofrecen información acerca de incertidumbre.

17.3. Futuras líneas de investigación

Se pueden proponer varias líneas de continuación a partir de los trabajos de investigación desarrollados en esta tesis:

1. Optimización automática de los modelos basados en redes neuronales con ajuste del número de neuronas, del factor de aprendizaje, de la selección de atributos entre los disponibles.
2. Desarrollo de un modelo híbrido que combine varios de los modelos desarrollados para ofrecer una única predicción como media ponderada de los obtenidos con cada modelo. Los coeficientes de ponderación podrían escogerse en función de los resultados obtenidos por cada uno de ellos en los últimos días con las mismas características (claro, nublado, etc.).
3. Desarrollo de modelos específicos en función de las características previstas para el horizonte de predicción (cielo claro, nublado, lluvioso, etc) y de sus horas precedentes.
4. Desarrollo de nuevos modelos que consideren el grado medio horario de suciedad (limpio, bajo, medio, alto) en las placas fotovoltaicas, estudiando el efecto que pueda tener el utilizar como entradas los valores pronosticados de variables meteorológicas, tanto en las horas precedentes como para el momento de predicción, como la precipitación, la velocidad del viento, la humedad relativa, etc.
5. Analizar nuevos modelos de predicción que tengan en cuenta el efecto que produce en la producción de la potencia eléctrica media horaria, el envejecimiento de las placas fotovoltaicas.
6. La ampliación del mecanismo MDM en el modelo HISIMI, como:
 - Incorporación de nuevas funciones de búsqueda, además de la gaussiana. Pueden añadirse funciones como: la triangular, sigmoidea, trapezoidal, etc. Para esto es necesario cambiar la estructura del tercer gen del cromosoma utilizado.
 - Utilización de un “factor de olvido”, dado que los errores de cualquier sistema de predicción se incrementan conforme aumenta el horizonte de predicción, las predicciones obtenidas con la herramienta NWP son tanto peores cuanto mayor sea el horizonte. La idea es dar menos peso a los datos de la base de conocimiento con un mayor horizonte de predicción (correspondiente a las previsiones de las variables meteorológicas) multiplicando la probabilidad resultante por un factor menor a 1, y tanto menor cuanto mayor sea el horizonte de predicción de dichos valores meteorológicos.
7. Una posible continuación del procedimiento de optimización en el modelo HISIMI:

- Aumento de la precisión de los datos numéricos utilizando más bits en la representación de cada parámetro o gen.
- Utilización combinada del algoritmo genético con otras estrategias de optimización de tipo evolutivo, como la optimización de enjambre de partículas. Se podrían desarrollar dos poblaciones distintas, evolucionando una de ellas con el algoritmo genético y la otra con la técnica de partículas. A partir de un número limitado de generaciones podrían converger en una única población evolucionando con una combinación de ambas técnicas.
- Incorporación del factor de olvido comentado anteriormente, como un gen más en el cromosoma que define el modelo de predicción.

8. Una mayor experimentación con nuevos datos, como:

- Reelaboración de los datos de predicciones meteorológicas con nuevas herramientas NWP. Los datos utilizados en las pruebas de investigación de esta tesis, fueron elaborados en ordenadores de la Universidad de La Rioja, con el modelo NWP de meso escala WRF. Estos nuevos modelos permitirán reducir los errores en la predicción de las variables meteorológicas, e incrementar así la precisión de los modelos de predicción de la generación eléctrica media horaria generada en una planta fotovoltaica, desarrollados en esta tesis.
- Crear nuevos modelos para horizontes de predicción mayores de 72 horas.

9. Otras líneas a experimentar serían:

- Una inteligente combinación de las diversas técnicas usadas en los trabajos de investigación de esta tesis, cada una de las cuales debe ser aplicada a aquellas horas del día que mejor se adecúen. Se pueden combinar estas técnicas con otras más clásicas como estadísticas, tratamiento de señal, wavelets, etc. Así se podrían conseguir resultados todavía mejores.

10. Otra posible línea podría ser:

- Analizar la posibilidad de generar predicciones probabilísticas en forma de *cuantiles* obtenidos a partir de las predicciones puntuales obtenidas por el conjunto de modelos (basados en las diferentes técnicas) desarrollados para cada uno de los horizontes. Este tipo de modelos apenas se han desarrollado para el campo de la predicción en plantas fotovoltaicas.

11. Las últimas líneas se podrían basar en:

- Probar los mismos modelos con datos de otras plantas de generación fotovoltaica, situadas en diferentes lugares de la Tierra, para poder comprobar la generalización de los modelos desarrollados.

- Extender los modelos a otro tipo de plantas generadoras basadas en recursos renovables tales como centrales solares de concentración o centrales micro-hidráulicas. La posibilidad de realización de pruebas en este sentido está obviamente condicionada a la disponibilidad de datos de generación de centrales de este tipo.

En definitiva, a partir de la presente tesis se abre un amplio abanico de posibilidades de continuación en este campo tan interesante en lo científico y en lo económico.

Apéndice A

Artículos científicos

A continuación, se hace referencia a los artículos científicos que parte de los trabajos de investigación de esta tesis, han sido publicados en la literatura internacional.

- L. A. Fernández-Jiménez, M. S. Terreros-Olarte, A. Falces, P. Zorzano-Santamaría, M. Mendoza-Villena, E. Zorzano-Alba, P. Lara-Santillán, E. García-Garrido, “Neuro-fuzzy short-term forecasting model for PV plants optimized with genetic algorithm”, *12 th WSEAS/IASME International Conference on Electric Power Systems, High Voltajes, Electric Machines (POWER'12)*, Praga, 2012.
- C. Monteiro, T. Santos, L. A. Fernández-Jiménez, I. J. Ramirez-Rosado, M. S. Terreros-Olarte, “Short-Term Power Forecasting Model for Photovoltaic Plants Based on Historical Similarity”, *Energies*, Vol. 6, 2624-2643, 2013.

Bibliografía

- [Aizerman 64] M. A. Aizerman, É. M. Braverman, L. I. Rozonoér, “Theoretical foundations of the potencial function method in pattern recognition learning”, *Automation and Remote Control*, Vol. 25, 821–837, 1964.
- [Alam 06] S. Alam, S. C. Kaushik, S. N. Garg, “Computation of beam solar radiation at normal incidence using artificial neural network”, *Renewable Energy*, Vol. 31, 1483–1491, 2006.
- [Alawi 98] S. M. Alawi, H. A. Hinai, “An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation”, *Renewable Energy*, Vol. 14, no. 1–4, 199–204, 1998.
- [Alessandrini 15] S. Alessandrini, L. D. Monache, S. Sperati, G. Cervone, “An analog ensemble for short-term probabilistic solar power forecast”, *Applied Energy*, Vol. 157, 95–110, 2015.
- [Al-Messabi 14] N. Al-Messabi, C. Goh, I. El-Amin, Y. Li, “Heuristically enhanced dynamic neural networks for structurally improving photovoltaic power forecasting”, *Proceedings of the International Joint Conference on Neural Networks*, 6889827, 2820–2825, 2014.
- [Anastasiades 13] G. Anastasiades, P. McSharry, “Quantile forecasting of wind power using variability indices”, *Energies*, Vol. 6, no. 2, 662–695, 2013.
- [Arakawa 77] A. Arakawa, V. Lamb, “Computational design of the basic dynamical processes in the UCLA general circulation model”, en J. Chang (Editor), “General circulation models of the atmosphere, methods in computational physics”, Academic press, New York, 17, 174–264, 1977.
- [Araujo 09] L. Araujo, C. Cervigón, “Algoritmos Evolutivos. Un enfoque práctico”, Ra-Ma, Madrid, 2009.
- [Armstrong 01] J. S. Armstrong (Editor), “Principles of forecasting: a handbook for researches and practitioners”, Kluwer Academic Publishers, Boston, 2001.
- [Ashrae 05] ASHRAE, “Handbook fundamentals”, Atlanta, 2005.

- [Atwater 79] M. A. Atwater, J. T. Ball, “Erratum”, *Solar Energy*, Vol. 23, 275, 1979.
- [Bacher 09] P. Bacher, H. Madsen, H. A. Nielsen, “Online short-term solar power forecasting”, *Solar Energy*, Vol. 83, no. 10, 1772–1783, 2009.
- [Bäck 91] T. Bäck, F. Hoffmeister, H. P. Schwefel, “A survey of evolution strategies”, Proc. of the Fourth International Conference on Genetic Algorithms, Morgan Kaufman, 2–9, 1991.
- [Badescu 02] V. Badescu, “3D isotropic approximation for solar diffuse irradiance on tilted surfaces”, *Renewable Energy*, Vol. 26, 221–223, 2002.
- [Badescu 08] V. Badescu, “Use of sunshine number for solar irradiance time series generation”, en V. Badescu (Editor), “Modeling solar radiation at the Earth surface”, Springer, Berlin, 327–355, 2008.
- [Badescu 12] V. Badescu, C. A. Gueymard, S. Cheval, C. Oprea, M. Baciú, A. Dumitrescu, F. Iacobescu, I. Milos, C. Rada, “Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models”, *Renewable and Sustainable Energy Reviews*, Vol. 16, 1636–1656, 2012.
- [Barbaro 79] S. Barbaro, S. Coppolino, C. Leone, E. Sinagra, “An atmospheric model for computing direct and diffuse solar radiation”, *Solar Energy*, Vol. 22, 225–228, 1979.
- [Bashahu 94] M. Bashahu, D. Laplace, “An atmospheric model for computing solar radiation”, *Renewable Energy*, Vol. 4, 455–458, 1994.
- [Belcher 07] B. N. Belcher, A. T. DeGaetano, “A revised empirical model to estimate solar radiation using automated surface weather observations”, *Solar Energy*, Vol. 81, 329–345, 2007.
- [Bellini 09] A. Bellini, S. Bifaretti, V. Iacovone, C. Cornaro, “Simplified model of a photovoltaic module”, *Applied Electronics*, 2009.
- [Bennett 92] K. P. Bennett, O. L. Mangasarian, “Robust linear programming discrimination of two linearly inseparable sets”, *Optimization Methods and Software*, Vol. 1, 23–34, 1992.
- [Benoit 97] R. Benoit, M. Desgagné, P. Pellerin, S. Pellerin, Y. Chartier, S. Desjardins, “The Canadian MC2: A semi-Lagrangian, semi-implicit wideband atmospheric model suited for fine scale process studies and simulation”, *Monthly Weather Review*, Vol. 125, 2382–2415, 1997.
- [Berberian 61] S. K. Berberian, “Introduction to Hilbert Space”, Oxford University Press, 1961.

- [Bergman 85] M. J. Bergman, J. W. Delleur, “Kalman filter estimation and prediction of daily streamflow: 1- review, algorithm, and simulation experiments; 2-application to the Potomac river”, *Water Resources Bulletin*, Vol. 21, no. 5, 815–832, 1985.
- [Biga 79] A. J. Biga, R. Rosa, “Contribution to the study of the solar radiation climate of Lisbon” *Solar Energy*, Vol. 23, 61–67, 1979.
- [Bird 80] R. E. Bird, R. L. Hulstrom, “Direct insolation models”, Solar Energy Research Institute, SERI/TR-335–344, 1980.
- [Bird 81] R. E. Bird, R. L. Hulstrom, “A simplified clear sky model for direct and diffuse insolation on horizontal surfaces”, Solar Energy Research Institute, SERI/TR-642–761, 1981.
- [Boser 92] B. E. Boser, I. M. Guyon, V. N. Vapnik, “A training algorithm for optimal margin classifiers”, en D. Haussler (Editor), *Proceedings of the Annual Conference on Computational Learning Theory*, ACM Press, Pittsburgh, 144–152, 1992.
- [Bouzerdoun 13] M. Bouzerdoun, A. Mellit, A. Massi Pavan, “A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant”, *Solar Energy*, Vol. 98, 226–235, 2013.
- [Bracale 13] A. Bracale, P. Caramia, G. Carpinelli, A. R. Di Fazio, G. A. Ferruzzi, “A Bayesian method for Short-Term probabilistic forecasting of photovoltaic generation in smart grid operation and control”, *Energies*, Vol. 6, no. 2, 733–747, 2013.
- [Breiman 84] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, “Classification and Regression Trees”, Wadsworth and Brooks/Cole, Monterey, 1984.
- [Bugler 77] J. W. Bugler, “The determination of hourly insolation on an inclined plane using a diffuse irradiance model based on hourly measured global horizontal insolation”, *Solar Energy*, Vol. 19, no. 6, 477–491, 1977.
- [Bühlmann 06] P. Bühlmann, “Boosting for high-dimensional linear models”, *Annals of Statistics*, Vol. 34, no. 2, 559–583, 2006.
- [Cao 05] S. Cao , J. Cao, “Forecast of solar irradiance using recurrent neural networks combined with wavelet analysis”, *Applied Thermal Engineering*, Vol. 25, 161–172, 2005.
- [Cao 06] J. C. Cao, S. H. Cao, “Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis”, *Energy*, Vol. 3, 13435–13445, 2006.

- [Carney 06] M. Carney, P. Cunningham, “Evaluating Density Forecasting Models”, Trinity College Dublin, Department of Computer Science, Dublin, TCD-CS-2006-21, 2006.
- [Caro 15] E. Caro, F. J. Cara, J. Juan, “Forecasting photovoltaic energy using MEWMA models”, International Conference on the European Energy Market, EEM, 7216655, 2015.
- [Carroll 85] J. J. Carroll, “Global transmissivity and diffuse fraction of solar radiation for clear and cloudy skies as measured and as predicted by bulk transmissivity models”, *Solar Energy*, Vol. 35, 105–118, 1985.
- [Cestnik 87] G. Cestnik, I. Kononenko, I. Bratko, “Assistant 86: A Knowledge Acquisition Tool for Sophisticated Users”, en I. Bratko, N. Lavrac (Editores), “Progress in Machine Learning”, Sigma Press, 1987.
- [Chandra 78] M. Chandra, “Dependence of solar radiation availability on atmospheric turbidity”, Proceedings ISES Conference, 430–434, 1978.
- [Charalambous 92] C. Charalambous, “Conjugate gradient algorithm for efficient training of artificial neural networks”, *IEEE Proceedings*, Vol. 139, no. 3, 301–310, 1992.
- [Chiron 03] A. Chiron de la Casinière, “Le Rayonnement solaire dans l’environnement terrestre”, Editions Publibook, Paris, 2003.
- [Chiron 08] A. Chiron de la Casinière, V. E. Cachorro, “La radiación solar en el sistema tierra-atmósfera”, Universidad de Valladolid, Valladolid, 2008.
- [Chiu 94] S. Chiu, “Fuzzy model identification based on cluster estimation”, *Journal of Intelligent and Fuzzy Systems*, Vol. 2, 267–278, 1994.
- [Chu 15] Y. Chu, B. Urquhart, S. M. I. Gohari, H. T. C. Pedro, J. Kleissl, C. F. M. Coimbra, “Short-term reforecasting of power output from a 48 MWe solar PV plant”, *Solar Energy*, Vol. 112, 68–77, 2015.
- [Clark 89] P. Clark, T. Niblett, “The CN2 induction algorithm”, *Machine Learning*, Vol. 3, no. 4, 261–283, 1989.
- [Clark 91] P. Clark, R. Boswell, “Rule induction with CN2: Some recent improvements”, Proceedings of the Fifth European Working Session on Learning, Springer, Berlin, 151–163, 1991.
- [Cohen 93] W. Cohen, “Efficient pruning methods for separate-and-conquer rule learning systems”, Proceedings of the 13th International Joint Conference on Artificial Intelligence, Morgan Kaufmann, 988–994, 1993.
- [Cuadras 88] C. M. Cuadras, B. Echeverría, J. Mateo, P. Sánchez, “Fundamentos de Estadística. Aplicación a las Ciencias Humanas”, Promociones Publicaciones Universitarias, Barcelona, 1988.

- [Daley 99] R. Daley, “Atmospheric data analysis”, Cambridge University Press, Cambridge, 1999.
- [Davies 79] J. A. Davies, T. C. Uboegbulam, “Parameterization of surface incoming radiation in tropical cloudy conditions”, *Atmosphere-Ocean*, Vol. 17, 14–23, 1979.
- [De Felice 15] M. De Felice, M. Petitta, P. M. Ruti, “Short-term predictability of photovoltaic production over Italy”, *Renewable Energy*, Vol. 80, 197–204, 2015.
- [De Giorgi 16] M. G. De Giorgi, M. Malvoni, P. M. Congedo, “Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine”, *Energy*, Vol. 107, 360–373, 2016.
- [De Lurgio 98] S. A. De Lurgio, “Forecasting principles and applications”, McGraw-Hill, Singapur, 1998.
- [Do 16] M. T. Do, T. Soubdhan, B. Robyns, “A study on the minimum duration of training data to provide a high accuracy forecast for PV generation between two different climatic zones”, *Renewable Energy*, Vol. 85, 959–964, 2016.
- [Dogniaux 76] B. Dogniaux, “Programme de calcul des eclairements solaires energetiques et lumineux de surfaces orientees et inclinees”, Rep. 14, IRM, Brussels, 1976.
- [Dorvio 02] A. S. S. Dorvio, J. A. Jervase, A. Al-Lawati, “Solar radiation estimation using artificial neural networks”, *Applied Energy*, Vol. 74, 307–319, 2002.
- [Dudhia 96] J. Dudhia, “A multi-layer soil temperature model for MM5”, Preprints, “The Sixth PSU/NCAR Mesoscale Model Users’ Workshop”, Boulder, Colorado, 49–50, 1996.
- [Elizondo 94] D. Elizondo, G. Hoogenboom, R. W. McClendon, “Development of a neural network model to predict daily solar radiation”, *Agricultural Forest Meteorology*, Vol. 71, 115–132, 1994.
- [Elman 90] J. L. Elman, “Finding structure in time”, *Cognitive Science*, Vol. 14, 179–211, 1990.
- [Elminir 07] H. K. Elminir, Y. A. Azzam, F. I. Younes, “Prediction of hourly and daily diffuse fraction using neural network, as compared to linear regression models”, *Energy*, Vol. 32, 1513–1523, 2007.
- [Eshelman 89] L. J. Eshelman, R. Caruana, J. D. Schaffer, “Biases in the crossover landscape”, Proc. of the 3rd International Conference on Genetic Algorithms, Morgan Kaufmann Publishers Inc., 10–19, 1989.

- [Esposito 97] F. Esposito, D. Malerba, G. Semerano, “A Comparative Analysis of Methods for Pruning Decision Trees”, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 19, no. 5, 476–491, 1997.
- [Esra 10] ESRA, “Helioclim”, *Lectures*, (<http://www.helioclim.net/publications/index.html#esra>), 2010.
- [Fernández 08] L. A. Fernández-Jiménez, “Modelos avanzados para la predicción a corto plazo de la producción eléctrica en parques eólicos”, Tesis doctoral, Universidad de La Rioja, Logroño, 2008.
- [Fernández 12a] L. A. Fernández-Jiménez, A. Muñoz-Jiménez, A. Falces, M. Mendoza-Villena, E. García-Garrido, P. M. Lara-Santillán, E. Zorzano-Alba, P. J. Zorzano-Santamaría, “Short-term power forecasting system for photovoltaic plants”, *Renewable Energy*, Vol. 44, 311–317, 2012.
- [Fernández 12b] L. A. Fernández-Jiménez, M. S. Terreros-Olarte, A. Falces, P. Zorzano-Santamaría, M. Mendoza-Villena, E. Zorzano-Alba, P. Lara-Santillán, E. García-Garrido, “Neuro-fuzzy short-term forecasting model for PV plants optimized with genetic algorithm”, 12th WSEAS/IASME International Conference on Electric Power Systems, High Voltages, Electric Machines (POWER’12), Praga, 2012.
- [Filipe 15] J. M. Filipe, R. J. Bessa, J. Sumaili, R. Tomé, J. N. Sousa, “A hybrid short-term solar power forecasting tool”, 18th International Conference on Intelligent System Application to Power Systems, ISAP 2015, 7325543, 2015.
- [Fletcher 89] R. Fletcher, “Practical Methods of Optimization”, John Wiley and Sons, New York, 1989.
- [Fonseca 12] J. G. Da Silva Fonseca, T. Oozeki, T. Takashima, G. Koshimizu, Y. Uchida, K. Ogimoto, “Use of support vector regression and numerically predicted cloudiness to forecast power output of a photovoltaic power plant in Kitakyushu, Japan”, *Progress in Photovoltaics: Research and Applications*, Vol. 20, no. 7, 874–882, 2012.
- [Frank 98] E. Frank, I. H. Witten, “Generating accurate rule sets without global optimization”, Proc. of the Fifteenth International Conference on Machine Learning, Morgan Kaufmann, 144–151, 1998.
- [Freitas 02] A. A. Freitas, “Data Mining and Knowledge Discovery with Evolutionary Algorithms”, “Natural Computing Series”, Springer-Verlag, Berlin, 2002.
- [Fuentes 00] J. L. Fuentes, “Iniciación a la Meteorología y la Climatología”, Mundi-Prensa, Madrid, 2000.
- [Funahashi 89] K. I. Funahashi, “On the approximate realization of continuous mappings by neural networks”, *Neural Networks*, Vol. 2, 183–192, 1989.

- [Gabriel 07] L. Gabriel, A. G. Christian, “Clear-sky solar luminous efficacy determination using artificial neural networks”, *Solar Energy*, Vol. 81, no. 7, 929–939, 2007.
- [García-Moya 02] J. A. García-Moya, “Los modelos numéricos de predicción del tiempo”, *TETHYS*, Vol. 2, no. 5, en <<http://www.tethys.cat/antics/num02/articles/art0205esp.htm>> [Consulta: 1 de mayo 2014], 2002.
- [Goldberg 89] D. E. Goldberg, “Genetic Algorithms in Search, Optimization and Machine Learning”, Addison-Wesley, 1989.
- [Golestaneh 15] F. Golestaneh, P. Pinson, H. B. Gooi, “Very Short-Term Nonparametric Probabilistic Forecasting of Renewable Energy Generation—With Application to Solar Energy”, *IEEE Transactions on Power Systems*, Vol. 31, 3850–3863, 2015.
- [Gomez 03] V. Gomez, A. Casanovas, “Fuzzy modeling of solar irradiance on inclined surfaces”, *Solar Energy*, Vol. 75, 307–315, 2003.
- [GOS] World Meteorological Organization. Global Observing System (GOS) en <http://www.wmo.int/pages/prog/www/OSY/GOS.html> [Consulta: 22 de Junio 2014]
- [Guessoum 98] A. Guessoum, S. Boubkeur, A. Maafi, “A global irradiation model using radial basis function neural network”, WREC, UK, 332–336, 1998.
- [Gueymard 84] C. Gueymard, “Physical modelling of the diffuse irradiance on tilted planes as a function of the aerosols anisotropic effect”, Meteorology and renewable energies conference, AFME, Valbonne, France, 303–314, 1984.
- [Gueymard 86] C. Gueymard, “Radiation on tilted planes: A physical model adaptable to any computational time-step”, Proceedings of INTERSOL85, Pergamon Press, Elmsford, NY, 2463–2467, 1986.
- [Gueymard 89] C. Gueymard, “A two-band model for the calculation of clear sky solar irradiance, illuminance, and photosynthetically active radiation at the earth’s surface”, *Solar Energy*, Vol. 43, 253–265, 1989.
- [Gueymard 04] C. A. Gueymard, “Direct solar transmittance and irradiance predictions with broadband models. Part 2: validation with high-quality measurements”, *Solar Energy*, Vol. 74, 381–395, 2003. Corrigendum: *Solar Energy*, Vol. 76, 515, 2004.
- [Gueymard 08] C. A. Gueymard, “REST2: high performance solar radiation model for cloudless-sky irradiance, illuminance and photosynthetically active radiation-validation with a benchmark dataset”, *Solar Energy*, Vol. 82, 272–285, 2008.

- [Gutiérrez 04] J. M. Gutiérrez, R. Cano, A. S. Cofiño, C. M. Sordo, “Redes Probabilísticas y Neuronales en las Ciencias Atmosféricas”, Centro de Publicaciones Secretaria General Técnica Ministerio de Medio Ambiente, Santander, 2004.
- [Hamilton 85] H. L. Hamilton, A. Jackson, “A shield for obtaining diffuse sky radiation from portions of the sky”, *Solar Energy*, Vol. 34, no. 1, 121–123, 1985.
- [Hammer 99] A. Hammer, D. Heinemann, E. Lorenz, B. Lückehe, “Short-term forecasting of solar radiation: A statistical approach using satellite data”, *Solar Energy*, Vol. 67, no. 1–3, 139–150, 1999.
- [Haque 13] A. U. Haque, M. H. Nehrir, P. Mandal, “Solar PV power generation forecast using a hybrid intelligent approach”, IEEE Power and Energy Society General Meeting, 6672634, 2013.
- [Hay 85] J. E. Hay, D. C. McKay, “Estimating solar irradiance on inclined surfaces: a review and assessment of methodologies”, *International Journal of Solar Energy*, Vol. 3, 203–240, 1985.
- [Haykin 99] S. Haykin, “Neural Networks. A Comprehensive Foundation”, Prentice-Hall, 1999.
- [Hecht-Nielsen 90] R. Hecht-Nielsen, “Neurocomputing”, Addison Wesley, 1990.
- [Hernández 04] J. Hernández, M. J. Ramírez, C. Ferri, “Introducción a la Minería de Datos”, Pearson Prentice Hall, Madrid, 2004.
- [Herrera 98] F. Herrera, M. Lozano, J. L. Verdegay, “Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis”, *Artificial Intelligence Review*, Vol. 12, no. 4, 265–319, 1998.
- [Heuklon 79] T. K. van Heuklon, “Estimating atmospheric ozone for solar radiation models”, *Solar Energy*, Vol. 22, 63–68, 1979.
- [Hocaoğlu 08] F. O. Hocaoğlu, Ö. N. Gerek, M. Kurban, “Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks”, *Solar Energy*, Vol. 82, no. 8, 714–726, 2008.
- [Hodur 97] R. M. Hodur, “The Naval Research Laboratory’s Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS)”, *Monthly Weather Review*, Vol. 125, 1414–1430, 1997.
- [Holland 75] J. H. Holland, “Adaptation in Natural and Artificial Systems”, University of Michigan Press, 1975.
- [Holmes 99] G. Holmes, M. Hall, E. Frank, “Generating Rule Sets from Model Trees”, Twelfth Australian Joint Conference on Artificial Intelligence, 1–12, 1999.

- [Hontoria 99] L. Hontoria, J. Riesco, P. Zufiria, J. Aguilera, “Improved generation of hourly solar radiation artificial series using neural networks”, EANN99, Warsaw, 1999.
- [Hontoria 00] L. Hontoria, J. Riesco, P. Zufiria, J. Aguilera, “Application of neural networks in the solar radiation field. Obtainment of solar radiation maps”, en “16th European photovoltaic for chemical engineers”, Elsevier, Amsterdam, Vol. 3, 385–408, 2000.
- [Hontoria 01] L. Hontoria, J. Aguilera, J. Riesco, P. Zufiria, “Recurrent neural supervised models for generating solar radiation synthetic series”, *Journal of Intelligent and Robotic Systems*, Vol. 31, 201–221, 2001.
- [Hontoria 02] L. Hontoria, J. Aguilera, P. Zufiria, “Generation of hourly irradiation synthetic series using the neural network multilayer Perceptron”, *Solar Energy*, Vol. 72, no. 5, 441–446, 2002.
- [Hontoria 05] L. Hontoria, J. Aguilera, P. Zufiria, “An application of the multilayer perceptron: solar radiation maps in Spain”, *Solar Energy*, Vol. 79, 523–530, 2005.
- [Hornik 89] K. Hornik, M. Stichcombe, H. White, “Multilayer feedforward networks are universal approximators”, *Neural Networks*, Vol. 2, 359–366, 1989.
- [Hottel 76] H. C. Hottel, “A simple model for estimating the transmittance of direct solar radiation through clear atmospheres”, *Solar Energy*, Vol. 18, 129–134, 1976.
- [Huang 10] Y. Huang, J. Lu, C. Liu, X. Xu, W. Wang, X. Zhou, “Comparative study of power forecasting methods for PV stations”, 2010 International Conference on Power System Technology: Technological Innovations Making Power Grid Smarter, POWERCON2010, 5666688, 2010.
- [Ideriah 81] F. J. K. Ideriah, “A model for calculating direct and diffuse solar radiation”, *Solar Energy*, Vol. 26, 447–452, 1981.
- [Ineichen 08] P. Ineichen, “A broadband simplified version of the Solis clear sky model”, *Solar Energy*, Vol. 82, 758–762, 2008.
- [Iqbal 83] M. Iqbal, “An introduction to solar radiation”, Academic Press, Canada, 1983.
- [Iqdour 05] R. Iqdour, A. Zeroual, “Prediction of daily global solar radiation using fuzzy systems”, *International Journal Sustainable Energy*, Vol. 26, no. 1, 19–29, 2005.
- [Jang 91] J. S. R. Jang, “Fuzzy modeling using generalized neural networks and Kalman filter algorithm”, Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI-91), 762–767, 1991.

- [Jang 93] J. S. R. Jang, “ANFIS: Adaptive-Network-based Fuzzy Inference Systems”, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 23, no. 3, 665–685, 1993.
- [Jang 97] J. S. R. Jang, C. T. Sun, E. Mizutani, “Neuro-Fuzzy and Soft Computing”, Prentice Hall, Upper Saddle River, 1997.
- [Janjai 10] S. Janjai, “Models for estimating direct normal and diffuse irradiances clear sky condition”, Technical report, Universidad de Nakhon Pathom, Tailandia, 2010.
- [Janjic 14] Z. Janjic, et al., “User’s Guide for the NMM Core of the Weather Research and Forecast (WRF) Modeling System Version 3”, Developmental Testbed Center / National Centers for Environmental Prediction, 2014.
- [Jimenez 86] J. I. Jimenez, Y. Castro, “National assembly of geophysics and geodesy”, Vol. 2, 805, 1986.
- [Jolliffe 02] I. T. Jolliffe, “Principal component analysis”, Springer Verlag, Nueva York, 2002.
- [Jolliffe 03] I. Jolliffe, D. Stephenson, “Forecast Verification: A Practitioner’s Guide in Atmospheric Science”, Wiley and Sons, Chichester, 2003.
- [Jorba 05] O. Jorba i Casellas, “Simulación de los campos de viento de la Península Ibérica y el área geográfica de Catalunya con alta resolución espacial para distintas situaciones meteorológicas típicas”, Tesis doctoral, Universidad Politécnica de Cataluña, 2005.
- [Josefsson 85] W. Josefsson, “Modeling direct and global radiation from hourly synoptic observations”, unpublished manuscript, 1985.
- [Jutglar 12] L. Jutglar, “Generación de energía solar fotovoltaica”, Marcombo, Barcelona, 2012.
- [Kalogirou 01] S. A. Kalogirou, “Artificial neural networks in renewable energy systems applications: a review”, *Renewable and Sustainable Energy Review*, 373–401, 2001.
- [Kalogirou 02] S. Kalogirou, S. C. Michaelides, F. S. Tymvios, “Prediction of maximum solar radiation using artificial neural networks”, World Renewable Energy Congress VII (WREC 2002), 2002.
- [Kambezidis 08] H. Kambezidis, B. E. Psiloglou, “The Meteorological Radiation Model (MRM): advancements and applications”, en V. Badescu (Editor), “Modeling solar radiation at the Earth’s surface. Recent advances”, Springer, Berlin, 357–392, 2008.

- [Kaplan 82] M. L. Kaplan, J. W. Zack, V. C. Wong, J. J. Tuccillo, “Initial results from mesoscale atmospheric simulation system and comparisons with an AVE-SESAME I data set”, *Monthly Weather Review*, Vol. 110, 1564–1590, 1982.
- [Kaplanis 07] S. Kaplanis, E. Kaplani, “A model to predict expected mean and stochastic hourly global solar radiation $I(h;nj)$ values”, *Renewable Energy*, Vol. 32, no. 8, 1414–1425, 2007.
- [Kaplanis 10] S. Kaplanis, E. Kaplani, “Stochastic prediction of hourly global solar radiation for Patra, Greece”, *Applied Energy*, Vol. 87, no. 12, 3748–3758, 2010.
- [Karavana 13] K. Karavana-Papadimou, B. E. Psiloglou, S. Lykoudis, H. D. Kambezidis, “Model for estimating atmospheric ozone content over Europe for use in solar radiation algorithms”, *Global NEST Journal*, Vol. 15, no. 2, 152–162, 2013.
- [Kardakos 13] E. G. Kardakos, M. C. Alexiadis, S. I. Vagropoulos, C. K. Simoglou, P. N. Biskas, A. G. Bakirtzis, “Application of time series and artificial neural network models in short-term forecasting of PV power generation”, Proceedings of the Universities Power Engineering Conference, 6714975, 2013.
- [Karush 39] W. Karush, “Minima of functions of several variables with inequalities as side constraints”, Master’s thesis, Dept. of Mathematics, University of Chicago, 1939.
- [Kasten 83] F. Kasten, “Parametrisierung der Globalstrahlung durch Bedeckungsgrad und Trübungsfaktor”, *Annalen der Meteorologie*, Vol. 20, 49–50, 1983.
- [Kearns 96] M. Kearns, Y. Mansour, “On the boosting ability of top-down decision tree learning algorithms”, *Journal of Computer and Systems Sciences*, Vol. 58, no. 1, 109–128, 1999. Also in Proceedings ACM Symposium on the Theory of Computing, ACM Press, 459–468, 1996.
- [Keerthi 01] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, K. R. K. Murty, “Improvements to Platt’s SMO algorithm for SVM classifier design”, *Neural Computation*, Vol. 13, 637–649, 2001.
- [Kemmoku 99] Y. Kemmoku, S. Orita, S. Nakagawa, T. Sakakibara, “Daily insolation forecasting using a multi-stage neural network”, *Solar Energy*, Vol. 66, no. 3, 193–199, 1999.
- [King 79] R. King, R. O. Buckius, “Direct solar transmittance for a clear sky”, *Solar Energy*, Vol. 22, 297–301, 1979.
- [Klucher 79] T. M. Klucher, “Evaluation of models to predict insolation on tilted surfaces”, *Solar Energy*, Vol. 23, no. 2, 111–114, 1979.

- [Kohonen 82a] T. Kohonen, “Self-organized formation of topologically correct feature maps”, *Biological Cybernetics*, Vol. 43, 59–69, 1982.
- [Kohonen 82b] T. Kohonen, “Analysis of a simple self-organizing process”, *Biological Cybernetics*, Vol. 44, 135–140, 1982.
- [Kohonen 89] T. Kohonen, “Self-Organization and Associative Memory”, 3th edition, Springer-Verlag, 1989.
- [Koronakis 86] P. S. Koronakis, “On the choice of the angle of tilt for south facing solar collectors in the Athens basin area”, *Solar Energy*, Vol. 36, no. 3, 217–225, 1986.
- [Koza 92] J. R. Koza, “Genetic Programming: On the Programming of Computers by Means of Natural Selection”, MIT Press, 1992.
- [Koza 94] J. R. Koza, “Genetic Programming II: Automatic Discovery of Reusable Programs”, MIT Press, 1994.
- [Krarti 06] M. Krarti, J. Huang, D. Seo, J. Dark, “Development of solar radiation models for tropical locations, final report for Research Project 1309-RP”, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers Inc., Atlanta, 2006.
- [Kuhn 51] H. W. Kuhn, A. W. Tucker, “Nonlinear programming”, Proc. 2nd Berkeley Symposium on Mathematical Statistics and Probabilistics, University of California Press, Berkeley, 481–492, 1951.
- [Lafore 98] J. P. Lafore, J. Stein, N. Asensio, P. Bougeault, V. Ducrocq, J. Duron, C. Fischer, P. Hérel, P. Mascart, V. Masson, J. P. Pinty, J. L. Redelsperger, E. Richard, J. Vilá-Guerau de Arellano, “The Meso-NH atmospheric simulation system. Part I: adiabatic formulation and control simulations”, *Annales Geophysicae*, Vol. 16, 90–109, 1998.
- [Laprise 97] R. Laprise, C. Caya, G. Bergeron, M. Giguere, “The formulation of the Andre Robert MC2 (Mesoscale Compressible Community) Model”, *Atmosphere-Ocean*, Vol. 35, 195–220, 1997.
- [Larson 16] D. P. Larson, L. Nonnenmacher, C. F. M. Coimbra, “Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest”, *Renewable Energy*, Vol. 91, 11–20, 2016.
- [Ledesma 11] M. Ledesma, “Principios de meteorología y climatología”, Paraninfo, Madrid, 2011.
- [Lezaun 02] M. Lezaun, “Predicciones del Tiempo y Matemáticas”, *Boletín Sociedad Española de Matemática Aplicada*, Vol. 22, no. 3, 59–98, 2002.

- [Li 09] Y. Z. Li, L. He, R. Q. Nie, “Short-term forecast of power generation for grid-connected photovoltaic system based on advanced Grey-Markov chain”, 2009 International Conference on Energy and Environment Technology, ICEET 2009, Vol. 2, 5367173, 275–278, 2009.
- [Li 13] Y. Li, Z. Wang, J. Niu, “Forecast of power generation for grid-connected photovoltaic system based on grey theory and verification Model”, Proceedings of the 2013 International Conference on Intelligent Control and Information Processing, ICICIP 2013, 6568054, 129–133, 2013.
- [Li 14a] Z. Li, Y. Zhou, C. Cheng, Y. Li, K. Lai, “Short term photovoltaic power generation forecasting using RBF neural network”, 26th Chinese Control and Decision Conference, CCDC 2014, 6852641, 2758–2763, 2014.
- [Li 14b] Y. Li, J. Zhang, J. Xiao, Y. Tan, “Short-term prediction of the output power of PV system based on improved grey prediction model”, International Conference on Advanced Mechatronic Systems, ICAMechS, 6911606, 547–551, 2014.
- [Li 15] Y. Li, X. Ren, J. Niu, “Application of rough sets theory in forecast of power generation for grid-connected photovoltaic system”, 27th Chinese Control and Decision Conference (CCDC), 2015.
- [Li 16] Z. Li, S. M. Mahbobur Rahman, R. Vega, B. Dong, “A hierarchical approach using machine learning methods in solar photovoltaic energy production forecasting”, *Energies*, Vol. 9, no. 1, 55, 2016.
- [Lighezzolo 14] R. A. Lighezzolo, “Integración de modelos numéricos de predicción meteorológica en sistemas de alerta temprana a emergencias”, Tesis de maestría, Universidad Nacional de Córdoba, 2014.
- [Lin 91] C. T. Lin, C. S. G. Lee, “Neural-network-based fuzzy logic control and decision system”, *IEEE Transactions on Computers*, Vol. 40, no. 12, 1320–1336, 1991.
- [Lin 12] M. Lin, Z. Wang, Y. Liu, M. Chen, F. Meng, “Power prediction model of grid-connected photovoltaic and power flow analysis”, Asia-Pacific Power and Energy Engineering Conference, APPEEC, 6307195, 2012.
- [Liu 62] B. Y. H. Liu, R. C. Jordan, “Daily insolation on surfaces tilted towards the equator”, *Trans ASHRAE*, 526–541, 1962.
- [Liu 98] H. Liu, H. Motoda (Editores), “Feature Selection for Knowledge Discovery and Data Mining”, Kluwer Academic Publishers, Boston, 1998.
- [Lo Brano 14] V. Lo Brano, G. Ciulla, M. Di Falco, “Artificial neural networks to predict the power output of a PV panel”, *International Journal of Photoenergy*, Vol. 2014, 193083, 2014.

- [Lopez 05] G. López, F. J. Batlles, J. Tovar-Pescador, “Selection of input parameters to model direct solar irradiance by using artificial neural networks”, *Energy*, Vol. 30, 1675–1684, 2005.
- [Lorenzo 06] E. Lorenzo, “Electricidad solar fotovoltaica. Volumen II: Radiación solar y dispositivos fotovoltaicos”, PROGENSA, Sevilla, 2006.
- [Ma 83] C. C. Y. Ma, M. Iqbal, “Statistical comparison of models for estimating solar radiation on inclined surfaces”, *Solar Energy*, Vol. 31, no. 3, 313–317, 1983.
- [Machler 85] M. A. Machler, M. Iqbal, “A modification of the ASHRAE clear sky irradiation model”, *ASHRAE Transactions Part IA*, 106–115, 1985.
- [Madsen 04] H. Madsen, G. Kariniotakis, H. A. Nielsen, T. S. Nielsen, P. Pinson, “A Protocol for Standardizing the Performance Evaluation of Short-Term Wind Power Prediction Models”, *Project ANEMOS*, no. ENK5-CT-2002-00665, 2004.
- [Makridakis 98] S. Makridakis, S. C. Wheelwright, R. J. Hyndman, “Forecasting. Methods and applications”, John Wiley and Sons, Hoboken (NJ), 1998.
- [Mandal 12] P. Mandal, S. T. S. Madhira, A. Ul haque, J. Meng, R. L. Pineda, “Forecasting power output of solar photovoltaic system using wavelet transform and artificial intelligence techniques”, *Procedia Computer Science*, Vol. 12, 332–337, 2012.
- [Mangasarian 69] O. L. Mangasarian, “Nonlinear Programming”, McGraw-Hill, New York, 1969.
- [Mao 13] M. Mao, W. Gong, L. Chang, “Short-term photovoltaic output forecasting model for economic dispatch of power system incorporating large-scale photovoltaic plant”, 2013 IEEE Energy Conversion Congress and Exposition, ECCE 2013, 6647308, 4540–4545, 2013.
- [Martín 01] N. Martín, J. M. Ruiz, “Material Solar Cells”, *Solar Energy*, Vol. 70, 25–38, 2001.
- [Martín 04] F. Martín, J. A. Quirantes, “Observación y identificación de nubes”, en “Unidad didáctica Meteorología y Climatología”, FECYT (Fundación española para la ciencia y la tecnología), 91–141, 2004.
- [Martín 06] B. Martín, A. Sanz, “Redes Neuronales y Sistemas Borrosos”, Ra-Ma, Madrid, 2006.
- [Martín 10] L. Martín, L. F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, M. Cony, “Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning”, *Solar Energy*, Vol. 84, no. 10, 1772–1781, 2010.

- [Maxwell 98] E. L. Maxwell, “METSTAT—The solar radiation model used in the production of the National Solar Radiation Data Base (NSRDB)”, *Solar Energy*, Vol. 62, 263–279, 1998.
- [McClelland 86] J. L. McClelland, D. E. Rumelhart, “Parallel Distributed Processing. Vol 2: Psychological and biological models”, MIT Press, 1986.
- [McCormick 83] G. P. McCormick, “Nonlinear Programming: Theory, Algorithms, and Applications”, John Wiley and Sons, New York, 1983.
- [Mellit 04] A. Mellit, M. Benghanem, A. Hadj Arab, A. Guessoum, “Modeling of global solar radiation data from sunshine duration and temperature using the Radial Basis Function networks”, Proceedings of the IAS-TED Modeling, Identification and Control Conference, Grindelwald, Switzerland, 2004.
- [Mellit 05] A. Mellit, M. Benghanem, M. Bendekhis, “Artificial neural network model for prediction solar radiation data: application for sizing stand-alone photovoltaic power system”, en Proceedings of IEEE Power Engineering Society, General Meeting, USA, Vol. 1, 40–44, 2005.
- [Mellit 06a] A. Mellit, M. Benghanem, S. A. Kalogirou, “An adaptive wavelet-network model for forecasting daily total solar radiation”, *Applied Energy*, Vol. 83, no. 7, 705–22, 2006.
- [Mellit 06b] A. Mellit, “Artificial neural network with discrete wavelet transforms for time series prediction: application for solar radiation data”, Proceedings of European modeling and simulation symposium, Greece, 2006.
- [Mellit 07] A. Mellit, A. Hadj Arab, N. Khorissi, H. Salhi, “An ANFIS-based forecasting for solar radiation data from sunshine duration and ambient temperature”, IEEE Power Engineering Society General Meeting, Florida, USA, 1–6, 2007.
- [Mellit 08] A. Mellit, S. A. Kalogirou, “Artificial intelligence techniques for photovoltaic applications: A review”, *Progress in Energy and Combustion Science*, Vol. 34, no. 5, 574–632, 2008.
- [Mellit 10] A. Mellit, H. Eleuch, M. Benghanem, C. Elaoun, A. M. Pavan, “An adaptive model for predicting of global, direct and diffuse hourly solar irradiance”, *Energy Conversion and Management*, Vol. 51, no. 4, 771–782, 2010.
- [Mellit 14] A. Mellit, A. Massi Pavan, V. Lughi, “Short-term forecasting of power production in a large-scale photovoltaic plant”, *Solar Energy*, Vol. 105, 401–413, 2014.

- [Mercer 09] J. Mercer, “Functions of positive and negative type and their connection with the theory of integral equations”, *Philosophical Transactions of the Royal Society*, Vol. 209, 415–446, 1909.
- [Micchelli 86] C. A. Micchelli, “Algebraic aspects of interpolation”, *Proceedings of Symposia in Applied Mathematics*, Vol. 36, 81–102, 1986.
- [Michalakes 01] J. Michalakes, S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, W. Skamarock, “Development of a Next Generation Regional Weather Research and Forecast Model”, en W. Zwiefelhofer, and N. Kreitz (Editores), “Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology”, World Scientific, Singapore, 269–276, 2001.
- [Michalewicz 94] Z. Michalewicz, “Genetic algorithms + Data Structures = Evolution Programs”, Springer-Verlag, 2nd edition, 1994.
- [Michalski 83] R. S. Michalski, J. Larson, “Incremental generation of vl_1 hypotheses: the underlying methodology and description of program aq11” *ISG*, 83–85, Dept. of Computer Science, Univ. of Illinois at Urbana-Champaign, Urbana, 1983.
- [Michalski 86] R. S. Michalski, I. Mozetic, J. Hong, N. Lavrac, “The AQ15 inductive learning system: an overview and experiments”, Proceedings of IMAL 1986, Orsay, France, Université de Paris-Sud, 1986.
- [Mihalakakou 00] G. Mihalakakou, M. Santamouris, D. N. Asimakopoulos, “The total solar radiation time series simulation in Athens, using neural networks”, *Theoretical and Applied Climatology*, Vol. 66, 185–197, 2000.
- [Mohandes 98] M. Mohandes, S. Rehman, T. O. Halawani, “Estimation of global solar radiation using artificial neural networks”, *Renewable Energy*, Vol. 14, no. 1–4, 179–184, 1998.
- [Mohandes 00] M. Mohandes, A. Balghonaim, M. Kassas, S. Rehman, T. O. Halawani, “Use of radial basis functions for estimating monthly mean daily solar radiation”, *Solar Energy*, Vol. 68, no. 2, 161–168, 2000.
- [Monteiro 13a] C. Monteiro, T. Santos, L. A. Fernández-Jiménez, I. J. Ramírez-Rosado, M. S. Terreros-Olarte, “Short-Term Power Forecasting Model for Photovoltaic Plants Based on Historical Similarity”, *Energies*, Vol. 6, 2624–2643, 2013.
- [Monteiro 13b] C. Monteiro, L. A. Fernandez-Jimenez, I. J. Ramirez-Rosado, A. Muñoz-Jimenez, P. M. Lara-Santillan, “Short-term forecasting models for photovoltaic plants: Analytical versus soft-computing techniques”, *Mathematical Problems in Engineering*, 767284, 2013.
- [Moody 89] J. E. Moody, C. J. Darken, “Fast learning networks of locally-tuned processing units”, *Neural Computation*, Vol. 1, no. 2, 281–294, 1989.

- [Moon 42] P. Moon, D. E. Spencer, "Illumination from a non-uniform sky", *Transactions from Illumination Engineering Society*, Vol. 37, 707–725, 1942.
- [Mubiru 08] E. J. Mubiru, K. B. Banda, "Estimation of monthly average daily global solar irradiation using artificial neural networks" *Solar Energy*, Vol. 82, 181–187, 2008.
- [Müller 90] B. Müller, J. Reinhardt, "Neural Networks. An Introduction", Springer-Verlag, 1990.
- [Muneer 87] T. Muneer, "Solar radiation modelling for the United Kingdom", PhD thesis, Council for National Academic Awards, London, 1987.
- [Muneer 90] T. Muneer, "Solar radiation model for Europe", *Building Services Engineering Research and Technology*, Vol. 11, no. 4, 153–163, 1990.
- [Muneer 97] T. Muneer, "Solar radiation and daylight models for the energy efficient design of buildings", Architectural Press, Oxford, 1997.
- [Muneer 98] T. Muneer, M. Gul, H. Kambezidis, "Evaluation of an all-sky meteorological radiation model against long-term measured hourly data", *Energy Conversion and Management*, Vol. 39, 303–317, 1998.
- [Murakami 14] Y. Murakami, Y. Takabayashi, Y. Noro, "Photovoltaic power prediction and its application to smart grid", 2014 IEEE Innovative Smart Grid Technologies - Asia, ISGT ASIA 2014, 6873762, 47–50, 2014.
- [Nijegorodov 97] N. Nijegorodov, J. A. Adedoyin, K. R. S. Devan, "A new analytical-empirical model for the instantaneous diffuse radiation and experimental investigation of its validity", *Renewable Energy*, Vol. 11, 341–350, 1997.
- [Nocedal 99] J. Nocedal, S. J. Wright, "Numerical Optimization", Springer-Verlag, New York, 1999.
- [Notton 06] G. Notton, C. Cristofari, P. Poggi, "Performance evaluation of various hourly slope irradiation models using Mediterranean experimental data of Ajaccio", *Energy Conversion and Management*, Vol. 47, 147–173, 2006.
- [Oldani 07] J. Oldani, "La meteorología conocer, prever el tiempo y comprender los fenómenos meteorológicos", De Vecchi, Barcelona, 2007.
- [Orlanski 75] I. Orlanski, "A Rational Subdivision of Scales for Atmospheric Processes", *Bulletin American Meteorological Society*, Vol. 56, no. 5, 527–530, 1975.
- [Orús 07] J. J. de Orús, M. A. Català, J. Núñez, "Astronomía esférica y mecánica celeste", Universidad de Barcelona, Barcelona, 2007.

- [Paltridge 76] G. M. Paltridge, C. M. R. Platt, “Radiative processes in meteorology and climatology”, Elsevier, New York, 1976.
- [Paoli 10] C. Paoli, C. Voyant, M. Muselli, M. L. Nivet, “Forecasting of preprocessed daily solar radiation time series using neural networks”, *Solar Energy*, Vol. 84, no. 12, 2146–2160, 2010.
- [Paulescu 03] M. Paulescu, Z. Schlett, “A simplified but accurate spectral solar irradiance model”, *Theoretical and Applied Climatology*, Vol. 75, 203–212, 2003.
- [Pawlak 95] Z. Pawlak, et al., “Rough sets Theory”, *Communication of ACM*, Vol. 38, no. 11, 89–95, 1995.
- [Pedro 12] H. T. C. Pedro, C. F. M. Coimbra, “Assessment of forecasting techniques for solar power production with no exogenous inputs”, *Solar Energy*, Vol. 86, no. 7, 2017–2028, 2012.
- [Perez 86] R. Perez, R. Stewart, R. Arbogast, J. Seals, J. Scott, “An anisotropic hourly diffuse radiation model for surfaces: description, performance validation, site dependency evaluation”, *Solar Energy*, Vol. 36, no. 6, 481–497, 1986.
- [Perez 87] R. Perez, R. Seals, P. Ineichen, R. Stewart, D. Menicucci, “A new simplified version of the Perez diffuse irradiance model for tilted surfaces”, *Solar Energy*, Vol. 39, no. 3, 221–231, 1987.
- [Perez 89] R. Perez, P. Ineichen, R. Seals, J. Michalsky, R. Stewart, “Modeling daylight availability and irradiance components from direct and global irradiance”, *Solar Energy*, Vol. 44, no. 5, 271–289, 1990.
- [Perrin 82] Ch. Perrin de Brichambaut, Ch. Vauge, “Le gisement solaire, evaluation de la ressource énergétique”, Lavoisier, Technique et Documentation, Paris, 1982.
- [Poggio 90] T. Poggio, F. Girosi, “Networks for approximation and learning”, *Proceedings of the IEEE*, Vol. 78, 1481–1497, 1990.
- [Principe 00] J. C. Principe, N. R. Euliano, W. C. Lefebvre, “Neural and Adaptive Systems. Fundamentals Through Simulations”, John Wiley, 2000.
- [Psiloglou 00] B. E. Psiloglou, M. Santamouris, D. N. Asimakopoulos, “Atmospheric broadband model for computation of solar radiation at the Earth’s surface. Application to mediterranean climate”, *Journal of Applied Geophysics*, Vol. 157, 829–860, 2000.
- [Quinlan 83] J. R. Quinlan, “Learning Efficient Classification Procedures and Their Application to Chess End Games”, en R. Michalski, J. Carbonell, T. Mitchell (Editores), “Machine Learning: An Artificial Intelligence Approach”, Morgan Kaufmann, San Mateo, CA, 1983.

- [Quinlan 86] J. R. Quinlan, “Induction of Decision Trees”, *Machine Learning*, Vol. 1, 81–106, 1986.
- [Quinlan 90] J. R. Quinlan, “Learning logical definitions from Relations”, *Machine Learning*, Vol. 5, 239–266, 1990.
- [Quinlan 92] J. R. Quinlan, “Learning with continuous classes”, Proceedings 5th Australian Joint Conference on Artificial Intelligence, Word Scientific, Singapore, 343–348, 1992.
- [Quinlan 93] J. R. Quinlan, “C4.5. Programs for Machine Learning”, Morgan Kaufmann, San Francisco, 1993.
- [Quinlan 96] J. R. Quinlan, “Learning first-order definitions from relations”, *Machine Learning*, Vol. 5, no. 3, 239–266, 1996.
- [Ramírez 12] C. A. Ramírez, J. C. Cuevas, “Sistema automático para la predicción de desastres meteorológicos en las ciudades”, *Ide@s CONCYTEG*, Vol. 7, no. 86, 1021–1048, 2012.
- [Rana 16] M. Rana, I. Koprinska, V. G. Agelidis, “Univariate and multivariate methods for very short-term solar photovoltaic power forecasting”, *Energy Conversion and Management*, Vol. 121, 380–390, 2016.
- [Rashkovska 15] A. Rashkovska, J. Novljan, M. Smolnikar, M. Mohorcic, C. Fortuna, “Online short-term forecasting of photovoltaic energy production”, 2015 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2015, 7131880, 2015.
- [Reddy 03] K. S. Reddy, R. Manish, “Solar resource estimation using artificial neural networks and comparison with other correlation models”, *Energy conversion and Management*, Vol. 44, 2519–2530, 2003.
- [Reindl 90] D. T. Reindl, W. A. Beckman, J. A. Duffie, “Evaluation of hourly tilted surface radiation models”, *Solar Energy*, Vol. 45, no. 1, 9–17, 1990.
- [Riaño 16] D. Riaño, “Árboles de regresión. M5”, Departament d’Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, en <<http://banzai-deim.urv.net/~riano/teaching/id3-m5.pdf>> [Consulta: 28 de octubre de 2016].
- [Ritter 91] H. Ritter, T. Martinetz, K. Schulten, “Neural Computation and Self-Organizing Maps”, Addison-Wesley, 1991.
- [Rodgers 78] G. G. Rodgers, C. G. Souster, K. J. Page, “The development of an interactive computer program SUN1 for the calculation of solar irradiances and daily irradiations on horizontal surfaces on cloudless days for given conditions of sky clarity and atmospheric water content”, Rep. BS28, Univ. of Sheffield, 1978.

- [Romero 10] M. Romero, “Energía solar fotovoltaica”, CEAC, Barcelona, 2010.
- [Rumelhart 86a] D. E. Rumelhart y J. L. McClelland (Editores), “Parallel Distributed Processing, Vol 1: Foundations”, MIT Press, 1986.
- [Rumelhart 86b] D. E. Rumelhart, G. E. Hinton, R. J. Williams, “Learning representations by backpropagating errors”, *Nature*, Vol. 323, 533–536, 1986.
- [Safi 02] S. Safi, A. Zeroual, M. Hassani, “Prediction of global daily solar radiation using higher order statistics”, *Renewable Energy*, Vol. 27, no. 4, 647–666, 2002.
- [Santos 09] T. F. Ferreira dos Santos, “Previsão de Potência Eólica com base em Modelos de Prospecção de Similaridade Histórica”, Dissertação realizada no âmbito do Mestrado Integrado em Engenharia Electrotécnica e de Computadores Major Energia, Faculdade de Engenharia da Universidade do Porto, 2009.
- [Sanz 09] J. F. Sanz, L. M. Navas, M. C. Rey, A. Corrêa, “Fundamentos de energía solar fotovoltaica para los grados de titulaciones científico-técnicas”, Universidad Europea Miguel de Cervantes, Valladolid, 2009.
- [Schulze 76] R. E. Schulze, “A physically based method of estimating solar radiation from suncards”, *Agricultural Meteorology*, Vol. 16, 85–101, 1976.
- [Sen 98] Z. Sen, “Fuzzy algorithm for estimation of solar irradiation from sunshine duration”, *Solar Energy*, Vol. 63, no. 1, 39–49, 1998.
- [Serrano 10] D. Serrano, M. J. Marín, M. P. Utrillas, F. Tena, J. A. Martínez-Lozano, “Medida y modelización de la irradiancia solar eritemática global sobre planos inclinados”, *Tethys*, Vol. 7, 57–66, 2010.
- [Sfetsos 00] A. Sfetsos, A. H. Coonick, “Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques”, *Solar Energy*, Vol. 68, no. 2, 169–78, 2000.
- [Sharma 65] M. R. Sharma, R. S. Pal, “Interrelationships between total, direct, and diffuse solar radiation in the tropics”, *Solar Energy*, Vol. 9, 183–192, 1965.
- [Shi 12] J. Shi, W. J. Lee, Y. Liu, Y. Yang, P. Wang, “Forecasting power output of photovoltaic systems based on weather classification and support vector machines”, *IEEE Transactions on Industry Applications*, Vol. 48, no. 3, 1064–1069, 2012.
- [Singh 13] V. P. Singh, V. Vijay, M. Siddhartha Bhatt, D. K. Chaturvedi, “Generalized neural network methodology for short term solar power forecasting”, 13th International Conference on Environment and Electrical Engineering, IEEEIC 2013 - Conference Proceedings, 6737883, 58–62, 2013.

- [Skartveit 86] A. Skartveit, J. A. Olseth, “Modeling slope irradiance at high latitudes”, *Solar Energy*, Vol. 36, no. 4, 333–344, 1986.
- [Smith 84] D. S. Smith, W. F. Dobson, “The heat budget at ocean weather station Bravo”, *Atmosphere-Ocean*, Vol. 22, 1–22, 1984.
- [Smola 04] A. J. Smola, B. Schölkopf, “A tutorial on support vector regression”, *Statistics and Computing*, Vol. 14, no. 3, 199–222, 2004.
- [Soubdhan 16] T. Soubdhan, J. Ndong, H. Ould-Baba, M. T. Do, “A robust forecasting framework based on the Kalman filtering approach with a twofold parameter tuning procedure: Application to solar and photovoltaic prediction”, *Solar Energy*, 131, 246–259, 2016.
- [Sozen 04a] A. Sozen, E. Arcaklýogblu, M. Ozalp, “Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data”, *Energy Conversion and Management*, Vol. 45, no. 18–19, 3033–3052, 2004.
- [Sozen 04b] A. Sozen, E. Arcaklýogclu, M. Ozalp, E. G. Kanýt, “Use of artificial neural networks for mapping the solar potential in Turkey”, *Applied Energy*, Vol. 77, 273–286, 2004.
- [Sozen 05] A. Sozen, E. Arcaklýogclub, M. Ozalpa, N. C. Agclarc, “Forecasting based on neural network approach of solar potential in Turkey”, *Renewable Energy*, Vol. 30, 1075–1090, 2005.
- [Sperati 16] S. Sperati, S. Alessandrini, L. Delle Monache, “An application of the ECMWF Ensemble Prediction System for short-term solar power forecasting”, *Solar Energy*, Vol. 133, 437–450, 2016.
- [Spiegel 01] M. R. Spiegel, J. Schiller, R. A. Srinivasan, “Probabilidad y Estadística”, Mc Graw Hill, 2001.
- [Stull 88] R. B. Stull, “An Introduction to Boundary Layer Meteorology”, Kluwer Academic Publishers, Dordrecht, 1988.
- [Suckling 77] P. W. Suckling, J. E. Hay, “A cloud layer-sunshine model for estimating direct, diffuse and total solar radiation”, *Atmosphere*, Vol. 15, 194–207, 1977.
- [Sugeno 85a] J. Sugeno, M. Nishida, “Fuzzy control of model car”, *Fuzzy Set and Systems*, Vol. 26, no. 2, 151–164, 1985.
- [Sugeno 85b] J. Sugeno, “Industrial applications of fuzzy control”, Elsevier, 1985.
- [Sugeno 88] M. Sugeno, G. T. Kang, “Structure identification of fuzzy model”, *Fuzzy Sets and Systems*, Vol. 28, 15–33, 1988.

- [Sumaili 14] J. Sumaili, R. J. Bessa, F. Rahman, R. Tomé, J. N. Sousa, “Electrical model parameter characterization for short-term solar power forecasting”, Proc. of the 4th International Workshop on Integration of Solar Power into Power Systems, Berlin, 2014.
- [Takagi 83] T. Takagi, M. Sugeno, “Derivation of fuzzy control rules from human operator’s control actions”, Proceedings of the IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis, 55–60, 1983.
- [Takagi 85] 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.
- [Tanguay 90] M. Tanguay, A. Robert, R. Laprise, “A semi-implicit semi-Lagrangian fully compressible regional forecast model”, *Monthly Weather Review*, Vol. 118, 1970–1980, 1990.
- [Temps 77] R. C. Temps, K. L. Coulson, “Solar radiation incident upon slopes of different orientation”, *Solar Energy*, Vol. 19, no. 2, 179–184, 1977.
- [Thunis 96] P. Thunis, R. Bornstein, “Hierarchy of Mesoscale Flow Assumptions and Equations”, *Journal of the Atmospheric Sciences*, Vol. 53, no. 3, 380–397, 1996.
- [Tou 74] J. T. Tou, R. C. González, “Pattern Recognition Principles”, Addison-Wesley Pub. Comp., 1974.
- [Tymvios 02] F. Tymvios, C. P. Jacovides, S. C. Michaelides, “The total solar energy on a horizontal level with the use of artificial neural networks”, en “Sixth hellenic conference of meteorology, climatology and atmospheric physics”, Ioannina, Greece, 2002.
- [Tymvios 05] F. S. Tymvios, C. P. Jacovides, S. C. Michaelides, C. Scouteli, “Comparative study of Angstroms and artificial neural networks methodologies in estimating global solar radiation”, *Solar Energy*, Vol. 78, 752–762, 2005.
- [Vanderbei 97] R. J. Vanderbei, “LOQO user’s manual—versión 3.10. Technical Report SOR-97-08”, en “Statistics and Operations Research”, Princeton University, 1997.
- [Vapnik 95] V. Vapnik, “The Nature of Statistical Learning Theory”, Springer, New York, 1995.
- [Vaz 16] A. G. R. Vaz, B. Elsinga, W. G. J. H. M. van Sark, M. C. Brito, “An artificial neural network to assess the impact of neighbouring photovoltaic systems in power forecasting in Utrecht, the Netherlands”, *Renewable Energy*, Vol. 85, 631–641, 2016.

- [Vilchez 09] R. Vilchez, “Estudio de alternativas para el suministro energético de una vivienda unifamiliar aislada situada en la población de Lligà d’Amunt. Anexo II: La Energía Solar Fotovoltaica”, Proyecto / Trabajo final de carrera, Universidad Politécnica de Cataluña, Escuela Superior de Ingeniería Industrial, Aeroespacial y Audiovisual de Terrassa, en upcommons.upc.edu/handle/2099.1/7538, 2009.
- [Voyant 11] C. Voyant, M. Muselli, C. Paoli, M. L. Nivet, “Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation”, *Energy*, Vol. 36, no. 1, 348–359, 2011.
- [Wang 92] L. X. Wang, J. M. Mendel, “Back-propagation fuzzy systems as non-linear dynamic system identifiers”, Proceedings of the IEEE International Conference on Fuzzy Systems, San Diego, 1992.
- [Wang 94] L. X. Wang, “Adaptative Fuzzy Systems and Control”, Prentice, 1994.
- [Wang 97] Y. Wang, I. H. Witten, “Induction of model trees for predicting continuous classes”, Proc. of the poster papers of the 9th European Conference on Machine Learning, Prague, Czech Republic, 128–137, 1997.
- [Wang 11] S. Wang, N. Zhang, Y. Zhao, J. Zhan, “Photovoltaic system power forecasting based on combined grey model and BP neural network”, Proceedings of the International Conference on Electrical and Control Engineering, ICECE 2011, 6057634, 4623–4626, 2011.
- [Wang 14] Z. Wang, L. He, X. Cheng, J. He, “Method for short-term photovoltaic generation power prediction base on weather patterns”, China International Conference on Electricity Distribution (CICED 2014), 6991696, 213–215, 2014.
- [Wassermann 93] P. D. Wasserman, “Advanced Methods in Neural Computing”, Van Nostrand Reinhold, 1993.
- [Watt 78] D. Watt, “On the nature and distribution of solar radiation”, Report for U. S. DOE, 1978.
- [Wei 90] W. Wei, “Time series analysis: univariate and multivariate methods”, Addison-Wesley, Reading, Massachusetts, 1990.
- [Widrow 60] B. Widrow, M. E. Hoff, “Adaptive switching circuits”, 1960 IRE WESCON Convention Record, Vol. 4, 96–104, 1960.
- [Widrow 88] B. Widrow, R. Winter, “Neural nets for adaptive filtering and adaptive pattern recognition”, *IEEE Computer*, Vol. 21, 25–39, 1988.
- [Williams 94] D. B. Williams, F. S. Zazueta, “Solar radiation estimation via neural network”, en ASAE (Editor), Sixth international conference on computers in agriculture, Cancun, Mexico, 140–146, 1994.

- [Williams 96] D. Williams, F. Zazueta, “Solar radiation estimation via neural network”, Sixth international conference on radiation, Computers in Agriculture, Cancun, Mexico, 1143–1149, 1996.
- [Wu 14] Y. K. Wu, C. R. Chen, H. Abdul Rahman, “A novel hybrid model for short-term forecasting in PV power generation”, *International Journal of Photoenergy*, 569249, 2014.
- [Xue 00] M. Xue, K. K. Droegemeier, V. Wong, “The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification”, *Meteorology and Atmospheric Physics*, Vol. 75, 161–193, 2000.
- [Yan 13] Z. Yan, W. Lei, G. Feng, . . . , X. Rui, H. Wang, “Integrated wind and solar power forecasting in China”, Proceedings of 2013 IEEE International Conference on Service Operations and Logistics, and Informatics, SOLI 2013, 6611466, 500–505, 2013.
- [Yan 14] C. Yan, J. Xiu, C. Liu, Z. Yang, “A high concentrated photovoltaic output power predictive model based on Fuzzy Clustering and RBF neural network”, CCIS 2014 - Proceedings of 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems, 7175765, 384–388, 2014.
- [Yang 01] K. Yang, G. W. Huang, N. Tamai, “A hybrid model for estimating global solar radiation”, *Solar Energy*, Vol. 70, 13–22, 2001.
- [Yang 13] Y. Yang, L. Dong, “Short-term PV generation system direct power prediction model on wavelet neural network and weather type clustering”, Proceedings - 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2013, Vol. 1, 6643868, 207–211, 2013.
- [Yang 14] Z. Yang, Y. Cao, J. Xiu, “Power generation forecasting model for photovoltaic array based on generic algorithm and BP neural network”, CCIS 2014 - Proceedings of 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems, 7175764, 380–383, 2014.
- [Yang 15] C. Yang, A. A. Thatte, L. Xie, “Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation”, *IEEE Transactions on Sustainable Energy*, Vol. 6, no. 1, 6945846, 104–112, 2015.
- [Yazdanbaksh 13] O. Yazdanbaksh, A. Krahn, S. Dick, “Predicting solar power output using complex fuzzy logic”, Proceedings of the 2013 Joint IFSA World Congress and NAFIPS Annual Meeting, IFSA/NAFIPS 2013, 6608579, 1243–1248, 2013.

- [Zamo 14] M. Zamo, O. Mestre, P. Arbogast, O. Pannekoucke, “A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production”, *Solar Energy*, Vol. 105, 792–803, 2014.
- [Zarzalejo 05] L. F. Zarzalejo, L. J. Ramirez, “Artificial intelligence techniques applied to hourly global irradiance estimation from satellite-derived cloud index”, *Energy*, Vol. 30, 1685–1697, 2005.
- [Zervas 08] P. L. Zervas, H. Sarimveis, J. A. Palyvos, N. C. G. Markatos, “Prediction of daily global solar irradiance on horizontal surfaces based on neural-network techniques”, *Renewable Energy*, Vol. 33, no. 8, 1796–1803, 2008.
- [Zhang 06] Q. Zhang, “Development of the typical meteorological database for Chinese locations”, *Energy and Buildings*, Vol. 38, 1320–1326, 2006.
- [Zhong 14] Z. Zhong, J. Tan, T. Zhang, L. Zhu, “PV power short-term forecasting model based on the data gathered from monitoring network”, *China Communications*, Vol. 11, no. 14, 61–69, 2014.
- [Zufiria 99] P. Zufiria, A. Vázquez, J. Riesco, J. Aguilera, L. Hontoria, “A neural network approach for generating solar radiation artificial series”, *Proceedings of the IWANN099*, Alicante, 1999.

