



UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Métodos de predicción de la generación agregada de energía eólica

Autor:

Miguel García Lobo

Director:

Dr. Ismael Sánchez Rodríguez-Morcillo

DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

Leganés, 2010

TESIS DOCTORAL

**MÉTODOS DE PREDICCIÓN DE LA
GENERACIÓN AGREGADA DE ENERGÍA EÓLICA**

Autor: Miguel García lobo

Director: Dr. Ismael Sánchez Rodríguez-Morcillo

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Leganés,

de

de

A mis padres y hermanas.

A Lidia, Lucía... y quien está por llegar.

Agradecimientos

La parte más simpática de escribir estas líneas es que uno nunca sabe qué decir. Cuando llega el momento de redactar esta primera página ya se han escrito otras doscientas, llenas de gráficos, ecuaciones y justificaciones sobrias, así que resulta difícil eso de cambiar el chip y de pronto expresarse en tono cálido. Como quiera que nunca se me dio muy bien eso de ser místico, trataré al menos de ser escueto, al tiempo de no dejarme a nadie en el tintero.

Tras unos cuantos años de incertidumbre, de estar perdido y encontrarme para volver a perderme y empezar de cero, de aprender tantas cosas nuevas, y sobre todo después de mucho trabajo, este capítulo de mi vida toca su fin. Y para que algo así ocurra tienen que reunirse muchas condiciones favorables:

Así, quiero dar las gracias a Red Eléctrica de España, y en particular a las personas con las que he colaborado, por prestarme los datos necesarios y sobre todo por permitirme aportar mi granito de arena en un tema que me parece tan interesante, y que ha acabado por convertirse en el tema principal de esta tesis doctoral. Obviamente, quiero hacer una mención especial a quienes me guiaron desde mis primeros pasos como becario, Gerardo González y Fernando Soto.

Siguiendo en el ámbito profesional, debo agradecer a los responsables del Departamento de Ingeniería Eléctrica de la Universidad Carlos III de Madrid que me hayan brindado las oportunidades y el tiempo necesarios para que este trabajo llegue a su fin. Y por supuesto a mi tutor, Ismael Sánchez, por enseñarme tantas cosas en un ámbito de la ciencia en el que siempre me he sentido como un simple ignorante.

Sin salir de la universidad, quiero agradecer a los que han sido mis compañeros, a la par que amigos, el gran apoyo que me han dado, haciendo que las interminables horas frente al ordenador hayan sido un poco más llevaderas. Queda demostrado que un proyecto de tal envergadura solo se puede llevar a cabo si se está bien acompañado.

A mis amigos en general y a los que han estado siempre ahí, aguantando estoicamente mis discursos de agobio. Por supuesto, debo mencionar el apoyo de Marta, que sorprendentemente ha aceptado acompañarme durante los últimos años, a pesar de mi casi total ausencia.

Pero el agradecimiento con mayúsculas, el que hace sombra a todos los demás, debo dárselo a mi familia, por haberme comprendido en los momentos difíciles y haberme soportado en esos momentos en que uno casi se vuelve insoportable. Y muy especialmente a mis padres, por habernos dado tanto cariño, y por haber dedicado a la educación y formación de sus hijos mil veces más tiempo y recursos que los que han dedicado a sí mismos. Enhorabuena, porque este logro es tan vuestro como mío.

Resumen

El fomento y desarrollo de las energías renovables, y como consecuencia la alta penetración eólica en los sistemas eléctricos, ha hecho necesario el uso de herramientas de predicción, para conocer con antelación suficiente la cantidad de energía de origen eólico que será inyectada en la red y poder coordinar el resto de fuentes de generación.

En ocasiones existe un especial interés en la predicción de generación agregada de varios parques en una región, ya sea por parte del operador del sistema eléctrico, o por los agentes que gestionan la energía que generan varios parques. Por otro lado, suele haber un suavizado espacial que reduce el error de las predicciones agregadas frente al error de predicción de los parques individuales. Estos motivos hacen de la predicción agregada un modo muy atractivo de predicción eólica.

En esta tesis se propone un método de predicción de la generación agregada de energía eólica en una región, mediante la búsqueda de similitudes entre el vector de viento previsto en algunas coordenadas de la región y otros vectores históricos de predicción de viento. El modelo propuesto se basa en modelos locales de suavizado de Media Ponderada y de Regresión Local Ponderada. Estos modelos ajustan de forma no paramétrica un modelo de predicción para cada punto de entrada al problema, que en nuestro caso será un vector de viento previsto.

Para hacer el suavizado de los datos, se proponen y comparan distintos procedimientos para tener en cuenta la medida de distancias entre vectores, el método de selección de datos cercanos, y las funciones de ponderación por distancia y antigüedad de los datos seleccionados.

Mediante un procedimiento de estimación adaptativa de parámetros, se conseguirá que el modelo de predicción se adecue a la evolución temporal de la compleja relación no lineal existente entre el viento previsto en la región y la potencia eólica total generada.

Las predicciones de potencia obtenidas con los modelos de Media Ponderada y Regresión Local Ponderada, se combinan para aprovechar las ventajas que ofrecen ambas aproximaciones en función de la complejidad del problema de predicción en cada momento.

El modelo propuesto se validará comparando sus predicciones con la agregación de predicciones que se obtienen con Sipleólico, una herramienta que calcula predicciones para cada uno de los parques en la región y después las suma.

Por último, para dar mayor valor a las predicciones agregadas, se propone un modelo de predicción probabilista condicionada a los vectores de viento previsto mediante estimación de densidades con Kernel, para el cual se utilizan las medidas de distancia y funciones de ponderación propuestas para el modelo de predicción de potencia.

Abstract

The promotion and development of renewable energy resources, and as a result, the high wind power penetration in the power systems, makes necessary the use of forecasting tools, in order to know in advance the amount of wind energy that will be injected into the network and to coordinate all the other generation sources.

Sometimes, there is a particular interest in predicting the aggregated wind power in a region, either by the system operator or other practitioners who manage the energy generated by several wind farms. On the other hand, there is usually a spatial smoothing effect that reduces the prediction error when making aggregated wind power predictions, instead of predicting individual wind farms. These reasons make the aggregated wind power prediction models very appealing.

This thesis proposes a method to make aggregated wind power predictions in a region, by seeking similarities between the next wind speed forecasting vector and a set of historical vectors of wind speed in some coordinates of the region. The proposed model is based on Weighted Average and Locally Weighted Regression smoothing models. These models fit a non-parametric prediction model for each query point, which in our case is a vector containing the wind speed forecasts for a set of coordinates in the region.

In order to smooth the data, different smoothing procedures are proposed and compared. These procedures must take into account the measure of distances between wind speed vectors, the nearby data selection method, and the distance and age weighting functions for the selected data.

An adaptive estimation procedure of the parameters will let the model to fit the time-varying evolution of the complex nonlinear relationship between the expected wind speed in the coordinates and the total wind power generated in the region.

The wind power predictions obtained with the alternative models, based on Weighted Average and Locally Weighted Regression, are combined in order to exploit the advantages of both approaches depending on the complexity of the prediction problem in every moment.

The proposed model is validated by comparing its predictions with the aggregation of predictions obtained with Siprolico, a wind power prediction tool that makes predictions for each of the wind farms in the region and adds them.

Finally, more value is added to the aggregated model by proposing a method to make probabilistic forecasts conditioned to the expected wind speed vectors using Kernel density estimation. This approach uses the measures of distance and the weighting functions proposed for the wind power prediction model.

Contenido

Agradecimientos	iii
Resumen	v
Abstract	vii
Contenido	ix
Índice de Figuras	xiii
Índice de Tablas	xxi
1. Introducción y Objetivos de la Tesis	1
1.1 Importancia de la generación eólica	3
1.2 La energía eólica en el sistema eléctrico español.....	5
1.3 Necesidad de predicción de energía eólica.....	8
1.4 Importancia de las predicciones para el Operador del Sistema Eléctrico	10
1.5 Objetivos de la tesis	11
1.6 Estructura y contenido de la tesis	12
2. Estado del Arte en Predicción de Energía Eólica	15
2.1 Introducción a la conversión de energía eólica.....	17
2.2 Modelos de previsión de viento	18
2.3 Revisión del estado del arte en modelos de predicción eólica.....	20
2.3.1 Horizontes de predicción	21
2.3.2 Modelos estadísticos de series temporales.....	22
2.3.3 Modelos físicos	24
2.3.4 Modelos estadísticos de regresión dinámica.....	26
2.4 La herramienta de predicción SIPREÓLICO	28
2.4.1 Datos de entrada	28
2.4.2 Modelos de predicción de Sipreólico.....	30
2.4.3 Combinación de predicciones obtenidas con viento de distintas agencias.....	34
2.4.4 Combinación de predicciones para el Total Peninsular	34
2.4.5 Informes de resultados y diagnóstico	35
2.5 El proyecto ANEMOS.....	36
2.6 Valor económico de la predicción eólica	37
2.7 Incertidumbre de la predicción eólica	39
2.8 Predicción agregada de energía eólica. Predicción Regional	41
2.8.1 Ventajas de la predicción agregada de energía eólica.....	43

Contenido

3. Método Propuesto para la Predicción Regional	45
3.1 Introducción	47
3.2 Modelos locales	48
3.3 Estructuras de modelos locales.....	51
3.3.1 Media Ponderada	53
3.3.2 Regresión Polinómica Local Ponderada	55
3.3.3 Ventajas e inconvenientes de la Media Ponderada y la Regresión Local Ponderada	58
3.4 Método propuesto para el cálculo de predicción eólica regional.....	59
3.5 Métodos de evaluación de los modelos de predicción	62
3.5.1 Sesgo del error de predicción.....	62
3.5.2 Error Medio Absoluto Normalizado (NMAE).....	63
3.5.3 Error Cuadrático Medio Normalizado (NMSE y NRMSE).....	63
3.5.4 Coeficiente de mejora de la predicción.....	64
4. Métodos de Suavizado.....	65
4.1 Funciones Kernel de suavizado	67
4.2 Método de suavizado propuesto	71
4.3 Medida de distancias	71
4.3.1 Distancia de velocidad media.....	72
4.3.2 Distancia absoluta	72
4.3.3 Distancia euclídea.....	73
4.3.4 Distancia de velocidad cúbica	73
4.4 Factores de escala. Distancias estandarizadas.....	74
4.4.1 Distancia absoluta estandarizada con la media	75
4.4.2 Distancia absoluta estandarizada con la desviación típica.....	75
4.4.3 Distancia euclídea estandarizada con la varianza	76
4.4.4 Distancia de velocidad cúbica estandarizada con la desviación típica	77
4.4.5 Distancia de Mahalanobis	77
4.5 Selección de datos en función de su distancia	78
4.5.1 Datos a una distancia inferior a un valor umbral fijo	79
4.5.2 M vecinos más cercanos	80
4.5.3 M como un porcentaje fijo de datos más cercanos del total de datos disponibles.....	81
4.5.4 Datos a una distancia inferior a un porcentaje fijo del rango de distancias	82
4.5.5 Datos a una distancia inferior a un porcentaje fijo del rango de distancias hasta la mediana.....	84
4.5.6 Método basado en la tangente a los datos ordenados por distancia	85
4.6 Funciones de ponderación por distancia	86
4.6.1 Corrección del parámetro α de ponderación en función de la distribución de distancias ..	89
4.7 Funciones de ponderación por antigüedad	91
5. Evaluación de los Métodos de Suavizado Propuestos	95
5.1 Datos utilizados	97
5.2 Evaluación de las funciones Kernel	97
5.2.1 Media Ponderada con funciones Kernel	98
5.2.2 Regresión Local Ponderada con funciones Kernel	104

5.2.3 Comparación entre Media Ponderada y Regresión Local Ponderada usando funciones Kernel.....	109
5.3 Análisis previo de las medidas de distancia propuestas.....	113
5.3.1 Dependencia del nivel de potencia prevista.....	114
5.3.2 Dependencia del tipo de medida de distancia elegido.....	119
5.4 Evaluación de los métodos de selección	129
5.4.1 Datos a una distancia inferior a un valor umbral fijo.....	130
5.4.2 M vecinos más cercanos.....	132
5.4.3 M como un porcentaje fijo de datos más cercanos del total de datos disponibles	134
5.4.4 Datos a una distancia inferior a un porcentaje fijo del rango de distancias.....	136
5.4.5 Datos a una distancia inferior a un porcentaje fijo del rango de distancias hasta la mediana	138
5.4.6 Método basado en la tangente a los datos ordenados por distancia.....	140
5.4.7 Resumen de la evaluación de métodos de selección	141
5.5 Evaluación de las medidas de distancia.....	144
5.5.1 Resumen de la evaluación de medidas de distancia	149
5.6 Evaluación de las funciones de ponderación por distancia.....	150
5.6.1 Comparación de las funciones de ponderación propuestas.....	153
5.6.2 Corrección del parámetro α en función de la distribución de distancias	156
5.6.3 Comparación de las funciones de ponderación con el parámetro α corregido	161
5.7 Comparación entre las funciones Kernel y el método de suavizado propuesto.....	162
5.8 Evaluación de la ponderación por antigüedad.....	164
5.9 Resumen general de la evaluación.....	168
6. Estimación Adaptativa o Evolutiva de Hiperparámetros	169
6.1 Introducción	171
6.2 Vectores de parámetros y cubo de predicciones alternativas	173
6.3 Error de las predicciones alternativas. Cubo de errores	176
6.4 Estimación adaptativa del NMSE mediante EWMA	177
6.5 Combinación de predicciones alternativas del cubo de predicciones	179
6.6 Desplazamiento de los vectores de parámetros	180
6.7 Evaluación de resultados obtenidos	183
7. Combinación Adaptativa de Predicciones	189
7.1 Introducción	191
7.2 Modelos de combinación	191
7.2.1 Combinación de promedio óptimo.....	192
7.2.2 Combinación de mínimo riesgo	194
7.3 Procedimiento de combinación en dos etapas	197
7.4 Evaluación de resultados obtenidos	199
7.4.1 Combinación de modelos de Media Ponderada y Regresión Local Ponderada	199
7.4.2 Combinación de predicciones obtenidas a partir de distintas fuentes de viento	202
7.4.3 Análisis de otras posibilidades de combinación	204

Contenido

8. Evaluación del Modelo Propuesto de Predicción de Potencia Eólica Agregada	209
8.1 Introducción	211
8.2 Modelo de predicción utilizado	212
8.2.1 Comprobación de la estimación adaptativa de hiperparámetros	213
8.2.2 Comprobación de la combinación adaptativa de predicciones	215
8.3 Comparación con otros modelos de predicción	218
9. Modelos de Predicción Probabilista	221
9.1 Introducción	223
9.2 Estimación de densidades con Kernel	225
9.3 Antigüedad de las observaciones de error	227
9.4 Predicción probabilista condicionada al nivel de potencia prevista	228
9.5 Predicción probabilista condicionada al vector de viento previsto	230
9.6 Métodos de evaluación de las predicciones probabilistas	232
9.6.1 Fiabilidad o calibración	232
9.6.2 Concentración o agudeza (sharpness)	234
9.6.3 Resolución	236
9.7 Evaluación de resultados obtenidos	237
9.7.1 Estimación de funciones de densidad basadas únicamente en observaciones históricas del error	237
9.7.2 Predicción probabilista condicionada al nivel de potencia prevista	242
9.7.3 Predicción probabilista condicionada al vector de viento previsto	244
9.7.4 Comparación de los modelos de predicción probabilista	245
10. Conclusiones Generales	251
10.1 Conclusiones y aportaciones originales	253
10.2 Desarrollos futuros	255
10.3 Publicaciones	257
Bibliografía	259

Índice de Figuras

Figura 1.1 : Evolución de la potencia eólica instalada en el mundo.....	4
Figura 1.2 : Evolución de la potencia eólica instalada en España	6
Figura 1.3 : Potencia instalada a 31 de diciembre de 2009. Fuente: REE. Avance del informe 2009.....	6
Figura 1.4 : Cobertura de la demanda, año 2009. Fuente: REE. Avance del informe 2009.....	7
Figura 2.1 : Curva potencia-viento característica de un aerogenerador.....	18
Figura 2.2 : Diagrama de funcionamiento de Sipleólico	29
Figura 2.3 : Esquema del módulo de predicción de Sipleólico.....	33
Figura 3.1 : Dispersión de puntos de la potencia generada en un parque eólico.....	48
Figura 3.2 : Ajuste de una curva paramétrica a toda la nube de puntos (izquierda) y estimación por intervalos (derecha).....	49
Figura 3.3 : Ajuste de un modelo local a los datos cercanos.....	50
Figura 3.4 : Media Ponderada.....	54
Figura 3.5 : Regresión Polinómica Local Ponderada	56
Figura 3.6 : Selección de un set de K coordenadas en la región, donde se tendrán las previsiones de velocidad del viento.....	60
Figura 3.7 : Esquema del algoritmo de predicción regional propuesto.....	61
Figura 4.1 : Ventana de datos a seleccionar según la distancia y el ancho de banda en la dimensión k.....	67
Figura 4.2 : Distancia relativa de los datos en la dimensión k en función del ancho de banda	68
Figura 4.3 : Funciones Kernel	70
Figura 4.4 : Contornos equidistantes usando factores de escala	74
Figura 4.5 : Contornos equidistantes usando la distancia de Mahalanobis.....	78
Figura 4.6 : Datos a distancia inferior a un valor umbral fijo.....	80
Figura 4.7 : Selección de M vecinos más cercanos	80
Figura 4.8 : Selección de M vecinos más cercanos	81
Figura 4.9 : Datos a una distancia inferior a un porcentaje fijo del rango de distancias	83
Figura 4.10 : Datos contenidos en la primera clase del histograma	83
Figura 4.11 : Datos a una distancia inferior a un porcentaje fijo del rango de distancias hasta la mediana	84
Figura 4.12 : Método de selección basado en la tangente.....	85
Figura 4.13 : Coeficiente de ponderación inversa a la distancia	88
Figura 4.14 : Coeficiente de ponderación por decaimiento exponencial con la distancia.....	88
Figura 4.15 : Histograma con alta concentración de datos cercanos	89
Figura 4.16 : Histograma con baja concentración de datos cercanos.....	90

Índice de Figuras

<i>Figura 4.17 : Coeficientes de ponderación por antigüedad</i>	<i>93</i>
<i>Figura 4.18 : Relación de coeficientes de ponderación por antigüedad</i>	<i>93</i>
<i>Figura 5.1 : Error de predicción en función del ancho de banda. Media Ponderada con Kernel uniforme</i>	<i>98</i>
<i>Figura 5.2 : Error de predicción en función del ancho de banda. Media Ponderada con Kernel triangular.....</i>	<i>100</i>
<i>Figura 5.3 : Error de predicción en función del ancho de banda. Media Ponderada con Kernel Epanechnikov.....</i>	<i>100</i>
<i>Figura 5.4 : Error de predicción en función del ancho de banda. Media Ponderada con Kernel quartic.....</i>	<i>101</i>
<i>Figura 5.5 : Error de predicción en función del ancho de banda. Media Ponderada con Kernel coseno.....</i>	<i>101</i>
<i>Figura 5.6 : Error de predicción en función del ancho de banda. Media Ponderada con Kernel gaussiano.....</i>	<i>102</i>
<i>Figura 5.7 : Comparación de errores NMAE de las funciones Kernel usando un modelo de Media Ponderada</i>	<i>103</i>
<i>Figura 5.8 : Comparación de errores NRMSE de las funciones Kernel usando un modelo de Media Ponderada</i>	<i>103</i>
<i>Figura 5.9 : Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel uniforme</i>	<i>104</i>
<i>Figura 5.10 : Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel triangular.....</i>	<i>105</i>
<i>Figura 5.11 : Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel Epanechnikov.....</i>	<i>106</i>
<i>Figura 5.12 : Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel quartic.....</i>	<i>106</i>
<i>Figura 5.13 : Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel coseno.....</i>	<i>107</i>
<i>Figura 5.14 : Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel gaussiano.....</i>	<i>107</i>
<i>Figura 5.15 : Comparación de errores NMAE de las funciones Kernel usando un modelo de Regresión Local Ponderada</i>	<i>108</i>
<i>Figura 5.16 : Comparación de errores NRMSE de las funciones Kernel usando un modelo de Regresión Local Ponderada</i>	<i>109</i>
<i>Figura 5.17 : Comparación de errores NMAE de los modelos de predicción de Media Ponderada y Regresión Local Ponderada</i>	<i>110</i>
<i>Figura 5.18 : Comparación de errores NRMSE de los modelos de predicción de Media Ponderada y Regresión Local Ponderada.....</i>	<i>110</i>
<i>Figura 5.19 : Comparación de sesgo de los modelos de predicción de Media Ponderada y Regresión Local Ponderada calculados para NMAE mínimo</i>	<i>111</i>
<i>Figura 5.20 : Comparación de sesgo de los modelos de predicción de Media Ponderada y Regresión Local Ponderada calculados para NRMSE mínimo</i>	<i>112</i>

<i>Figura 5.21 : Histograma de medidas de potencia total peninsular registrada durante el año 2007..</i>	113
<i>Figura 5.22 : Selección de instantes a representar con varios niveles de potencia correspondientes a los percentiles 10, 25, 50, 75 y 90 %.....</i>	115
<i>Figura 5.23 : Dispersión de distancias para potencia esperada muy baja, baja, media, alta y muy alta.....</i>	117
<i>Figura 5.24 : Histograma de distancias para potencia esperada muy baja, baja, media, alta y muy alta.....</i>	118
<i>Figura 5.25 : Dispersiones e histogramas para distancias de velocidad media.....</i>	120
<i>Figura 5.26 : Dispersiones e histogramas para distancia absoluta.....</i>	121
<i>Figura 5.27 : Dispersiones e histogramas para distancia absoluta estandarizada con la media.....</i>	122
<i>Figura 5.28 : Dispersiones e histogramas para distancia absoluta estandarizada con la desviación típica.....</i>	123
<i>Figura 5.29 : Dispersiones e histogramas para distancia euclídea.....</i>	124
<i>Figura 5.30 : Dispersiones e histogramas para distancia euclídea estandarizada con la varianza.....</i>	125
<i>Figura 5.31 : Dispersiones e histogramas para distancia cúbica.....</i>	126
<i>Figura 5.32 : Dispersiones e histogramas para distancia cúbica estandarizada con la desviación típica.....</i>	127
<i>Figura 5.33 : Dispersiones e histogramas para distancia de Mahalanobis.....</i>	128
<i>Figura 5.34 : Errores NMAE y NRMSE para el método de selección de distancia umbral con modelo de Media Ponderada.....</i>	130
<i>Figura 5.35 : Errores NMAE y NRMSE para el método de selección de distancia umbral con modelo de Regresión Local Ponderada.....</i>	131
<i>Figura 5.36 : Errores NMAE y NRMSE para el método de selección de M vecinos más cercanos con modelo de Media Ponderada.....</i>	132
<i>Figura 5.37 : Errores NMAE y NRMSE para el método de selección de M vecinos más cercanos con modelo de Regresión Local Ponderada.....</i>	133
<i>Figura 5.38 : Errores NMAE y NRMSE para el método de selección de porcentaje de datos más cercanos con modelo de Media Ponderada.....</i>	134
<i>Figura 5.39 : Errores NMAE y NRMSE para el método de selección de porcentaje de datos más cercanos con modelo de Regresión Local Ponderada.....</i>	135
<i>Figura 5.40 : Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias con modelo de Media Ponderada.....</i>	136
<i>Figura 5.41 : Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias con modelo de Regresión Local Ponderada.....</i>	137
<i>Figura 5.42 : Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias hasta la mediana con modelo de Media Ponderada.....</i>	138
<i>Figura 5.43 : Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias hasta la mediana con modelo de Regresión Local Ponderada.....</i>	139

Índice de Figuras

Figura 5.44 : Errores NMAE y NRMSE para el método de selección basado en la tangente con modelo de Media Ponderada	140
Figura 5.45 : Errores NMAE y NRMSE para el método de selección basado en la tangente con modelo de Regresión Local Ponderada.....	141
Figura 5.46 : Comparación de errores NMAE y NRMSE para todos los métodos de selección con un modelo de Media Ponderada	142
Figura 5.47 : Comparación de errores NMAE y NRMSE para todos los métodos de selección con un modelo de Regresión Local Ponderada.....	143
Figura 5.48 : Evaluación inicial de las medidas de distancia	144
Figura 5.49 : Evaluación inicial de las medidas de distancia para selección de porcentaje de datos y umbral como porcentaje de distancia hasta la mediana, con modelo de Media Ponderada	145
Figura 5.50 : Evaluación inicial de las medidas de distancia para selección de porcentaje de datos y umbral como porcentaje de distancia hasta la mediana, con modelo de Regresión Local Ponderada	146
Figura 5.51 : Coeficiente de Variación del viento en las coordenadas.....	148
Figura 5.52 : Comparación de errores para distancia absoluta estandarizada con la media y estandarizada con la varianza, con modelo de Regresión Local Ponderada	149
Figura 5.53 : Errores NMAE para las funciones de ponderación inversa y de decaimiento exponencial con la distancia, en función del parámetro α	151
Figura 5.54 : Error NMAE para la función de ponderación de decaimiento exponencial con la distancia, en función de los parámetros p y α , con modelo de Media Ponderada	152
Figura 5.55 : Error NMAE para la función de ponderación inversa con la distancia, en función de los parámetros p y α , con modelo de Media Ponderada	152
Figura 5.56 : Error NMAE para la función de ponderación de decaimiento exponencial con la distancia, en función de los parámetros p y α , con modelo de Regresión Local Ponderada	153
Figura 5.57 : Error NMAE para la función de ponderación inversa con la distancia, en función de los parámetros p y α , con modelo de Regresión Local Ponderada	153
Figura 5.58 : Comparación de errores para ponderación inversa y decaimiento exponencial con la distancia, con método de selección de un porcentaje de datos y modelo de Media Ponderada	154
Figura 5.59 : Parámetros α con los que se obtiene el menor error NMAE, con método de selección de un porcentaje de datos y modelo de Media Ponderada	154
Figura 5.60 : Comparación de errores para ponderación inversa y decaimiento exponencial con la distancia, con método de selección de un porcentaje de distancia hasta la mediana y modelo de media Ponderada.....	155
Figura 5.61 : Parámetros α con los que se obtiene el menor error NMAE, con método de selección de un porcentaje de distancia hasta la mediana y modelo de Media Ponderada.....	155
Figura 5.62 : Comparación de errores para ponderación inversa y decaimiento exponencial con la distancia, con método de selección de un porcentaje de datos y modelo de Regresión Local Ponderada	156

<i>Figura 5.63 : Parámetros α con los que se obtiene el menor error NMAE, con método de selección de un porcentaje de datos y modelo de Regresión Local Ponderada.....</i>	<i>156</i>
<i>Figura 5.64 : Dispersiones e histogramas, media y mediana de las distancias para situaciones de baja y alta potencia prevista.....</i>	<i>157</i>
<i>Figura 5.65 : Comparación de errores para ponderación con decaimientos exponencial sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada.....</i>	<i>158</i>
<i>Figura 5.66 : Parámetros α con los que se obtiene el menor error NMAE, para ponderación con decaimiento exponencial sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada</i>	<i>158</i>
<i>Figura 5.67 : Comparación de errores para ponderación inversa sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada</i>	<i>159</i>
<i>Figura 5.68 : Parámetros α con los que se obtiene el menor error NMAE, para ponderación inversa sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada.....</i>	<i>159</i>
<i>Figura 5.69 : Comparación de errores para ponderación inversa sin estandarizar y estandarizada con la mediana de distancias, con modelo de Regresión Local Ponderada</i>	<i>160</i>
<i>Figura 5.70 : Parámetros α con los que se obtiene el menor error NRMSE, para ponderación inversa sin estandarizar y estandarizada con la mediana de distancias, con modelo de Regresión Local Ponderada.....</i>	<i>160</i>
<i>Figura 5.71 : Comparación de errores NMAE y NRMSE entre ponderación inversa estandarizada con la mediana y ponderación de decaimiento exponencial estandarizado con la mediana, modelo de Media Ponderada.....</i>	<i>161</i>
<i>Figura 5.72 : Comparación de errores NMAE y NRMSE entre ponderación inversa estandarizada con la mediana y ponderación de decaimiento exponencial estandarizado con la mediana, modelo de Regresión Local Ponderada</i>	<i>162</i>
<i>Figura 5.73 : Comparación de error NMAE entre función de suavizado Kernel y método de ponderación propuesto</i>	<i>163</i>
<i>Figura 5.74 : Comparación de error NRMSE entre función de suavizado Kernel y método de ponderación propuesto</i>	<i>164</i>
<i>Figura 5.75 : Valores de parámetros p y α que minimizan los errores NMAE y NRMSE, con modelos de Media ponderada y Regresión Local Ponderada</i>	<i>165</i>
<i>Figura 5.76 : Valores del parámetro λ que minimizan los errores NMAE y NRMSE, con modelos de Media ponderada y Regresión Local Ponderada.....</i>	<i>166</i>
<i>Figura 5.77 : Comparación de los errores de predicción sin ponderación por antigüedad y con ponderación por antigüedad.....</i>	<i>167</i>
<i>Figura 5.78 : Coeficiente de mejora o disminución de los errores NMAE y NRMSE con el uso de ponderación por antigüedad de los datos.....</i>	<i>168</i>
<i>Figura 6.1 : Parámetros p y α que proporcionan el menor error NRMSE en distintos periodos.....</i>	<i>172</i>
<i>Figura 6.2 : Parámetros p, α y λ que proporcionan el menor error NRMSE en el periodo de julio a septiembre de 2007.....</i>	<i>184</i>
<i>Figura 6.3 : NRMSE para distintos valores de memoria asintótica en la estimación evolutiva de parámetros, con modelo de Media Ponderada.....</i>	<i>186</i>

Índice de Figuras

Figura 6.4 : NRMSE para distintos valores de memoria asintótica en la estimación evolutiva de parámetros, con modelo de Regresión Local Ponderada	186
Figura 6.5 : Disminución de los errores NMAE y NRMSE medios en función del periodo de actualización de los vectores de parámetros y el cubo de errores	187
Figura 6.6 : Evolución del valor central de los vectores de parámetros p , α y λ durante el periodo de octubre a diciembre de 2007	188
Figura 7.1 : Procedimiento de combinación en dos etapas	198
Figura 7.2 : Combinación en dos etapas de predicciones basadas en modelos de Media Ponderada y Regresión Local Ponderada	200
Figura 7.3 : Coeficiente de mejora o disminución del error NRMSE en función de la memoria asintótica del modelo de combinación	201
Figura 7.4 : Combinación en dos etapas de predicciones basadas en modelos de Media Ponderada y Regresión Local Ponderada, obtenidas con una segunda fuente de viento ...	203
Figura 7.5 : Combinación en dos etapas de predicciones obtenidas a partir de distintas fuentes de viento	204
Figura 7.6 : Combinación de predicciones obtenidas a partir de distintas medidas de distancia entre vectores de viento	205
Figura 7.7 : Combinación de predicciones obtenidas a partir de distintos métodos de selección de datos cercanos.....	206
Figura 7.8 : Combinación de predicciones obtenidas a partir de modelos que utilizan distintos métodos de suavizado	207
Figura 8.1 : Error de las predicciones obtenidas con la agencia de Viento A, con modelos de Media Ponderada y Regresión Local Ponderada, y estimaciones evolutiva y no evolutiva de hiperparámetros	213
Figura 8.2 : Error de las predicciones obtenidas con la agencia de Viento B, con modelos de Media Ponderada y Regresión Local Ponderada, y estimaciones evolutiva y no evolutiva de hiperparámetros	214
Figura 8.3 : Error de las predicciones obtenidas con la agencia de Viento A, con modelos de Media Ponderada y Regresión Local Ponderada, y la combinación de ambos.....	216
Figura 8.4 : Error de las predicciones obtenidas con la agencia de Viento B, con modelos de Media Ponderada y Regresión Local Ponderada, y la combinación de ambos.....	216
Figura 8.5 : Error de las predicciones obtenidas con las agencias de Viento A y Viento B, y la combinación de ambas.....	217
Figura 8.6 : Comparación del modelo propuesto con Sipleólico y otra agencia de predicción de potencia agregada.....	218
Figura 8.7 : Mejora del modelo propuesto frente a Sipleólico	219
Figura 9.1 : Desviación estándar del error de predicción en función del nivel de potencia	229
Figura 9.2 : Posibles formas que puede adquirir el PIT	234
Figura 9.3 : Ajuste del PIT de la estimación de la función de densidad en función del ancho de banda.....	238
Figura 9.4 : Histogramas PIT de la estimación de la función de densidad para varios anchos de banda.....	239

Índice de Figuras

<i>Figura 9.5 : Ajuste del PIT de la estimación de la función de densidad en función de la memoria asintótica de las observaciones de error</i>	240
<i>Figura 9.6 : Histogramas PIT de la estimación de la función de densidad para varias memorias asintóticas de las observaciones de error</i>	241
<i>Figura 9.7 : Ajuste del PIT de la estimación de la función de densidad condicionada al nivel de potencia prevista en función del ancho de banda</i>	243
<i>Figura 9.8 : Ajuste del PIT de la estimación de la función de densidad condicionada al nivel de potencia prevista en función del porcentaje de datos seleccionados</i>	244
<i>Figura 9.9 : Ajuste del PIT de la estimación de la función de densidad condicionada al vector de viento previsto en función del parámetro α_{prob}</i>	245
<i>Figura 9.10 : Histogramas PIT para los distintos métodos de estimación de la función de densidad expuestos</i>	246
<i>Figura 9.11 : Tamaño medio de los intervalos de predicción en función de su cobertura</i>	247
<i>Figura 9.12 : Desviación estándar de los intervalos de predicción en función de su cobertura</i>	248
<i>Figura 9.13 : Desvío de la cobertura de los intervalos de predicción en función de su cobertura nominal</i>	249

Índice de Tablas

<i>Tabla 5.1 : Porcentaje de predicciones no resueltas por falta de datos con Kernel uniforme.....</i>	<i>99</i>
<i>Tabla 5.2 : Parámetros estadísticos básicos de la serie de medidas de potencia media horaria peninsular durante el año 2007</i>	<i>113</i>
<i>Tabla 5.3 : Percentiles de medidas de potencia media horaria peninsular durante el año 2007.....</i>	<i>114</i>
<i>Tabla 5.4 : Datos seleccionados a varios niveles de potencia correspondientes a los percentiles 10, 25, 50, 75 y 90 %</i>	<i>115</i>

1. Introducción y Objetivos de la Tesis

Resumen

En este capítulo se analiza el panorama actual en cuanto a generación eólica se refiere, tanto a nivel mundial como a nivel peninsular español. Se plantea la necesidad de herramientas para su predicción a corto plazo y su utilidad para el operador del sistema eléctrico. Finalmente, se fijan los objetivos y se resume la estructura y contenido de la tesis.

1.1 Importancia de la generación eólica

El continuo desarrollo industrial de los últimos 150 años ha venido ligado a un aumento progresivo en el consumo de energía. Como contrapartida, debido a que las principales fuentes de energía que se han utilizado para hacer posible este desarrollo han sido combustibles fósiles y sus derivados, se han emitido grandes cantidades de gases de efecto invernadero, que se consideran responsables del calentamiento global que está experimentando el planeta.

Por otra parte, el continuo aumento de la demanda de energía en los países desarrollados y en vías de desarrollo está acarreando una fuerte dependencia energética para muchos países, principalmente europeos.

La fuerte dependencia energética, la presión social y la toma de conciencia por parte de los gobiernos en la lucha contra el cambio climático, han dado lugar a que se adopten marcos regulatorios que favorecen el desarrollo y el uso de recursos energéticos renovables, más limpios y sostenibles. Estos recursos renovables son obtenidos a partir de recursos naturales y desechos, e incluyen, entre otras, la energía hidráulica, eólica, solar, biomasa, geotérmica, y el aprovechamiento de los residuos sólidos urbanos e industriales.

Fruto de estos marcos regulatorios, a nivel europeo se planteó en el año 2001 un ambicioso objetivo general [1], consistente en aumentar hasta un 12% la aportación de las fuentes de energía renovables en la energía primaria demandada en la Unión Europea para el año 2010, aumentando este valor hasta un 22% con respecto a la energía eléctrica consumida. La electricidad se convierte así en una forma de energía clave para la integración de las energías renovables.

Pero además de los beneficios medioambientales, el impulso a las energías renovables trae consigo otros beneficios. Por un lado, debido a la gran dispersión en la ubicación de las instalaciones generadoras que utilizan recursos renovables, su fomento ha propiciado la creación de empleos locales, el desarrollo regional y un sustancial rendimiento económico para los municipios en los que están instaladas las plantas generadoras. Por otro lado, este tipo de generación dispersa, que en la mayoría de los casos es de potencias relativamente reducidas, conlleva otras ventajas, como la disminución de las pérdidas por transporte en la red, gracias a que la generación se acerca más al consumo [2],[3].

El fuerte impulso que se ha dado a las fuentes de energía de origen renovable ha propiciado que en los últimos años se haya producido un gran aumento de la potencia instalada, y entre ellas, la energía eólica ha sido la que ha tenido un mayor auge, debido principalmente a la gran madurez de la tecnología, resultando en un coste por megavatio instalado cada vez menor. Por otra parte, la gran ventaja de la energía eólica, que ha impulsado en mayor medida su desarrollo respecto de otras fuentes de energía renovables, es su relativamente elevada disponibilidad geográfica, pues en mayor o menor medida hay corrientes de viento en casi cualquier región del planeta.

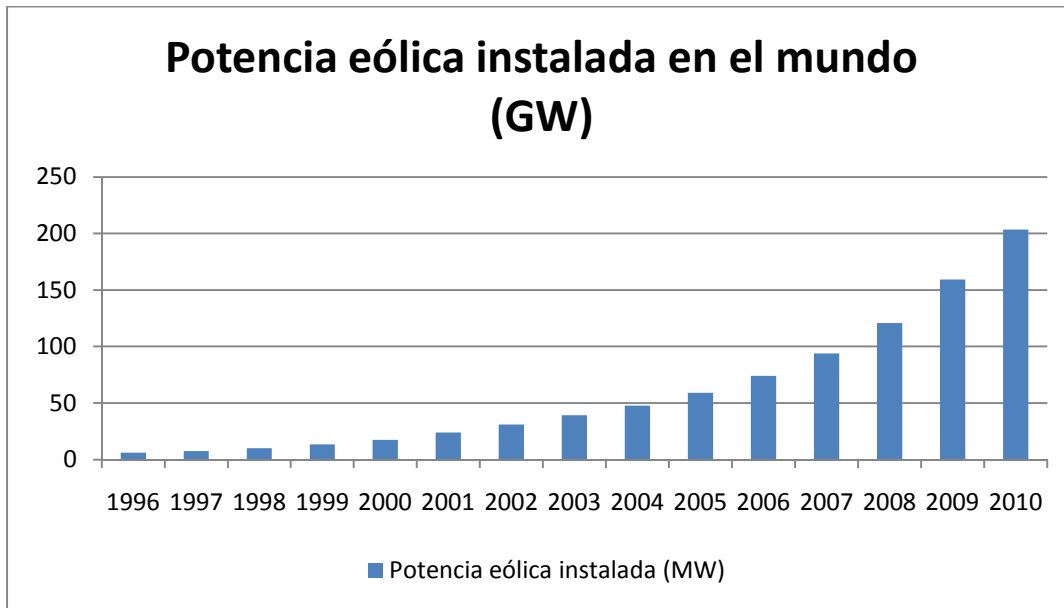


Figura 1.1: Evolución de la potencia eólica instalada en el mundo

Como contrapartida, la generación eólica conlleva ciertos inconvenientes. Por un lado, la gran variabilidad del viento, y por tanto de la generación eólica, complica su integración en el sistema eléctrico. Además, algunos generadores pueden desconectarse del sistema ante perturbaciones como los huecos de tensión, descensos bruscos en la tensión que tienen lugar cuando se produce un cortocircuito en el sistema [4],[5]. Estas perturbaciones pueden provocar una pérdida súbita de generación eólica en el sistema, que puede poner en riesgo la seguridad del suministro. Sin embargo, la reciente incorporación de mejoras tecnológicas en el comportamiento de los aerogeneradores frente a perturbaciones en la red comienza a permitir un alto grado de penetración de la energía eólica sin comprometer en demasía la seguridad del abastecimiento eléctrico.

Gracias a las mejoras en el comportamiento de los aerogeneradores, a la continua bajada en el coste por megavatio instalado de potencia eólica, y a los avances en los métodos de apoyo a la programación y gestión de esta energía, la eólica se ha convertido en la fuente de energía renovable que más se ha desarrollado en los últimos años [6], pasando de los 7.5 GW instalados en 1997 en todo el mundo hasta los 160 GW instalados en 2009, esperándose alcanzar los 203 GW a finales de 2010 y hasta los 330 GW sobre 2013 (Figura 1.1). Esto ha llevado a la eólica a representar el 2% de la energía eléctrica generada en todo el mundo durante 2009 [7].

1.2 La energía eólica en el sistema eléctrico español

En lo que respecta al caso español, y dentro del marco regulatorio que impulsa las energías renovables en España mediante el Plan de Energías Renovables en España (PER) 2005-2010 [8], se marcaron como objetivos que el 12,1% del consumo de energía primaria en el año 2010 sea abastecido por las energías renovables y que la producción eléctrica con fuentes renovables sea del 30,3% del consumo bruto de electricidad. Destaca la importante contribución prevista de la energía eólica, que eleva hasta 20.155 MW el objetivo de potencia instalada en 2010, con una producción estimada de 45.000 GWh para este año.

Además de la legislación favorable a conseguir una mayor penetración eólica, tanto a nivel nacional como regional, los factores que propician un mayor impulso del sector eólico en España son varios. Por una parte, aún existe un amplio potencial eólico en nuestro territorio todavía sin explotar, pero además, el eólico es un sector maduro que en España cuenta con amplia experiencia, capacidad y competitividad tecnológica.

De este modo, entre las ventajas del aprovechamiento del recurso eólico en España, y además de destacar la contribución a la reducción de emisiones contaminantes, la eólica ha contribuido a la diversificación del balance energético, al progreso de la industria nacional, a la creación de nuevas empresas y a la consolidación de empleo en el entorno rural.

Al finalizar el año 2009, con 18 GW instalados, España se situaba como la cuarta potencia eólica mundial y segunda europea en términos de potencia eólica instalada, tras Estados Unidos con 35 GW, Alemania con 26 GW, y China, que se ha convertido en el país en el que se ha registrado mayor incremento de eólica en los últimos años, pasando de los 6 GW instalados en 2007 a los 26 GW instalados a finales de 2009 [6],[7].

La potencia eólica instalada en el sistema eléctrico español a finales del año 2009 alcanzó los 18.119 MW [9], con más de 600 parques instalados en el territorio peninsular. Si tenemos en cuenta que durante los últimos 6 años la potencia eólica instalada en España ha crecido a razón de aproximadamente 2.000 MW al año, el objetivo del PER 2005-2010 de llegar a los 20.155 MW será previsiblemente alcanzado a finales del presente año 2010 (Figura 1.2).

En cuanto a capacidad de generación, la eólica representó el 19% de la potencia eléctrica total instalada en España en el año 2009 (Figura 1.3), incrementándose desde años anteriores en los que la potencia instalada fue del 17% o el 16% en los años 2008 y 2007 respectivamente.

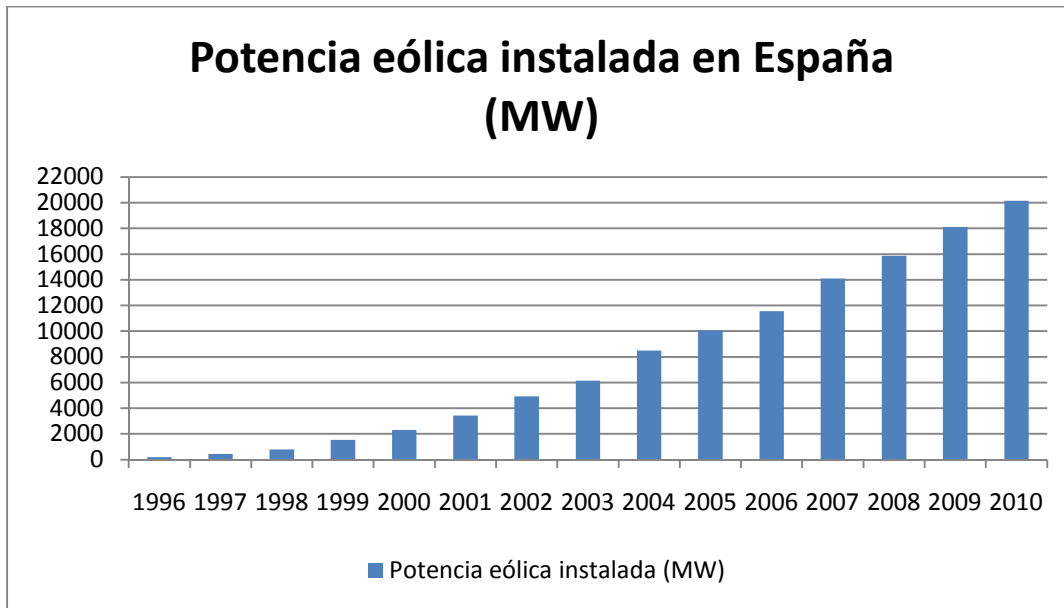


Figura 1.2: Evolución de la potencia eólica instalada en España

Entretanto, las previsiones auguran que buena parte del incremento esperado a medio plazo en la participación de las renovables en el mix de energía sea debida al gran aumento de la potencia eólica instalada. En España, el potencial neto total, sumando tierra firme y dominio marítimo de nuestras costas, es superior a los 40.000 MW [8], muy por encima de las máximas expectativas de instalación de potencia eólica hasta el año 2010.

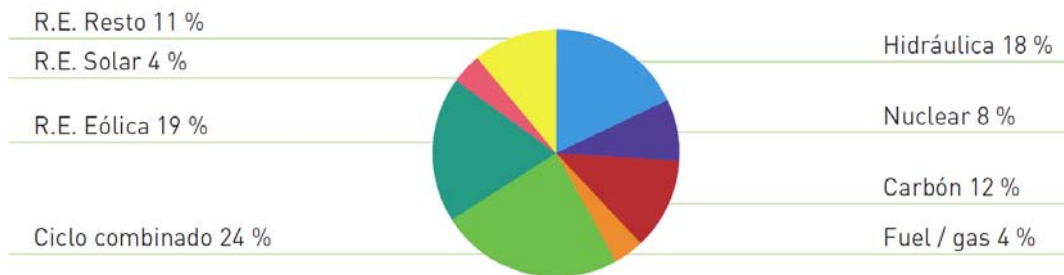


Figura 1.3: Potencia instalada a 31 de diciembre de 2009. Fuente: REE. Avance del informe 2009

Respecto a la cobertura de la demanda, las energías renovables en nuestro país cubrieron durante 2009 el 26% de la demanda eléctrica, creciendo desde un 24% en 2008, y el objetivo para 2010 es que sea del 30,3%. La aportación propia de la eólica ha sido de un 13% en 2009 (Figura 1.4), creciendo desde un 11% en 2008 [10]. El ascenso en la potencia instalada de energías renovables ha contribuido a reducir las emisiones de CO₂ del sector eléctrico español, que en 2009 se redujeron un 15,5 % respecto del año anterior [9].

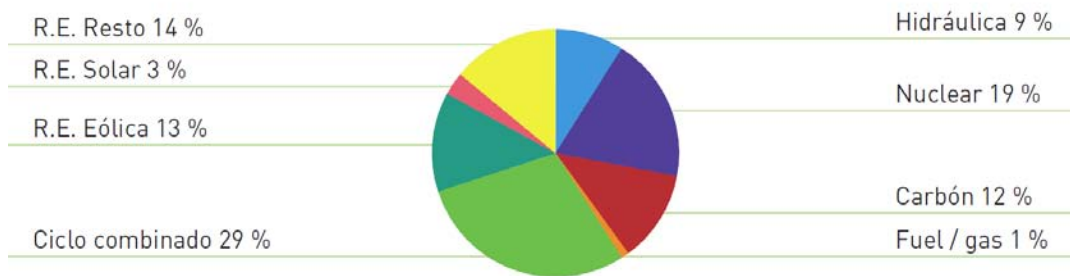


Figura 1.4: Cobertura de la demanda, año 2009. Fuente: REE. Avance del informe 2009

En términos generales, se espera un elevado incremento en la contribución de todas las energías renovables en el sector eléctrico español, hasta el punto de que la aportación a la demanda eléctrica pueda superar el 42% para el año 2020, lo que supondría un 22% del total de energía primaria consumida en nuestro país [11],[12].

Un dato que da una idea de la importancia que está adquiriendo la energía eólica en España es que en noviembre de 2009 la energía eólica cubrió un 22,7% de la demanda del mes [9]. Como contrapartida, la variabilidad que caracteriza la energía eólica ha dado lugar a situaciones extremas, como la producida el 8 de noviembre en la que el 44,9% de la demanda de ese día fue cubierta por generación eólica, llegándose durante algunos momentos de la madrugada al 53,7%, mientras que el 27 de agosto de 2009 durante la mañana la eólica apenas cubría el 1%.

Estos extremos pueden llegar a provocar vulnerabilidades en el sistema. La capacidad de interconexión con los países vecinos ayuda a mejorar el aprovechamiento de las energías renovables, pues ayudan a responder a las variaciones inevitables de este tipo de energía. Sin embargo, desde un punto de vista eléctrico, el sistema español se puede considerar casi aislado, debido a que tiene unas interconexiones relativamente débiles con el resto del sistema eléctrico europeo. Esto lo convierte en un sistema aún más vulnerable ante perturbaciones como los huecos de tensión o variaciones bruscas en la generación, y hace que sea más importante si cabe conocer y gestionar con mayor cautela la generación eólica. En un sistema poco interconectado como el nuestro cobra especial importancia una buena predicción de generación eólica.

Una falta de previsión puede acarrear problemas de sobreproducción no prevista, que debe ser compensada con reserva a bajar o incluso con la reducción o limitación de generación eólica. Para facilitar este cometido, Red Eléctrica de España, el operador del sistema eléctrico español que opera y coordina la generación y transporte en el sistema, cuenta con la herramienta de previsión eólica SIPREÓLICO [13], que será descrita en detalle más adelante. Además, se ha creado un centro específico de control de la generación de energías renovables, el CECRE [14]. En este centro de control se supervisan y controlan los generadores de régimen especial, principalmente eólicos, recibiendo información sobre las unidades de producción en tiempo real y calculando, a partir de esa información, la producción eólica máxima que puede integrarse en cada

momento en el sistema eléctrico en función de las características de los generadores y del estado del sistema.

El operador del sistema y el centro de control de energías renovables se han convertido así en agentes clave para facilitar el desarrollo de la energía eólica en España, integrándola en el sistema eléctrico en las mayores condiciones de seguridad, consiguiendo así que la producción de energía eólica siga creciendo a un ritmo elevado y que esta tecnología ocupe un papel cada vez más importante en la cobertura de la demanda.

1.3 Necesidad de predicción de energía eólica

Un problema de predicción estadística consiste básicamente en analizar los valores pasados de una variable y de otras variables relacionadas para buscar patrones significativos, con el objetivo de poder conocer o extrapolar los valores que tomará dicha variable en el futuro.

En las últimas décadas ha tenido lugar un gran avance en los sistemas de predicción, debido al enorme incremento en potencia de cálculo de los ordenadores actuales, que permiten almacenar, analizar y relacionar grandes cantidades de variables y sus valores en un tiempo muy reducido. Es por ello que los algoritmos de predicción son una herramienta muy desarrollada y utilizada hoy en día.

La importancia de las predicciones radica en la ayuda que aportan para planificar y anticiparse a los valores futuros que afectarán a un sistema, ayudando a gestionar la adquisición de los recursos necesarios con antelación suficiente, o sirviendo como herramienta para maximizar la rentabilidad mediante la toma de decisiones que maximicen los beneficios de una actividad.

Los primeros esfuerzos por crear algoritmos de predicción en el ámbito de los sistemas eléctricos se dedicaron a la previsión de la demanda eléctrica [15]. Desde el punto de vista del operador del sistema eléctrico o TSO, las previsiones de demanda han ayudado a la toma de decisiones como la puesta en funcionamiento con antelación suficiente de unidades de producción eléctrica o la programación del mantenimiento de elementos del sistema eléctrico. Desde el punto de vista de los agentes de mercado, la previsión de demanda ha sido una herramienta utilizada para optimizar sus ofertas de energía al mercado eléctrico con el objetivo de maximizar la rentabilidad.

Pero el gran auge que han tenido las fuentes de energía renovables durante los últimos años ha obligado a plantearse un nuevo problema de previsión, el de conocer con la antelación suficiente la energía que van a generar estas centrales. Dado que el mayor auge en el ámbito de las energías renovables lo ha experimentado la energía eólica, es en este sector en el que se han venido aplicando mayores esfuerzos por crear herramientas de predicción de energía fiables y eficientes, que ayuden a integrar esta forma de energía en la red [16],[17].

La eólica es una forma de generación no programable, ya que solo se produce energía cuando sopla el viento, que puede llegar a ser muy variable incluso en el corto plazo, con posibilidad de intermitencia y grandes cambios en intervalos cortos de tiempo. Por ello, es difícil conocer con antelación y precisión suficiente la cantidad de energía eólica con la que podremos contar en cada momento. Esta variabilidad hace especialmente compleja su operación, por lo que su producción futura tiene que ser estimada o prevista, viniendo esta previsión de potencia futura inevitablemente afectada por un error o incertidumbre de predicción.

Si el viento disminuye, la potencia generada en los parques eólicos también disminuye, y esa falta de potencia debe ser reemplazada por otras fuentes de generación con una reserva suficiente en magnitud y velocidad de respuesta para que la demanda eléctrica no se vea afectada. En otras ocasiones, puede ocurrir que no se pueda integrar en el sistema toda la producción eólica disponible, ya que la energía eólica no se genera de acuerdo a las necesidades de consumo, y sea necesario reducir el suministro de esta fuente de energía. Por todo esto, la predicción de generación eólica se ha convertido en un tema clave para hacer factible el desarrollo e implantación de la energía eólica, y su integración en el sistema eléctrico.

Desde el punto de vista de generación eólica, o de cualquier otra fuente de energía renovable, su previsión resulta útil tanto para el operador del sistema como para los agentes del mercado o los propietarios de parques. Así, el operador del sistema eléctrico necesita conocer con antelación suficiente la cantidad de energía eólica que será inyectada en la red para gestionar la potencia que deberán generar las centrales convencionales, con el objetivo de cubrir la demanda total del sistema. Entretanto, los agentes de mercado estarán interesados en conocer con la mayor certeza posible la potencia que generarán sus parques eólicos con el objetivo de seguir las estrategias que resulten más rentables en el mercado de energía eléctrica. Además, los propietarios de parques eólicos también estarán interesados en conocer en qué periodos se esperan menores potencias generadas en sus instalaciones, para afrontar labores programadas de mantenimiento.

El valor de la previsión de generación eólica en términos económicos tiene dos perspectivas. Por un lado, tenemos la reducción de costes de operación en el sistema originada por la reducción de reserva necesaria. Por otro lado están las posibles penalizaciones económicas que se aplican a los agentes, debidas a los desvíos en sus compromisos de generación adquiridos en el mercado de energía eléctrica.

En [18] se comprueba que con un modelo de predicción se pueden reducir las penalizaciones por desvíos en la oferta de venta de energía al mercado eléctrico español. El uso de predicciones eólicas contribuye así a minimizar tanto los costes de operación del sistema como a maximizar los beneficios o minimizar las penalizaciones de los agentes de mercado.

1.4 Importancia de las predicciones para el Operador del Sistema Eléctrico

El operador de un sistema eléctrico debe mantener un equilibrio dinámico continuo entre la energía que se genera y la que se consume, ya que la energía eléctrica no se puede almacenar de forma masiva. Por tanto, se debe hacer una planificación diaria, y con antelación suficiente, de la energía eléctrica que se va a generar y el tipo de centrales que participarán en ese mix de energía, determinando además la reserva de generación necesaria para mitigar los posibles desvíos.

Hace apenas una década la potencia eólica instalada no era muy elevada, insignificante en comparación con la potencia total demandada en el sistema, de modo que la energía eólica generada se consideraba como una demanda negativa, casi despreciable para el sistema. Así, se programaba la generación con centrales convencionales de tanta energía eléctrica como fuese necesaria para cubrir la demanda prevista.

Sin embargo, en algunos países la penetración eólica en el sistema en términos de energía ha alcanzado valores entre el 15% y el 20%. Conforme ha ido aumentando la potencia eólica instalada, la consideración de la eólica como una ligera perturbación negativa en la demanda ha adquirido valores cada vez más significativos, que afectan de forma decisiva en la curva de demanda de energía eléctrica prevista, alejándola de la demanda real [19].

En aquellos países en los que la potencia eólica instalada ha adquirido valores importantes, se ha optado por desligar la aportación eólica de la previsión de demanda, y se ha hecho necesaria por tanto una predicción específica para conocer la potencia eólica que habrá en el sistema a corto plazo.

La predicción eólica se hace imprescindible para los operadores del sistema eléctrico por los siguientes motivos:

- Gestión de congestiones. La predicción eólica ayuda a identificar posibles cuellos de botella en elementos de la red en la evacuación de este tipo de energía, y programar las centrales que ayuden a descongestionarlos. Estas congestiones pueden producirse tanto en las líneas de un sistema eléctrico como en las interconexiones entre sistemas de distintos países [20].
- Identificación de posibles inestabilidades en la red, como desconexiones masivas inadmisibles debidas a huecos de tensión.
- Estimación de reservas. Dada la variabilidad del viento, la generación eólica no ofrece garantía de potencia, lo que obliga a disponer de una reserva de potencia adicional de otras fuentes de generación para hacer frente a una posible ausencia de producción eólica.
- Planificación de intercambios internacionales con sistemas eléctricos vecinos, teniendo en cuenta la disponibilidad o exceso de generación eólica que no se espera consumir.

Por todo ello, se puede concluir que no toda la generación eólica se puede integrar en la red. La producción eólica estará limitada por las posibles congestiones o sobrecargas en las líneas provocadas por la evacuación de la generación eólica, la estabilidad del sistema frente a posibles desconexiones masivas ante perturbaciones como los huecos de tensión, la viabilidad en los balances de potencia para cubrir la demanda teniendo en cuenta las restricciones técnicas de las unidades gestionables, y los excedentes de generación eólica no integrables en el sistema.

El hecho de que el operador del sistema disponga de un sistema de predicción de la generación eólica que permita estimar la energía horaria que se generará en los parques eólicos durante las siguientes horas permite disminuir los desvíos en los programas de generación y reserva y minimizar el impacto en la operación, haciendo que ésta sea más fiable y segura.

La predicción ayuda por tanto a avanzar hacia un mayor grado de penetración de la generación eólica en el sistema eléctrico, que sea compatible con la seguridad de operación y que permita integrar la mayor cantidad de energía renovable, eólica en este caso, en el sistema. Los avances en los modelos de predicción contribuyen así a lograr una mayor penetración de la energía eólica en el balance energético.

1.5 Objetivos de la tesis

El objetivo fundamental que persigue esta tesis consiste en desarrollar un modelo de predicción a corto plazo de la generación agregada de energía eólica en una región con varios parques eólicos instalados.

Dada la habitual carencia en la disponibilidad de datos específicos, como las características individuales de cada parque y las telemidas de potencia o viento real registrado en cada parque eólico, se desea un sistema de predicción que necesite la menor cantidad de datos posible. En este sentido, el modelo de predicción propuesto necesitará únicamente un conjunto histórico de predicciones de velocidad de viento en algunas coordenadas distribuidas por la región, y las correspondientes medidas de potencia eólica total registrada. Con esta información, el modelo podrá proporcionar predicciones de potencia eólica total en la región cada vez que reciba un nuevo vector de predicciones de viento en las coordenadas.

El algoritmo de predicción se basará en dos métodos no paramétricos de modelado local por suavizado, la Media Ponderada y la Regresión Local Ponderada, basándose para el cálculo en medidas de distancia entre vectores de predicción de viento. Para llevar a cabo el suavizado de los datos, se propondrán y evaluarán varias formas de medir la distancia entre vectores de viento, diferentes procedimientos de selección de datos cercanos, y varias funciones de ponderación en base a la distancia y a la antigüedad de los datos. Uno de los trabajos principales de la tesis consistirá precisamente en evaluar todas estas alternativas, y seleccionar aquellas que se adapten mejor al problema de predicción regional eólica.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Con el objetivo de que el algoritmo de predicción pueda ser implementado en un entorno real de funcionamiento sin supervisión, se propondrá un método de estimación adaptativa de los parámetros del modelo de predicción de forma automática. Frente a una estimación global de parámetros con un periodo de entrenamiento, esta estimación adaptativa también puede mejorar las predicciones, gracias a la adaptación del modelo a la evolución de las condiciones climáticas a lo largo del año, o a los cambios en las características de los parques instalados en la región.

Debido a la complejidad del problema de predicción, y a las ventajas e inconvenientes que proporcionan las distintas configuraciones alternativas de modelado local, se establecerá un proceso de combinación en dos etapas de las predicciones obtenidas con los distintos modelos alternativos. Este proceso de combinación también podrá ser utilizado en el caso de disponer de varias fuentes de predicción de viento como entradas al modelo.

Habitualmente, los modelos de predicción de potencia eólica suelen dar un valor único de potencia prevista para cada horizonte de predicción. Frente a este enfoque de predicción puntual, las predicciones probabilistas proporcionan mayor información, ya que dan una idea de cómo será la incertidumbre de la potencia esperada. En esta tesis se propondrá un método para obtener la función de distribución de las predicciones, basado en la estimación por suavizado de la función de densidad de probabilidad condicionada a las distancias entre vectores de predicción de viento. El modelo de predicción probabilista se basará en la idea de que las distancias entre vectores pueden proporcionar una estimación de la incertidumbre, asociada a la inestabilidad de la situación atmosférica, representada en este caso por el vector de predicciones de viento en varias coordenadas de la región. A partir de la función de distribución calculada se podrá obtener de forma inmediata cualquier percentil o intervalo de predicción.

1.6 Estructura y contenido de la tesis

Para introducir al lector en el problema a tratar, el capítulo 2 presenta el estado del arte en materia de predicción de generación eólica, comenzando por los modelos de previsión de viento y exponiendo a continuación las distintas aproximaciones, físicas y estadísticas, para la previsión de energía eólica. Se describen las características de algunas de las herramientas utilizadas en la actualidad para este fin y se enumeran varias aplicaciones para las que resulta valiosa la predicción de energía eólica y la estimación de su incertidumbre. Por último, se argumenta la utilidad y se resume el estado del arte en el ámbito de las predicciones regionales agregadas.

En el capítulo 3 se exponen las bases del modelado local por suavizado, describiendo los modelos locales de Media Ponderada y de Regresión Local Ponderada, así como sus ventajas e inconvenientes dependiendo de la complejidad del problema de predicción a tratar. En base a la estructura de este tipo de modelos locales, se presenta el algoritmo de predicción regional de generación eólica propuesto en esta tesis,

basado en la medida de distancias entre vectores de predicción de viento. Por último, se describen las medidas de error que serán utilizadas para evaluar el funcionamiento del modelo de predicción propuesto.

En el capítulo 4 se describen las funciones Kernel de suavizado clásicas, utilizadas comúnmente en el modelado local. Frente al enfoque de utilizar estos Kernel clásicos, se propone un procedimiento adaptado al problema que nos ocupa, en el que se especifican de forma independiente la medida de distancia, el proceso de selección de datos en función de su distancia al punto a predecir, y la función de ponderación de los datos seleccionados en función de su distancia y su antigüedad.

En el capítulo 5 se hace una evaluación exhaustiva de las distintas propuestas descritas en el capítulo 4, utilizando datos de potencia eólica total medida y predicciones de viento repartidas en algunas coordenadas del sistema eléctrico peninsular español. El objetivo de esta evaluación es establecer cuáles son la medida de distancia, el método de selección, y las funciones de ponderación más adecuadas para afrontar el problema de predicción regional eólica con un modelo de suavizado local.

El capítulo 6 se destina a la estimación adaptativa de los parámetros del modelo. Con el fin de que el modelo de predicción propuesto pueda estar en funcionamiento en un entorno real y sin supervisión experta, se propone un método para estimar de forma adaptativa los parámetros óptimos de selección y ponderación por distancia y antigüedad de los datos, de forma que el modelo de predicción se adapte automáticamente a los cambios en la relación entre viento y potencia, y a la evolución estacional de las condiciones meteorológicas. Al final del capítulo se comprueba la bondad de los resultados obtenidos con la estimación adaptativa de parámetros frente a un modelo con parámetros fijados con un periodo de ensayo.

En el capítulo 7 se describe y comprueba un proceso de combinación de predicciones, debido a que se tienen dos modelos de predicción alternativos, basados en la Media Ponderada y la Regresión Local Ponderada, con ciertas ventajas e inconvenientes cada uno, que proporcionan sendas predicciones de alternativas. El proceso de combinación se realiza en dos etapas, aprovechando las ventajas de dos enfoques clásicos de combinación: la combinación de promedio óptimo, y la combinación de mínimo riesgo. Se evalúa al final la utilidad de este proceso de combinación en dos etapas, tanto para el problema de combinación de modelos de predicción, como para combinar predicciones calculadas con distintas fuentes de predicción de viento.

En el capítulo 8 se realiza una evaluación del modelo de predicción propuesto utilizando datos de un periodo distinto al usado en capítulos anteriores, con el fin de comprobar los resultados fuera de la muestra de entrenamiento. Las predicciones del modelo agregado propuesto se comparan finalmente con las predicciones de la herramienta Sipreólico y de otro modelo de predicción de potencia agregada.

En el capítulo 9 se propone un modelo de predicción probabilista, basado en técnicas no paramétricas de estimación de densidades con Kernel. Se plantea un método de estimación de densidad condicionada a la distancia entre vectores de velocidad de

Métodos de Predicción de la Generación Agregada de Energía Eólica

viento, concepto en el que se ha basado el modelo de predicción de potencia propuesto en esta tesis. Por último, se evalúa la utilidad de las predicciones probabilistas obtenidas frente a las que proporcionan otros modelos no condicionados o condicionados únicamente al nivel de potencia prevista.

Finalmente, el capítulo 10 resume las conclusiones generales y aportaciones originales de este trabajo de investigación, y se identifican y argumentan posibles perspectivas de investigación futura.

2. Estado del Arte en Predicción de Energía Eólica

Resumen

En este capítulo se presentan las bases de la predicción de energía eólica. Se comienza planteando la dependencia existente entre el viento y la potencia que genera una turbina eólica, y se relaciona la complejidad de los modelos de predicción meteorológica con la dificultad de calcular predicciones de potencia.

Se incluye una revisión de los distintos enfoques y modelos existentes de predicción de potencia eólica y se argumenta el valor económico de las predicciones de generación eólica. También se exponen algunos desarrollos recientes en el área de la estimación de incertidumbre en la predicción. Finalmente se justifica la utilidad y estado del arte en el ámbito de las predicciones regionales agregadas.

2.1 Introducción a la conversión de energía eólica

El viento es una masa de aire en movimiento, y como tal, contiene cierta cantidad de energía cinética, que depende de la velocidad de la masa de aire. Esta última, vendrá determinada por la temperatura y presión atmosféricas o, en definitiva, por la densidad del aire:

$$E = \frac{1}{2} m_{air} v^2 \quad (2.1)$$

donde m_{air} es la masa del aire y v es su velocidad.

Un aerogenerador convierte en energía eléctrica la energía cinética contenida en el viento. Teniendo en cuenta la energía cinética contenida en una masa de aire en movimiento que atraviesa un área determinada por el diámetro del rotor de un aerogenerador, se puede llegar a la relación entre la velocidad del viento y la potencia eléctrica entregada por el aerogenerador como la energía que lo atraviesa por unidad de tiempo:

$$p_e = \frac{1}{2} C_p \rho_{air} A_r v^3 \quad (2.2)$$

donde p_e es la potencia eléctrica entregada por el aerogenerador, ρ_{air} es la densidad del aire, A_r es el área que atraviesa, definida por el diámetro del rotor o el tamaño de las palas, y v es la velocidad de la corriente de aire. C_p es el coeficiente de potencia, que indica el grado de aprovechamiento de la energía contenida en la corriente de aire, con un valor teórico máximo de 0.59 según el límite de Betz, aunque su valor habitual dada la tecnología actual y las pérdidas en los elementos mecánicos es bastante inferior.

La curva habitual que describe la relación entre la potencia eléctrica entregada por un aerogenerador y la velocidad del viento incidente en la turbina viene representada en la Figura 2.1, en lo que se conoce como curva de potencia o curva P-V de un aerogenerador.

Para valores de velocidad de viento inferiores a unos 4 o 5 m/s, el aerogenerador no produce potencia. Ésta es la llamada velocidad de arranque del aerogenerador, a partir de la cual la potencia generada presenta un crecimiento cúbico con la velocidad del viento, hasta valores de viento de unos 12 a 15 m/s. A partir de esta velocidad del viento, la potencia eléctrica generada se mantiene en el valor nominal para el cual fue diseñada la máquina, o decrece levemente dependiendo de si se trata de un aerogenerador de velocidad variable o fija. Para valores de velocidad de viento a partir de 20 o 25 m/s, en lo que se conoce como velocidad de corte, se produce la parada del aerogenerador para protegerlo frente a vientos demasiado severos.

Métodos de Predicción de la Generación Agregada de Energía Eólica

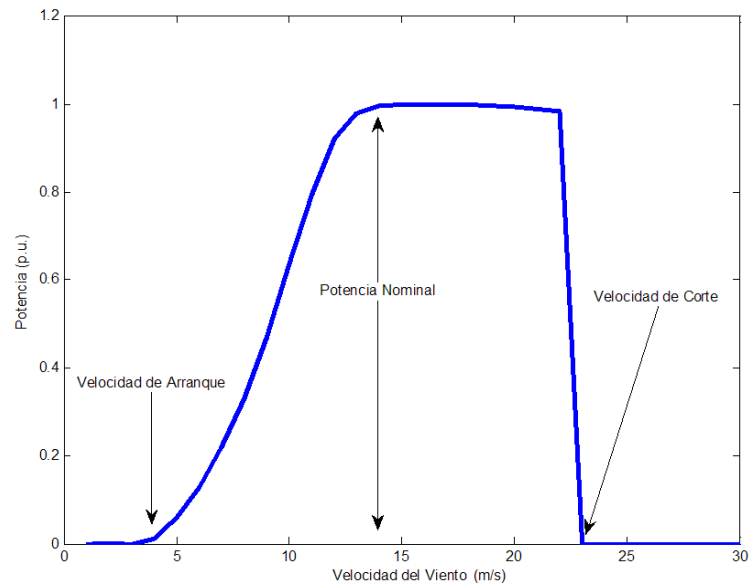


Figura 2.1: Curva potencia-viento característica de un aerogenerador

En el punto de funcionamiento habitual, correspondiente a velocidades de viento entre los 5 y 12 m/s, la curva de potencia presenta una fuerte pendiente, de modo que pequeñas variaciones de la velocidad del viento incidente en el rotor provocan grandes variaciones en la potencia eléctrica generada, es decir, la potencia eléctrica es muy sensible a los cambios de velocidad del viento.

Debido a esto y a que el viento es una variable que puede tener grandes cambios en escalas temporales muy cortas, del orden de apenas algunos minutos o incluso segundos, las variaciones en la potencia entregada por un aerogenerador a lo largo del tiempo pueden ser muy pronunciadas.

Esto último da una idea de la importancia de la precisión de las previsiones de velocidad de viento proporcionadas por las agencias meteorológicas para el correcto funcionamiento de las herramientas de predicción de potencia eólica, como veremos más adelante.

2.2 Modelos de previsión de viento

Se pueden encontrar dos clases de modelos de predicción de potencia eólica atendiendo a si usan o no predicciones meteorológicas (Numerical Weather Predictions o NWP). Normalmente, para calcular predicciones de generación eólica es necesario disponer previamente de predicciones meteorológicas, salvo para los modelos de predicción de potencia basados exclusivamente en extrapolar la tendencia de la producción eólica.

Así, nos encontramos con que habitualmente para la obtención de predicciones de energía eólica a corto plazo entran en juego dos sistemas, el modelo atmosférico de previsión meteorológica, y el propio modelo de predicción de potencia a partir de las previsiones meteorológicas.

Los modelos de previsión meteorológica se basan en la resolución de las ecuaciones físicas que describen el estado de la atmósfera, aplicadas a una serie de puntos de una malla o rejilla tridimensional que la representa y que, partiendo de unos valores iniciales de medidas del estado de la atmósfera, calculan cómo será la evolución de la misma durante las siguientes horas. Típicamente, un modelo de previsión meteorológica se ejecuta cada 6 o 12 horas, generando predicciones para las siguientes 48 o 72 horas.

Dependiendo de la escala espacial o tamaño de rejilla que utilicen los modelos, podemos diferenciar entre modelos de macro-escala, modelos meso-escalares, y modelos de micro-escala. Con el mismo esfuerzo computacional, un modelo de macro-escala podrá realizar predicciones a mayor horizonte y abarcando una región mucho más amplia que un modelo de micro-escala, aunque con mucha menos resolución que este último.

Los modelos atmosféricos de macro-escala globales utilizan una aproximación hidrostática, es decir, los movimientos verticales de las masas de aire se suponen despreciables frente a los movimientos horizontales, y por lo tanto no tienen en cuenta la posible evolución vertical de la atmósfera. Al abarcar toda la superficie terrestre tienen su malla definida en coordenadas esféricas. Dos ejemplos de modelos de macro-escala globales son el GFS (Global Forecast System) del servicio meteorológico norteamericano NCEP (National Centers for Environmental Prediction) [21], y el perteneciente al centro europeo de predicción ECMWF (European Centre for Medium-Range Weather Forecasts), actualmente con resoluciones espaciales entre 0.5° y 0.25° .

Dado que las regiones de mayor interés en la predicción meteorológica no abarcan todo el planeta, también existen otros modelos de macro-escala no globales que centran su predicción en las regiones de interés. Estos modelos están limitados a una región específica, con lo que con la misma potencia de cálculo pueden tener una resolución mayor que los modelos globales. Para su ejecución, además de condiciones iniciales necesitan condiciones de contorno de la región, que son obtenidas de los modelos globales. El modelo HIRLAM (High Resolution Limited Area Model) es un ejemplo de modelo atmosférico regional, que actualmente ofrece predicciones para parte de Europa con resolución espacial entre 0.05° y 0.16° .

Debido a la limitación computacional de los ordenadores, la separación entre dos puntos consecutivos de la malla que usan los modelos atmosféricos macro-escalares como los descritos es superior al tamaño característico de un parque eólico. Dado que los puntos de la malla para los que se calcula la previsión de viento no suelen coincidir con la ubicación de los parques eólicos, la predicción de viento en el emplazamiento de un parque se obtiene mediante una interpolación entre los valores de los puntos de la rejilla más cercanos resueltos en el modelo atmosférico. Además,

Métodos de Predicción de la Generación Agregada de Energía Eólica

esta malla tampoco suele tener la resolución suficiente como para representar con un detalle adecuado la orografía del terreno, con lo que las predicciones de viento a baja altura y en terrenos complejos suelen ser bastante deficientes.

Para tratar de paliar estos inconvenientes, existen modelos de menor escala que, partiendo de las previsiones meteorológicas de un modelo de macro-escala en los puntos de la malla cercanos a la ubicación del parque eólico, vuelven a resolver las ecuaciones físicas sin usar una aproximación hidrostática, teniendo en cuenta la aceleración vertical provocada por la orografía del terreno. Esta vez se define una malla mucho más fina, localizada en el entorno del emplazamiento del parque eólico y representando con gran detalle la orografía del terreno en coordenadas cartesianas. Ejemplos de modelos meso-escalares son MM5 y RAMS [22],[23], y WAsP es un caso representativo de modelo de micro-escala [24].

Con estas técnicas se pueden mejorar y adaptar a una zona localizada las previsiones meteorológicas de un modelo atmosférico global o regional. Así, por medio de una representación más detallada de la topografía se pueden mejorar los resultados de los modelos de predicción de viento en una zona concreta [25].

Por último, se debe tener en cuenta que las condiciones iniciales de cálculo del modelo numérico de previsión meteorológica determinan la evolución de la previsión de viento. Esto se convierte en una posible fuente de error, pues unas condiciones iniciales incorrectas, consecuencia de errores en el proceso de medida o interpretación de las variables de estado de partida, darán lugar a una evolución distinta a la real, con diferencias entre los valores de las variables meteorológicas previstas y las reales.

Todas estas aproximaciones introducen fuentes de error en la predicción del viento en el emplazamiento del parque, que posteriormente se traducirán de forma amplificada en una incertidumbre en la predicción de potencia debido al crecimiento cúbico de la potencia con el viento en la zona de trabajo de los aerogeneradores, ecuación (2.2). Por ello, en general se considera que la mayor fuente de error en las previsiones de potencia eólica se debe al paso previo, el de la previsión meteorológica [26].

2.3 Revisión del estado del arte en modelos de predicción eólica

Existen dos aproximaciones básicas para la predicción de la energía eólica a corto plazo, los modelos físicos y los modelos estadísticos.

Los modelos físicos tienen en cuenta consideraciones físicas para adaptar las predicciones de viento en una zona a las condiciones concretas del emplazamiento del parque. Para hacer esta adaptación se utilizan modelos de meso-escala o micro-escala que, partiendo de las condiciones iniciales y de contorno obtenidas de un modelo atmosférico de mayor escala, calculan la velocidad del viento incidente en las turbinas del parque para posteriormente calcular la predicción de potencia por medio de la curva de potencia.

Por otro lado, de entre los modelos estadísticos podemos encontrar la familia de las series temporales, que solo utilizan valores pasados de las variables como datos de entrada del modelo, y los que además de valores pasados utilizan como entradas los valores de predicción meteorológica de modelos atmosféricos, relacionándolos con los valores de potencia histórica u otros valores históricos medidos.

El modelo de predicción podría ser una combinación de ambos, es decir, tener en cuenta consideraciones físicas para estimar con detalle el viento en el emplazamiento de las turbinas y usar modelos estadísticos avanzados que utilicen esta información localizada, además de las medidas de potencia, para minimizar el error.

A continuación, se resumen los principales aspectos de los sistemas más representativos utilizados en la actualidad para calcular predicciones de generación eólica, algunos de ellos basados en modelos físicos, otros en modelos estadísticos, y combinaciones de ambos.

Aunque en este texto se resumirán los principales sistemas de previsión de generación eólica y las líneas de trabajo actuales, pueden encontrarse varias revisiones bibliográficas sobre el estado del arte de la predicción a corto plazo, como por ejemplo en [27] y [17], o una revisión más reciente en [28]. También en [29], como etapa previa al comienzo del proyecto ANEMOS, se describieron la mayoría de los progresos y publicaciones relacionadas con el tema de predicción eólica, incluyendo tanto una relación de los modelos y líneas de investigación más utilizadas en la actualidad como otras en desuso.

2.3.1 Horizontes de predicción

El horizonte de predicción es el periodo de tiempo que determina el momento futuro para el que realizamos las predicciones, y para el tema que nos ocupa en esta tesis las clasificaremos en muy corto plazo, corto plazo, medio y largo plazo.

La predicción a muy corto plazo, hasta 4 o 6 horas de horizonte, desde el punto de vista del operador del sistema tiene como objetivo la mejora de la seguridad dinámica de la red y la gestión de la regulación secundaria. Desde el punto de vista del propietario de un parque eólico, la predicción del viento a muy corto plazo, en el orden incluso de apenas algunos minutos o segundos, resulta útil para el control de las turbinas eólicas, y por motivos de seguridad para la toma de decisiones en cuanto a paradas o arranques de los aerogeneradores.

Las predicciones a corto plazo, que abarcan horizontes hasta las 48 o 72 horas, son las que tienen mayor aplicación en la determinación de necesidades de generación para los mercados diario e intradiarios, para la previsión de reserva secundaria y terciaria, y en general la adopción de las medidas necesarias que minimicen los riesgos en la estabilidad del sistema ante posibles contingencias futuras [30].

En las predicciones a medio y largo plazo, los horizontes pueden variar entre varios días o incluso llegar a meses o años. Las predicciones a varios días tienen aplicación

Métodos de Predicción de la Generación Agregada de Energía Eólica

en la previsión de generación de centrales eléctricas convencionales o para programar labores de mantenimiento en los aerogeneradores aquellos días en los que las expectativas de generación sean escasas o nulas. Las previsiones de potencia basadas en previsiones de viento con horizontes temporales en el plazo de años tienen utilidad en la selección de emplazamientos para ubicar parques eólicos.

El horizonte que en general suscita mayor interés en la literatura por su utilidad es el corto plazo, hasta 48 o 72 horas, y será el horizonte de predicción objetivo de esta tesis. Abarca tanto las necesidades de predicción para la participación por parte de los productores en los mercados diarios e intradiarios como las necesidades del operador del sistema para su gestión. Por otro lado, la resolución habitual de las predicciones de potencia suele ser de una hora, ya que la mayoría de los sistemas y mercados eléctricos se programan siguiendo una escala horaria.

2.3.2 Modelos estadísticos de series temporales

Un modelo de predicción basado en series temporales trata de extrapolar los valores futuros que tomará una variable mediante el conocimiento y análisis de los valores pasados de dicha variable u otras variables explicativas.

Este es el enfoque propuesto entre otros en los modelos ARIMA o Box-Jenkins [31], que resulta útil para la predicción en determinados procesos industriales [32], y en el contexto de la predicción de energía eólica proporciona resultados razonablemente buenos para horizontes hasta 6 horas.

Sin embargo, el uso de modelos basados en series temporales clásicas resulta muy poco útil para predicciones a horizontes superiores, hasta las 48 o 72 horas, que son los que suscitan más interés para la operación del sistema y la participación en los mercados de energía.

Existen otras aproximaciones alternativas al modelo clásico de serie temporal, en las que se propone la predicción de potencia eléctrica en parques eólicos por medio del uso de modelos basados en redes neuronales. Las redes neuronales resultan una solución interesante debido a su flexibilidad de uso para multitud de problemas, si bien presentan el inconveniente de presentarse como una caja negra.

Como ejemplo, en [33] se utilizan redes neuronales recurrentes con varias capas dedicadas a distintos horizontes de predicción, hasta 10 horas. También se han utilizado las redes neuronales multicapa para la predicción de la velocidad del viento hasta 1 hora para su aplicación en el control de un parque eólico usando como entrada los últimos valores de la serie temporal de viento [34],[35].

En cualquier caso, ya sea mediante el uso de modelos ARIMA o con redes neuronales, las predicciones que solo tienen en cuenta los valores pasados de las variables no suelen proporcionar buenos resultados más que para el muy corto plazo, y se hace necesario contar con otras fuentes de información, como previsiones

meteorológicas obtenidas con modelos atmosféricos, para poder calcular con precisión suficiente predicciones de potencia para horizontes a partir de 4 ó 6 horas.

Modelos de referencia

En el campo de la predicción, se han definido algunos modelos sencillos que sirven como referencia a la hora de evaluar la bondad de los nuevos modelos de predicción desarrollados. Estos modelos de referencia están basados en modelos sencillos de series temporales.

Uno de los modelos de referencia más utilizados es el modelo de Persistencia. Según este modelo las predicciones de potencia a futuro, para cualquier horizonte, coinciden con el valor actual de la variable, en este caso la potencia eólica:

$$\hat{P}_{t+h|t} = P_t \quad (2.3)$$

donde $\hat{P}_{t+h|t}$ es la predicción para el horizonte de tiempo $t+h$, calculada en el instante actual t , y P_t es la medida actual de potencia registrada en el parque o la región de estudio.

Este modelo de predicción para parques solamente tiene un buen funcionamiento para predicciones hasta un horizonte muy reducido, de no más de 2 ó 3 horas, aunque si se trata de predicción regional, puede llegar a dar buenos resultados para horizontes algo mayores, hasta 4 ó 6 horas, ya que los cambios atmosféricos en una región tienen un orden temporal bastante superior al de los cambios que afectan a un parque individual.

El modelo de persistencia es un caso concreto del modelo de media móvil de las últimas n observaciones, con $n=1$

$$\hat{P}_{t+h|t} = \frac{1}{n} \sum_{\tau=0}^{n-1} P_{t-\tau} \quad (2.4)$$

siendo $P_{t-\tau}$ la medida de potencia en el instante $t-\tau$, de antigüedad τ .

Cuando el orden del modelo de media móvil tiende a infinito, entonces tenemos la media global o incondicional de todo el histórico de potencias eólicas registradas en el parque o región, también llamada media climática, lo que sería un buen modelo de predicción a muy largo plazo:

$$\hat{P}_{t+h|t} = \bar{P}_t \quad (2.5)$$

Métodos de Predicción de la Generación Agregada de Energía Eólica

Fruto de combinar la bondad de las predicciones del modelo de persistencia para el muy corto plazo, y la relativa bondad de las predicciones de un modelo de media global de potencia para predicciones a largo plazo, nace un nuevo modelo de referencia, propuesto en [36], que resulta de una ponderación de ambos métodos, cuyos coeficientes de ponderación dependen del horizonte de predicción:

$$\hat{P}_{t+h|t} = a_h P_t + (1 - a_h) \bar{P}_t \quad (2.6)$$

donde a_h es el factor de ponderación que habría que estimar, con valores entre 0 y 1 dependiendo del horizonte de predicción, y que tomará valores cercanos a la unidad para horizontes de predicción muy cortos, dando mucho peso al modelo de persistencia, y valores cercanos a cero para horizontes de predicción muy elevados, en los que la potencia esperada se aproximará más a la media de potencia registrada históricamente.

Aunque el nuevo modelo de referencia propuesto en la ecuación (2.6) es más elaborado y proporciona mejores resultados que el modelo de persistencia, en la práctica el modelo de referencia más utilizado sigue siendo el modelo de persistencia, debido a su sencillez y a que no requiere la estimación de parámetro alguno.

2.3.3 Modelos físicos

Los modelos físicos se basan en el modelado físico del emplazamiento donde se encuentran los aerogeneradores, y tratan de predecir el viento que incidirá en cada una de las turbinas para, mediante el conocimiento de las características de las turbinas y su curva de potencia, poder dar una predicción de potencia eólica.

En primer lugar, dado el viento previsto por algún modelo global o regional en las proximidades del parque eólico, se debe hacer una adaptación de la predicción de viento al emplazamiento concreto del parque usando un modelo de meso-escala o micro-escala, para estimar el viento en la posición y altura donde se encuentran las turbinas eólicas. Posteriormente, se transforma esta velocidad del viento en un valor de potencia mediante el uso de la curva de potencia propia de las máquinas instaladas.

Una de las ventajas de los modelos físicos es que para su implementación no se necesitan datos históricos, a menos que se le añada un módulo estadístico para corregir los posibles sesgos o defectos de escala. De hecho, la mayoría de los modelos físicos incorporan un módulo estadístico MOS, o Model Output Statistics, para considerar los posibles efectos que no tienen en cuenta los modelos físicos y realizar las correcciones oportunas [37]. Así pues, en la práctica casi todos los modelos físicos se podrían considerar híbridos, pues suelen incorporar un módulo estadístico de mayor o menor complejidad.

El modelo Prediktor, desarrollado por Landberg en el Risø National Laboratory de Dinamarca para la operadora del sistema eléctrico Elkraft es un ejemplo de modelo

físico [38]. Utiliza WAsP (Wind Atlas Analysis and Application Program) para convertir las predicciones meteorológicas del modelo atmosférico HIRLAM en viento sobre las turbinas de un parque, y calcula la predicción de potencia mediante el uso de la curva de potencia. El efecto de estela de las turbinas se tiene en cuenta mediante el uso de PARK, que modela este efecto basándose en la información de la posición relativa de las turbinas dentro del parque. Los demás efectos que no tienen en cuenta los modelos físicos se corrigen con un modelo estadístico o MOS, ajustando los resultados con las medidas de potencia históricas.

En la actualidad, el modelo Prediktor forma parte, junto con el modelo estadístico Wind Power Prediction Tool (WPPT), de la herramienta Zephyr [39], que tiene el objetivo de aunar las ventajas de un modelo estadístico y un modelo físico. El modelo estadístico proporciona buenos resultados para las primeras horas mientras que el modelo físico proporciona mejores resultados para horizontes mayores, incluso ante la ausencia de medidas de potencia.

Previento [40], desarrollado por Focken en la universidad de Oldenburg, es otro ejemplo representativo de modelo físico, que en este caso utiliza predicciones meteorológicas LokalModell del servicio meteorológico alemán DWD [41]. Posteriormente calcula predicciones regionales totales a partir de las predicciones individuales de algunos parques seleccionados.

En Estados Unidos, la herramienta EWind [42], desarrollada por TrueWind, usa el modelo de meso-escala ForeWind para adaptar las predicciones a condiciones locales, tomando como condiciones iniciales y de contorno las predicciones meteorológicas de un modelo de mayor ámbito [43]. Los errores sistemáticos también se atenúan con el uso de un módulo estadístico.

CENER y CIEMAT han desarrollado Localpred [44],[45], pensado para hacer predicciones en terrenos complejos y que está basado en un modelo de predicción de velocidad de viento local por la interpolación y corrección de puntos cercanos HIRLAM con regresión múltiple. Incluye predicciones a muy corto plazo basadas en modelos de series temporales, e incorpora un modelado físico del terreno en alta resolución con el modelo de meso-escala MM5 para mejorar las predicciones meteorológicas en el entorno del parque eólico y transformarlas en predicciones de potencia mediante una curva de potencia.

Por último, también se han propuesto herramientas de predicción de potencia que están basadas directamente en los modelos de previsión atmosféricos. Un ejemplo de esta integración directa del modelo de predicción meteorológica a un modelo de predicción de potencia eólica es el HIRLAM Power Prediction Model, o HIRPOM [46].

2.3.4 Modelos estadísticos de regresión dinámica

Los modelos estadísticos no usan tanta información específica de los parques eólicos como los modelos físicos, ya sea la modelización del relieve del emplazamiento donde se encuentran las turbinas o la disposición, características o curva de potencia de las mismas. Simplemente tratan de relacionar los datos de potencia histórica registrada con variables como la velocidad del viento o la dirección entre otras.

De algún modo, se pueden interpretar como modelos de regresión dinámicos que tratan de estimar los parámetros de una función que relaciona el valor futuro de potencia con una serie de variables explicativas, como puede ser el viento previsto o la potencia medida en el pasado. Los modelos estadísticos además presentan la ventaja de que permiten evaluar posibles desajustes sistemáticos, y por tanto eliminarlos.

La herramienta Wind Power Prediction Tool (WPPT) es un ejemplo de modelo estadístico, desarrollada en Dinamarca para el operador Elsam/Eltra por el Instituto de Informática y Modelado Matemático (IMM) de la universidad DTU. Inicialmente contaba con un modelo autorregresivo y daba predicciones hasta 36 horas, aunque con una calidad razonable solo hasta las primeras 12. Posteriormente, se añadieron predicciones meteorológicas HIRLAM que mejoraron las predicciones hasta un horizonte de 39 horas [47],[48].

Las predicciones de la herramienta WPPT utilizan modelos autorregresivos con entradas exógenas, con un término armónico que tiene en cuenta la estacionalidad diaria observada en la producción. Además utiliza redes neuronales y una transformación de variables con la raíz cuadrada. Los coeficientes de la serie temporal van variando conforme se van adquiriendo datos, lo que otorga al sistema WPPT capacidad de aprendizaje conforme va recibiendo datos medidos y previsiones meteorológicas [49].

Aunque WPPT calcula predicciones de potencia para parques, está pensada para poder calcular predicciones para una región completa mediante la división de la región en varias sub-áreas, calculando la predicción de potencia en ellas según dos aproximaciones [50]. La primera implica la selección y predicción de potencia de algunos parques representativos en cada sub-área, de los que se tengan medidas de potencia en tiempo real y previsiones de viento. La potencia total de cada sub-área se calculará con un reescalado de las potencias de los parques de referencia. En la segunda aproximación se calcula la potencia directamente en las sub-áreas utilizando datos off-line de potencia total medida y viento previsto. Finalmente se calcula el total regional como suma de las predicciones en las sub-áreas y se combinan las dos aproximaciones de cálculo.

Con la colaboración del ISET (Instituto de Investigación en energía Solar y Eólica en Alemania), se desarrolló el Advanced Wind Power Prediction Tool (AWPT) [51], un modelo de predicción de energía eólica en Alemania para ayudar a la compañía E.On a integrar la eólica instalada en la zona PEN (Preussen Elektra Net). El sistema se basaba en un modelo estadístico que utiliza técnicas de redes neuronales artificiales.

Utilizando datos históricos de potencia, medidas de potencia en tiempo real y previsiones de viento Lokalmodell, calcula previsiones eólicas para algunos parques representativos que se extrapolaban al total de la región.

Posteriormente, ISET ha desarrollado mejoras, con el Wind Power Management System (WPMS) [52], dando predicciones para la práctica totalidad de parques instalados en Alemania. Se mejoran las previsiones meteorológicas de LokalModell en la localización de los parques usando un modelo atmosférico de meso-escala KLIMM (KLImaModell Mainz), y se calcula la predicción utilizando una red neuronal artificial.

Otro enfoque en cuanto a modelos estadísticos se refiere es el uso de lógica borrosa o fuzzy-logic, que aunque tradicionalmente se había utilizado para control, ahora también se utiliza para predicción de demanda y de potencia eólica. Los modelos fuzzy-logic utilizan reglas del tipo “si... entonces...” para la toma de decisiones, mediante entrenamiento con largos periodos de datos.

Este es el tipo de modelo usado en el Armines Wind Power Prediction System (AWPPS), desarrollado en el ámbito del proyecto MORE-CARE [53], en el que participan la Escuela de Minas de Paris (ARMINES), y el Rutherford Appleton Laboratory de Reino Unido (RAL). Se desarrolló un modelo para predicción de potencia eólica a corto plazo basado en el uso de telemidas en tiempo real y predicciones meteorológicas. La predicción es el resultado de la combinación de un modelo para muy corto plazo basado en series temporales y modelos a corto plazo basados en redes neuronales difusas [54]. El software MORE-CARE está instalado en Creta, Madeira e Irlanda [55].

En el caso español, el operador del sistema eléctrico Red Eléctrica de España cuenta desde el año 2002 con la herramienta de predicción Sipleólico [13],[56], entre cuyos desarrolladores se encuentran el autor de esta tesis y su director. Está basada en modelos estadísticos y calcula predicciones para los parques eólicos ubicados en el territorio peninsular, utilizando como datos de entrada las medidas de potencia en los parques y predicciones de viento de una o varias agencias meteorológicas. Finalmente se pueden calcular predicciones agregadas por regiones o para el total nacional como suma de las predicciones individuales. También existe una versión de Sipleólico adaptada a las necesidades del archipiélago canario.

Se incluye a continuación una descripción más detallada de esta herramienta de predicción, debido a que se utilizarán las predicciones agregadas de potencia proporcionadas por Sipleólico para el total peninsular en la comparación con las predicciones obtenidas mediante el método de predicción agregada propuesto en esta tesis.

2.4 La herramienta de predicción SIPREÓLICO

La herramienta Sipleólico, fruto de la colaboración entre la Universidad Carlos III de Madrid y Red Eléctrica de España, está en funcionamiento on-line desde febrero de 2002, y fue desarrollada con la intención de mitigar el impacto de la generación eólica instalada en el territorio peninsular español sobre la operación del sistema eléctrico [57].

Empleando métodos estadísticos, proporciona predicciones a corto plazo, hasta 48 horas, de la producción eólica para cada una de las unidades de predicción recogidas en una tabla de parques. Para ello se vale de previsiones meteorológicas obtenidas de modelos atmosféricos y de medidas de potencia registradas en las unidades de predicción (Figura 2.2).

Las unidades de predicción pueden representar parques individuales o pequeñas agrupaciones de parques cercanos que vierten su potencia a un mismo nudo y que, por tanto, comparten la información de ubicación geográfica y de medidas de generación eléctrica.

Posteriormente al cálculo de predicciones para los parques, se obtienen predicciones agrupadas por zonas, empresas, nudos, etc, mediante la agregación de las predicciones de varias unidades de predicción. Del mismo modo, también se calcula la predicción de potencia total peninsular como la suma de las predicciones de todos los parques.

Todas las predicciones, individuales o agregadas, se pueden reportar mediante informes de salida adaptados a las necesidades del usuario final de las predicciones.

2.4.1 Datos de entrada

Sipleólico calculará las predicciones usando modelos estadísticos en base al uso de unos datos de entrada, que son la previsión de viento en el emplazamiento o proximidades de los parques, y la potencia en tiempo real registrada en los mismos.

Dado que Sipleólico calcula predicciones para cada una de las unidades de predicción de forma individual, necesita una tabla o base de datos que contenga los datos estáticos de los parques. Esta tabla, entre otra información, contiene la ubicación de los parques en coordenadas de longitud y latitud, que será usada para asignarles predicciones de viento, y la potencia eólica total instalada en los mismos.

Las medidas de potencia en tiempo real consisten en medias cuartohorarias o veinteminutales de la potencia generada por los parques eólicos. Estos valores se obtienen tras el procesado de las telemidas recibidas cada 12 segundos por el centro de control de Red Eléctrica.

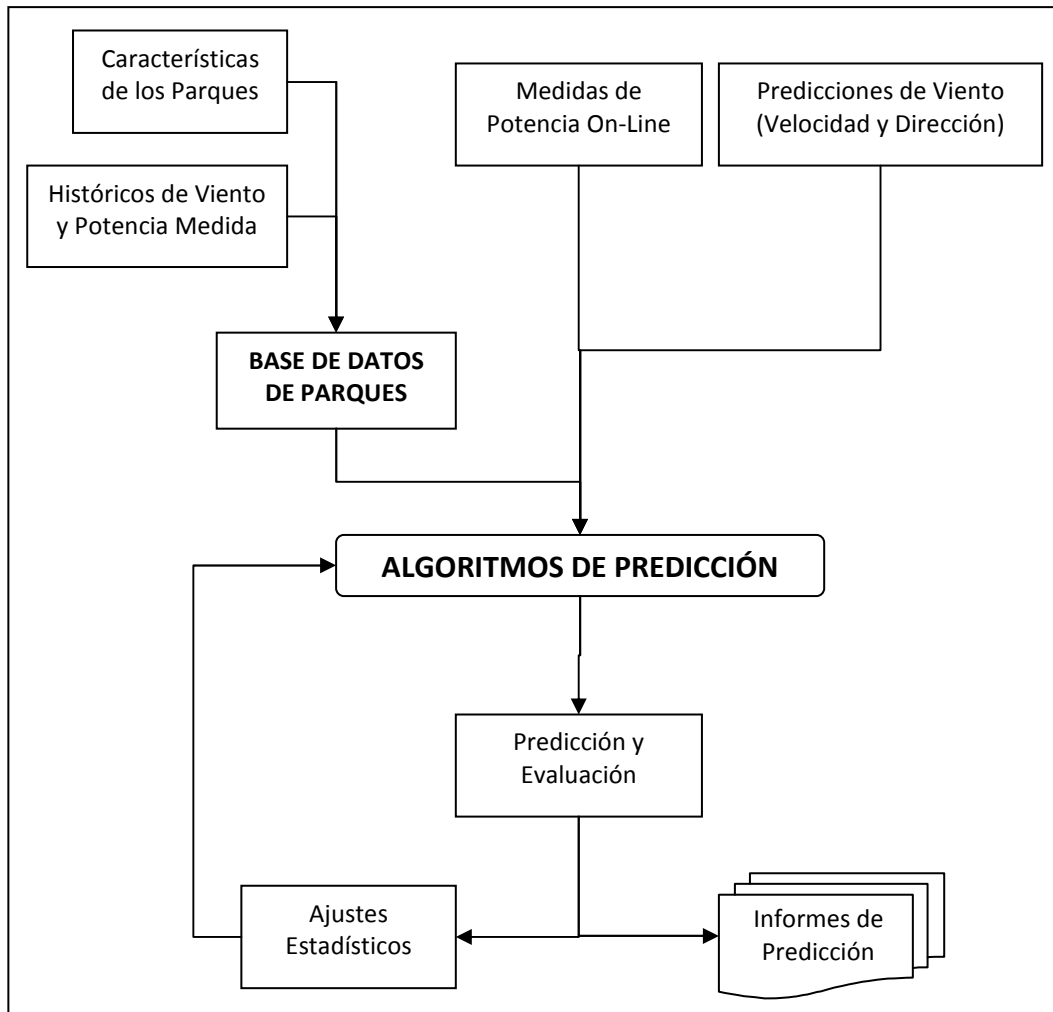


Figura 2.2: Diagrama de funcionamiento de Sipreólico

Las telemidas son las que definen el concepto de unidad de predicción, siendo en ocasiones telemidas de un solo parque, o englobando en otras ocasiones en una misma medida la agrupación de varios parques cercanos conectados a un mismo nudo y con una sola telemida común.

En la actualidad, aproximadamente el 92 % de la potencia eólica instalada en el Sistema Eléctrico Peninsular proporciona telemidas en tiempo real [58], por lo que Sipreólico puede realizar cálculos completos para la casi totalidad de los parques eólicos, con la necesidad de hacer estimaciones de predicción mediante reescalado para tan solo el restante 8 %.

Sipreólico también necesita predicciones meteorológicas para el cálculo de la predicción de potencia. Estas predicciones meteorológicas incluyen las previsiones de velocidad y dirección del viento en una serie de coordenadas geográficas distribuidas por el territorio peninsular, y cercanas a los emplazamientos de los parques eólicos. Debido a que muchos de los parques eólicos se encuentran cercanos unos a otros, no

Métodos de Predicción de la Generación Agregada de Energía Eólica

es necesaria una predicción de viento específica para las coordenadas de cada parque, sino que se pueden compartir coordenadas de previsión de viento.

Actualmente, se utilizan previsiones meteorológicas de dos agencias: Las provenientes de la Agencia Estatal de Meteorología (AEMET), que proporciona cada 6 horas predicciones de velocidad y dirección del viento con un horizonte de 56 horas, calculadas por el modelo atmosférico HIRLAM de 0.16 grados de resolución espacial e interpoladas a las coordenadas de interés, y las previsiones de viento generadas cada 12 horas por el modelo atmosférico del Centro Europeo de Predicciones (ECMWF), que proporciona predicciones con un horizonte de hasta 4 días.

2.4.2 Modelos de predicción de Sipleólico

Existe una relación no lineal entre la velocidad del viento y la potencia que entregan los parques eólicos que no resulta fácilmente modelable, pues en el emplazamiento de un parque eólico existen perturbaciones que afectan a la relación entre la velocidad y la potencia, como puede ser una orografía compleja, la presencia de obstáculos o el efecto de estela dependiendo de la disposición de las turbinas, lo que hace que esta relación sea incluso dependiente de la dirección del viento.

Además, el hecho de que las curvas de potencia deban ser estimadas y construidas con predicciones de viento proporcionadas por institutos meteorológicos de no muy alta resolución, e interpoladas a coordenadas más o menos cercanas al emplazamiento de las turbinas, hace que la construcción de la curva de potencia sea aún más compleja y con mayores fuentes de error posibles.

Por otro lado, la relación existente entre potencia y viento es variable a lo largo del tiempo, siendo dependiente de otras variables como la temperatura, presión y humedad del aire, o incluso con una componente estacional o diaria, lo que hace que sea una relación dinámica y cambiante con el tiempo.

Para poder afrontar estos problemas y situaciones posibles, la herramienta de predicción Sipleólico emplea varios modelos de predicción, que pueden clasificarse en dos tipos: paramétricos y no paramétricos.

Los modelos paramétricos son aquellos en los que la relación entre el viento y la potencia se realiza mediante modelos basados en polinomios de diferente grado y cuyos coeficientes se estiman de forma recursiva. En Sipleólico estos modelos incluyen una componente autorregresiva para recoger la inercia de las variables, y otras componentes para tener en cuenta la velocidad y dirección del viento.

También se utilizan modelos no paramétricos. En este tipo de modelos, la curva que relaciona la potencia y el viento no es prefijada de antemano, sino que es estimada con la información que proporcionan las observaciones próximas a cada punto de la relación mediante un suavizado.

Sipreólico cuenta con ocho modelos diferentes, paramétricos y no paramétricos, agrupados en cuatro bloques que recogen las distintas posibilidades que se plantean en cuanto a la estimación de la relación entre potencia y viento, y enfocados tanto a horizontes de muy corto plazo como para horizontes de corto plazo.

El primer bloque de modelos (M1 y M2) son predictores univariantes autorregresivos, es decir, solamente interviene la potencia medida en el parque. El predictor es una ecuación que relaciona el valor futuro de potencia con valores recientes de la energía medida.

La segunda categoría de modelos (M3 y M4) tienen en cuenta, además de la serie de potencia medida para la parte autorregresiva, la información de la velocidad prevista del viento. Esta velocidad se puede incluir mediante polinomios de diferente orden.

El tercer bloque de modelos (M5 y M6) emplea la información de la dirección del viento previsto, además de la información de potencias medidas y predicciones de velocidad de viento de los modelos de la segunda categoría mencionada, también con la posibilidad de incluir la velocidad mediante polinomios de diferente orden.

El último bloque (M7 y M8) lo forman modelos que incluyen una estimación no paramétrica de la curva de potencia usando la información de la velocidad y dirección prevista del viento, y también pueden tener una componente autorregresiva basada en medidas de potencia.

Los modelos utilizados para generar predicciones son los siguientes, donde \hat{P}_{t+h} representa la predicción calculada en la hora t para la hora $t+h$:

- M1 = Persistencia:

$$\hat{P}_{t+h} = P_t \quad (2.7)$$

donde P_t es la medida de potencia de la hora t .

- M2 = Autorregresivo:

$$\hat{P}_{t+h} = \alpha_{1t}P_t + \alpha_{2t}P_{t-1} + \dots + \alpha_{qt}P_{t-q+1} \quad (2.8)$$

donde los coeficientes $\alpha_{1t} \dots \alpha_{qt}$ se estiman de forma adaptativa, con lo que su valor varía con el tiempo.

Métodos de Predicción de la Generación Agregada de Energía Eólica

- M3 = Autorregresivo y velocidad lineal:

$$\hat{P}_{t+h} = \alpha_{0t} + \alpha_{1t}P_t + \beta_{1t}\hat{v}_{t+h} \quad (2.9)$$

donde \hat{v}_{t+h} es la predicción de velocidad de viento proporcionada por la agencia meteorológica para el periodo $t+h$. Al igual que en el modelo autorregresivo, los parámetros se estiman de forma adaptativa.

- M4 = Autorregresivo con velocidad cuadrática:

$$\hat{P}_{t+h} = \alpha_{0t} + \alpha_{1t}P_t + \beta_{1t}\hat{v}_{t+h} + \beta_{2t}\hat{v}_{t+h}^2 \quad (2.10)$$

- M5 = Autorregresivo con velocidad lineal y dirección de viento:

$$\hat{P}_{t+h} = \alpha_{0t} + \alpha_{1t}P_t + \gamma_{1t} \sin\left(\frac{2\pi\phi_{t+h}}{360}\right) + \gamma_{2t} \cos\left(\frac{2\pi\phi_{t+h}}{360}\right) \quad (2.11)$$

donde ϕ_{t+h} es la predicción de dirección de viento proporcionada por la agencia meteorológica para el periodo $t+h$.

- M6 = Autorregresivo con velocidad cuadrática y dirección de viento:

$$\hat{P}_{t+h} = \alpha_{0t} + \alpha_{1t}P_t + \beta_{1t}\hat{v}_{t+h} + \beta_{2t}\hat{v}_{t+h}^2 + \gamma_{1t} \sin\left(\frac{2\pi\phi_{t+h}}{360}\right) + \gamma_{2t} \cos\left(\frac{2\pi\phi_{t+h}}{360}\right) \quad (2.12)$$

- M7 = Autorregresivo y predicción no paramétrica basada en velocidad y dirección

$$\hat{P}_{t+h} = \alpha_{0t} + \alpha_{1t}P_t + \delta_{1t}G_{t+h} \quad (2.13)$$

donde G_{t+h} es la predicción de potencia basada en métodos no paramétricos que utilizan como variables explicativas la velocidad y dirección de viento.

- M8 = Predicción no paramétrica basada en velocidad y dirección.

$$\hat{P}_{t+h} = \alpha_{0t} + \delta_{1t}G_{t+h} \quad (2.14)$$

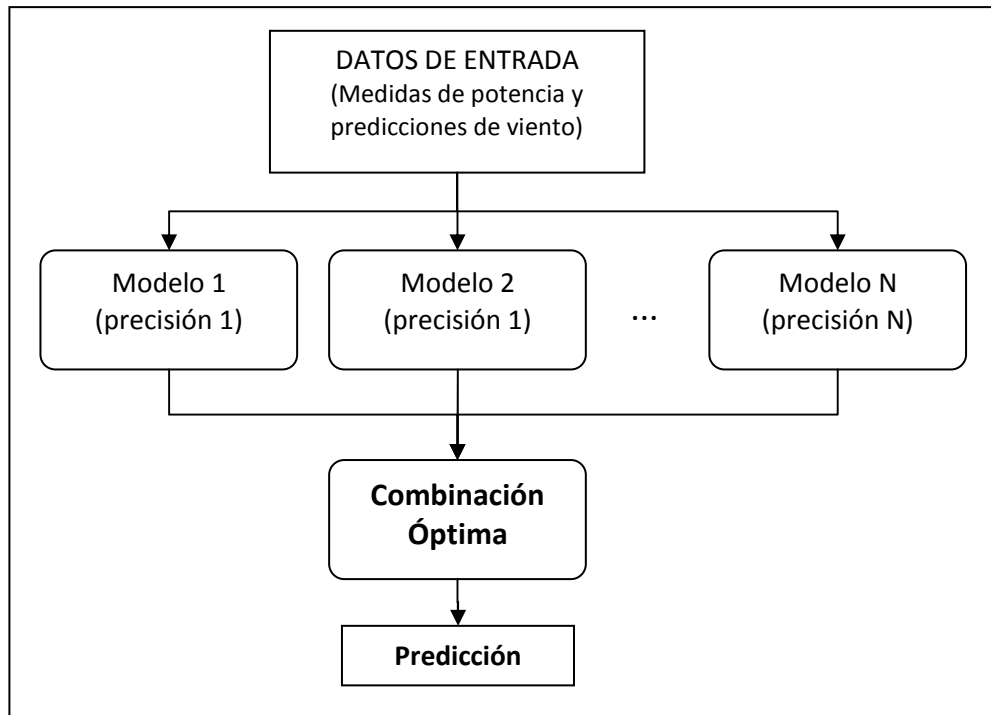


Figura 2.3: Esquema del módulo de predicción de Sipreólico

La estimación de los parámetros de los modelos se hace de forma adaptativa, ya que las curvas que relacionan la producción eólica y el viento son variables en el tiempo en función de circunstancias cambiantes como la situación meteorológica. Por ello, los modelos y sus parámetros no pueden considerarse fijos y deben adaptarse en el tiempo en función de la nueva información de que se disponga. Además, los parámetros serán distintos en función del horizonte de predicción.

Por tanto, cada vez que se recoge nueva información, como velocidades de viento o medidas de potencia, se actualizan los parámetros de los diferentes modelos. Esta adaptación se realiza mediante un procedimiento óptimo de elección del grado de velocidad de ajuste y de la flexibilidad de la adaptación a los nuevos datos [59].

La predicción final se obtiene combinando las predicciones generadas por los distintos modelos (Figura 2.3). A medida que se van recopilando datos, los diferentes modelos tendrán una precisión diferente que será la que determine su peso en la predicción final para cada horizonte [60].

2.4.3 Combinación de predicciones obtenidas con viento de distintas agencias

Cuando se reciben predicciones meteorológicas de más de una agencia, Sipleólico elabora predicciones de potencia horaria usando el viento previsto por cada una de ellas. Después, se elabora una predicción de potencia final mediante la combinación de ambas predicciones.

La combinación de las predicciones calculadas con el viento de cada agencia se realiza en dos etapas [61]:

- En la primera etapa se utilizan dos tipos de combinaciones basadas en métodos de propiedades distintas, generándose dos predicciones combinadas alternativas. El primer método consiste en promediar las predicciones iniciales utilizando la teoría de combinación óptima (método del promedio óptimo). Este método persigue que la combinación mejore a las predicciones individuales. En la práctica, esta mejora sólo ocurrirá en determinadas condiciones, siendo posible que la predicción combinada pueda ser peor que alguna de las predicciones originales. El segundo método utilizado en esta primera etapa tiene como objetivo dar la mejor predicción en cada momento, y no busca por tanto mejorar a las predicciones alternativas. El fundamento teórico de este método se basa en la minimización del riesgo (método del riesgo mínimo).
- En la segunda etapa de la combinación se combinan las dos predicciones combinadas obtenidas en la primera etapa. Estas dos predicciones se combinan usando sólo el método del riesgo mínimo. De esta forma, lo que se obtiene es la mejor combinación en cada momento, lo que a su vez implica que la combinación final tenderá a ser mejor o igual que usar en cada momento el mejor predictor.

2.4.4 Combinación de predicciones para el Total Peninsular

Cada 15 o 20 minutos, dependiendo de la frecuencia en la recepción de medidas, Sipleólico ejecuta un cálculo de predicción eólica con resolución horaria para las siguientes 48 horas, si bien cada 15 o 20 minutos solo se recalculan o actualizan las predicciones para las siguientes 4 horas, y las predicciones hasta el horizonte 48 solo se calculan una vez cada hora.

En cada ejecución se calculan predicciones según los modelos descritos para cada una de las unidades de predicción que dispongan de medidas en tiempo real. Debido a que hay algunos parques eólicos que no disponen de telemidas, su predicción se calcula por medio de un reescalado por zonas, teniendo en cuenta la potencia instalada en el parque, la potencia total prevista en los parques telemidados de la zona donde se encuentre, y la potencia total instalada en dichos parques telemidados.

Con esto, se tiene la predicción de potencia para todos los parques eólicos y se pueden calcular las predicciones de agrupamientos por zonas, empresas, subestaciones, comunidades autónomas, provincias, etc, y por consiguiente el total peninsular.

Una vez que se tiene la predicción de potencia total, ésta se combina con otras previsiones de potencia total, como puede ser la previsión de generación eólica ofertada en el mercado, u otras predicciones de generación total suministradas por otras herramientas de predicción distintas de Sipleólico.

Esta combinación de predicciones totales de nuevo se realiza en dos etapas, utilizando los métodos basados en el promedio óptimo y métodos que buscan el mínimo riesgo, mediante el mismo procedimiento que se siguió en la combinación de predicciones obtenidas de distintas agencias meteorológicas. Al final de este proceso, se tendrá una predicción total que tenderá a ser la mejor predicción combinada para cada caso y cada horizonte, aprovechando las bondades de cada predictor.

2.4.5 Informes de resultados y diagnóstico

Una vez calculada la predicción para todos los parques, agrupaciones de parques, y total peninsular, se generan una serie de informes de resultados, adaptando el formato a los distintos usuarios posibles. Estos informes contienen las predicciones de generación horaria para cada horizonte de predicción, habitualmente hasta un horizonte de 48 horas, especializados a la generación de las 24 horas del día siguiente, o incluso conteniendo predicciones a 10 días para el total peninsular.

También se generan informes con intervalos de predicción, habitualmente correspondientes a los percentiles 15 y 85, para la predicción total peninsular. Esta predicción probabilista es utilizada por el operador de sistema para programar reservas de generación. Por ejemplo, en horas de baja demanda se utiliza el percentil 85, que da el valor de generación que será superado con un 15% de probabilidad. Para horas de gran demanda se utiliza el percentil 15, que da una idea del valor de potencia que será superado con un 85% de probabilidad.

Los intervalos de la predicción probabilista pueden ser calculados de dos modos: Con métodos estadísticos, mediante la estimación de la distribución de errores basada en las observaciones de error de las predicciones de Sipleólico en función del horizonte de predicción, o mediante la adaptación a la predicción total peninsular de la banda de predicción proporcionada por un modelo de predicción agregado que utiliza conjuntos o ensembles de predicción meteorológica para calcular la predicción de potencia.

Por último, Sipleólico también incorpora un módulo de evaluación y diagnóstico de las predicciones. Además de los informes de predicción de potencia, se generan informes de diagnóstico donde se evalúa de forma continua la precisión de las predicciones realizadas. Cada hora, se generan ficheros que incluyen información sobre la media del error absoluto con respecto a la potencia instalada, la media del error con respecto a la energía medida, o el coeficiente de determinación R^2 .

Métodos de Predicción de la Generación Agregada de Energía Eólica

Todas estas estimaciones de error son calculadas de forma recursiva en base a diferentes periodos de tiempo, que varían entre las medias del último mes, media de la última semana, o el error medio referido a las últimas 24 horas.

Actualmente, el error absoluto medio de Sipleólico con respecto a la potencia total peninsular instalada para horizontes hasta 4 horas es inferior al 4%, creciendo hasta el 5.5% o 6.5% para horizontes entre 24 y 48 horas [58].

2.5 El proyecto ANEMOS

En 2002 comenzó el proyecto ANEMOS [62], dentro del quinto programa marco de la Unión Europea, liderado tanto por entidades de investigación del ámbito de la predicción eólica e institutos meteorológicos, como por usuarios finales de las predicciones, operadores de sistema y autoridades reguladoras, de un total de siete países: Francia, Irlanda, España, Alemania, Grecia, Dinamarca y Reino Unido.

Su objetivo fue desarrollar una plataforma de predicción eólica que integrase las bondades de todos los métodos de predicción existentes hasta el momento, basados tanto en modelos físicos como estadísticos, con el objetivo de dar predicciones para parques situados tanto en tierra firme como en el mar (off-shore), y en todo tipo de emplazamientos independientemente de la complejidad de la orografía [63].

El horizonte de predicción establecido es de 48 horas, pues es el horizonte temporal habitualmente utilizado por los operadores del sistema eléctrico para la planificación de la generación de energía, aunque se contempla la predicción hasta 7 días como herramienta de utilidad para programar labores de mantenimiento en parques o instalaciones. En cuanto a alcance espacial, abarca tanto predicción de parques individuales como predicción regional o nacional. También incluye incertidumbre en la predicción y estimación de riesgo.

El proyecto ANEMOS fue el primer banco de pruebas de varios sistemas de predicción que ya estaban funcionando en varios países europeos, comparando los sistemas AWPPS, Prediktor, WPPT, Previento, LocalPred, Sipleólico, y los modelos de RAL, ARIA y NTUA [26], utilizando para la comparación los protocolos de evaluación descritos en [64].

Para ello, se tomaron como casos de estudio algunos parques localizados en España, Alemania, Dinamarca, Irlanda, Grecia y Francia, incluyendo incluso algún parque offshore. La selección de parques incluye así emplazamientos en terrenos llanos, complejos y muy complejos.

Se propusieron varios sistemas de predicción meteorológica en los emplazamientos de los parques eólicos, entre ellos HIRLAM, SKIRON, ALADIN o DM, y también se tuvo en cuenta el uso de información meteorológica de alta resolución, como la obtenida del modelo de meso-escala MM5, para ser comparada con los modelos macro-escalares, más globales y de menor resolución. Se comprobó que se pueden obtener mejores predicciones de potencia eólica si se combinan varias fuentes de

viento, y además que el uso de modelos de mayor resolución proporcionaba mejoras en la predicción para los parques situados en los terrenos especialmente complejos.

En los resultados del estudio, se comprobó que el valor medio del Error Medio Absoluto de las predicciones de los modelos para un horizonte de 12 horas oscila entre el 10% y el 21% [26], dependiendo de si se trata de un parque eólico ubicado en terreno llano o en el mar, o si está situado en un terreno montañoso complejo. En [65-67] se puede encontrar una relación más detallada de los resultados de los modelos estadísticos, físicos y de predicción offshore.

También se comprobó que, de entre los modelos de predicción comparados, ninguno era claramente mejor que los demás en todos los horizontes. Por eso, la plataforma de predicción ANEMOS incluye varios módulos y funciones de predicción. Replicando la estructura de combinación de Sipleólico, en la plataforma ANEMOS se combinan las predicciones de los distintos modelos de forma óptima en dos etapas para obtener la mejor predicción para cada horizonte, aprovechando las bondades de cada modelo de predicción independiente que tiene integrado [61].

2.6 Valor económico de la predicción eólica

La predicción de la generación eólica puede tener importantes repercusiones económicas para el operador del sistema, con una reducción de los costes de operación gracias a la reducción de la necesidad de reserva. Como ejemplo, en [68] se trata un modelo para la determinación óptima de reserva rodante en el sistema teniendo en cuenta la posible generación eólica en los próximos minutos.

Una evaluación sobre el impacto que tiene la precisión de las predicciones sobre los costes de operación del sistema puede encontrarse en [69]. Usando un modelo de evaluación de costes, se estudió el impacto de la precisión de las predicciones eólicas en los costes de operación del sistema, concluyendo que a menor precisión en las predicciones, mayores son los costes de operación por la necesidad de aumentar la reserva rodante. También se comprueba que los beneficios de una compañía con potencia eólica instalada se ven mermados debido a los errores de predicción en la generación eólica, al tener que comprar a otras compañías la energía eléctrica que fue prevista pero que no fue generada.

Así, además de ayudar a la operación del sistema eléctrico, la predicción de potencia eólica se ha convertido en una herramienta de peso para los agentes de mercado o los propietarios de parques eólicos que venden la energía de sus instalaciones en el mercado de electricidad. La explosión en el desarrollo de la energía eólica ha coincidido con la liberalización de los mercados eléctricos en muchos países, y existen penalizaciones económicas frente a los desvíos de generación con respecto de la cantidad de energía pactada en el mercado.

En [70] se presenta el valor de la predicción en términos de ahorro de combustible para una compañía con parques eólicos. El gasto de combustible es comparado con el

Métodos de Predicción de la Generación Agregada de Energía Eólica

gasto que se realizaría si se conociese la generación en el parque eólico, usando tanto un modelo simple de persistencia como teniendo en cuenta las predicciones de un modelo atmosférico.

Un estudio para el caso del mercado Danés se presenta en [71], donde se examinó el impacto de la potencia eólica en el mercado Nordpool, simulando el uso de una herramienta de predicción bajo ciertas modificaciones en las reglas del mercado.

En otro estudio descrito en [72] y [73] se utiliza el sistema de predicción de Armines AWPPS sobre el mercado de electricidad alemán APX, y se plantea la posibilidad de utilizar modelos de incertidumbre para diseñar estrategias en las ofertas de energía teniendo en cuenta la diferencia de los precios de los desvíos por exceso o por defecto de generación. Se obtienen mayores beneficios teniendo en cuenta la incertidumbre de las predicciones para la estrategia de oferta de potencia eólica, que ofertando al mercado una predicción puntual [74].

Atendiendo al caso concreto del sistema eléctrico español, se presenta en [18] un estudio que utiliza predicciones del modelo Sipleólico para un parque eólico que envía ofertas de venta de energía al mercado eléctrico. La reducción de ingresos debida a las posibles penalizaciones por desvíos en la energía comprometida en el mercado resulta inferior si se utiliza el modelo de predicción. En [75] se trata la estrategia de ofertas óptimas, utilizando además un modelo de predicción de precios.

Las penalizaciones en los desvíos, se aplican tanto si el parque eólico genera menos potencia de la acordada en el mercado como si tiene un exceso de producción sobre el valor comprometido, pues en ambos casos se debe hacer uso de las bandas de regulación. Para mitigar estos desvíos, se ha planteado la posibilidad de hacer ofertas conjuntas de generación hidráulica y eólica con el objetivo de minimizar las penalizaciones, encargándose la hidráulica de compensar los desvíos en la generación eólica prevista.

En [76] se presentan ofertas óptimas de generación, teniendo en cuenta la incertidumbre de la generación eólica para minimizar el riesgo de penalizaciones durante el día siguiente en el sistema eléctrico español. Además, se plantea el uso de un generador hidráulico para minimizar las penalizaciones por desvíos. En [77] y [78] se propone usar centrales de bombeo junto con un parque eólico para diseñar una estrategia óptima de operación y maximizar los beneficios en el mercado portugués.

El posible uso del recurso hidráulico para almacenar energía y compensar las variaciones de la potencia eólica ha dado lugar a estudios que plantean el tamaño óptimo que deberían tener estas centrales hidráulicas de bombeo para recuperar la energía eólica no integrable en la red [79], o para compensar las variaciones de eólica teniendo en cuenta la incertidumbre asociada [80], convirtiendo incluso el servicio de almacenamiento hidráulico en un mercado rentable.

2.7 Incertidumbre de la predicción eólica

Dado que es imposible dar una predicción exacta de la potencia eólica que se va a generar en un parque o conjunto de parques eólicos, resulta útil proporcionar información sobre la incertidumbre asociada al valor de la predicción, que servirá para conocer cómo se distribuye el error de predicción y poder manejar de forma óptima la producción eólica.

Un modo de dar información sobre la fiabilidad del valor de predicción eólica, y dar una estimación del error asociado a la predicción, es proporcionar intervalos de predicción, que dan una estimación de la probabilidad de que la producción real caiga dentro de cierta banda de generación prevista. Los intervalos de predicción, junto con los índices de riesgo [81], resultan muy útiles como herramienta complementaria en las estrategias de mercado de electricidad.

Tradicionalmente, los métodos de cálculo de intervalos de predicción se basaban en la asunción de que el error de predicción sigue una distribución gaussiana. Sin embargo, debido a la forma de la curva de potencia que relaciona el viento con la potencia generada, esta afirmación no es correcta en la mayoría de las situaciones.

La curva característica que relaciona la velocidad del viento con la potencia eólica generada tiene una elevada pendiente en la zona habitual de funcionamiento, como se muestra en la Figura 2.1. Por ello, pequeñas variaciones en el valor de la velocidad del viento provocan grandes variaciones en la potencia eólica generada para valores de viento intermedios. Sin embargo, para velocidades próximas a la velocidad de arranque o en el rango de velocidades correspondientes a la zona de potencia nominal de las turbinas, la potencia generada varía poco con las variaciones de velocidad del viento.

Un modo de tener en cuenta la relación no lineal entre potencia y viento se describe en [82], donde se propone un modelo paramétrico que presupone que la distribución de error de la velocidad del viento sí puede tener una distribución gaussiana, posteriormente esta distribución gaussiana de velocidades se transforma en una distribución no gaussiana mediante la curva de potencia.

En el ámbito del proyecto ANEMOS, se recoge un estudio sobre la dependencia entre la distribución del error de predicción y el nivel de potencia prevista, teniendo en cuenta los cuatro primeros momentos de la distribución (sesgo, desviación típica, asimetría y apuntamiento) [83]. Se comprobó que las distribuciones de error presentan fuertes asimetrías tanto a potencias bajas como para potencias altas, con lo que los intervalos no son simétricos respecto del valor de predicción.

Teniendo en cuenta la forma asimétrica de la distribución de error debida a la forma de la curva de potencia, en [84],[85] se plantea usar una distribución Beta para describir la forma del error, de la que se deben estimar los parámetros con datos históricos de errores en función del nivel de potencia.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Pero los intervalos también se pueden estimar de forma no paramétrica. La ventaja de estos métodos frente a los paramétricos es que no parten de ninguna hipótesis sobre el tipo de distribución que debe seguir el error. Un método no paramétrico se describe en [86], donde Pinson propone un algoritmo para estimar intervalos de predicción a partir de valores históricos del error. Los intervalos de error de potencia prevista se calculan como distribuciones condicionadas, en función de la velocidad de viento, usando un sistema de inferencia difusa para seleccionar los datos. Mediante un método de remuestreo con los datos seleccionados se determina la distribución y se obtienen los intervalos.

También en [81] se presenta un método no paramétrico para estimar la incertidumbre de las predicciones de potencia. Mediante un método de remuestreo se generan bandas de predicción y además se introduce el índice “meteo-risk” como un indicador de la probabilidad de que haya grandes errores de predicción de acuerdo a la estabilidad atmosférica, dando una idea de lo dispersas que fueron las predicciones meteorológicas usadas para ajustar las bandas de predicción.

Otro método para proporcionar intervalos de predicción se propone en [87] y [88], por medio de regresión cuantílica, en la que el objetivo es calcular algún percentil de la distribución, a diferencia de la regresión por mínimos cuadrados cuyo objetivo es estimar la media. Entre otras conclusiones, en este estudio se observó que la principal variable que afecta a los intervalos es el nivel de potencia prevista.

Como complemento a los métodos de predicción basados en las predicciones meteorológicas deterministas tradicionales que proporcionan un único valor de viento para un horizonte dado, existe la posibilidad de proporcionar intervalos de potencia mediante el uso de un conjunto de previsiones meteorológicas, conocidas en la literatura como “ensembles”. Estas previsiones dan una colección de posibles valores de velocidad de viento para cada horizonte de predicción.

Para conseguir las ensembles se ejecuta el modelo numérico de predicción atmosférica (NWP) varias veces incluyendo pequeñas perturbaciones en las condiciones iniciales, ya que la evolución de la atmósfera, y por tanto las previsiones de viento, serán distintas dependiendo de la inicialización de modelo meteorológico. De este modo se obtendrán distintas evoluciones posibles de la atmósfera y del viento, que podrán ser parecidas o muy distintas entre sí dependiendo de la estabilidad atmosférica. Al procesar esta colección de previsiones de viento con el modelo de predicción de potencia se obtendrá una colección de predicciones de potencia que podrán ser parecidas o muy distintas. Finalmente, a partir de la colección de predicciones de potencia se obtienen los intervalos de predicción.

Ejemplos del uso de ensembles se encuentran en [89] y [90], donde se utilizan ensembles de los centros de predicción meteorológica americano NCEP y europeo ECMWF, con el modelo de predicción Zephir y un modelo de estimación de curva de potencia.

2.8 Predicción agregada de energía eólica. Predicción Regional

Desde el punto de vista de un operador del sistema eléctrico, adquiere gran importancia poder conocer la cantidad total de energía de origen eólico que será inyectada en el sistema, con el fin de gestionar el resto de unidades de generación, y pierde relevancia la potencia concreta que inyectará cada parque individual. Es por ello que desde este punto de vista cobra gran importancia la predicción de energía total o agregada en una región, conocida en la literatura como “regional forecasting” [91].

Pero el operador del sistema no es el único interesado en conocer la potencia agregada total que generarán los parques. Los agentes de mercado que gestionan gran cantidad de potencia instalada entre varios parques eólicos también estarán interesados en predecir la potencia eólica total que generarán sus instalaciones para diseñar estrategias en su participación en los mercados diarios e intradiarios de energía.

Para poder calcular predicciones de un parque eólico concreto pueden ser necesarios una serie de datos de la instalación, como el número y características de las turbinas, la localización, la potencia, o la orografía del terreno donde se encuentran instaladas. Además de estos datos, que se consideran estáticos, puede ser necesario contar con otros datos dinámicos o cambiantes con el tiempo, como la potencia eólica registrada en el parque, el viento medido en los anemómetros, o el viento previsto en el parque. Pero lo habitual es que el operador de sistema o los agentes no tengan acceso a toda esta información para todos los parques eólicos de una región, de modo que resulta complejo calcular la previsión de potencia eólica de las instalaciones de las que se tiene poca o ninguna información.

Un modo habitual de predicción regional de potencia eólica consiste en calcular la potencia total en la región por medio de una extrapolación o reescalado, a partir de la predicción de algunos parques eólicos de los que sí se tiene información suficiente, comúnmente conocidos como parques de referencia. Esta idea se basa en el hecho de que generalmente no se tienen datos suficientes de todos los parques eólicos, y aunque se tuvieran, la carga computacional de calcular predicciones para todos ellos puede ser muy elevada.

Dependiendo del modo en que los modelos agregados calculan la predicción de potencia total en la región, y en su caso, combinan los modelos básicos referidos a la predicción de parques individuales, podemos encontrar tres configuraciones o aproximaciones alternativas en los modelos de predicción regional [91]:

- **Configuración directa:** El modelo agregado calcula la predicción de potencia total en la región directamente, sin pasos intermedios, pudiendo utilizar como entradas las previsiones meteorológicas en la región y la medida de potencia total.

Métodos de Predicción de la Generación Agregada de Energía Eólica

- **Configuración en cascada:** En este caso, el modelo regional utiliza como entrada las predicciones de potencia de parques individuales, que habrán sido calculadas con modelos de predicción de potencia para parques, ya sean físicos o estadísticos. La potencia total se calcula como adición y, en su caso, reescalado de las predicciones de los parques individuales.
- **Configuración en clúster:** Según esta configuración, el modelo regional usa como entradas las predicciones de potencia para sub-regiones o sub-áreas, calculadas con algún otro modelo. Estas predicciones sub-regionales podrán haber sido calculadas de forma directa, en cascada, o con otro modelo en clúster. La predicción de potencia total regional se calculará como suma de las predicciones en cada sub-región.

En cualquier caso, estas tres configuraciones son bastante genéricas, y se pueden utilizar para describir en mayor o menor medida los modelos de predicción agregada existentes, combinando en ocasiones varias configuraciones.

La predicción agregada con Previento se calcula en base a un modelo físico de predicción de potencia en parques individuales [40]. Utiliza una configuración en cascada para calcular la potencia en 45 sub-áreas, seleccionando en cada sub-área un parque de referencia para el que se calcula la predicción y aplicando un reescalado. Posteriormente usa una configuración en clúster para calcular la predicción total de la región a partir de las predicciones en las sub-áreas.

El modo de cálculo de predicciones para una región utilizado en la aplicación Zephyr utiliza dos aproximaciones o ramas paralelas para el cálculo de potencia en sub-áreas que posteriormente son agregadas para calcular el total regional [50].

La primera rama es similar a lo descrito anteriormente para Previento, calculando la potencia eólica para una selección de parques de referencia que disponen de predicciones de viento y medidas de generación en tiempo real. Estas medidas proporcionan la ventaja de mejorar las predicciones a muy corto plazo debido a la autocorrelación de la producción. Las predicciones eólicas calculadas para los parques de referencia en cada sub-área se suman y reescalan para calcular la predicción total en cada sub-área, usando por tanto una aproximación en cascada.

La segunda rama predice la potencia en cada sub-área de forma directa usando un modelo que relaciona las medidas de potencia total en el sub-área, obtenidas en tiempo diferido, con las predicciones de viento promediadas en el sub-área. Esta rama aprovecha las ventajas del suavizado de la producción total debido a la dispersión de la generación.

Una vez calculada la predicción de potencia usando cada una de las dos ramas para cada sub-área, se calcula la potencia total de la región para cada rama como la suma de potencias de las sub-áreas, usando por tanto una configuración en clúster. Finalmente se combinan ambas predicciones de potencia total.

También hay experiencias de predicción agregada en Alemania con el Advanced Wind Power Prediction Tool (AWPT) desarrollado por ISET. Utilizando datos históricos de potencia, medidas de potencia en tiempo real y previsiones de viento se calculaban previsiones eólicas en algunos parques representativos que posteriormente se extrapolaban al total de la región PEN.

Las mejoras incluidas en este sistema llegaron con el Wind Power Management System (WPMS), que calcula predicciones para casi todos los parques instalados en Alemania. En la zona de E.On Netz la potencia total se estima mediante el reescalado de 50 parques de referencia, que representan cerca del 30% de la potencia instalada [52].

En [92] se propone también un método de cálculo de predicciones regionales. Comparte con otros métodos de predicción regional la idea de seleccionar algunos parques de referencia, pero en esta ocasión no se hace un reescalado de potencia, sino que se utiliza una red neuronal difusa para relacionar la potencia en los parques seleccionados con la potencia total en la región.

Una aproximación distinta para la predicción regional es la propuesta en [93], mediante una descomposición espacial de la generación eólica en Alemania con un análisis de componentes principales para obtener el patrón principal de variabilidad. Con ello se consigue un tiempo de cálculo muy reducido, lo que permite utilizar varias fuentes de viento para combinar sus predicciones y mejorarlas.

Otra herramienta de predicción agregada desarrollada por CENER/CIEMAT es RegioPred [45], que proporciona una predicción regional a partir de predicciones individuales calculadas con LocalPred, mediante un módulo de reescalado de predicciones calculadas para una selección de parques representativos. A este respecto se publicó en [94] la influencia de tomar agrupaciones de parques para obtener predicciones agregadas de potencia, usando 17 parques distribuidos en tres zonas de la península ibérica, concluyendo que solo era necesario tomar un parque de referencia en cada zona para estimar la potencia total.

2.8.1 Ventajas de la predicción agregada de energía eólica

Cuando lo que importa es la predicción agregada de generación eólica, como ocurre en el caso de los operadores del sistema eléctrico o de grandes compañías con varios parques, el uso de métodos agregados de predicción es un modo a tener en cuenta para estimar la potencia eólica total que será generada. Además, la predicción regional o agregada presenta una serie de ventajas con respecto a la predicción individual de parques eólicos.

Una ventaja de los modelos de predicción agregados es que no necesitan gran cantidad de información sobre los parques, sino que utilizan solamente algunos datos más fácilmente disponibles para calcular la predicción de potencia total en la región. Además, cuando el número de parques eólicos en una región es elevado, la predicción

Métodos de Predicción de la Generación Agregada de Energía Eólica

regional es un método de predicción mucho más rápido que obtener la potencia total haciendo predicciones para cada uno de los parques eólicos en la región y sumarlas.

Pero la rapidez de cálculo o la simplicidad de los datos de entrada no son las únicas ventajas del uso de predicciones agregadas. La variabilidad de la potencia generada con respecto a la instalada es mucho más acusada a nivel de parque que a nivel agregado, con lo que predecir la potencia a nivel de parque es más difícil que hacerlo para una región. Por tanto, la predicción a un nivel agregado puede ser más precisa que la suma de predicciones individuales, pues evita acumular el error asociado al cálculo de la potencia de cada parque gracias a efectos de suavizado del error de predicción debido a la habitual dispersión geográfica de los parques eólicos. Debido a que la correlación entre los errores de predicción de distintos parques eólicos suele ser bastante débil, más aún conforme aumenta la distancia entre parques, la incertidumbre en la predicción de una región será mucho inferior al error de predicción en parques individuales.

Como ejemplo, la posibilidad de reducir del error de predicción usando un modelo agregado se estudia en [95] y [96], comparando el error de predicción de un parque individual con el obtenido en la predicción de un grupo de parques. Debido a la dispersión espacial el modelo agregado proporciona menor error en sus predicciones. La reducción del error es tanto mayor cuanto mayor sea el tamaño de la región, no siendo necesario calcular la predicción de todos los parques para estimar la predicción total en la región. En el estudio se encontró que había un nivel de saturación, a partir del cual la inclusión de nuevos parques para calcular la predicción agregada no mejoraba sustancialmente las predicciones, con lo que se concluye que con solo predecir la potencia de una porción de parques se puede estimar la potencia en toda la región con un error reducido. Por tanto la reducción del error es mucho más sensible al tamaño de la región y a lo dispersos que estén los parques eólicos, que al número de parques que se utilicen para calcular la predicción.

3. Método Propuesto para la Predicción Regional

Resumen

En este capítulo se introducen las bases de los modelos de aprendizaje basados en la memoria, y se describen las estructuras de los modelos locales por suavizado de Media Ponderada y Regresión Local Ponderada, junto con sus ventajas e inconvenientes dependiendo de la aplicación para la que se utilicen. En base a este tipo de modelado local, se plantea el esquema y se resumen las etapas del algoritmo de predicción regional propuesto. Finalmente, se describen los modos de evaluación que serán utilizados para comprobar la bondad del modelo de predicción propuesto.

3.1 Introducción

La predicción regional se justifica debido a la gran carga computacional que implica calcular predicciones de generación eólica en regiones con gran cantidad de parques eólicos instalados. A medida que aumenta el número de parques o aerogeneradores individuales instalados en una región extensa, se puede hacer inmanejable calcular la predicción eólica para todos ellos, debido al tiempo necesario para calcular predicciones de potencia para cada unidad de predicción individual. Por otro lado, en ocasiones puede haber dificultad para disponer de telemidas de todos los parques o unidades de generación individuales, información de gran importancia para poder calcular predicciones de potencia a nivel individual.

Como se ha comprobado en la revisión del estado del arte, existen varios métodos de predicción regional propuestos en la literatura, que en ocasiones utilizan aproximaciones en cascada para calcular la predicción total mediante la adición de las predicciones individuales de varios parques, o aproximaciones tipo clustering que dividen la región en sub-regiones para las que se calcula la predicción, ya sea de forma directa o de nuevo con una aproximación en cascada, sumando finalmente la predicción de las sub-regiones.

Frente a estos enfoques en cascada o tipo clustering, el método descrito en esta tesis es un método de estimación directa de la potencia eólica regional. Es decir, no se realiza ninguna división inicial en sub-regiones, ni se utilizan parques de referencia para los que se calcula la previsión de potencia para posteriormente estimar la potencia agregada por medio de algún reescalado. La predicción total regional o nacional en nuestro caso se calcula directamente, sin pasos intermedios a nivel de parque o sub-regionales, lo que confiere a este método una gran velocidad de cálculo frente a la carga computacional de otras configuraciones.

El método agregado tratado en esta tesis propondrá calcular predicciones de potencia total agregada por medio de una comparación entre vectores de predicción de velocidad de viento en un conjunto de coordenadas seleccionadas en la región. Básicamente, cada vez que se reciba un nuevo vector con valores de velocidad de viento prevista en las coordenadas para un horizonte concreto, el algoritmo lo comparará con un histórico almacenado de vectores de velocidad de viento, y generará una predicción de potencia basándose en los resultados de dicha comparación.

Como gran ventaja, el método de predicción que se propone no necesita ningún tipo de información sobre las características de cada uno de los parques eólicos en la región, ya sea la orografía del terreno donde estén ubicados, la curva de potencia de las máquinas, o un sistema de telemidas a nivel de parque que opere en tiempo real. La única información necesaria para el funcionamiento del algoritmo de predicción será un conjunto almacenado de n pares de datos históricos de la forma (\mathbf{v}_i, P_i) , con $\mathbf{v}_i = [v_{1,i}, \dots, v_{K,i}]^T$ conteniendo valores de velocidad de viento en algunos lugares de la región y P_i la potencia eólica total registrada.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Este conjunto de datos históricos nos proporciona una especie de mapa meteorológico en cada instante del pasado, además de la potencia total registrada en la región bajo cada una de esas condiciones meteorológicas. Así, la idea del método de predicción será buscar similitudes entre el mapa meteorológico previsto dentro de unas horas por el modelo atmosférico, y los mapas o situaciones meteorológicas registradas en el pasado.

3.2 Modelos locales

Cuando se modela la curva de potencia de un aerogenerador o de un parque eólico, se busca modelar la relación existente entre una variable de entrada, que es la velocidad del viento incidente en la turbina o viento medio en el emplazamiento del parque, y una variable de salida o variable respuesta, que en este caso será la potencia eólica generada.

Generalmente, el modelado de la curva de potencia se basa en la recolección de pares de datos de velocidad del viento y potencia registrada, obteniéndose una dispersión de puntos como la de la Figura 3.1.

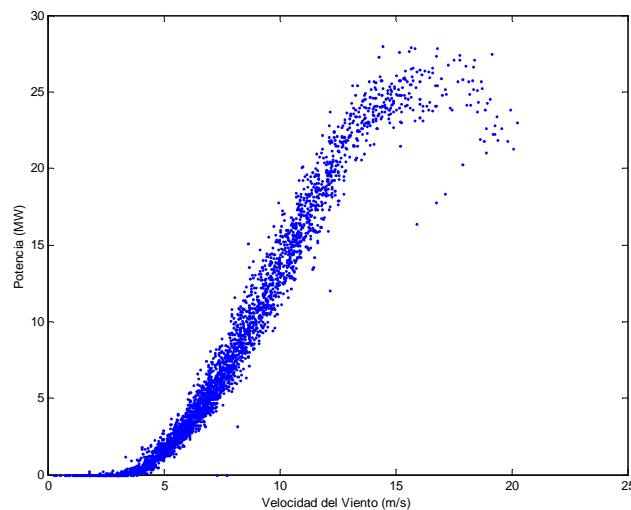


Figura 3.1: Dispersión de puntos de la potencia generada en un parque eólico

A partir de esta dispersión de puntos se podría tratar de ajustar una curva paramétrica mediante un polinomio de cierto grado, que relacionase la variable de salida “potencia generada” con la variable de entrada “velocidad del viento”. Esta curva representaría un modelo global que abarcaría todo el rango de valores posibles de velocidad de viento, de modo que cada vez que se quisiera predecir la potencia que será generada por el parque cuando la variable de entrada “velocidad del viento” tenga un valor, bastaría con sustituir dicho valor de viento en la ecuación ajustada (Figura 3.2, izquierda).

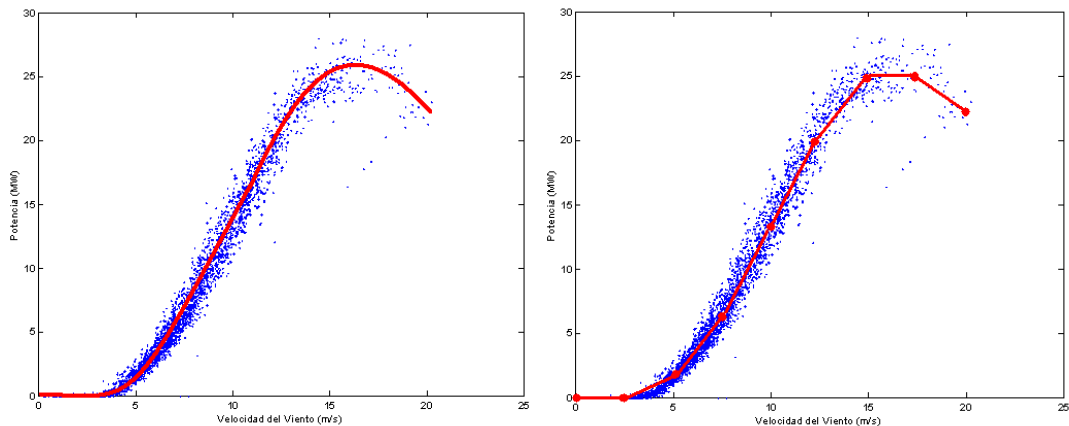


Figura 3.2: Ajuste de una curva paramétrica a toda la nube de puntos (izquierda) y estimación por intervalos (derecha)

También se puede estimar la curva de potencia por intervalos o bins. Este procedimiento consiste en agrupar los puntos de la curva por intervalos de velocidad, normalmente de la misma amplitud, de modo que se puede estimar la potencia esperada para un intervalo de velocidad concreto como la media de las potencias registradas dentro de dicho intervalo. Esto nos llevaría a obtener una curva de potencia a escalones o lineal por tramos, a partir de la cual se podría calcular la predicción de potencia sustituyendo el valor de viento previsto en la ecuación lineal del tramo correspondiente (Figura 3.2, derecha).

Tanto si se decide ajustar un polinomio a toda la nube de puntos, como si se divide la nube en intervalos o bins, tras procesar la dispersión de datos se obtendrá en ambos casos un modelo paramétrico que relacionará la variable de entrada “viento” con la variable respuesta “potencia”.

Pero también se pueden obtener predicciones de potencia directamente a partir de la dispersión de puntos, sin necesidad de haber modelado previamente una curva, sea o no lineal por tramos, a todos los datos de la dispersión. Este sería un método no paramétrico, según el cual, cada vez que se tuviese un nuevo valor de viento previsto, se obtendría la predicción de potencia recurriendo directamente a la nube de puntos almacenada, ajustando un modelo local relativamente sencillo a los puntos pertenecientes a los datos cuyo viento es similar al viento para el que se quiere calcular la predicción (Figura 3.3).

Métodos de Predicción de la Generación Agregada de Energía Eólica

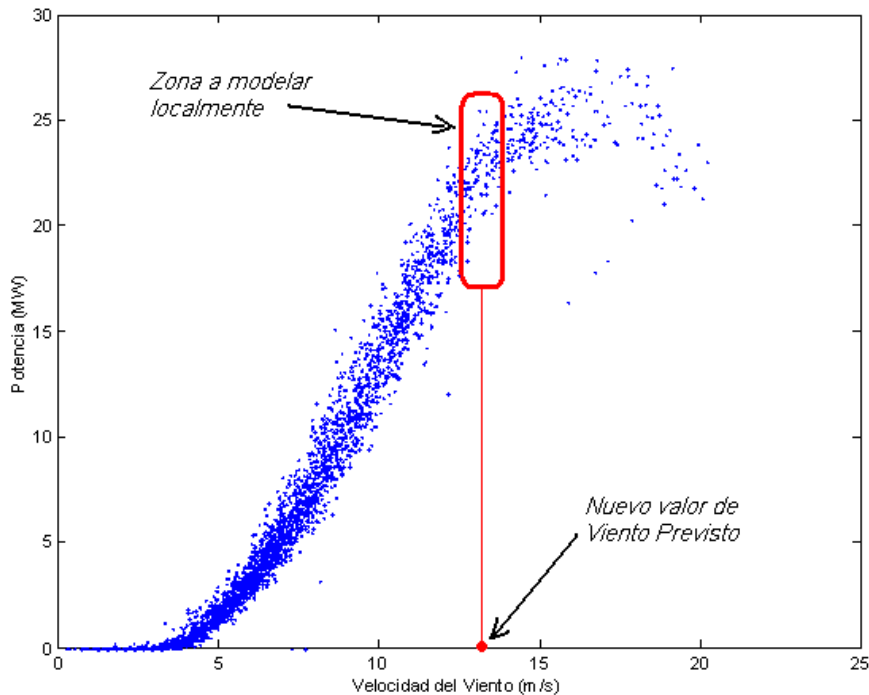


Figura 3.3: Ajuste de un modelo local a los datos cercanos

Cuando el sistema a modelar es sencillo, en el que la curva que relaciona la variable respuesta con la variable de entrada no es muy compleja, cabe la posibilidad de plantearse si merece o no la pena recurrir a un modelo paramétrico para estimar dicha curva de forma global. Sin embargo, a medida que la curva, o superficie en el caso de dos variables de entrada, es más compleja, también será más complicado ajustar una curva o superficie a toda la nube de puntos. Esta dificultad además será mucho mayor conforme aumente el número de variables de entrada al sistema, y por tanto el número de dimensiones del espacio de entrada.

En el caso del modelado de la potencia total que será generada en una región con varios parques eólicos instalados, en la que se tendrán diferentes valores de velocidad de viento en distintos lugares de la región, la dimensión del espacio de entrada puede ser muy elevada, con tantas dimensiones como coordenadas con velocidades de viento distintas se tengan en cuenta. En estos casos es inviable la construcción de modelos paramétricos globales que se ajusten adecuadamente a toda la superficie que relaciona las variables de entrada, o velocidades de viento en varios lugares de la región, con la variable de salida o potencia eólica total generada. Para este tipo de sistemas más complejos, una buena opción es acudir al modelado local.

En esta tesis se propone este tipo de modelado local, que es conocido en la literatura como aprendizaje basado en la memoria o “lazy learning” [97], ya que aplaza el procesamiento de los datos de entrenamiento hasta que se necesita dar una respuesta a una pregunta. En el caso que nos ocupa, la pregunta o entrada al sistema será un vector con valores de velocidad de viento previsto en algunas coordenadas de la región, y la

respuesta o salida será la potencia eólica total esperada en la región. Para dar respuesta a la pregunta, se acudirá a un histórico de datos almacenados en la memoria de la forma (\mathbf{v}_i, P_i) , formado por vectores con velocidades de viento en algunas coordenadas de la región y la potencia eólica total registrada en la región para esos vectores de viento.

3.3 Estructuras de modelos locales

En los modelos de aprendizaje basados en la memoria, el punto para el que se busca la predicción ya es conocido durante el procesado de los datos de entrenamiento. Por tanto, lo que se busca es modelar nuestro problema exclusivamente en el entorno local del vector de entrada para el que se desea una respuesta. Así, los modelos locales tratan de ajustar los datos de entrenamiento solamente en el entorno cercano al punto para el que se desea calcular la predicción, combinando los datos cercanos al punto que se quiere predecir para obtener un modelo (Figura 3.3).

Este tipo de modelado implica almacenar grandes cantidades de datos de entrenamiento en memoria, y acceder a ellos cada vez que se quiere calcular una predicción. Para que el algoritmo no tenga que tratar todos los datos históricos disponibles, se seleccionan los datos más relevantes de entre la gran colección almacenada. Esta relevancia suele medirse con alguna función de distancia, que evalúa la similitud entre los vectores de entrada históricos y el vector que define la nueva entrada para la que se desea una predicción. Así, la modelización local se basa en la idea de que entradas similares en un sistema proporcionarán salidas o respuestas similares.

Por tanto, el objetivo del ajuste de un modelo local es encontrar la relación existente entre dos variables, que en nuestro caso serán el viento y la potencia, usando para ello pares de datos (\mathbf{v}_i, P_i) observados. En general, la relación se modela como:

$$P_i = g(\mathbf{v}_i) + e_i, \quad i = 1, \dots, n, \quad (3.1)$$

donde e_i es el error aleatorio alrededor de la regresión o ajuste $g(\cdot)$.

Dado un conjunto de pares de datos (\mathbf{v}_i, P_i) , correspondientes a las observaciones de potencia y viento, el objetivo de un modelo local por suavizado es obtener la relación funcional entre las variables, dada por $g(\cdot)$ cuando se utiliza el modelo de la ecuación (3.1).

Se puede proporcionar una predicción a partir de los datos almacenados mediante una simple selección y promedio de datos cercanos sin tener en cuenta la distancia de cada dato concreto, o se puede tener en cuenta la distancia entre el nuevo dato y los almacenados para dar una predicción mediante una ponderación local.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Para dar mayor importancia en el cálculo a los datos almacenados más cercanos a la pregunta o punto para el que se desea hacer una predicción se utilizan funciones de ponderación, que tienen en cuenta la distancia asignando más relevancia o peso a los datos históricos que se encuentran más cerca del punto para el que se desea calcular la predicción.

La forma en que se seleccionen y combinen los datos, el uso de funciones de distancia y ponderación, y las asunciones sobre la forma de la superficie que define la distribución de la variable respuesta, nos llevarán al planteamiento de distintas estructuras de modelos locales. Algunos ejemplos de estos modelos locales pueden ser la selección del vecino más cercano, la Media Ponderada o la Regresión Local Ponderada [98]:

- **Selección del vecino más cercano:** Un modelo de selección del vecino más cercano simplemente seleccionaría el dato registrado en el pasado más parecido al dato para el que se busca una predicción, y ofrece como predicción el mismo valor de la variable respuesta que tuvo este punto más cercano. En el caso eólico consistiría en dar como predicción la potencia medida durante la situación meteorológica más parecida registrada en el pasado. Sin embargo, pese a su atractiva sencillez, este tipo de modelado tiene el problema de presentar gran variabilidad en la variable respuesta, además de ser muy sensible a los datos erróneos o atípicos almacenados. Una extensión a este tipo de modelos son los modelos K-NN (K-Nearest Neighbor), en los que la predicción es la media de la variable respuesta correspondiente al conjunto K de datos más cercanos. Estos modelos evitan la variabilidad propia de los modelos de vecino más cercano, gracias al suavizado que se deriva de promediar varios valores históricos.
- **Media Ponderada:** Los modelos locales de Media Ponderada promedian los valores almacenados de la variable respuesta, en nuestro caso las potencias eólicas medidas de unos cuantos datos cercanos, ponderándolos teniendo en cuenta la distancia al punto para el que se desea calcular la predicción.
- **Regresión Local Ponderada:** La Regresión Local Ponderada trata de ajustar una curva o superficie a los puntos cercanos, usando para ello una regresión paramétrica ponderada con las distancias de los datos. Esta superficie ajustada localmente puede ser una recta o plano, en cuyo caso se trata de una regresión lineal, o puede tratar de ajustarse una superficie más compleja, en la que se tengan en cuenta curvaturas y por tanto términos de mayor orden no lineal.

En cualquier caso, tanto si la estructura de los modelos locales a utilizar es de Media Ponderada como de Regresión Local Ponderada, se deben tener en cuenta ciertas características que determinarán el suavizado de la respuesta, entre las que se encuentran las funciones de distancia para evaluar la similitud entre vectores de entrada, el “ancho de banda” o criterio que define los datos que serán seleccionados en función de su distancia, y las funciones de ponderación de los datos seleccionados.

3.3.1 Media Ponderada

Un modelo muy sencillo para obtener la predicción de potencia podría estar basado en la media simple de los valores históricos de potencia correspondiente a los datos más cercanos seleccionados, ajustando una constante como modelo local, sin utilizar factores de ponderación que tengan en cuenta la distancia concreta de cada dato seleccionado:

$$\hat{P}_{t+h|t} = \frac{\sum_{m=1}^M P_m}{m} \quad (3.2)$$

donde:

$\hat{P}_{t+h|t}$ = Potencia total estimada en la región en la hora t para la hora $t+h$

P_m = Potencia histórica medida en la región en el instante m

$m = 1, \dots, M$ son los datos cercanos seleccionados, según algún criterio de distancia umbral o ancho de banda

La expresión de la ecuación (3.2) es la de una media simple en la que no se está teniendo en cuenta la distancia de cada dato seleccionado. La Media Ponderada sí tendrá en cuenta la distancia de los datos. La aproximación de ajustar constantes usando un criterio de entrenamiento ponderado localmente se conoce en la literatura estadística como Regresión Kernel [99]. Nadaraya y Watson ya propusieron usar una Media Ponderada para un conjunto o set de vecinos cercanos usando Kernels dentro del contexto de regresión [100],[101].

La ponderación se puede interpretar como un modo de dar más peso a los datos relevantes y descartar los datos irrelevantes. Esta relevancia se mide por medio de una medida de distancia, que en el caso de predicción eólica sería $d_{i,t+h}=d(\mathbf{v}_i, \mathbf{v}_{t+h})$, la distancia entre el nuevo vector de viento \mathbf{v}_{t+h} y cada vector \mathbf{v}_i de vientos históricos almacenados.

Para calcular el peso de los vectores históricos teniendo en cuenta su distancia al nuevo vector de datos de viento, se utiliza una función de ponderación de la forma $\omega_{i,t+h}=\omega(d(\mathbf{v}_i, \mathbf{v}_{t+h}))$. Esta función de ponderación tendrá un valor máximo para distancia cero, es decir, para el punto exacto en el que se quiere calcular la predicción, y su valor irá decayendo progresivamente conforme aumente la distancia, hasta hacerse cero para distancias superiores a un ancho de banda prefijado, de modo que los datos a partir de esa distancia no se tengan en cuenta. El peso calculado es utilizado finalmente para el cálculo de la Media Ponderada.

Métodos de Predicción de la Generación Agregada de Energía Eólica

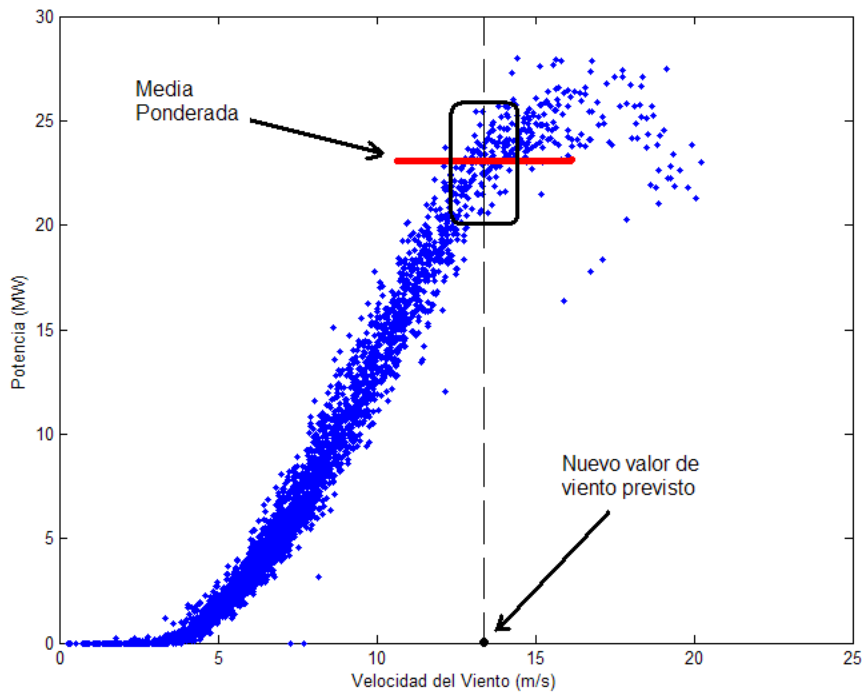


Figura 3.4: Media Ponderada

El modelo de Media Ponderada se define como:

$$\hat{P}_{t+h} = \frac{\sum_{i=1}^n P_i \cdot \omega_{i,t+h}}{\sum_{i=1}^n \omega_{i,t+h}} \quad (3.3)$$

donde:

P_i = Potencia medida en la región en el instante histórico i

$\omega_{i,t+h} = \omega(d(\mathbf{v}_i, \mathbf{v}_{t+h}))$ = Coeficiente de ponderación por distancia entre el vector de viento previsto para la hora $t+h$, y el vector i almacenado.

Es importante insistir en que la estimación de la potencia prevista depende de la localización del nuevo punto, definida por el vector de viento previsto \mathbf{v}_{t+h} para el que se desea calcular la predicción, pues los coeficientes de ponderación dependen a su vez de la posición de dicho punto, lo que nos recuerda que estamos haciendo un modelado local (Figura 3.4).

3.3.2 Regresión Polinómica Local Ponderada

En la Regresión Local Ponderada el objetivo es ajustar modelos locales a una nube de puntos, sean curvas o superficies, utilizando los datos cercanos por medio de un suavizado. Se trata de ajustar un modelo simple al punto local de interés, en lugar de utilizar un complejo modelo global para tratar de aproximarlos a todos los datos existentes [102],[103]. Así, para un punto dado, definido por un vector de datos de entrada, se ajusta una función paramétrica a los datos cercanos a dicho punto. En el caso del modelado de la potencia eólica en una región se tendría:

$$E(P_i|v_i) = g(v_i), \quad i = 1, \dots, n, \quad (3.4)$$

donde P_i son las observaciones de una variable respuesta, en nuestro caso la potencia eólica, y los vectores v_i contienen las observaciones de K variables independientes, que en el caso que nos ocupa son velocidades de viento en K coordenadas de una región.

En regresión local se procura que tanto los modelos locales como el criterio de entrenamiento sean muy simples. Esto nos lleva a buscar modelos que sean lineales en los parámetros a estimar, como los modelos polinómicos locales lineales (Figura 3.5), que son mucho más fáciles de entrenar que otros modelos de mayor orden. Un polinomio lineal localizado en el entorno de v_{t+h} sería de la forma:

$$g(v) \cong \beta_0 + \beta_1(v_1 - v_{1,t+h}) + \beta_2(v_2 - v_{2,t+h}) + \dots + \beta_K(v_K - v_{K,t+h}) \quad (3.5)$$

$$g(v) \cong \beta_0 + \sum_{k=1}^K \beta_k(v_k - v_{k,t+h}) \quad (3.6)$$

donde para simplificar el modelo de regresión, el origen de coordenadas de los datos históricos de entrada al modelo se ha desplazado al punto para el que se desea calcular la predicción. De este modo, la predicción coincidirá con el valor que tome la regresión en el origen de coordenadas, y será por tanto el valor del término independiente en la regresión, β_0 .

Se podrían utilizar modelos locales de mayor orden, como cuadráticos o cúbicos, que aún se ajustan mejor a áreas de gran curvatura con picos y valles, pero su entrenamiento resulta mucho más costoso, y más aún si el número de dimensiones es elevado, como es nuestro caso. Algunos de estos modelos de mayor orden se analizan en [104].

Para estimar los parámetros del polinomio se utiliza el criterio de mínimos cuadrados. Los métodos de estimación de mínimos cuadrados para modelos lineales locales se obtienen a partir de los procedimientos de regresión de modelos lineales globales.

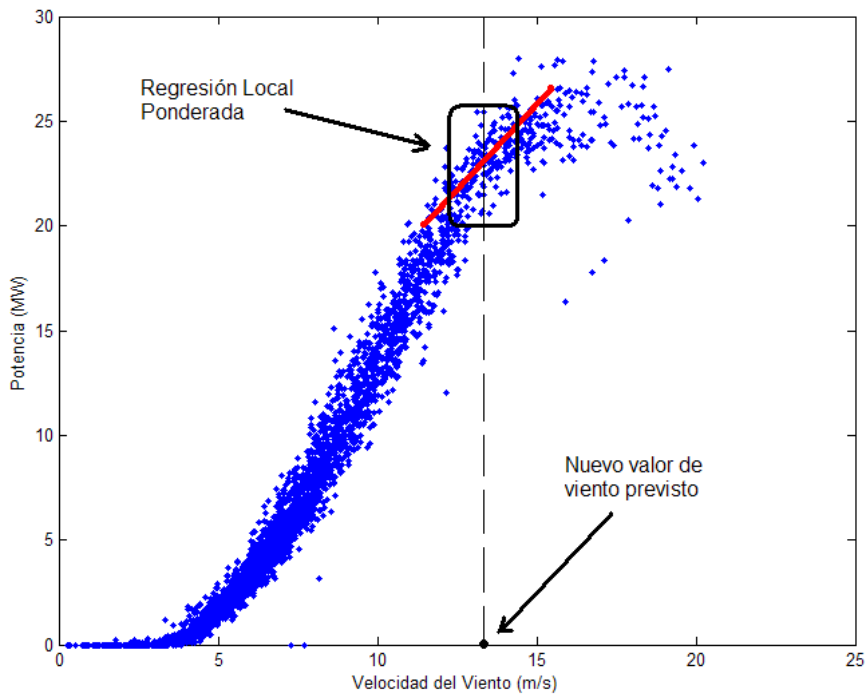


Figura 3.5: Regresión Polinómica Local Ponderada

Un modelo local que es lineal en los parámetros $\beta = [\beta_0, \beta_1, \dots, \beta_k]'$ se puede expresar, a partir de la expresión (3.6) como [97],[105]:

$$P_i = g(v_i) + e_i = \mathbf{x}_i' \boldsymbol{\beta} + e_i \quad (3.7)$$

siendo $\mathbf{x}_i' = [1, (v_i' - v_{t+h}')] = [1, (v_{1,i} - v_{1,t+h}), \dots, (v_{K,i} - v_{K,t+h})]$, el vector histórico de viento con el origen desplazado al punto definido por el nuevo vector de viento para el que se quiere calcular la predicción, y al que se ha añadido la constante 1 para incluir un término independiente en la regresión. El término de error e_i engloba al resto de variables que pueden influir en el valor de P_i . Si ponemos todos los datos de entrenamiento en forma matricial, se tiene:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad (3.8)$$

donde \mathbf{X} es una matriz cuya i -ésima fila es el vector \mathbf{x}_i' , y el vector $\mathbf{y}' = [P_1, \dots, P_n]$ contiene las medidas históricas de potencia P_i correspondientes. Así, la matriz \mathbf{X} es de dimensiones $n \times (K+1)$, donde n es el número de datos históricos almacenados y K es el número de variables de entrada, que en nuestro caso es el número de coordenadas geográficas en la región para las que se tienen predicciones de viento.

Para estimar los parámetros β de un modelo $\hat{P}_i = g(x_i, \hat{\beta})$ usando el criterio de mínimos cuadrados se busca minimizar el criterio:

$$Criterio(\beta) = \sum_i (\hat{P}_i - P_i)^2 = \sum_i (x_i' \beta - P_i)^2 \quad (3.9)$$

de modo que la estimación por mínimos cuadrados de los parámetros β del polinomio que minimizan $Criterio(\beta)$ se calculan como [98]:

$$\hat{\beta} = (X'X)^{-1}X'y \quad (3.10)$$

Al ajustar una recta, plano o hiperplano, dependiendo del número de dimensiones del espacio de entrada, a un conjunto de puntos, una regresión no ponderada como la de la ecuación (3.10) otorgaría el mismo peso a los puntos distantes que a los puntos cercanos al origen, o punto para el que se quiere calcular la predicción, con lo que no se tendría una estimación suficientemente localizada de la regresión.

Dado que nuestro objetivo consiste en calcular la predicción de una variable respuesta en un problema no lineal, necesitamos linealizar únicamente en el entorno del dato de entrada para el que se busca la predicción. El modelo lineal local puede especializarse al punto objetivo, dando más peso a los datos cercanos y menor peso o incluso ignorando los puntos más alejados. Para ello, se calcula la distancia de cada dato almacenado al nuevo dato de viento, y se asigna un peso por medio de una función de ponderación $\omega_{i,t+h} = \omega(d(v_i, v_{t+h}))$, que tenga en cuenta la distancia entre vectores de viento.

Para ponderar cada dato histórico con el peso correspondiente se utiliza la matriz diagonal Ω , en la que los elementos de la diagonal son $\Omega_{ii} = \omega_{i,t+h}$, de modo que la estimación de β se obtiene como:

$$\hat{\beta} = (X'\Omega X)^{-1}X'\Omega y \quad (3.11)$$

de donde se obtendrá la predicción de potencia como la estimación del término independiente en la regresión, que es el punto de corte de la recta o plano con el origen de coordenadas de las variables de entrada, es decir, la estimación de su valor a distancia cero:

$$\hat{P}_{t+h} = \hat{\beta}_0 \quad (3.12)$$

Es importante observar que cada problema de regresión local define una función $g(v_{t+h})$ diferente para cada nuevo punto v_{t+h} , de modo que cada vez que se hace la regresión los pesos cambian y también cambian los parámetros $\hat{\beta}$ estimados, es decir, se está haciendo una estimación local de parámetros.

3.3.3 Ventajas e inconvenientes de la Media Ponderada y la Regresión Local Ponderada

Los métodos de modelado ponderados localmente en general minimizan el coste computacional de entrenamiento, pues los nuevos datos sencillamente se almacenan en la memoria, y el coste computacional radica en la búsqueda de un modelo adaptado a una selección de datos cercanos. Tanto el modelo de Regresión Polinómica Local Ponderada como el modelo de Media Ponderada almacenan los datos de entrenamiento de forma explícita, y no los procesan hasta que se desea obtener una predicción para un nuevo dato, o lo que es lo mismo, hasta que no se necesita dar respuesta a una pregunta.

La diferencia entre ambos radica en la carga computacional del cálculo necesario para estimar la predicción. Un modelo de regresión polinómica local requiere unos procedimientos de regresión relativamente complejos, con lo que su coste computacional será muy superior al necesario para los modelos de vecino más cercano o de Media Ponderada, ya que para cada punto que se desea predecir es necesario construir un modelo de regresión lineal distinto con parámetros locales distintos, en el que el número de parámetros a estimar es relativamente elevado.

En general, para datos distribuidos de forma regular, la regresión Kernel o Media Ponderada, y la Regresión Local Ponderada son equivalentes [106], ya que la Media Ponderada no es más que un caso concreto de Regresión Polinómica Local en el que el polinomio es de orden cero, ajustando una constante. Sin embargo, si los datos se distribuyen de forma irregular, la regresión polinómica local aporta cierta ventaja sobre la Media Ponderada, ya que ajusta una superficie a la nube de datos, en vez de dar un valor único en el entorno del punto objetivo [107]. Por tanto, los modelos locales constantes o de Media Ponderada resultan muy útiles y sencillos para manejar dispersiones de puntos uniformes, mientras que los modelos polinómicos locales lineales se comportan mejor en zonas con mayor curvatura, ya que pueden modelar mejor datos no lineales.

A medida que se aumenta el orden del modelo se reduce el sesgo, pero aumenta la complejidad y la varianza de las estimaciones [108],[103], pues se trata de estimar mayor número de parámetros con la misma cantidad de datos. Así, el orden del modelo local influye en el suavizado, siendo menos suave el ajuste conforme mayor sea el orden del modelo local. Esto nos lleva a plantearnos un compromiso entre reducir el sesgo o reducir varianza, y por tanto a un compromiso en la elección del orden del modelo local a ajustar.

Por otro lado, si en general para cualquier proceso estadístico es necesaria una colección suficiente de datos, en los modelos locales este requerimiento es crucial, y aumenta conforme aumenta el número de dimensiones del espacio. Un problema potencial es que en el entorno del punto que se desea predecir no haya bastantes datos históricos, de modo que si estamos utilizando un modelo de Regresión Local Ponderada, la matriz $X'\Omega X$ de la ecuación (3.11) puede ser difícilmente invertible

debido a que no haya suficientes datos para estimar los parámetros β . En esas ocasiones podría funcionar mucho mejor un modelo de Media Ponderada.

Debido al elevado número de dimensiones de nuestro problema, en el que tenemos tantas dimensiones como coordenadas de predicción de viento seleccionadas en la región, no se plantearán modelos de regresión de mayor orden que los modelos polinómicos lineales. Dado que la complejidad del problema no es conocida a priori, en esta tesis se utilizarán tanto modelos locales de Media Ponderada como Regresión Local Ponderada con modelos polinómicos lineales para calcular la predicción, puesto que ambas aproximaciones tienen ciertas ventajas e inconvenientes. El uso de estas dos aproximaciones en paralelo permitirá adaptar el modelado a la evolución del problema, de modo que finalmente se pueda hacer una combinación adaptativa de las predicciones obtenidas mediante ambas aproximaciones.

3.4 Método propuesto para el cálculo de predicción eólica regional

Como hemos visto, para calcular la predicción de potencia eólica regional siguiendo un método de aprendizaje basado en la memoria, se necesita un conjunto histórico de datos. Estos pares de datos históricos (v_i, P_i) estarán formados por vectores históricos con valores de velocidad de viento $v_i' = [v_{1,i}, \dots, v_{K,i}]$, para un set de K coordenadas geográficas seleccionadas en la región (Figura 3.6), y la potencia eólica total P_i , generada en la región, durante los $i=1, \dots, n$, instantes correspondientes registrados en el periodo histórico.

El método propuesto se podría resumir del siguiente modo (Figura 3.7):

- En el instante t , llega un nuevo vector $v_{t+h} = [v_{1,t+h}, \dots, v_{K,t+h}]'$, con predicciones de velocidad de viento en algunas coordenadas seleccionadas en la región para un determinado horizonte de predicción h , donde $v_{k,t+h}$, $k=1, \dots, K$, es la velocidad de viento prevista en la coordenada k para el instante $t+h$. Estas predicciones de velocidad de viento serán suministradas por una agencia de previsión meteorológica.
- Este nuevo vector de predicción v_{t+h} se compara con los vectores históricos de viento v_i almacenados, por medio de una medida de distancia, $d_{i,t+h}$, que resume en un escalar la similitud entre el nuevo vector de predicción de viento v_{t+h} y cada uno de los vectores de viento históricos v_i almacenados.
- Una vez calculadas las distancias, se hace una selección de los M datos más cercanos.

Métodos de Predicción de la Generación Agregada de Energía Eólica

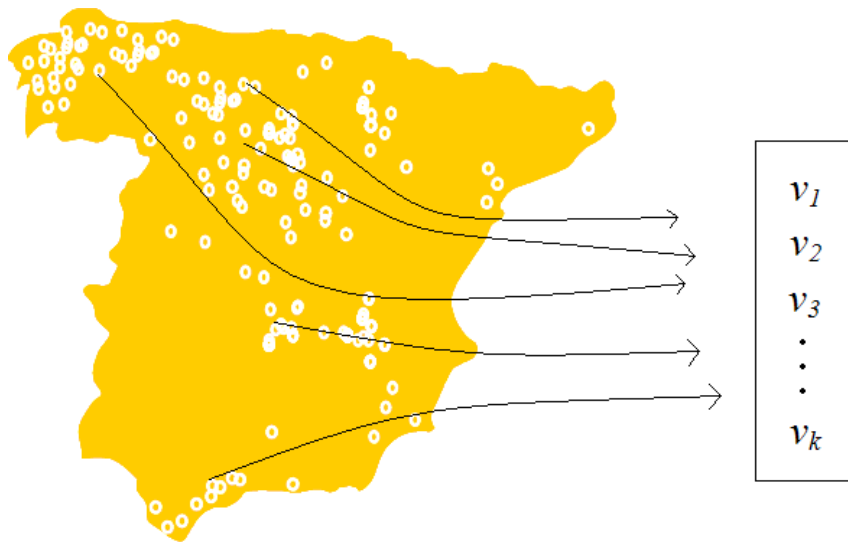


Figura 3.6: Selección de un set de K coordenadas en la región, donde se tendrán las previsiones de velocidad del viento

- Posteriormente, unas funciones de ponderación calculan los coeficientes $\omega_{m,t+h}$, que tienen en cuenta la distancia de los datos históricos y también la antigüedad de los mismos. Estas funciones de ponderación estarán afectadas por una serie de hiperparámetros.
- Finalmente, usando los valores históricos medidos de potencia total P_m , correspondientes a los $m=1, \dots, M$ vectores de viento más cercanos seleccionados, la potencia agregada prevista para la región se calcula mediante dos procedimientos alternativos: Media Ponderada y Regresión Polinómica Local Ponderada, afectados por los coeficientes $\omega_{m,t+h}$ calculados anteriormente.
- Estos dos procedimientos proporcionarán sendas predicciones de potencia, $\hat{P}_{t+h|t}^{(MP)}$ y $\hat{P}_{t+h|t}^{(RLP)}$, que serán combinadas para obtener la predicción final de potencia agregada $\hat{P}_{t+h|t}$, calculada en el instante t para el instante $t+h$.
- Por último, un módulo de estimación adaptativa servirá para ajustar los hiperparámetros del modelo de predicción.

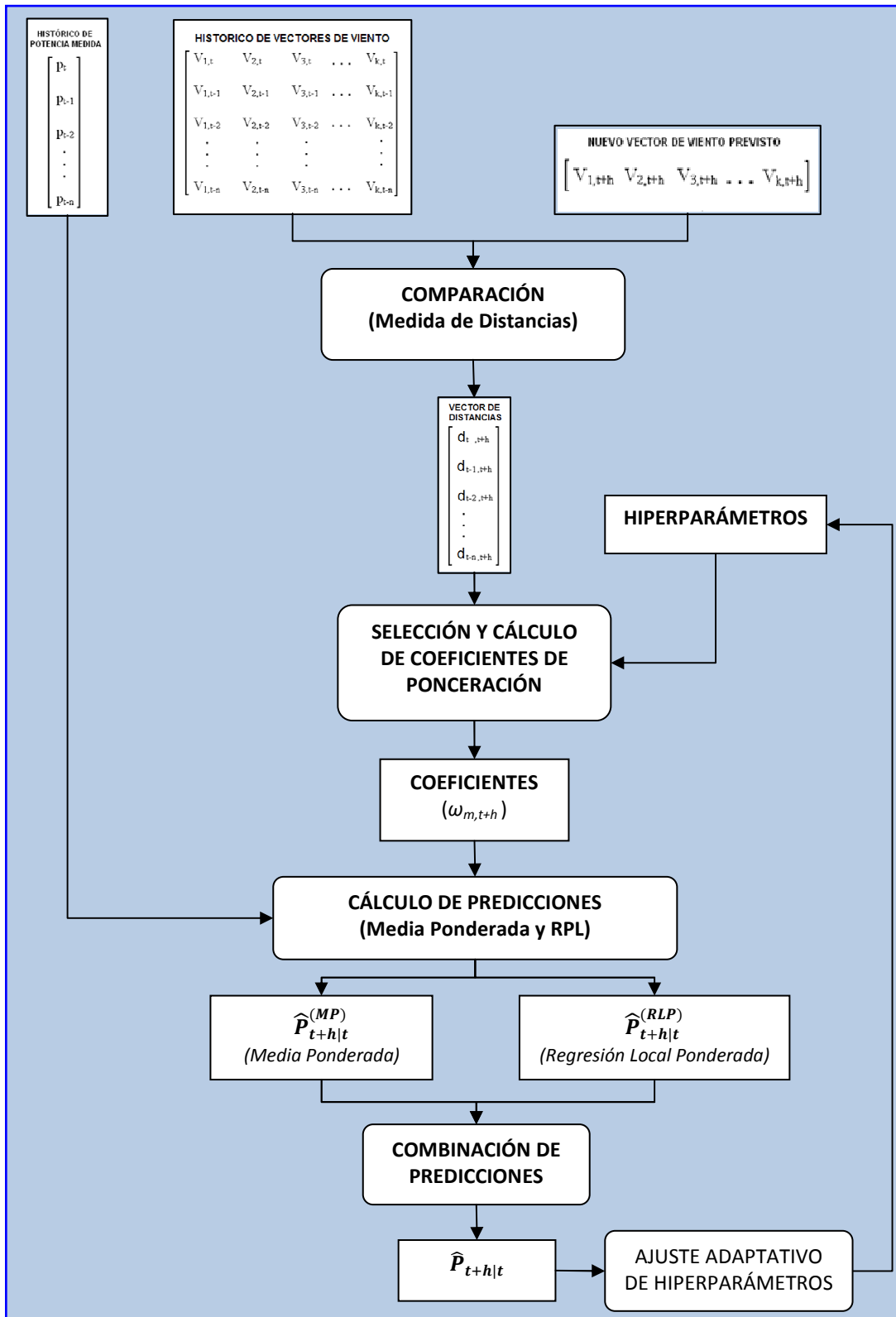


Figura 3.7: Esquema del algoritmo de predicción regional propuesto

3.5 Métodos de evaluación de los modelos de predicción

Existen varios modos de evaluación ya propuestos en la literatura para la comparación entre diferentes modelos con distintos casos de estudio [109]. Para conocer la bondad de los diferentes procedimientos de cálculo de predicciones agregadas que se proponen y poder comparar su precisión, se calcularán los errores de las predicciones calculadas para horizontes desde 1 hasta 48 horas.

El error de predicción se define como la diferencia entre la potencia medida y la potencia prevista por el modelo. Para que su valor sea independiente del valor de potencia instalada, se normaliza el error con la potencia instalada en el parque eólico o en la región de estudio en cada momento, de este modo se puede comparar la bondad de modelos entre parques eólicos o regiones con potencias instaladas diferentes.

$$e_{t+h|t} = \frac{P_{t+h} - \hat{P}_{t+h|t}}{P_{nom,t}} \quad (3.13)$$

donde:

$e_{t+h|t}$ = Error de la predicción hecha en el instante t para el instante $t+h$

P_{t+h} = Potencia real medida en la región en el instante $t+h$

$\hat{P}_{t+h|t}$ = Potencia prevista por el modelo agregado en el instante t para el instante $t+h$

$P_{nom,t}$ = Potencia total instalada en la región en el instante t

Existen varias medidas para describir la precisión de los modelos de predicción, aunque los más utilizados son el Error Medio Absoluto Normalizado (NMAE) y la Raíz del Error Cuadrático Medio Normalizado (NRMSE). A continuación se describen éstos y otros criterios de evaluación empleados en el contexto de la predicción de potencia eólica.

3.5.1 Sesgo del error de predicción

El error puede ser descompuesto en una componente sistemática y una componente aleatoria. Idealmente, el error de un modelo de predicción debería ser no sesgado y tendría por tanto una componente sistemática nula. El sesgo del error a un horizonte dado se definirá por tanto como la media del error cometido por el modelo de predicción a ese horizonte:

$$\bar{e}(h) = \frac{1}{N} \sum_{t=1}^N e_{t+h|t} = \frac{1}{N} \sum_{t=1}^N \left(\frac{P_{t+h} - \hat{P}_{t+h|t}}{P_{nom,t}} \right) \quad (3.14)$$

donde:

$\bar{e}(h)$ = Sesgo o media del error de predicción para el horizonte de predicción h

N = número de predicciones calculadas con el modelo en el periodo ensayado

3.5.2 Error Medio Absoluto Normalizado (NMAE)

El error medio absoluto normalizado, asociado al primer momento de la distribución de error de la predicción, es la media de los errores en valor absoluto de las predicciones calculadas respecto de las medidas de potencia real. Se normaliza con la potencia nominal o instalada con el fin de poder comparar el error de un modelo independientemente de la potencia instalada en la región o parque en que se esté probando. De este modo, el error absoluto normalizado tendrá un valor entre 0 y 1. Se define como:

$$NMAE(h) = \frac{1}{N} \sum_{t=1}^N |e_{t+h|t}| = \frac{1}{N} \sum_{t=1}^N \frac{|P_{t+h} - \hat{P}_{t+h|t}|}{P_{nom,t}} \quad (3.15)$$

donde $NMAE(h)$ es el Error Medio Absoluto Normalizado calculado para horizonte de predicción h

3.5.3 Error Cuadrático Medio Normalizado (NMSE y NRMSE)

El error cuadrático medio normalizado se define como:

$$NMSE(h) = \frac{1}{N} \sum_{t=1}^N (e_{t+h|t})^2 = \frac{1}{N} \sum_{t=1}^N \left(\frac{P_{t+h} - \hat{P}_{t+h|t}}{P_{nom,t}} \right)^2 \quad (3.16)$$

donde $NMSE(h)$ es el Error Cuadrático Medio Normalizado calculado para horizonte de predicción h

El error cuadrático medio está asociado a la varianza o el segundo momento de la distribución de error, y penaliza los instantes con errores grandes en mayor medida que el error medio absoluto. Esto lo convierte en una medida de error muy útil en el ámbito de la predicción eólica, pues es en los momentos de mayor error de predicción de generación cuando el sistema eléctrico se pone en mayor riesgo.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Alternativamente se puede utilizar la raíz del error cuadrático medio normalizado (NRMSE), con el fin de mantener las unidades del error, como:

$$NRMSE(h) = \sqrt{NMSE(h)} \quad (3.17)$$

donde $NRMSE(h)$ es la raíz del Error Cuadrático Medio Normalizado calculado para horizonte de predicción h

3.5.4 Coeficiente de mejora de la predicción

Para cuantificar la mejora en el error de predicción que presenta un modelo evaluado con respecto a otro que se tome como referencia, se realiza una comparación relativa entre los valores de alguno de los indicadores de error definidos anteriormente. El coeficiente de mejora se define entonces como:

$$mejora_{eval,ref}(h) = \frac{CE_{ref}(h) - CE_{eval}(h)}{CE_{ref}(h)} \quad (3.18)$$

donde $mejora_{eval,ref}(h)$ representa la reducción porcentual del error del modelo evaluado con respecto al modelo de referencia. “CE” denota el Criterio de Error, que podría ser por ejemplo el NMAE o NRMSE, “eval” denota el modelo que está siendo evaluado, y “ref” denota el modelo de referencia contra el que se esté haciendo la comparación.

4. Métodos de Suavizado

Resumen

En este capítulo se describen las características de las funciones clásicas de suavizado por Kernel que suelen ser utilizadas para la estimación de predicciones por modelado local. Como alternativa a estas funciones Kernel, y dada la complejidad del problema multivariante que nos ocupa, se propone un método de suavizado en el que se tratan de forma independiente los tres problemas a resolver en el suavizado: la forma en que se miden las distancias entre vectores de viento, el ancho de banda que define la ventana de selección de datos, y las funciones de ponderación de los datos seleccionados en función de su distancia.

Se proponen varias alternativas de medidas de distancia, pudiendo tener en cuenta factores de escala que distorsionen el espacio de entrada. También se sugieren distintos procedimientos de selección de datos, tanto de ancho de banda fijo como de ancho de banda variable. Por último, se proponen y resumen las características de varias funciones de ponderación de los datos para tener en cuenta tanto la distancia al punto a predecir como la antigüedad de los mismos.

4.1 Funciones Kernel de suavizado

Un método de ponderación utilizado muy habitualmente para el modelado local es el suavizado por Kernel. El suavizado por Kernel parte de una ventana de la forma $[v_{t+h} - \delta, v_{t+h} + \delta]$, donde δ se conoce como ancho de banda, y determina el tamaño de una ventana de selección (Figura 4.1). Esta ventana determina el conjunto de datos que serán utilizados para calcular el ajuste del modelo local en el entorno del punto objetivo v_{t+h} para el que se busca la predicción P_{t+h} .

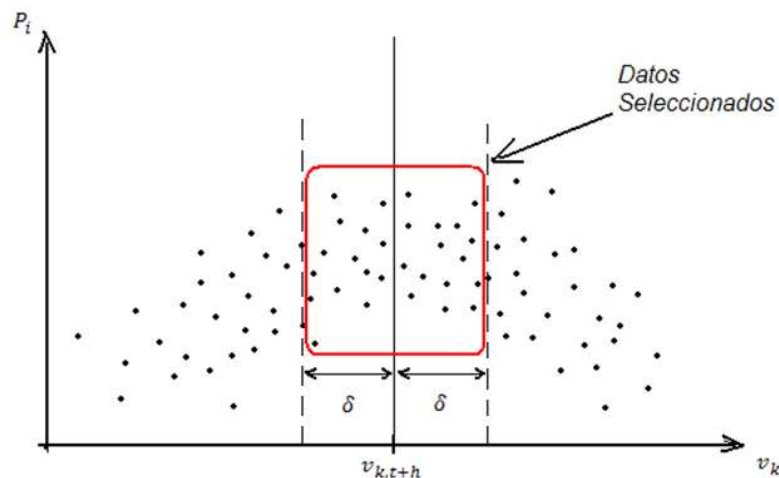


Figura 4.1: Ventana de datos a seleccionar según la distancia y el ancho de banda en la dimensión k

Una vez establecido el ancho de banda o ventana de selección, quedan definidos los datos que son seleccionados, de modo que a mayor ancho de banda, mayor número de datos será seleccionado, y por tanto mayor será el suavizado. En las funciones Kernel, la distancia relativa de los datos de entrenamiento almacenados respecto del punto para el que se busca una predicción se mide en cada dimensión mediante u_k (Figura 4.2). El valor u_k se calcula para cada dato histórico i en cada coordenada k como la diferencia entre el valor de viento $v_{k,i}$ y el valor $v_{k,t+h}$, dividida por el ancho de banda δ :

$$u_k = \frac{v_{k,i} - v_{k,t+h}}{\delta} \quad (4.1)$$

La función Kernel es una función de ponderación que depende de la distancia u_k calculada en la ecuación anterior, y se suele elegir de modo que el mayor peso se otorgue a las observaciones más cercanas al punto para el que se hace el ajuste, y decaiga conforme aumenta la distancia.

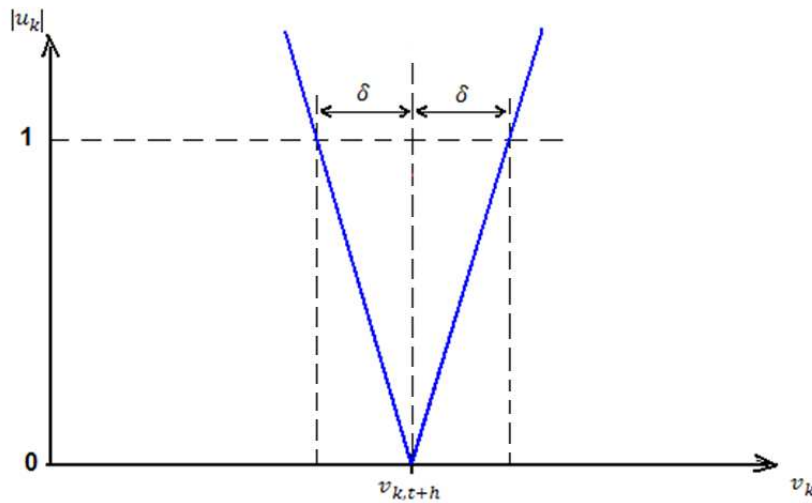


Figura 4.2: Distancia relativa de los datos en la dimensión k en función del ancho de banda

Las funciones Kernel son no negativas y simétricas con respecto al origen. Por otro lado, en general nos interesará que las funciones de ponderación sean distintas de cero solamente en un intervalo, de modo que se asignen pesos nulos a los datos que queden fuera de ese intervalo o ancho de banda, en lugar de asignarles pesos muy pequeños. El motivo principal de este requisito es la velocidad de cálculo, pues al ignorar completamente los datos lejanos, estos no serán utilizados en absoluto en la regresión y el algoritmo de cálculo será mucho más rápido. Además, se suele pedir que el área bajo la curva definida por la función Kernel sea la unidad.

$$\int_{-\infty}^{\infty} K(u_k) du_k = 1 \quad (4.2)$$

Existen múltiples posibilidades en funciones Kernel de ponderación. Una de las funciones Kernel más sencilla es el Kernel uniforme, usado en [110]:

$$K(u_k) = \begin{cases} 1/2, & |u_k| \leq 1 \\ 0, & |u_k| > 1 \end{cases} \quad (4.3)$$

La función Kernel uniforme asigna un peso idéntico de $1/2$ a todas las observaciones o datos cuya distancia al punto para el cual se desea estimar la predicción no es mayor que el ancho de banda δ . De esta forma, los puntos que estén muy alejados del punto a predecir obtendrán un peso nulo, mientras que todos los puntos dentro del intervalo tendrán igual ponderación. El inconveniente de la función uniforme es que puede llevar a estimaciones con ruido, ya que con pequeñas variaciones del punto de entrada se seleccionan o deseleccionan datos históricos de forma brusca, de modo que estos cambios bruscos se pueden transformar en ruido en la estimación.

Parece más razonable asignar un peso mayor a los datos observados que se encuentren más cerca del punto objetivo de nuestra regresión local, teniendo en cuenta la distancia u_k calculada en la ecuación (4.1). Algunos de los Kernel más utilizados y que tienen en cuenta esta distancia son [108]:

Kernel Triangular:

$$K(u_k) = \begin{cases} 1 - |u_k|, & |u_k| \leq 1 \\ 0, & |u_k| > 1 \end{cases} \quad (4.4)$$

Kernel Cuadrático o Epanechnikov [111]:

$$K(u_k) = \begin{cases} 3/4 (1 - u_k^2), & |u_k| \leq 1 \\ 0, & |u_k| > 1 \end{cases} \quad (4.5)$$

Kernel Quartic:

$$K(u_k) = \begin{cases} 15/16 (1 - u_k^2)^2, & |u_k| \leq 1 \\ 0, & |u_k| > 1 \end{cases} \quad (4.6)$$

Kernel Coseno:

$$K(u_k) = \begin{cases} \pi/4 \cos\left(\frac{\pi}{2} u_k\right), & |u_k| \leq 1 \\ 0, & |u_k| > 1 \end{cases} \quad (4.7)$$

Kernel Gaussiano:

$$K(u_k) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u_k^2\right) \quad (4.8)$$

Se puede observar que el Kernel Gaussiano no es truncado, es decir, su ponderación no es nula para datos que quedan fuera del ancho de banda. En este caso, el valor de la función de ponderación decae suavemente conforme aumenta la distancia. En la Figura 4.3 se muestra la forma de todas estas funciones Kernel en función de la distancia u_k .

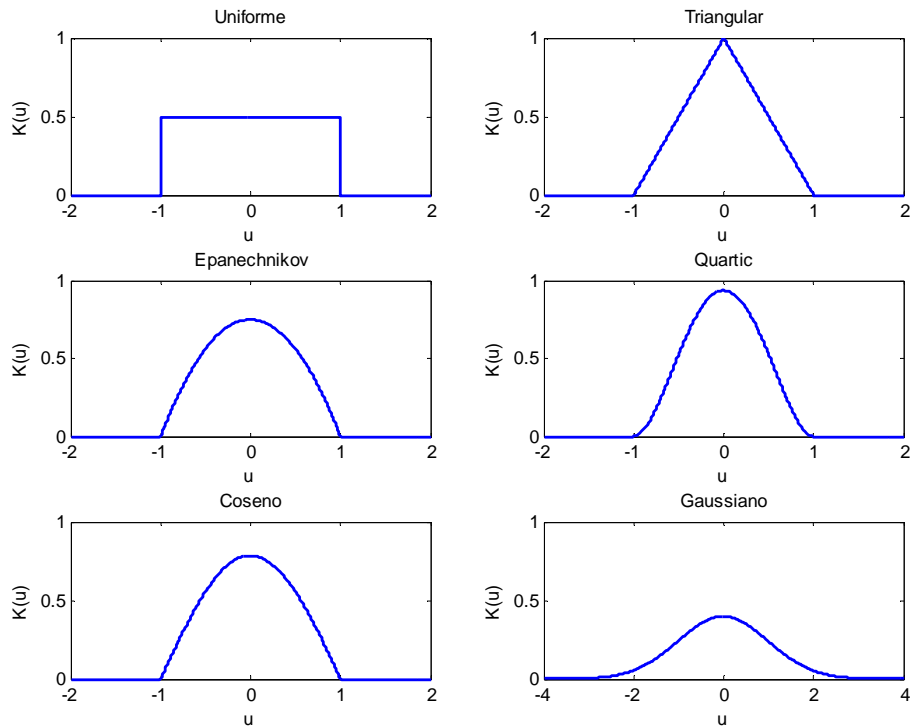


Figura 4.3: Funciones Kernel

Las funciones Kernel expuestas son univariantes, de modo que solo tienen en cuenta una dimensión del espacio. Cuando el problema a resolver es multivariante, como es nuestro caso, en el que tenemos un total de K dimensiones o coordenadas seleccionadas en la región de donde se tienen predicciones de velocidad de viento, la función Kernel multivariante será de la forma $\mathcal{K}(u_1, \dots, u_K)$. Una estructura multivariante para $\mathcal{K}(\mathbf{u})$ puede ser muy compleja. Por esta razón, se suele construir un Kernel multivariante de forma multiplicativa a partir de los Kernel univariantes [108].

$$\mathcal{K}(\mathbf{u}) = K\left(\frac{v_{1,i} - v_{1,t+h}}{\delta}\right) \cdot \dots \cdot K\left(\frac{v_{K,i} - v_{K,t+h}}{\delta}\right) = \prod_{k=1}^K K(u_k) \quad (4.9)$$

En general, en la literatura no se encuentran evidencias de que el uso de una función Kernel u otra sea claramente relevante [103],[112], salvo por la posible sensibilidad al ruido del Kernel uniforme. Sin embargo, el ancho de banda es uno de los principales parámetros de control para el suavizado de las predicciones obtenidas, y por tanto para el ajuste de las funciones Kernel.

El ajuste del ancho de banda δ se deberá hacer siguiendo el criterio de minimizar alguna función de pérdidas. En el caso de la predicción eólica trataremos de minimizar el error cuadrático medio para ajustar el valor del ancho de banda apropiado.

4.2 Método de suavizado propuesto

En la resolución de un problema de predicción por suavizado utilizando funciones Kernel, y una vez elegida una función de entre las muchas alternativas que podemos encontrar en la literatura, el único parámetro a ajustar en nuestro problema será el ancho de banda δ . Este parámetro decide al mismo tiempo dos factores en la ponderación: el conjunto de datos que será seleccionado, y la pendiente o severidad con que la función Kernel disminuye su peso conforme nos alejamos del punto que queremos predecir.

Existe en cambio la posibilidad de proceder de otras formas para hacer el suavizado, calculando medidas de distancias multidimensionales específicas para el problema a tratar y decidiendo el ancho de banda o número de elementos a seleccionar de forma independiente al cálculo de distancias. Por último, a los datos seleccionados de este modo se les pueden aplicar unos pesos calculados con alguna función de ponderación adaptada al problema.

Como parte de esta tesis, se ha hecho un estudio en profundidad de los tres factores que se deben tener en cuenta a la hora de hacer un modelado local del problema de predicción eólica que nos ocupa, que son la forma en que serán medidas las distancias entre vectores de viento, el ancho de banda o proceso de selección de los datos cercanos, y las funciones de ponderación para asignar los pesos.

4.3 Medida de distancias

Los modelos ponderados localmente necesitan una medida de relevancia para calcular los pesos que son aplicados a los datos históricos. Esta relevancia se evalúa mediante una medida de distancia, de modo que los puntos cercanos resulten más relevantes que los puntos más alejados al punto para el que se desea calcular la predicción.

La función de medida de distancia entre dos puntos en un espacio multidimensional necesitará como entrada los dos vectores que representan los dos puntos en el espacio, y como salida se obtendrá un escalar que representa la distancia entre ambos puntos. En nuestro caso, dado que utilizamos vectores de velocidad de viento, la medida de distancia entre una observación de vientos almacenados en la memoria y el nuevo viento previsto será de la forma $d(\mathbf{v}_i, \mathbf{v}_{t+h})$.

En esta sección se proponen y presentan las características de diferentes modos de calcular la distancia entre vectores de velocidad de viento.

4.3.1 Distancia de velocidad media

Este tipo de distancia, el más sencillo de todos los propuestos, calcula la diferencia entre valores medios de velocidad de viento en toda la región. Se define como:

$$d_{i,t+h} = \bar{v}_i - \bar{v}_{t+h} \quad (4.10)$$

$$i = 1, \dots, n$$

donde:

$d_{i,t+h}$ = distancia entre el vector de viento \mathbf{v}_{t+h} previsto para el instante $t+h$ y el vector histórico de viento \mathbf{v}_i correspondiente al instante i almacenado

\bar{v}_i es la velocidad media de viento en la región en el instante histórico i , calculada como la media de las velocidades en todas las coordenadas seleccionadas:

$$\bar{v}_i = \frac{1}{K} \sum_{k=1}^K v_{k,i} \quad (4.11)$$

\bar{v}_{t+h} es la velocidad media de las predicciones de viento en la región para el instante $t+h$, para el que queremos calcular la predicción de potencia:

$$\bar{v}_{t+h} = \frac{1}{K} \sum_{k=1}^K v_{k,t+h} \quad (4.12)$$

K = número de elementos de los vectores de viento, que se corresponde con número de coordenadas seleccionadas de la región para las que se tendrán predicciones de viento

h = horizonte para el que se desea calcular la predicción de potencia eólica

t = instante actual

$v_{k,i}$ = viento en la coordenada k de la región en el instante histórico i

$v_{k,t+h}$ = nueva predicción de viento para la coordenada k en el instante $t+h$

4.3.2 Distancia absoluta

La distancia absoluta entre dos vectores de velocidad de K elementos se definirá como la suma de los valores absolutos de la diferencia entre los módulos de velocidad de cada uno de los elementos de los dos vectores, correspondientes a las K coordenadas seleccionadas en la región.

Para que el orden de magnitud de la distancia sea independiente del número de elementos de los vectores de velocidad de viento, es decir, del número de coordenadas seleccionadas en la región, se normaliza con $1/K$. Así, la distancia entre dos vectores \mathbf{v}_i y \mathbf{v}_{t+h} será:

$$d_{i,t+h} = \frac{1}{K} \sum_{k=1}^K |v_{k,i} - v_{k,t+h}| \quad (4.13)$$

4.3.3 Distancia euclídea

La distancia euclídea entre dos puntos del espacio se define como la longitud de la recta que los une. Si se usa la distancia euclídea para medir similitud entre dos variables multidimensionales, representadas en este caso por vectores de viento de dimensión K , la distancia entre dos puntos definidos por los vectores \mathbf{v}_i y \mathbf{v}_{t+h} será:

$$d_{i,t+h} = \sqrt{\frac{1}{K} \sum_{k=1}^K (v_{k,i} - v_{k,t+h})^2} = \sqrt{\frac{1}{K} (\mathbf{v}_i - \mathbf{v}_{t+h})' \cdot (\mathbf{v}_i - \mathbf{v}_{t+h})} \quad (4.14)$$

La menor distancia en este caso será la que minimice la suma de cuadrados de las diferencias de velocidades de viento. Por tanto, frente a la distancia absoluta, la distancia euclídea consigue penalizar a aquellos vectores históricos de velocidad que contienen alguna coordenada para la que la predicción de viento es muy distinta a la del nuevo vector de velocidad de viento prevista.

4.3.4 Distancia de velocidad cúbica

Como se vio en el capítulo 2, el flujo de energía o potencia eólica disponible en una corriente de aire en movimiento que atraviesa el área de barrido de una turbina depende del cubo de la velocidad del aire, tal y como se mostraba en la ecuación (2.1). Atendiendo a esto, se propone una distancia que depende de las diferencias de velocidad cúbica, de la forma:

$$d_{i,t+h} = \sqrt[3]{\frac{1}{K} \sum_{k=1}^K |v_{k,i}^3 - v_{k,t+h}^3|} \quad (4.15)$$

4.4 Factores de escala. Distancias estandarizadas

Se puede hacer que las funciones de distancia sean asimétricas. De este modo, la importancia relativa de las variables de entrada a la hora de generar la medida de distancia dependerá de cómo es la escala de dichas variables. Así, si por ejemplo en un problema con dos variables de entrada una variable tiene una amplitud de valores posibles muy inferior al rango de valores posibles de la otra variable, es deseable que la medida de distancia se vea afectada en mayor medida por una misma variación de la variable que tiene menor amplitud de valores posible.

Modificar la función de distancias distorsionando el espacio de entrada podría servir para obtener predicciones más precisas. Esta distorsión se consigue aplicando factores de escala η_k diferentes sobre las dimensiones del espacio [97]. El factor de escala η_k será el que se aplique a la dimensión k -ésima.

El factor de escala η_k deberá ser mayor sobre aquellas dimensiones en las que deba darse mayor importancia a las diferencias observadas entre los datos, y menor en las dimensiones en las que las variaciones se consideren menos importantes. Estos η_k se suelen aglomerar en una matriz diagonal Γ , donde $\Gamma_{kk} = \eta_k$.

En la Figura 4.4 se muestran los contornos equidistantes para un espacio de entrada bidimensional, en el que las variaciones en el eje horizontal (eje 1) se consideran más importantes que las variaciones en el eje vertical (eje 2), es decir: $\eta_1 > \eta_2$. Esto se traduce en una distorsión del espacio de entrada, consistente en un aplastamiento sobre el eje horizontal y un estiramiento sobre el eje vertical.

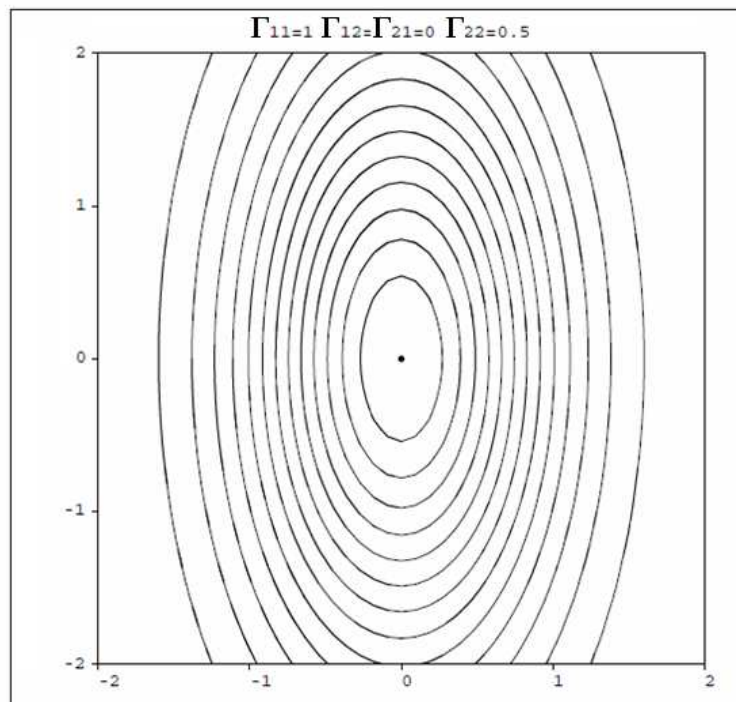


Figura 4.4: Contornos equidistantes usando factores de escala

En un caso extremo, si alguno de los factores de escala se establece en cero, entonces la dimensión correspondiente es ignorada por la función de distancia, y el modelo local se convierte en global para esa dirección, pues todos los puntos alineados en esa dimensión obtienen el mismo peso, es decir, que su mayor o menor distancia sobre esa dimensión resultaría irrelevante. También se podrían utilizar factores de escala muy elevados o infinitos en una de las dimensiones, de modo que cualquier diferencia en esa dimensión entre el dato histórico y el dato actual para el que se quiere calcular la predicción implique que ese dato histórico sea descartado en la ponderación [113],[114].

4.4.1 Distancia absoluta estandarizada con la media

La distancia absoluta por sí sola simplemente calcula la diferencia entre los valores de velocidad en cada coordenada. Sin embargo, parece razonable pensar que esa diferencia entre valores de velocidad tendrá más importancia en aquellas coordenadas en las que las velocidades adquieran habitualmente unos valores menores, y menor importancia cuando la media de velocidades en la coordenada sea superior.

Así, por ejemplo, una diferencia de 1 m/s en una coordenada donde la velocidad media sea de 5 m/s, tendrá mayor peso que la misma diferencia de 1 m/s en una coordenada donde la velocidad habitualmente adquiera valores en el entorno de los 10 m/s.

De este modo, se define la velocidad absoluta estandarizada con la media, utilizando un factor de escala de la forma $\eta_k = \frac{1}{\bar{v}_k}$, como:

$$d_{i,t+h} = \frac{1}{K} \sum_{k=1}^K \eta_k |v_{k,i} - v_{k,t+h}| = \frac{1}{K} \sum_{k=1}^K \left| \frac{v_{k,i} - v_{k,t+h}}{\bar{v}_k} \right| \quad (4.16)$$

$$i = 1, \dots, n$$

donde \bar{v}_k es la media de la velocidad de viento histórico registrado en la coordenada k , calculada con los n datos históricos almacenados hasta el instante t :

$$\bar{v}_k = \frac{1}{n} \sum_{i=1}^n v_{k,i} \quad (4.17)$$

$$k = 1, \dots, K$$

4.4.2 Distancia absoluta estandarizada con la desviación típica

La desviación típica es una medida de dispersión que indica cuánto tienden a alejarse los valores puntuales del promedio de una distribución, de forma que a mayor desviación típica, mayor tenderá a ser el rango de los datos. Al igual que para el cálculo de la distancia absoluta estandarizada con la media, se puede tener en cuenta

Métodos de Predicción de la Generación Agregada de Energía Eólica

que una misma diferencia en el valor de velocidad en una coordenada tendrá más importancia en aquellas coordenadas donde la varianza o desviación típica de la velocidad de viento sea menor, y menor importancia en aquellas coordenadas donde la varianza sea mayor.

Así, por ejemplo, una diferencia de velocidad de viento de 1 m/s en una coordenada donde la desviación típica sea de 2 m/s tendrá mayor importancia que la misma diferencia de 1 m/s en otra coordenada donde la desviación típica sea de 4 m/s.

Esta estandarización estará muy relacionada con la estandarización de la distancia absoluta utilizando la media expuesta anteriormente, ya que suele observarse que el grado de dispersión de los datos sobre el valor promedio de la velocidad en una coordenada es mayor en aquellas coordenadas con una velocidad media de viento mayor, y viceversa.

De este modo, se define la velocidad absoluta estandarizada con la desviación típica, utilizando un factor de escala de la forma $\eta_k = \frac{1}{s_k}$, como:

$$d_{i,t+h} = \frac{1}{K} \sum_{k=1}^K \left| \frac{v_{k,i} - v_{k,t+h}}{s_k} \right| \quad (4.18)$$

donde s_k es la desviación típica de la velocidad de viento en la coordenada k :

$$s_k^2 = \frac{1}{n-1} \sum_{i=1}^n (v_{k,i} - \bar{v}_k)^2 \quad (4.19)$$

4.4.3 Distancia euclídea estandarizada con la varianza

Al igual que se hizo con la distancia absoluta estandarizada con la desviación típica, se puede estandarizar la distancia euclídea con la varianza para tener en cuenta la variabilidad de los valores de velocidad en cada coordenada, de modo que tenga mayor peso la diferencia de velocidad de viento en aquellas coordenadas donde la varianza de la velocidad sea menor.

Utilizando el factor de escala $\eta_k = \frac{1}{s_k}$ ordenado en una matriz diagonal $\mathbf{\Gamma}$, donde $\mathbf{\Gamma}_{kk} = \eta_k$, se tendría:

$$d_{i,t+h} = \sqrt{\frac{1}{K} (\mathbf{v}_i - \mathbf{v}_{t+h})' \mathbf{\Gamma}' \mathbf{\Gamma} (\mathbf{v}_i - \mathbf{v}_{t+h})} = \sqrt{\frac{1}{K} \sum_{k=1}^K \frac{(v_{k,i} - v_{k,t+h})^2}{s_k^2}} \quad (4.20)$$

Siendo s_k^2 la varianza de la velocidad de viento en la coordenada k , calculada con los n datos históricos de viento almacenados.

4.4.4 Distancia de velocidad cúbica estandarizada con la desviación típica

Se puede aplicar un factor de escala a la distancia de velocidad cúbica. Se define así la velocidad cúbica estandarizada con la desviación típica usando el factor de escala $\eta_k = \frac{1}{s_k}$, como:

$$d_{i,t+h} = \sqrt[3]{\frac{1}{K} \sum_{k=1}^K \left| \frac{v_{k,i}^3 - v_{k,t+h}^3}{s_k^3} \right|} \quad (4.21)$$

4.4.5 Distancia de Mahalanobis

Dado que las velocidades de viento en puntos geográficos cercanos dentro de una región no son independientes, es habitual que si se tienen velocidades de viento elevadas en una coordenada, también se tengan velocidades de viento elevadas en otra coordenada situada en un punto cercano dentro de la región de estudio. Es decir, habrá cierta correlación entre las velocidades de viento de coordenadas geográficamente cercanas.

La distancia de Mahalanobis tendrá en cuenta esta dependencia, y por tanto se diferencia de la distancia euclídea estandarizada con la varianza en que, además de las varianzas, tiene en cuenta las covarianzas, calculadas con los registros almacenados de velocidad de viento en las coordenadas. Se define como:

$$d_{i,t+h} = \sqrt{\frac{1}{K} (\mathbf{v}_i - \mathbf{v}_{t+h}) \mathbf{\Gamma}^{-1} (\mathbf{v}_i - \mathbf{v}_{t+h})^T} \quad (4.22)$$

donde $\mathbf{\Gamma}$ en este caso es la matriz de covarianzas calculada con el conjunto de vectores históricos de viento almacenados hasta el instante t .

Esta distancia se calcularía de una forma similar a la distancia euclídea con factores de escala, en los que la matriz $\mathbf{\Gamma}$ ya no es diagonal sino que puede tener valores distintos de cero en cualquier lugar. Se puede observar que la distancia euclídea estandarizada con la varianza, presentada previamente en (4.20), sería un caso particular simplificado de la distancia de Mahalanobis, en el que sólo se tendría en cuenta la diagonal de la matriz de covarianzas $\mathbf{\Gamma}$.

En la Figura 4.5 se representan los contornos equidistantes en un espacio de entrada bidimensional, en el que el factor de escala en el eje horizontal es mayor que el factor de escala en el eje vertical, pero además, se tiene en cuenta la covarianza en las variables, representada por los elementos $\mathbf{\Gamma}_{12}$ y $\mathbf{\Gamma}_{21}$ de la matriz de covarianzas.

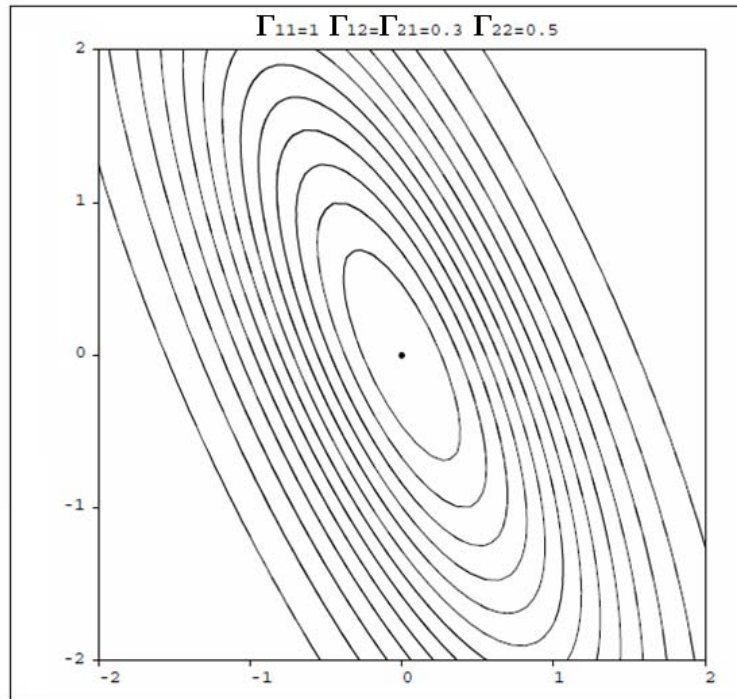


Figura 4.5: Contornos equidistantes usando la distancia de Mahalanobis

4.5 Selección de datos en función de su distancia

Una vez calculada la distancia entre el nuevo vector de velocidad de viento previsto \mathbf{v}_{t+h} y el resto de vectores de viento \mathbf{v}_i almacenados en el set de datos históricos, mediante alguno de los métodos de cálculo de distancia expuestos, se ha de seleccionar un conjunto de datos históricos en función de dichas distancias para estimar la potencia eólica prevista por el modelo agregado mediante algún modelo local.

El parámetro de selección es el ancho de banda δ , que define el alcance de la ventana de selección de datos que serán ponderados. En cada punto para el que se desea calcular una predicción, definido por el vector de entrada, el ancho de banda establece un vecindario del espacio K-dimensional de variables independientes, basándose en alguna medida de distancia. Así, el ancho de banda indica la ventana $[\mathbf{v}_{t+h} - \delta, \mathbf{v}_{t+h} + \delta]$ sobre la que se aplicará el suavizado, de modo que a medida que aumenta el ancho de banda, aumenta el suavizado.

Existe la posibilidad de utilizar un ancho de banda fijo, en el que δ sea un valor constante. Aunque un ancho de banda fijo es fácil de interpretar, lo que le puede hacer muy atractivo, el volumen de datos dentro del intervalo seleccionado por el ancho de banda será mayor o menor dependiendo de la densidad de datos cercanos, lo que puede acarrear grandes variabilidades y ruido en la predicción en zonas con baja densidad de datos, al ser la estimación menos precisa. La probabilidad de que pueda

ocurrir este problema aumenta a medida que aumenta la dimensión del espacio de entrada. En los suavizados con un ancho de banda fijo incluso podría darse el caso de que no se encontrase ningún dato dentro del dominio cercano a determinados vectores para los que se desea una predicción, con lo que no se podría estimar predicción alguna.

También se puede hacer una selección con ancho de banda variable. Un ejemplo sería la selección de M elementos cercanos, en la que de un modo implícito se varía el ancho de banda de modo que el valor de δ se ajusta a la distancia del M -ésimo dato seleccionado. De este modo, los cambios en la escala de las distancias o en la concentración de datos cercanos se compensan con cambios en el ancho de banda. En este caso, cuanto más densamente estén distribuidos los datos observados \mathbf{v}_i alrededor del nuevo punto definido por \mathbf{v}_{t+h} , más pequeño será el ancho de banda necesario en la ventana. Así, en general un ancho de banda variable y dependiente del punto a predecir de la forma $\delta(\mathbf{v}_{t+h})$ se adaptará mejor a la distribución de datos.

En esta tesis se han analizado los resultados obtenidos utilizando tanto un ancho de banda fijo como variable en la selección de datos. Además, se han introducido otras variantes originales para este parámetro de selección, tratando de adaptar la estimación del ancho de banda a la distribución de los datos y sus distancias. A continuación se exponen y analizan estos métodos de selección propuestos.

4.5.1 Datos a una distancia inferior a un valor umbral fijo

Este sería el caso clásico de un ancho de banda fijo. De este modo se seleccionarán los elementos que se encuentren a una distancia inferior a un valor δ fijo propuesto, lo que dará lugar a un número variable M de elementos cercanos. Una vez calculadas las distancias con cada uno de los vectores históricos, la predicción de potencia será calculada con los datos de potencia registrada en los M instantes correspondientes a los M vectores de viento históricos a una distancia inferior a una distancia umbral fija. Esta distancia umbral δ es el parámetro a optimizar en este método de selección.

Si se representa la dispersión de distancias calculadas entre el nuevo vector de viento y los vectores históricos, los datos seleccionados serán aquellos que se encuentren a la izquierda de un valor umbral de distancia. En la Figura 4.6 se representa un ejemplo en el que el ancho de banda de selección es $\delta = 1\text{ m/s}$.

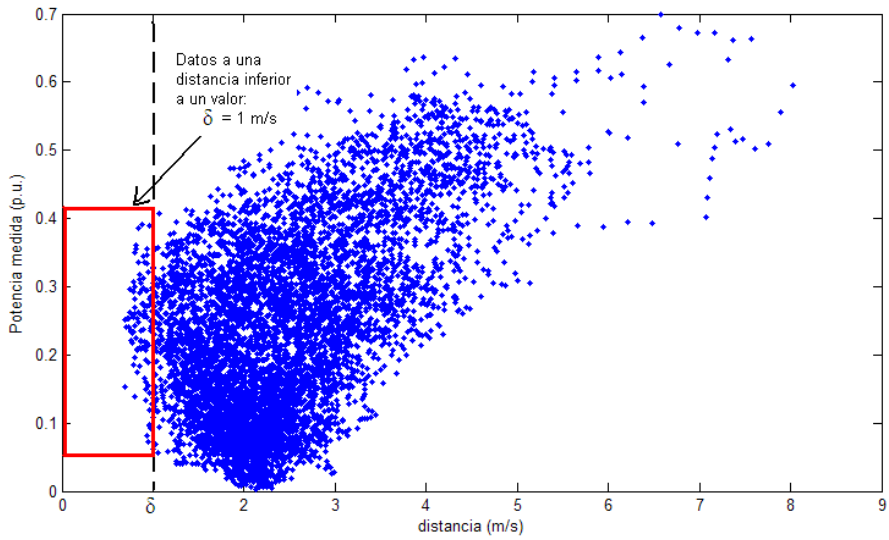


Figura 4.6: Datos a distancia inferior a un valor umbral fijo

4.5.2 M vecinos más cercanos

Según este método de selección, una vez calculadas las distancias con cada uno de los vectores históricos, se seleccionará un número fijo M de elementos cercanos. Se trata por tanto de un método clásico de ancho de ventana variable, que se suele encontrar referenciado en la literatura como K-NN (K-Nearest Neighbor) [115].

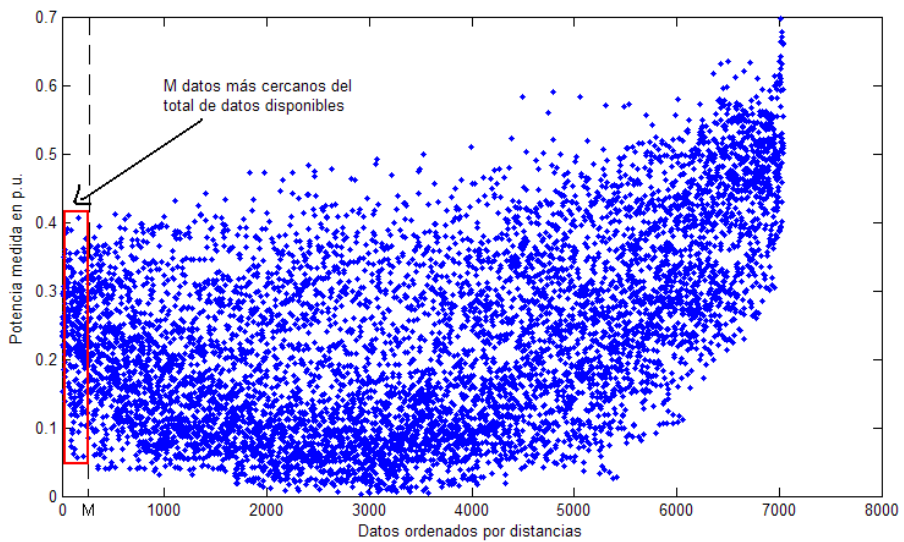


Figura 4.7: Selección de M vecinos más cercanos

Si se ordenan por distancias todos los datos almacenados, se tendrá una dispersión como la de la Figura 4.7, en la que cada dato corresponde a una distancia inferior a la calculada para el dato siguiente. En la parte izquierda de esta dispersión, tal y como se observa con más detalle en la Figura 4.8, se seleccionará un número M fijo de datos, y sus potencias observadas serán usadas para calcular la nueva potencia prevista con el modelo local.

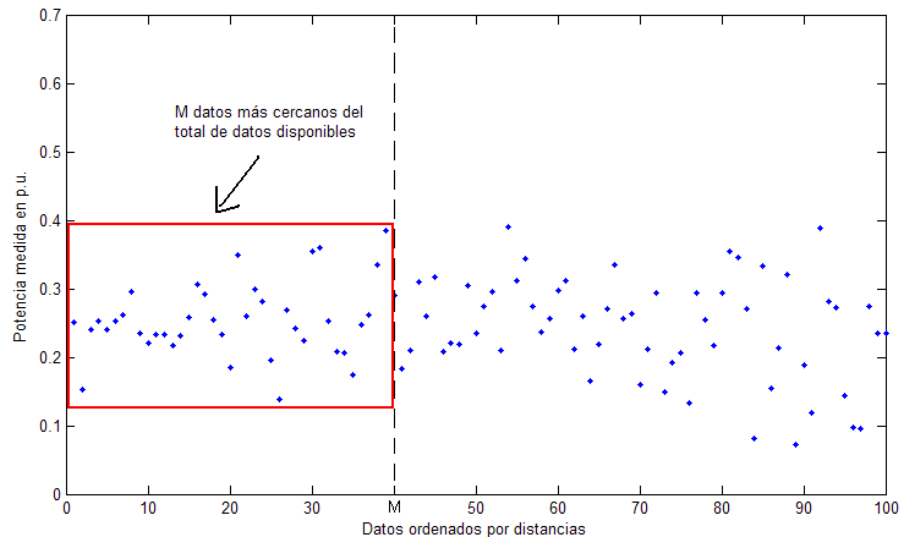


Figura 4.8: Selección de M vecinos más cercanos

4.5.3 M como un porcentaje fijo de datos más cercanos del total de datos disponibles

Mediante este procedimiento, similar al anterior expuesto, se seleccionará un determinado porcentaje de datos p , siendo p el parámetro a optimizar, que dará lugar a un número creciente de elementos cercanos conforme el número de datos observados n vaya aumentando. Una vez calculadas las distancias con cada uno de los vectores históricos, la predicción de potencia será una combinación de la potencia registrada en los M instantes correspondientes a los vectores de viento históricos más cercanos, calculados como un porcentaje de datos de la forma:

$$M = \frac{p \cdot n}{100} \quad (4.23)$$

En este caso, además del ancho de banda δ , también el número M de elementos seleccionados es variable, debido a que conforme se van observando medidas de potencia y valores de viento en las coordenadas, el número de pares de datos históricos almacenados en memoria va aumentando, con lo que el valor de M aumentaría continuamente con el tiempo.

4.5.4 Datos a una distancia inferior a un porcentaje fijo del rango de distancias

Dependiendo del horizonte de predicción y del nivel de potencia esperado, que vendrá dado por el nivel de viento en las coordenadas, la dispersión de las distancias será distinta, de forma que el rango de las distancias puede tener cierta variabilidad.

Este método propone seleccionar los elementos que se encuentren por debajo de una distancia umbral, que variará dependiendo del rango de las distancias obtenidas en cada instante. De este modo, cuando la amplitud del rango de distancias sea elevada, será mayor la ventana que marca la distancia umbral de selección. En cambio, en los casos en que el rango de distancias calculadas sea pequeño, la distancia umbral será también menor. Esta distancia umbral se define como:

$$\delta = d_{min} + \Delta d \quad (4.24)$$

donde:

δ = Distancia umbral de selección o ancho de banda

d_{min} = Distancia del elemento histórico más cercano

Δd = Incremento de distancia respecto de la distancia mínima.

El incremento Δd no tendrá un valor fijo, sino que dependerá de la distribución de distancias calculadas, de la forma:

$$\Delta d = p \cdot (d_{max} - d_{min}) \quad (4.25)$$

siendo:

d_{max} = Distancia del elemento histórico más alejado

p = Parámetro de selección a optimizar, es el porcentaje del rango total de distancias que define la distancia a incrementar respecto de la distancia mínima.

Una vez calculadas las distancias con cada uno de los vectores históricos, el ancho de banda quedará determinado por la distancia mínima calculada y el incremento Δd variable, definido por un porcentaje p fijo del rango de distancias. De este modo se seleccionará un número variable M de elementos a una distancia inferior a una distancia umbral δ también variable.

Si se representa la dispersión de distancias calculadas entre el nuevo vector de viento y los vectores históricos, los datos seleccionados serán aquellos que se encuentren a una distancia inferior a la distancia umbral, calculada a partir de la distancia mínima y el incremento Δd (Figura 4.9).

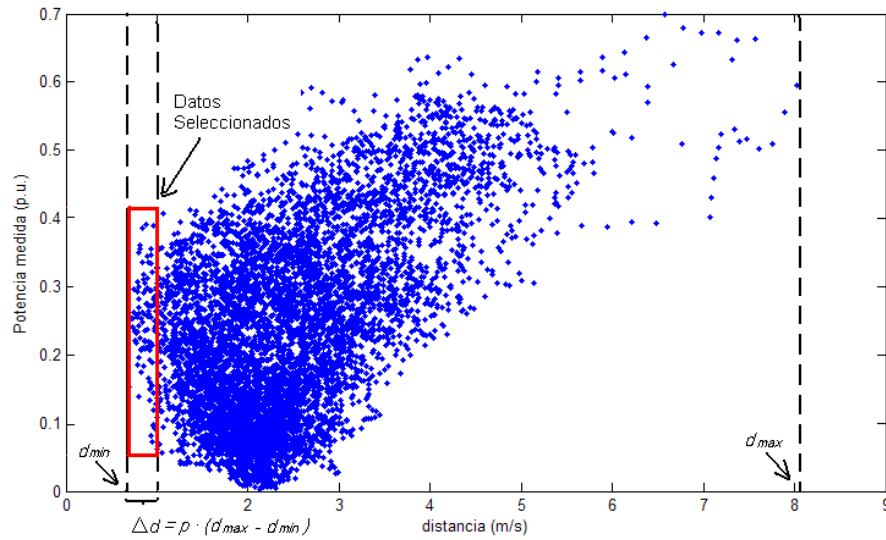


Figura 4.9: Datos a una distancia inferior a un porcentaje fijo del rango de distancias

Este método sería equivalente a dividir el rango de distancias observadas en un número fijo de clases C , de forma que cada clase represente una división equivalente al porcentaje a seleccionar, y seleccionar los datos incluidos en la primera clase del histograma. El número de clases a dividir el histograma será la inversa del porcentaje del rango de distancias a seleccionar: $C = 1/p \times 100$. Así, en el ejemplo de la Figura 4.9, si se desean seleccionar los datos dentro de un rango de distancias inferior al $p = 5\%$, se habrán de seleccionar los datos que estén dentro de la primera clase del histograma dividido en $C=20$ clases.

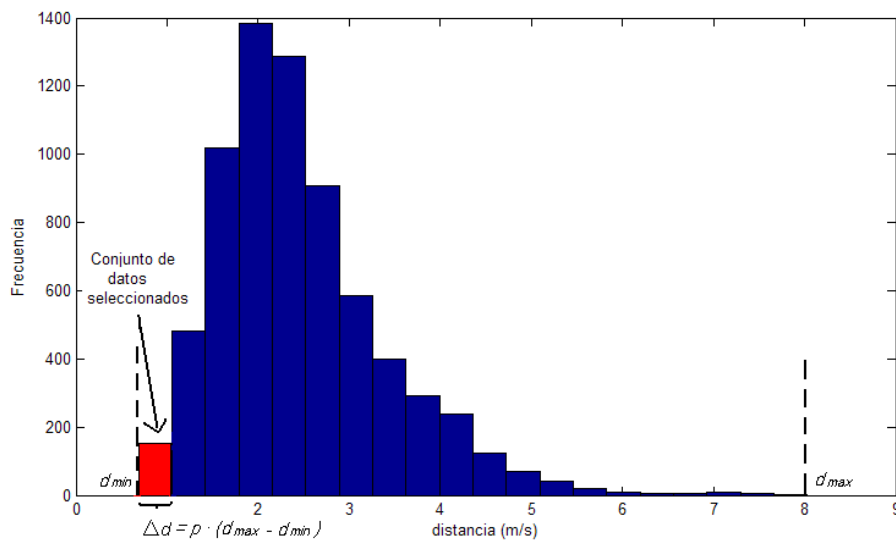


Figura 4.10: Datos contenidos en la primera clase del histograma

4.5.5 Datos a una distancia inferior a un porcentaje fijo del rango de distancias hasta la mediana

Alternativamente al método anterior, se puede seleccionar como rango de distancias el comprendido entre la distancia mínima y la distancia situada en la mediana (Figura 4.11). De este modo se evitaría el problema derivado de la posibilidad de que en determinadas ocasiones la mayor distancia calculada corresponda a algún vector aislado y a una distancia atípicamente elevada, provocada por una distribución de distancias con una asimetría positiva muy marcada. Este rango se define como:

$$\Delta d = p \cdot (d_{med} - d_{min}) \quad (4.26)$$

donde:

d_{med} = Distancia del elemento histórico situado en la mediana de las distancias calculadas

p = Parámetro de selección a optimizar

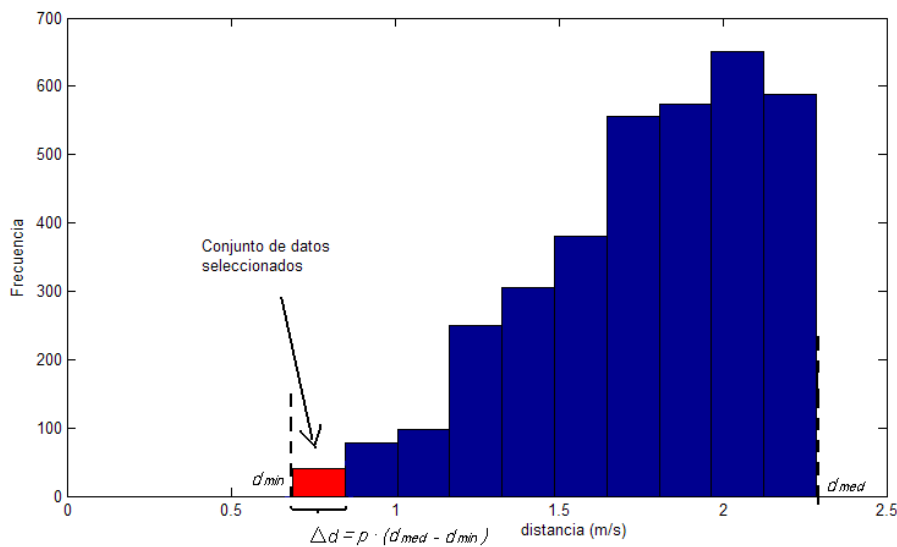


Figura 4.11: Datos a una distancia inferior a un porcentaje fijo del rango de distancias hasta la mediana

Con este método se estará teniendo en cuenta de forma implícita la forma de la distribución de las distancias calculadas y la concentración de datos a poca distancia. De este modo, cuando se encuentren muchos datos a poca distancia, la distancia del punto en la mediana será menor, y por tanto el ancho de banda también será menor. Cuando por el contrario haya poca concentración de datos cercanos, la distancia del punto en la mediana será mayor, con lo que el ancho de banda también será mayor, coherentemente con la baja concentración de datos cercanos.

4.5.6 Método basado en la tangente a los datos ordenados por distancia

Al igual que el método anterior descrito, este método busca un valor variable del ancho de banda o distancia umbral δ por debajo del cual se seleccionará un número variable de M datos para construir el modelo local, tratando de adaptarse a la forma de la distribución de las distancias calculadas con los datos históricos.

Si se ordenan las distancias calculadas con los vectores de viento históricos almacenados, y se representan los datos ordenados y las distancias calculadas, se obtendrá una curva monótona creciente como la de la Figura 4.12.

En este caso, el objetivo es seleccionar el valor de la distancia umbral δ a partir de la cual un pequeño incremento de dicho valor significaría un gran incremento en el número M de datos seleccionados. El método basado en la tangente que se propone tratará de encontrar el “codo” de la curva representada, a partir del cual, un pequeño incremento en la distancia, representada en el eje de abscisas, lleva asociado un gran incremento en el número de datos contenidos a la izquierda de esa distancia, representado en el eje de ordenadas.

La recta tangente que marcará la pendiente de referencia $s=1$ será aquella recta que une el origen de coordenadas con el valor situado en la mediana. El motivo de elegir la mediana como punto para definir la pendiente de referencia es su relativa estabilidad respecto a las asimetrías y a las distancias atípicamente elevadas, ya que como puede observarse en la figura pueden existir algunos datos aislados a distancias elevadas.

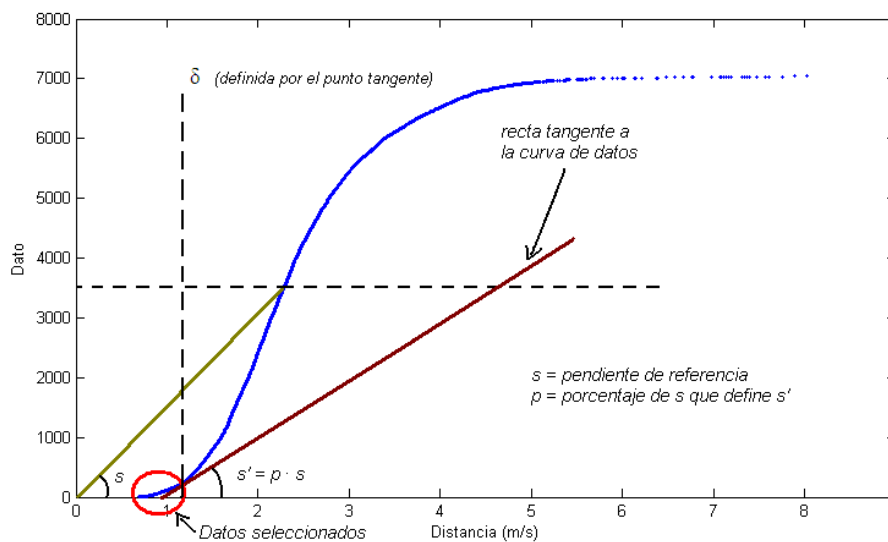


Figura 4.12: Método de selección basado en la tangente

Finalmente, la recta tangente del modelo que marcará la distancia umbral de selección será aquella con una pendiente s' , determinada por un porcentaje fijo p de la pendiente de referencia s . El parámetro p será también en este caso el parámetro a optimizar.

4.6 Funciones de ponderación por distancia

Para calcular el valor de predicción de potencia a partir de las potencias históricas medidas se deben tener en cuenta determinados coeficientes para dar más o menos peso a los elementos en función de su cercanía al punto a predecir. Hablaríamos entonces de una ponderación en distancias.

La utilidad de la función de ponderación en el modelo de predicción local es valorar la relevancia de los datos históricos observados para calcular la predicción, y asignarles unos pesos o coeficientes en función de su distancia.

Una función de ponderación debe asignar un peso de valor máximo a una distancia cero, con decaimiento continuo y suave conforme aumente la distancia. De igual modo, si las funciones de ponderación se hacen cero para los datos que no han sido seleccionados de acuerdo al ancho de banda utilizado, los datos que estén más alejados del ancho de banda de selección serán ignorados en el cálculo de la predicción.

Así, las funciones de ponderación por distancia serán distintas de cero solamente para los $m = 1, \dots, M$ datos que hayan sido seleccionados, y dependerán de la distancia entre los vectores de viento.

El coeficiente de ponderación por distancia entre el vector de viento $\mathbf{v}_{t+h|t}$ previsto en la hora t para la hora $t+h$, y el vector \mathbf{v}_m de los M seleccionados será de la forma:

$$\omega_m^d = \omega(d(\mathbf{v}_m, \mathbf{v}_{t+h|t})) \quad (4.27)$$

$$m = 1, \dots, M$$

En adelante, y por simplicidad, se omitirá en los coeficientes de ponderación descritos el subíndice $t+h|t$, que indica que el coeficiente es calculado en el instante t con vectores de predicción de viento para el horizonte h .

Una función de ponderación sencilla es la que simplemente eleva la distancia a un exponente negativo. La magnitud del exponente α determina cómo de local será la ponderación, de forma que cuanto mayor sea el exponente, más localizado será el modelado [116],[117].

Se define así la función de ponderación inversa a la distancia como:

$$\omega_m^d = \frac{1}{d_m^\alpha} \quad (4.28)$$

donde:

ω_m^d = Coeficiente de ponderación por distancia para el dato histórico m de los M seleccionados

d_m = distancia entre el vector histórico m y el nuevo vector de viento para el que se desea calcular la predicción de potencia

α = parámetro a optimizar

El posible inconveniente de esta función de ponderación es que tiende a infinito cuando algún dato observado se aproxima mucho al punto a predecir, de modo que fuerza a la predicción a tomar el valor de esa observación. El suavizado obtenido por los modelos de estimación local filtran de forma natural el ruido siempre y cuando la función de ponderación no tenga un valor infinito a distancia cero, en cuyo caso las predicciones para puntos muy cercanos a un dato se verán muy afectadas por ese dato.

Si los datos almacenados tienen ruido o hay datos erróneos, esta interpolación exacta no es deseable, por lo que una función de ponderación limitada (que no tienda a infinito) podría ser más deseable. Como función para limitar el valor de ponderación, se propone la función con decaimiento exponencial a la distancia [118]:

$$\omega_m^d = e^{-\alpha \cdot d_m} \quad (4.29)$$

El parámetro α en las dos funciones anteriores es ajustado con los datos que se usan para entrenar el modelo de predicción, seleccionando aquel valor que proporcione mejores resultados en términos de error de predicción.

En la Figura 4.13 y Figura 4.14 se representa el valor que obtendría el coeficiente ω^d para un rango de distancias entre 0 y 2 m/s, usando las dos funciones de ponderación por distancia propuestas en (4.28) y (4.29), con distintos valores del parámetro α . Con $\alpha=0$ se tendría el mismo peso para todos los elementos seleccionados, independientemente de la distancia con el nuevo vector de viento. A medida que aumenta α , mayor es la influencia de la distancia de los datos seleccionados en la ponderación.

Métodos de Predicción de la Generación Agregada de Energía Eólica

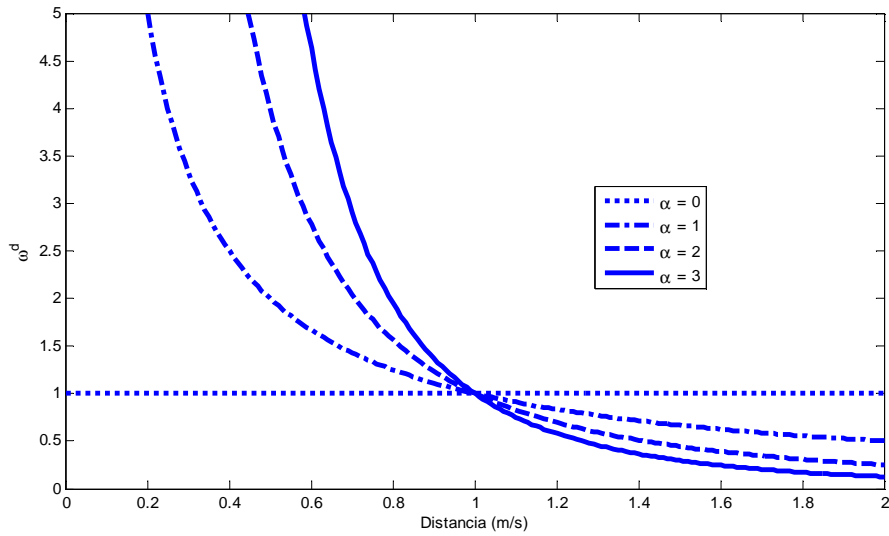


Figura 4.13: Coeficiente de ponderación inversa a la distancia

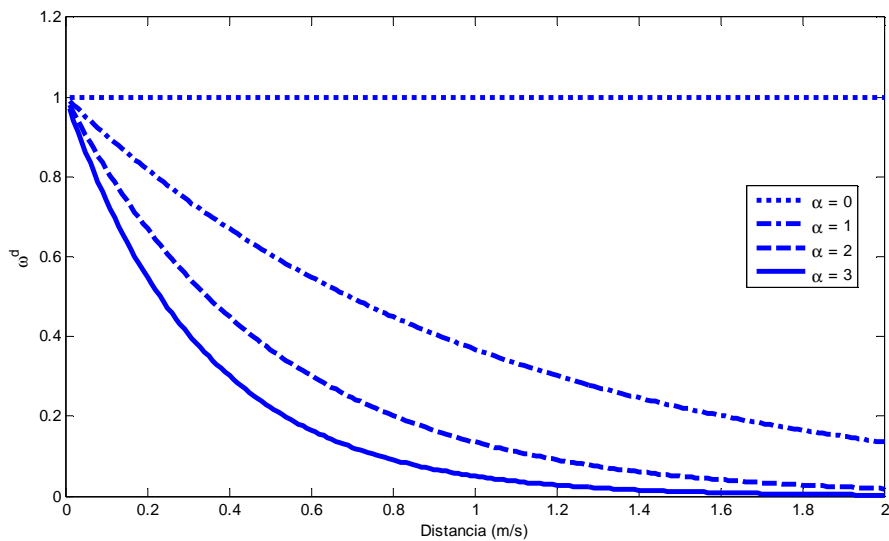


Figura 4.14: Coeficiente de ponderación por decaimiento exponencial con la distancia

En la Figura 4.13, donde se representa la función de ponderación inversamente proporcional a la distancia, se puede comprobar que los coeficientes pueden tender a infinito para distancias tendiendo a cero, lo que constituye una posible pérdida en el suavizado de los datos cuando exista alguna distancia igual a cero, en cuyo caso la predicción coincidirá con el valor de potencia de ese dato observado.

4.6.1 Corrección del parámetro α de ponderación en función de la distribución de distancias

En la sección dedicada a la selección de datos, se propusieron algunos métodos de selección que tienen en cuenta la forma de la distribución de las distancias. Trataremos de tener en cuenta esta distribución también en las funciones de ponderación por distancias, de modo que la variación de los coeficientes de ponderación sea más severa con la distancia en aquellos casos en los que se encuentre una gran concentración de datos cercanos, y más suave en aquellos casos en los que la concentración de datos cercanos sea menor. De este modo se conseguiría una mayor selección cuando la concentración de datos cercanos sea elevada, y una ponderación más suave cuando haya pocos datos cercanos.

Si se representan las distribuciones de distancias calculadas para un mismo tipo de distancia y un mismo horizonte, pero con vectores de entrada muy distintos, se obtendrán distribuciones como las de la Figura 4.15 y Figura 4.16, en las que se aprecian claras diferencias en la concentración de datos a poca distancia, con histogramas más concentrados en el lado izquierdo o en el lado derecho de la distribución.

Para tener en cuenta esta concentración, se estudiará el efecto de modificar las funciones de ponderación por distancias, afectando los parámetros de ponderación α con alguna medida que represente la concentración de datos cercanos. Esto se puede conseguir dividiendo, en cada instante, el parámetro de ponderación α entre un parámetro μ que sea representativo del centro de la distribución de distancias.

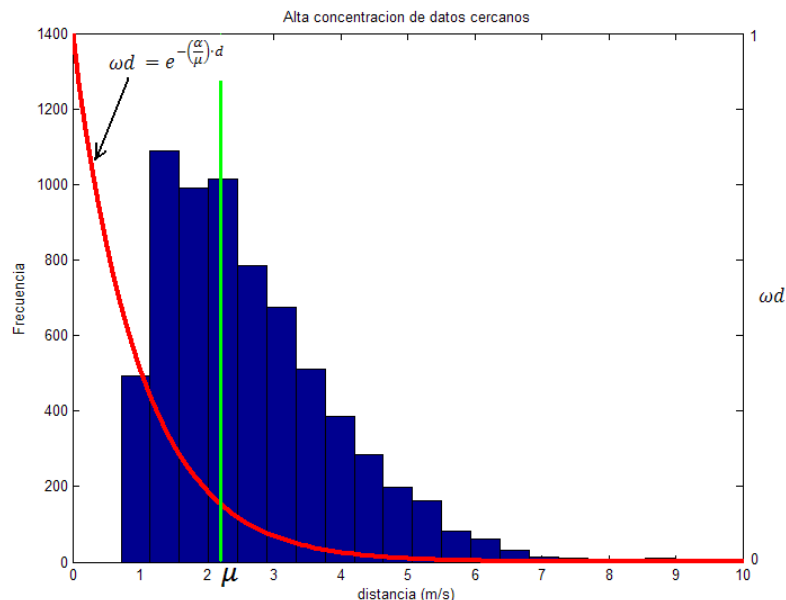


Figura 4.15: Histograma con alta concentración de datos cercanos

Métodos de Predicción de la Generación Agregada de Energía Eólica

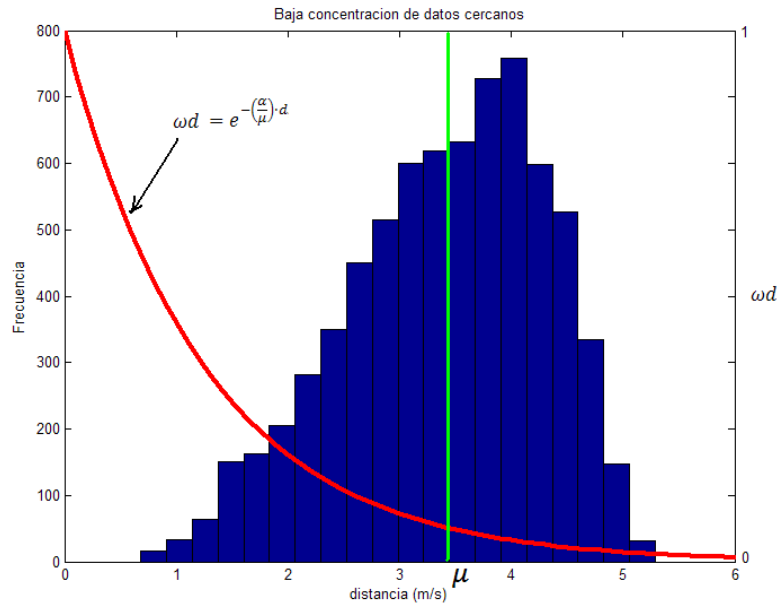


Figura 4.16: Histograma con baja concentración de datos cercanos

Los valores de la media o la mediana de las distribuciones de distancia dependen de la forma de la distribución, y son buenos candidatos para la misión de corregir el parámetro de ponderación, tomando valores relativamente bajos cuando la concentración de datos cercanos sea muy elevada, y valores más elevados cuando haya poca concentración de datos cercanos.

De este modo, si en las funciones de ponderación exponencial e inversa propuestas en las ecuaciones (4.28) y (4.29) se sustituye el parámetro α por otro parámetro α^* , de modo que sea $\alpha^* = \alpha/\mu$, se tendrán los nuevos coeficientes de ponderación como:

$$\omega_m^d = \frac{1}{d_m^{\alpha^*}} = \frac{1}{d_m^{(\alpha/\mu)}} \quad (4.30)$$

para la función de ponderación inversa a la distancia, y para la función de ponderación de decaimiento exponencial será:

$$\omega_m^d = e^{-\alpha^* \cdot d_m} = e^{-\left(\frac{\alpha}{\mu}\right) \cdot d_m} \quad (4.31)$$

donde μ puede ser la media o la mediana de las distancias.

De esta forma, en cada instante se estará teniendo en cuenta la forma de la distribución de distancias, determinada por su centro μ , para calcular los coeficientes de ponderación.

En el caso de que el valor de μ sea pequeño, que corresponde a instantes como el de la Figura 4.15, en los que la distribución presenta una gran concentración de elementos a poca distancia, el cociente α/μ será elevado, de modo que el decaimiento de la función de ponderación será más severo, coherentemente con la gran concentración de elementos cercanos a ponderar.

Mientras, en el caso de que el valor de μ sea elevado, que corresponde a instantes en los que la distribución presenta pocos elementos a una distancia cercana (Figura 4.16), el cociente α/μ será pequeño, de modo que el decaimiento de la función de ponderación será menos acentuado y se tendrán por tanto en cuenta suficientes datos con una ponderación más suavizada.

4.7 Funciones de ponderación por antigüedad

Además de las funciones de ponderación por distancias, también se pueden tener en cuenta ciertos coeficientes de ponderación que contemplen la antigüedad de los datos históricos, para dar más peso a los elementos más recientes del conjunto de datos históricos almacenados y restar importancia a los que ocurrieron hace mucho tiempo. El coeficiente de ponderación final a utilizar en (3.3) y (3.11), correspondiente a los modelos de Media Ponderada y Regresión Local ponderada respectivamente, sería entonces el producto de los coeficientes de ponderación por distancias y por antigüedad:

$$\omega_m = \omega_m^d \cdot \omega_m^\tau \quad (4.32)$$

donde:

ω_m^d = Coeficiente de ponderación por distancia entre el vector de viento previsto y el vector m de los M seleccionados

ω_m^τ = Coeficiente de ponderación por antigüedad del dato m de los M seleccionados

La ponderación por antigüedad de los datos confiere al algoritmo de predicción cierta adaptabilidad ante los cambios. Estos cambios pueden ser debidos a variaciones estacionales en la relación entre potencia y viento en los parques eólicos, en los que por ejemplo se registrarán distintas potencias para los mismos valores de velocidad del viento en las coordenadas dependiendo de la temperatura media del aire y por tanto de la densidad del viento incidente. Además, pueden existir otros motivos por los que la relación entre las variables de entrada y salida del sistema pueden variar a lo largo del tiempo, como podría ser la instalación de nuevos parques eólicos en la región o la sustitución de aerogeneradores en algún parque.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Esta ponderación tendrá especial importancia en horizontes de predicción bajos, ya que la potencia esperada para la región a corto plazo estará muy influenciada por los últimos valores de potencia registrados en la región, debido a la persistencia del viento.

Por tanto, de entre los elementos seleccionados a ponderar, se dará más peso a aquellos que se observaron en un instante más reciente o cercano en el tiempo al instante actual. Para este fin, los coeficientes de ponderación que suelen utilizarse son:

$$\omega_m^\tau = \lambda^{\tau_m} \quad (4.33)$$

y

$$\omega_m^\tau = e^{(\lambda-1)\cdot\tau_m} \quad (4.34)$$

donde:

ω_m^τ = coeficiente de ponderación por antigüedad del dato m de los M seleccionados

τ_m = antigüedad en horas, u otro intervalo de tiempo, del elemento m seleccionado

λ = factor de olvido, es el parámetro a ajustar

El valor del parámetro λ , también conocido como factor de olvido, será un número entre 0 y 1. Se suelen usar valores elevados, del orden de 0.995 ó 0.9999. En cualquier caso, será un parámetro a optimizar con el objeto de minimizar los errores de predicción, usando para ello un periodo de entrenamiento.

En la Figura 4.17 se muestra el valor de los coeficientes ω_m^τ en función de la antigüedad τ_m de los datos, para diferentes valores del parámetro λ , utilizando las dos funciones de ponderación por antigüedad propuestas.

La forma de las curvas representadas es exactamente igual para las dos funciones de ponderación. Esto es debido a que para valores elevados de λ , cercanos a 1, y valores no demasiado elevados de antigüedad τ_m , las ecuaciones (4.33) y (4.34) proporcionan resultados similares.

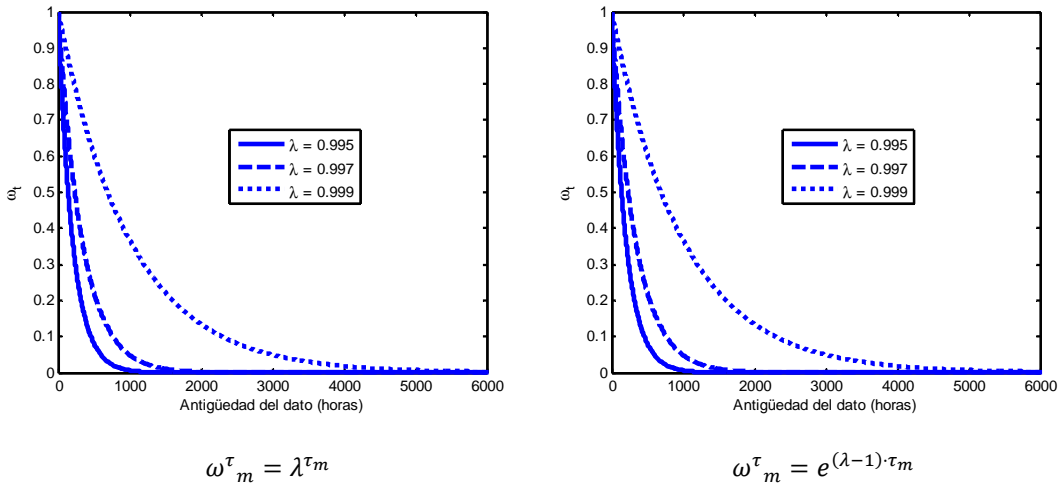


Figura 4.17: Coeficientes de ponderación por antigüedad

En la Figura 4.18 se muestra la relación $\frac{\omega^{\tau}_{m,1}}{\omega^{\tau}_{m,2}} = \frac{\lambda^{\tau_m}}{e^{(\lambda-1)\cdot\tau_m}}$ entre ambos coeficientes de ponderación, para algunos valores típicos del factor de olvido y una antigüedad τ_m hasta 10.000 horas, lo que equivaldría a más de 1 año de datos históricos.

Se puede apreciar que para que se perciban diferencias entre los dos coeficientes de ponderación tenemos que irnos a valores muy elevados de antigüedad τ_m , con factores de olvido λ más bien reducidos. Sin embargo, a medida que el valor de λ es menor, el peso de los datos más antiguos apenas será relevante respecto del peso de los datos más recientes. Por tanto, se puede concluir que los dos métodos de cálculo del coeficiente de ponderación por antigüedad o distancia temporal son equivalentes.

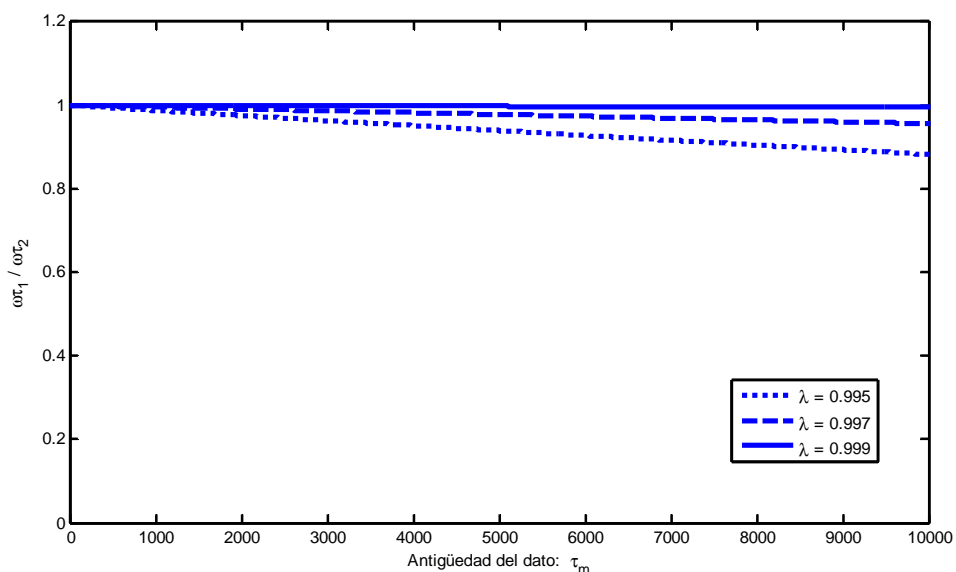


Figura 4.18: Relación de coeficientes de ponderación por antigüedad

Métodos de Predicción de la Generación Agregada de Energía Eólica

Esta equivalencia también se puede deducir si se desarrollan ambas funciones de ponderación. Si se sustituye $(\lambda - 1)$ por una variable x en ambos métodos de cálculo de los coeficientes de ponderación, y se desarrollan en serie para valores de x cercanos a cero, lo que equivale a valores de λ próximos a 1, se tiene que:

Si $|x| < 1$, condición que se cumple para los valores en que está definido λ , se puede desarrollar la ecuación (4.33) como la serie binomial:

$$\omega^{\tau_{m,1}} = \lambda^{\tau_m} = (1 + x)^{\tau_m} = \sum_{j=0}^{\infty} \binom{\tau_m}{j} \cdot x^j = \sum_{j=0}^{\infty} \left(\left(\prod_{k=1}^j \frac{\tau_m - k + 1}{k} \right) \cdot x^j \right) \quad (4.35)$$

que desarrollando resulta en:

$$\omega^{\tau_{m,1}} = 1 + \tau_m \cdot x + \frac{\tau_m(\tau_m - 1)}{2} x^2 + \frac{\tau_m(\tau_m - 1)(\tau_m - 2)}{3!} x^3 + \dots \quad (4.36)$$

Mientras, la serie de la función exponencial (4.34) resulta:

$$\omega^{\tau_{m,2}} = e^{(\lambda-1)\tau_m} = e^{x \cdot \tau_m} = 1 + \tau_m \cdot x + \frac{\tau_m^2}{2} x^2 + \frac{\tau_m^3}{3!} x^3 + \frac{\tau_m^4}{4!} x^4 + \dots \quad (4.37)$$

Ambas series coinciden en los dos primeros términos y difieren en el resto de términos, en los que la potencia x^j tiende a anularlos debido a que el valor de x es muy próximo a cero al ser λ muy próximo a 1.

Por tanto, debido a la similitud en los resultados, se selecciona uno de los métodos de cálculo de los coeficientes de ponderación temporal propuestos. La función de ponderación temporal elegida en esta tesis para su análisis posterior es la correspondiente a la ecuación (4.33).

Por último, si se sumasen los coeficientes de ponderación de todos los elementos históricos disponibles, desde el instante actual hasta el instante más antiguo registrado, se tendría la serie:

$$\sum_{\tau_m=0}^n \omega^{\tau_m} = \sum_{\tau_m=0}^n \lambda^{\tau_m} = \lambda^0 + \lambda^1 + \lambda^2 + \dots + \lambda^n \quad (4.38)$$

Esta serie, para $|\lambda| < 1$, tiende a un valor que se conoce como “memoria asintótica”, y está definida por el factor de olvido λ del siguiente modo:

$$\gamma = \frac{1}{1 - \lambda} \quad (4.39)$$

La memoria asintótica γ , da una idea del número de datos históricos equivalentes que corresponden a un factor de olvido dado.

5. Evaluación de los Métodos de Suavizado Propuestos

Resumen

En este capítulo se hace una evaluación de las distintas aproximaciones propuestas en el capítulo anterior para el suavizado de los datos en su uso con los modelos locales de Media Ponderada y Regresión Local Ponderada.

En primer lugar se analizarán los resultados obtenidos utilizando como funciones de suavizado las distintas alternativas de funciones Kernel.

Posteriormente se hará una evaluación exhaustiva de las diferentes medidas de distancia propuestas, así como las alternativas en los métodos de selección y funciones de ponderación por distancia y antigüedad. El objetivo será identificar la medida de distancia, el método de selección y las funciones de ponderación que más se adecúen a nuestro problema.

Finalmente, los resultados obtenidos con las funciones Kernel clásicas se compararán con los que se obtienen con el método de suavizado propuesto, que trata de forma independiente los problemas de medida de distancias, selección y ponderación de datos cercanos.

5.1 Datos utilizados

Para evaluar los modelos propuestos en esta tesis se utilizarán datos de producción eólica total en el sistema eléctrico peninsular español. Los datos de potencia agregada consisten en la integral horaria de la energía eólica registrada durante el año 2007 (8760 horas), facilitados por Red Eléctrica de España.

Los datos de viento utilizados consisten en la velocidad del viento previsto por una agencia meteorológica durante el año 2007 en 83 coordenadas seleccionadas repartidas por el territorio peninsular, cercanas a los emplazamientos de los parques eólicos.

Las predicciones de viento en las coordenadas se actualizan cada 12 horas, tienen una resolución de 6 horas y un horizonte de predicción superior a 48 horas, horizonte máximo para el que haremos la evaluación en esta tesis. Para obtener la predicción de viento a escala horaria se ha realizado una interpolación lineal de las predicciones originales de 6 horas de resolución. Dado que no se tienen medidas de viento en las coordenadas, en los datos del periodo histórico se considerarán como medidas de viento en las coordenadas las predicciones de menor horizonte que se tengan para cada hora.

Como datos históricos iniciales se utilizan los seis primeros meses del año 2007, entre la primera hora del mes de enero y la última hora del mes de junio. La evaluación se realiza para el segundo semestre del año. Para ello se calcula el Error Medio Absoluto Normalizado (NMAE) y la Raíz del Error Cuadrático Medio Normalizado (NRMSE), en función del horizonte de predicción, durante el periodo comprendido entre la primera hora del mes de julio y la última hora del mes de diciembre.

Se considera además que, cada vez que el modelo calcula las predicciones de potencia en una hora determinada, la medida de potencia y las predicciones de viento en las coordenadas para esa hora pasan a formar parte del registro histórico almacenado para ser utilizadas en la hora siguiente como datos de entrenamiento. Así, en la primera hora del mes de julio se tendrá un histórico de 6 meses, en la hora segunda se tendrá un registro histórico de 6 meses y 1 hora, y así sucesivamente hasta que en la última hora del año se tenga un registro histórico previo almacenado de 8759 horas.

Todos los algoritmos de predicción descritos se han programado y ejecutado utilizando el programa MATLAB.

5.2 Evaluación de las funciones Kernel

Aplicando un Kernel multivariante (4.9), en el que la ponderación para cada dato se calcula como la multiplicación de las ponderaciones resultantes para cada una de las coordenadas de viento o dimensiones del problema, se calculará la predicción de

potencia agregada en el territorio peninsular español para la segunda mitad del año 2007, utilizando los métodos de Media Ponderada y de Regresión Local Ponderada.

El objetivo de esta evaluación será identificar qué función Kernel es la más adecuada para el problema que nos ocupa, y conocer la magnitud del ancho de banda apropiado para minimizar el error de predicción. Por simplicidad sólo se representan las curvas de error para algunos horizontes de predicción.

Por otro lado, se desea comparar la bondad de las predicciones utilizando los métodos de Media Ponderada y de Regresión Local Ponderada, de modo que se compruebe cuál de las dos configuraciones proporciona menores errores.

5.2.1 Media Ponderada con funciones Kernel

Kernel uniforme

En la Figura 5.1 se representan los errores de predicción NMAE y NRMSE utilizando una Media Ponderada con Kernel uniforme para los horizontes $h=8$ y 48 horas. Si bien pudiera parecer que el error de predicción disminuye con el ancho de banda incluso para valores de δ aún más bajos que los representados, se debe tener en cuenta que para anchos de banda inferiores a 6 m/s no todas las predicciones de potencia pudieron ser calculadas.

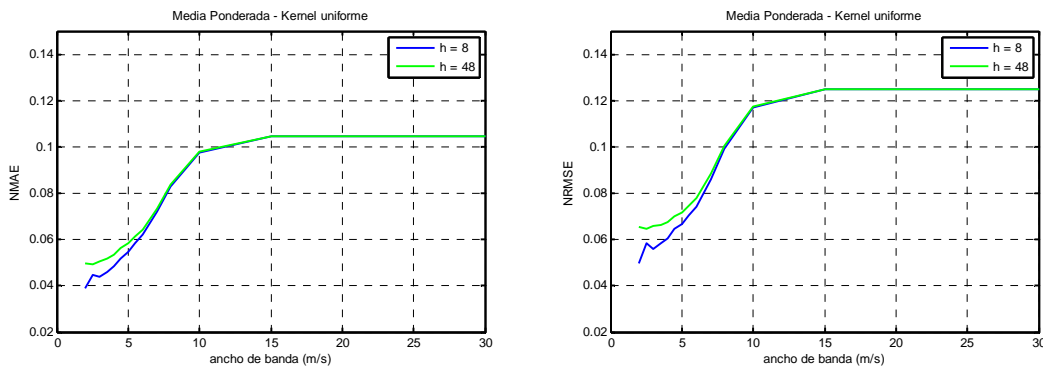


Figura 5.1: Error de predicción en función del ancho de banda. Media Ponderada con Kernel uniforme

Esto se debe a que, en algunas ocasiones, el empleo de un ancho de banda inferior implicó que ninguno de los datos históricos se encontrase dentro del ancho de banda para alguna de las coordenadas. Si tenemos en cuenta que el Kernel multivariante es el resultado de multiplicar los Kernel o ponderaciones de todas las dimensiones, nos encontraríamos con que todos los datos históricos tendrían al menos una coordenada con la velocidad fuera del ancho de banda, y por tanto la ponderación final sería nula para todos los datos, de modo que no podría calcularse predicción alguna.

Tabla 5.1: Porcentaje de predicciones no resueltas por falta de datos con Kernel uniforme

Ancho de Banda δ (m/s)	2	2.5	3	3.5	4	4.5	5	5.5	$\delta \geq 6$
Predicciones no resueltas (%)	78.5	33.7	13	3.5	1.5	0.5	0.2	0.02	0

En la Tabla 5.1 se muestra el porcentaje de predicciones que quedaron sin resolver en función del ancho de banda. Como ejemplo, según los datos de la tabla se tiene que utilizando un ancho de banda de 3 m/s, el 13% de las predicciones que intentaron calcularse entre la primera hora del mes de julio y la última hora del mes de diciembre de 2007 no obtuvieron predicción alguna porque no encontraron ningún dato que cumpliera el ancho de banda para todas las dimensiones o coordenadas del vector.

Por tanto, para anchos de banda inferiores a 6 m/s, aunque el valor de los errores medios sea inferior, se debe tener en cuenta que las medias de error fueron calculadas con tan solo un porcentaje de las predicciones, ya que el resto de predicciones quedaron sin ser calculadas al no encontrarse ningún dato dentro de la ventana de selección.

Esto puede ser aplicado al resto de funciones de ponderación Kernel, en las que cualquier ancho de banda inferior a 6 m/s implicaría que no se puedan calcular algunas de las predicciones, y por tanto no sería adecuado.

Por último, se observa en la figura que para valores de ancho de banda a partir de 15 m/s el error de predicción permanece constante. No hay diferencia en la predicción pese a utilizar un ancho de banda mayor, lo que nos indica que a partir de este ancho de banda todos los datos históricos son seleccionados en cada cálculo. Dado que el Kernel uniforme asigna el mismo peso a todos los datos seleccionados, se obtendrán las mismas predicciones independientemente de que aumente aún más el ancho de banda, pues todos los datos entran dentro de la ventana de selección, por grande que sea, con el mismo peso.

Kernel triangular

En la Figura 5.2 se muestran los errores en función del ancho de banda para una Media Ponderada utilizando un Kernel triangular, representando los horizontes de predicción $h = 8, 12, 18, 24, 36$ y 48 horas. En esta ocasión sí que se observa una zona que minimiza los errores para los horizontes de predicción representados, situándose para anchos de banda en el entorno de los 8 a 10 m/s. Para estos valores todas las predicciones pudieron ser calculadas, al ser el ancho de banda superior a los 6 m/s que se determinaron anteriormente con el Kernel uniforme.

Se puede apreciar que el ancho de banda óptimo, que minimiza los errores, aumenta ligeramente conforme aumenta el horizonte de predicción. Esto se debe a que la variabilidad de la predicción de viento aumenta conforme aumenta el horizonte de predicción, con lo que para lograr un mejor suavizado se hace necesario un mayor número de datos seleccionados, y por tanto un mayor ancho de banda.

Métodos de Predicción de la Generación Agregada de Energía Eólica

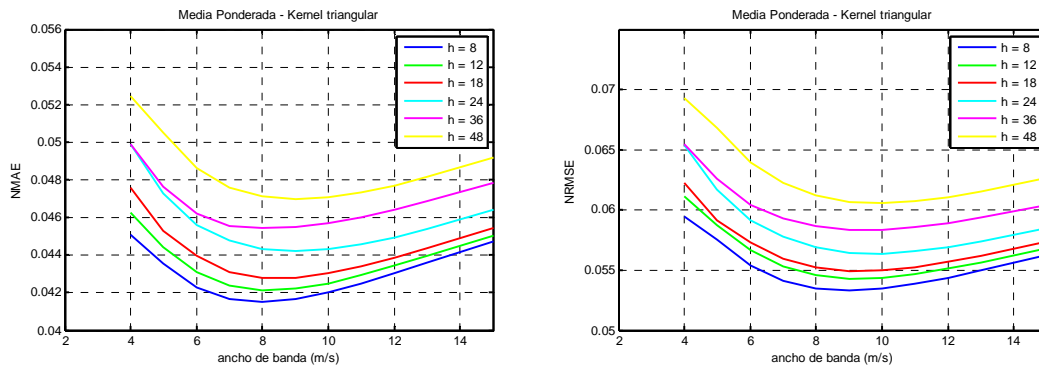


Figura 5.2: Error de predicción en función del ancho de banda. Media Ponderada con Kernel triangular

Además, como era esperable, la media de error aumenta a medida que aumenta el horizonte de predicción, pasando por ejemplo de valores NMAE en el entorno de un 4.2% para un horizonte de 8 ó 12 horas, hasta un valor cercano al 4.7% para un horizonte de 48 horas.

Kernel Epanechnikov

Utilizando un Kernel Epanechnikov, el ancho de banda óptimo resultante para nuestro problema está en el entorno de los 5 a 6 m/s, dependiendo del horizonte de predicción y de si lo que se minimiza es el error NMAE o NRMSE. Esto nos indica que, tal y como se vio en la Tabla 5.1 de predicciones no resueltas para el Kernel uniforme, hay aproximadamente un 0.2 % de predicciones que no pudieron ser calculadas para estos valores de ancho de banda.

En cualquier caso, y tal y como se comprobará más adelante, los valores mínimos de error obtenidos son ligeramente superiores a los obtenidos con un Kernel triangular para todos los horizontes de predicción.

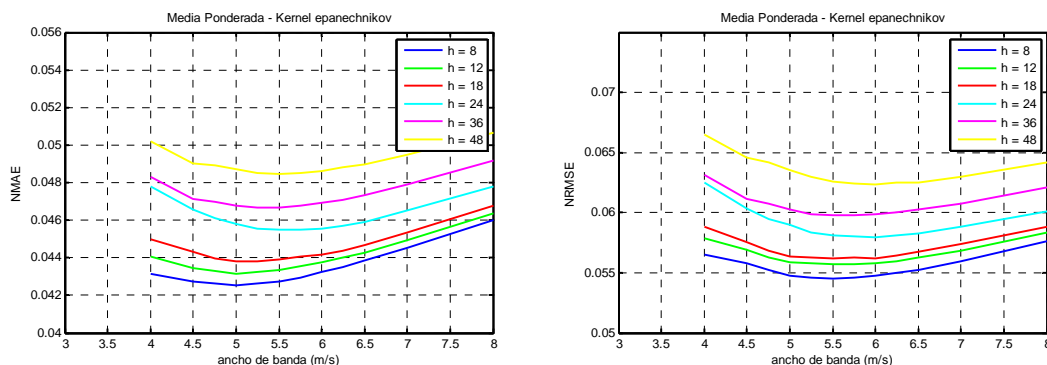


Figura 5.3: Error de predicción en función del ancho de banda. Media Ponderada con Kernel Epanechnikov

Kernel quartic

Utilizando un Kernel quartic como función de ponderación aplicada a una Media Ponderada se obtienen los errores representados en la Figura 5.4, en función de los horizontes de predicción y del ancho de banda elegido. En esta ocasión, el ancho de banda que minimiza los errores se sitúa en el entorno de los 7 a 8 m/s, no quedando por tanto ninguna predicción de potencia sin calcular por falta de datos seleccionados dentro de los anchos de banda.

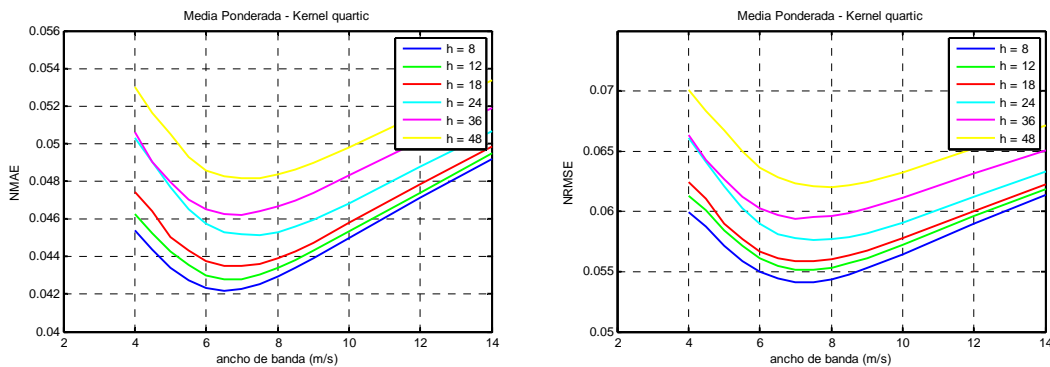


Figura 5.4: Error de predicción en función del ancho de banda. Media Ponderada con Kernel quartic

Kernel coseno

Al igual que ocurría con el Kernel Epanechnikov, los valores de ancho de banda que minimizan los errores cuando se utiliza una función Kernel coseno para la ponderación se encuentran por debajo de 6 m/s en algunos horizontes, con lo que algunas de las predicciones no habrán sido calculadas y por tanto este tipo de Kernel debería ser descartado para hacer predicciones con Media Ponderada. En todo caso, de nuevo se aprecia que para todos los horizontes los valores mínimos de error obtenido son superiores a los valores mínimos de error conseguidos con un Kernel triangular.

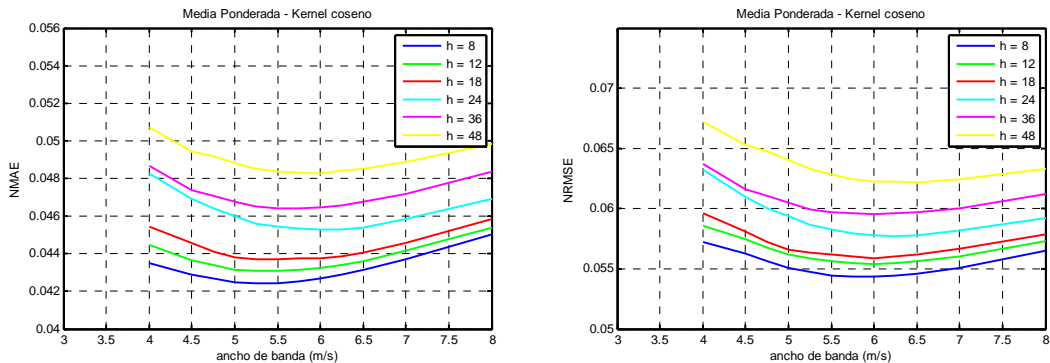


Figura 5.5: Error de predicción en función del ancho de banda. Media Ponderada con Kernel coseno

Kernel gaussiano

Al contrario que el resto de funciones Kernel expuestas, el Gaussiano es un Kernel no truncado, que asigna pesos cada vez más bajos a los datos alejados del ancho de banda, pero nunca se hacen nulos. Por tanto, cuando se utiliza un Kernel Gaussiano no se da el caso de que no se puedan calcular algunas predicciones por ausencia de datos dentro de la ventana de selección en alguna de las coordenadas.

En este caso, la predicción con menores errores se consigue con anchos de banda en el entorno de los 3 a 3.5 m/s, tal y como se aprecia en la Figura 5.6.

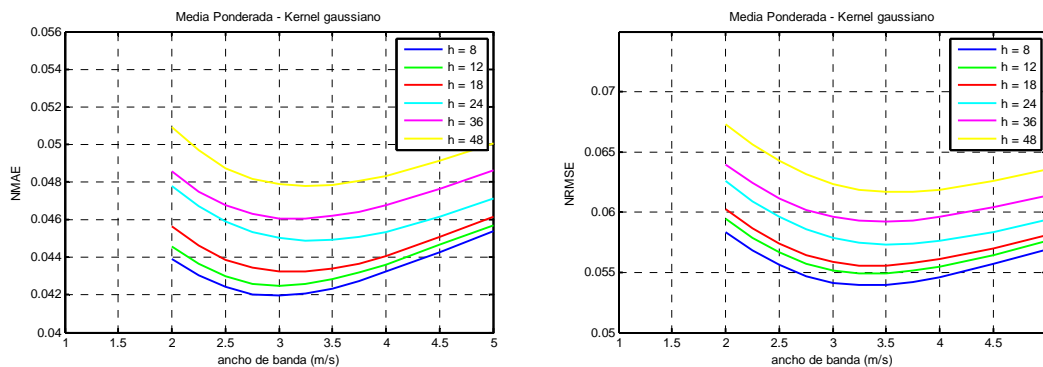


Figura 5.6: Error de predicción en función del ancho de banda. Media Ponderada con Kernel gaussiano

Comparación de funciones Kernel utilizando Media Ponderada

Como resumen comparativo de las funciones Kernel expuestas, se dibujan en la Figura 5.7 y Figura 5.8 los mínimos errores NMAE y NRMSE conseguidos para todas las funciones Kernel en función del horizonte de predicción, utilizando en cada caso los valores óptimos de ancho de banda y un modelo de Media Ponderada como método de cálculo de predicción, según la ecuación (3.3).

Tanto si se representa el error medio absoluto NMAE como la raíz del error cuadrático medio NRMSE, la función Kernel que proporciona mejores resultados en la ponderación para todos los horizontes de predicción analizados es el Kernel triangular. Tras el triangular, el Kernel gaussiano es el siguiente que presenta mejores resultados, y posteriormente el Kernel cuártic.

Los peores resultados se obtienen con las funciones Kernel Epanechnikov y coseno, que además son aquellas cuyos valores óptimos de ancho de banda estaban en el entorno de 5 a 6 m/s, lo que implica la desventaja añadida de no poder calcular alguna de las predicciones por falta de datos, como se comprobó con el Kernel uniforme.

Evaluación de los Métodos de Suavizado Propuestos

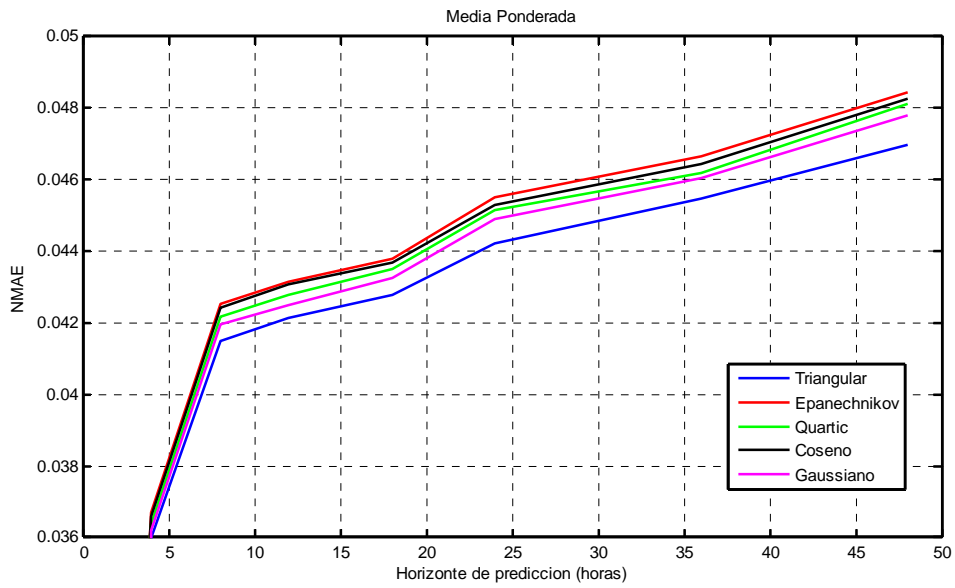


Figura 5.7: Comparación de errores NMAE de las funciones Kernel usando un modelo de Media Ponderada

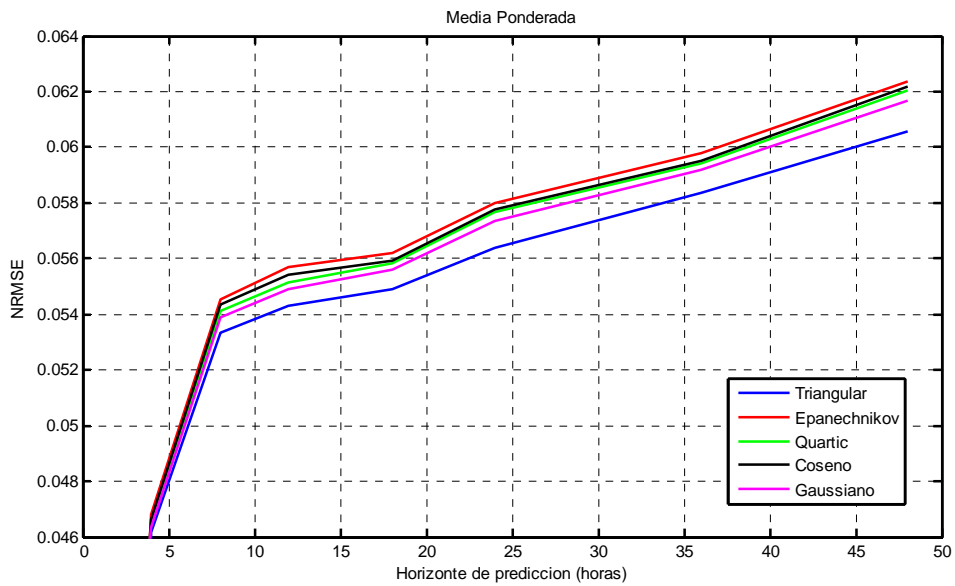


Figura 5.8: Comparación de errores NRMSE de las funciones Kernel usando un modelo de Media Ponderada

5.2.2 Regresión Local Ponderada con funciones Kernel

Se analizan a continuación los resultados obtenidos por las distintas funciones Kernel analizadas, utilizándolas para calcular los coeficientes de ponderación en un modelo de Regresión Local Ponderada, basado en el estimador (3.11).

Kernel uniforme

En la Figura 5.9 se representan el NMAE y NRMSE de predicción utilizando una Regresión Local Ponderada con Kernel uniforme para los horizontes $h=8$ y 48 horas.

En esta ocasión, nos encontramos con que el ancho de banda mínimo para poder calcular todas las predicciones es de $\delta = 8$ m/s. Con anchos de banda inferiores el algoritmo de cálculo es incapaz de calcular la inversa $(\mathbf{X}'\mathbf{\Omega}\mathbf{X})^{-1}$ en la ecuación de regresión (3.11) para alguno de los vectores de predicción de viento, ya que el número de datos seleccionados es insuficiente.

No debemos olvidar que, en la resolución del sistema de ecuaciones de la expresión (3.11), se pretenden estimar los $K+1$ parámetros de la regresión, siendo estos parámetros $\hat{\beta} = [\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_K]$, con lo que un valor mínimo razonable de elementos seleccionados para estimarlos sería $K+1$, de modo que al menos se tuvieran tantas ecuaciones como incógnitas. Por ello, si el ancho de banda no es suficientemente amplio, puede darse el caso de que no haya suficientes datos para estimar la regresión al no encontrarse datos suficientes que cumplan el criterio de ancho de banda para todas las coordenadas.

De nuevo, al igual que se observó cuando se utilizaba la Media Ponderada para calcular la predicción, en la Figura 5.9 se aprecia que los errores de predicción permanecen constantes con valores de ancho de banda a partir de 15 m/s, ya que a partir de este ancho de banda se seleccionan todos los datos históricos y se obtendrán las mismas predicciones, pues el Kernel uniforme asigna el mismo peso a todos los datos seleccionados.

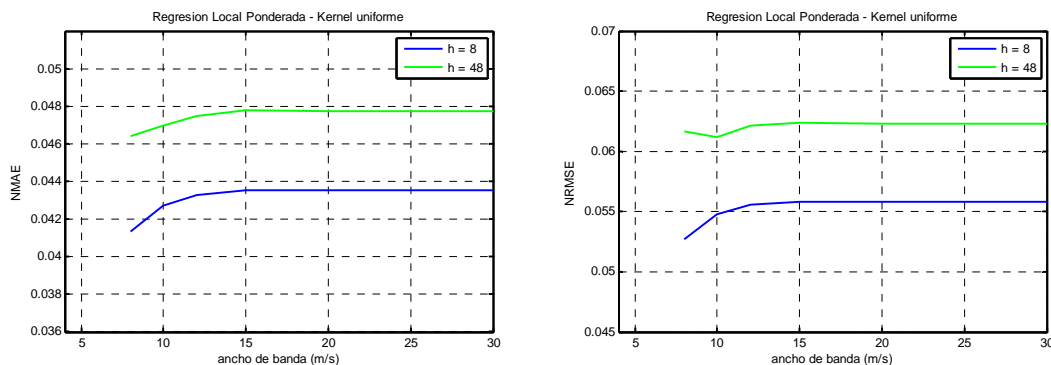


Figura 5.9: Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel uniforme

Kernel triangular

En la Figura 5.10 se muestran los errores en función del ancho de banda para una Regresión Local Ponderada utilizando un Kernel triangular como función de ponderación. El valor óptimo de ancho de banda se sitúa en el entorno de los 35 a 50 m/s, dependiendo del horizonte de predicción, siendo mayor el ancho de banda que minimiza el error conforme aumenta el horizonte de predicción, ya que la mayor variabilidad de las predicciones de viento a horizontes elevados hacen necesario un mayor suavizado, y por tanto un mayor ancho de banda, para calcular una buena predicción.

Otra apreciación importante se deriva de la gran amplitud del ancho de banda, de entre 35 y 50 m/s, que es claramente superior a los 15 m/s a partir de los cuales el Kernel uniforme proporcionaba predicciones similares, y que por tanto indica que se están seleccionando todos los datos históricos.

Utilizando un Kernel triangular, con anchos de banda que duplican el ancho de banda a partir del cual se seleccionan todos los datos, se puede concluir que para calcular las predicciones de potencia se están utilizando siempre todos los datos históricos. Sin embargo, aunque se utilizan todos los datos, el hecho de que el Kernel multivariante se calcule como producto de todas las ponderaciones Kernel univariantes según (4.9), hace que el peso relativo asignado a la mayoría de las observaciones históricas sea casi nulo, ya que se eleva un número menor que la unidad a una potencia K , que es el número de coordenadas o dimensiones en nuestro problema, 83 con los datos utilizados en nuestro caso. De este modo, en términos efectivos, es como si se estuviera haciendo una selección de datos.

En cuanto a los valores mínimos de error, se aprecia una disminución con respecto a los errores mínimos que se obtuvieron con el método de Media Ponderada, con un NMAE que pasa del 3.8 % para horizonte de 8 horas hasta un valor de 4.5 % para un horizonte de 48 horas.

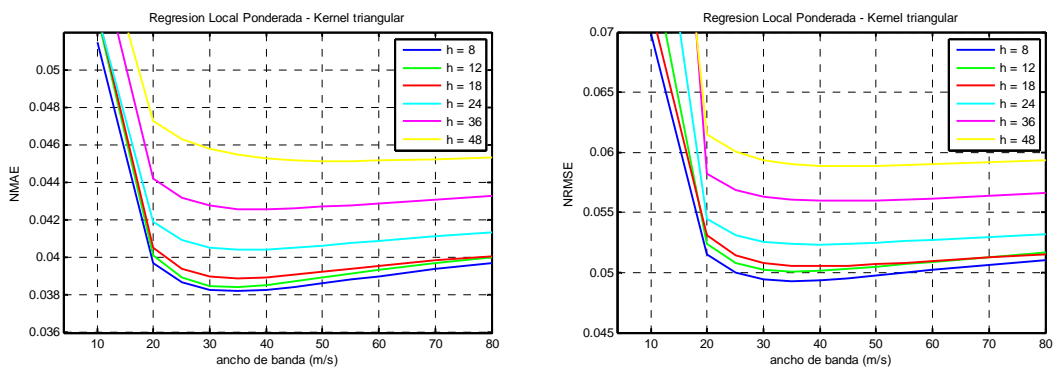


Figura 5.10: Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel triangular

Kernel Epanechnikov

Utilizando un Kernel Epanechnikov se obtienen los errores representados en la Figura 5.11, en función de los horizontes de predicción y del ancho de banda. En esta ocasión el ancho de banda que minimiza los errores se sitúa en el entorno de los 14 a 16 m/s.

Con estos valores de ancho de banda de nuevo tenemos garantizada la selección de la práctica totalidad de los datos históricos para el cálculo de la predicción, si bien el factor de ponderación final de cada dato, al ser fruto de la multiplicación de las ponderaciones en cada dimensión, puede llegar a ser muy próximo a cero para muchos datos.

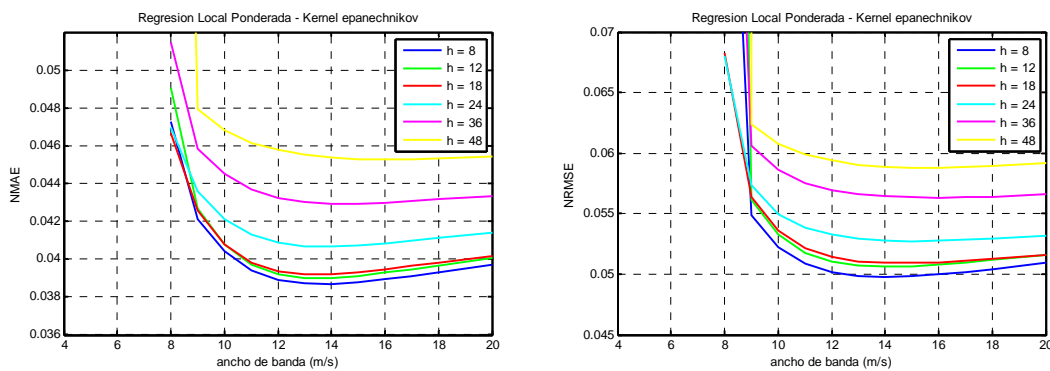


Figura 5.11: Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel Epanechnikov

Kernel quartic

Utilizando un Kernel quartic en la Regresión Local Ponderada, el ancho de banda óptimo resultante está en el entorno de los 18 a 23 m/s, dependiendo del horizonte de predicción y de si lo que se minimiza es el error NMAE o NRMSE, aumentando conforme aumenta el horizonte de predicción al igual que ocurre con el resto de funciones Kernel analizadas.

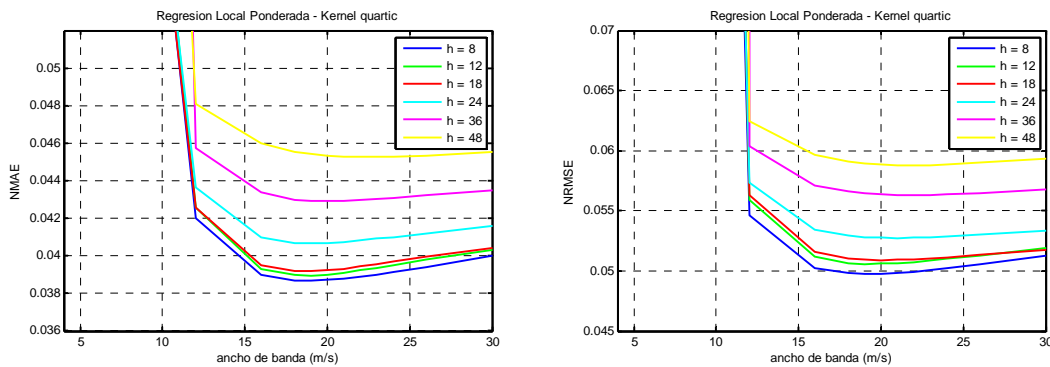


Figura 5.12: Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel quartic

Kernel coseno

Los valores de ancho de banda que minimizan los errores utilizando un Kernel Coseno se encuentran entre los 15 y 18 m/s, que de nuevo implican la selección de la práctica totalidad de observaciones históricas en el cálculo de la regresión (Figura 5.13). Los niveles de error observados son muy similares a los obtenidos con un Kernel Epanechnikov o quartic.

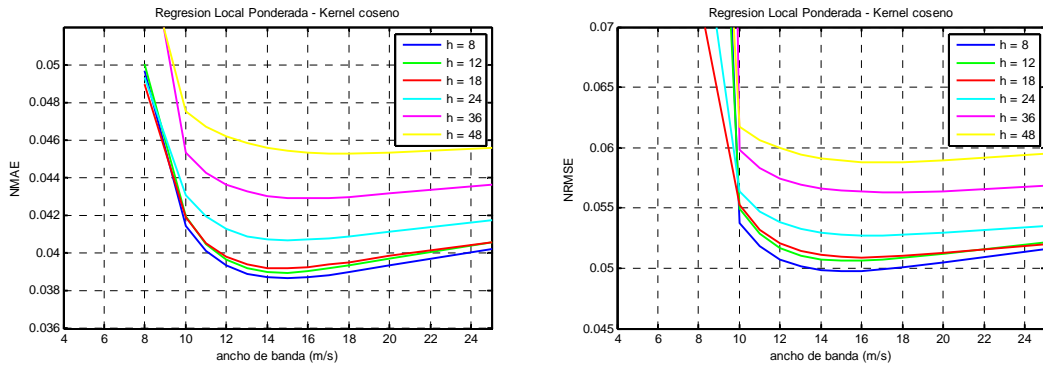


Figura 5.13: Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel coseno

Kernel gaussiano

En este caso, la predicción con menores errores se consigue con anchos de banda en el entorno de los 9 a 11 m/s (Figura 5.14). Pese a que estos valores sí son inferiores a 15 m/s, de nuevo nos encontramos con que el Kernel gaussiano también selecciona todos los datos, ya que es un Kernel no truncado, y siempre asignará cierto peso a todos los datos observados. De nuevo se observan niveles de error similares a los obtenidos con las funciones Kernel anteriores.

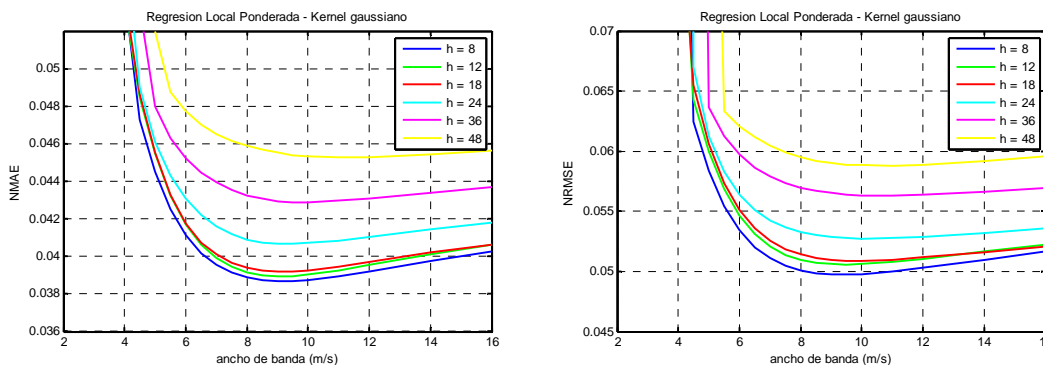


Figura 5.14: Error de predicción en función del ancho de banda. Regresión Local Ponderada con Kernel gaussiano

Comparación de funciones Kernel utilizando Regresión Local Ponderada

Como resumen comparativo de las funciones Kernel expuestas utilizando una Regresión Local Ponderada como método de cálculo de predicción, se dibujan en la Figura 5.15 y Figura 5.16 los errores NMAE y NRMSE en función del horizonte de predicción para horizontes $h = 4, 8, 12, 18, 24, 36$ y 48 horas para todas las funciones Kernel.

Al igual que ocurría con el método de cálculo de Media Ponderada, la función Kernel que proporciona mejores resultados en todos los horizontes de predicción analizados es el Kernel triangular, independientemente del tipo de error representado. El resto de funciones Kernel proporcionan errores similares, ligeramente superiores al obtenido con el Kernel triangular.

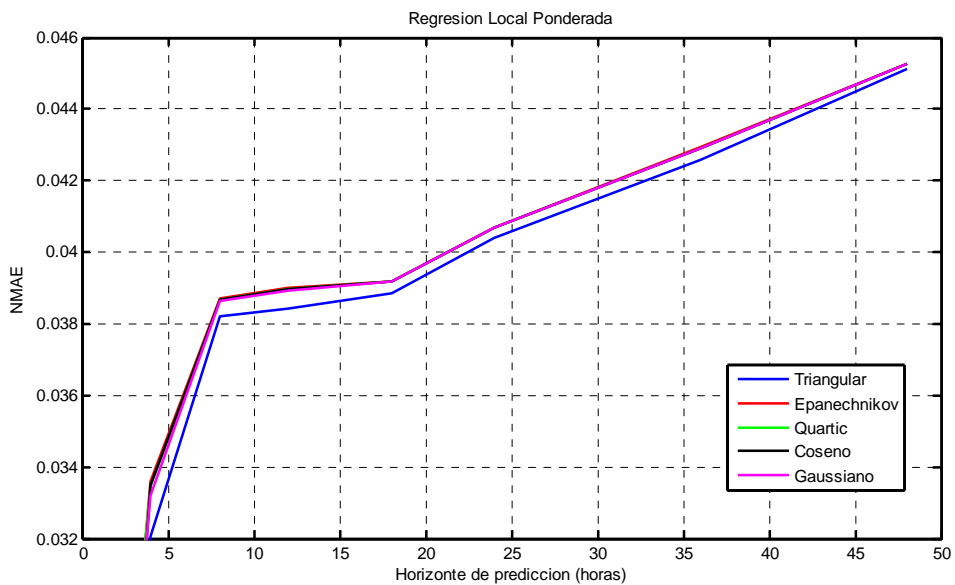


Figura 5.15: Comparación de errores NMAE de las funciones Kernel usando un modelo de Regresión Local Ponderada

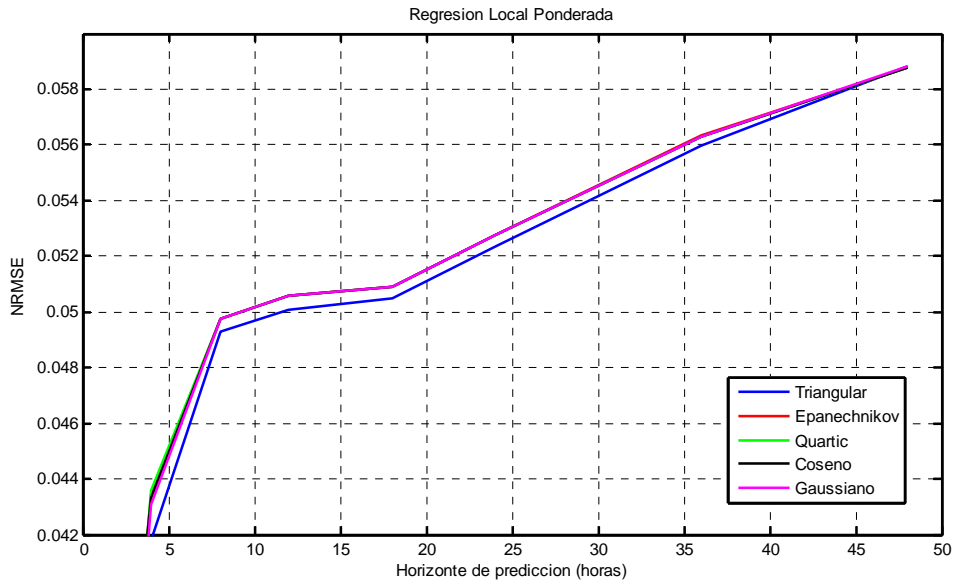


Figura 5.16: Comparación de errores NRMSE de las funciones Kernel usando un modelo de Regresión Local Ponderada

5.2.3 Comparación entre Media Ponderada y Regresión Local Ponderada usando funciones Kernel

Como análisis final, se representan en la Figura 5.17 y Figura 5.18 los errores NMAE y NRMSE obtenidos al calcular las predicciones con los métodos de Media Ponderada y Regresión Local Ponderada, utilizando un Kernel triangular como función de ponderación, pues es la función Kernel que ha presentado mejores resultados en ambos casos para todos los horizontes analizados. En este caso, se han calculado y representado los errores de predicción para todos los horizontes entre $h=1$ y $h=48$ horas.

Independientemente del error de predicción representado, en las figuras se observa que la Regresión Local Ponderada proporciona menor error para todos los horizontes de predicción.

Métodos de Predicción de la Generación Agregada de Energía Eólica

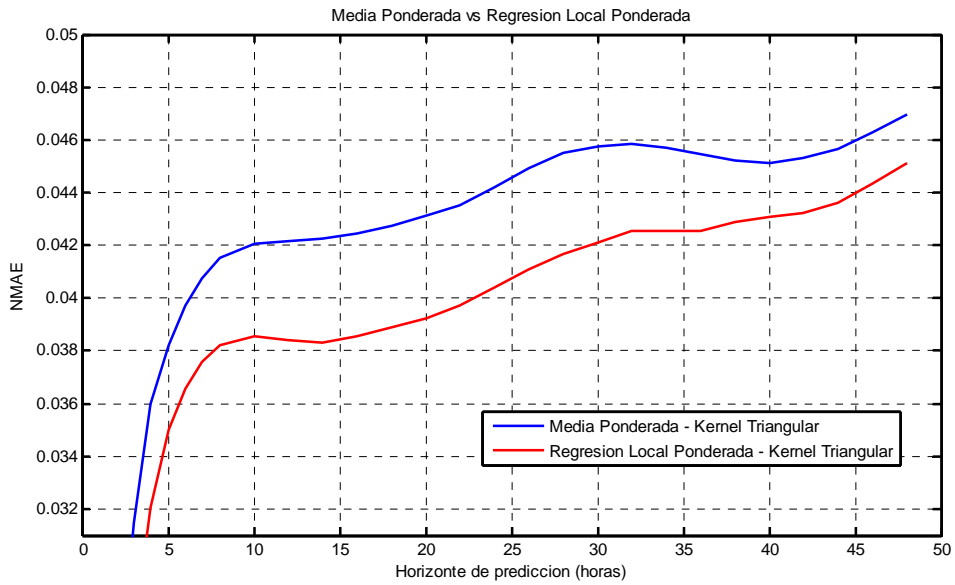


Figura 5.17: Comparación de errores NMAE de los modelos de predicción de Media Ponderada y Regresión Local Ponderada

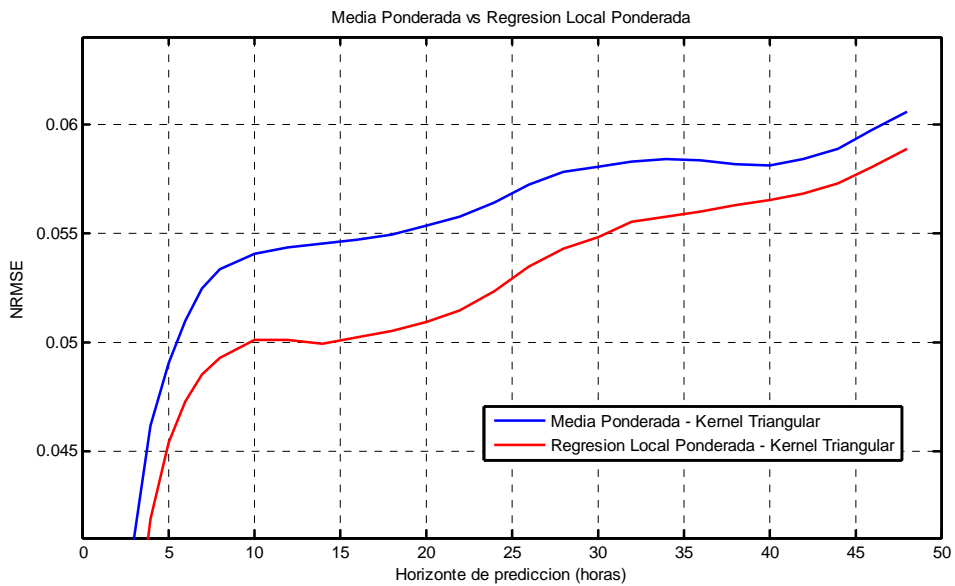


Figura 5.18: Comparación de errores NRMSE de los modelos de predicción de Media Ponderada y Regresión Local Ponderada

Esta disminución del error del modelo de Regresión Local Ponderada frente a la Media Ponderada también se puede apreciar si además se representa el sesgo de las predicciones, tanto si se seleccionan los anchos de banda que minimizan el NMAE (Figura 5.19) como si se seleccionan los que minimizan el NRMSE (Figura 5.20). El sesgo es menor cuando se utiliza una Regresión Local Ponderada que cuando se utiliza una Media Ponderada para estimar las predicciones, pasando de sesgos del orden de 0.002 p.u. a valores que rondan el 0.012 p.u.

Efectivamente, como se adelantó en el capítulo de estructuras de modelos locales, cuando la superficie a modelar tiene cierta curvatura, el método de Media Ponderada es un estimador más sesgado que el de Regresión Local Ponderada, lo que deriva en mayores errores de predicción.

Por tanto, cuando las funciones de ponderación utilizadas son las funciones Kernel, parece más adecuado utilizar una Regresión Local Ponderada que una Media Ponderada para calcular las predicciones en nuestro problema, pues presentan menor sesgo, y esto se traduce en un menor error de predicción.

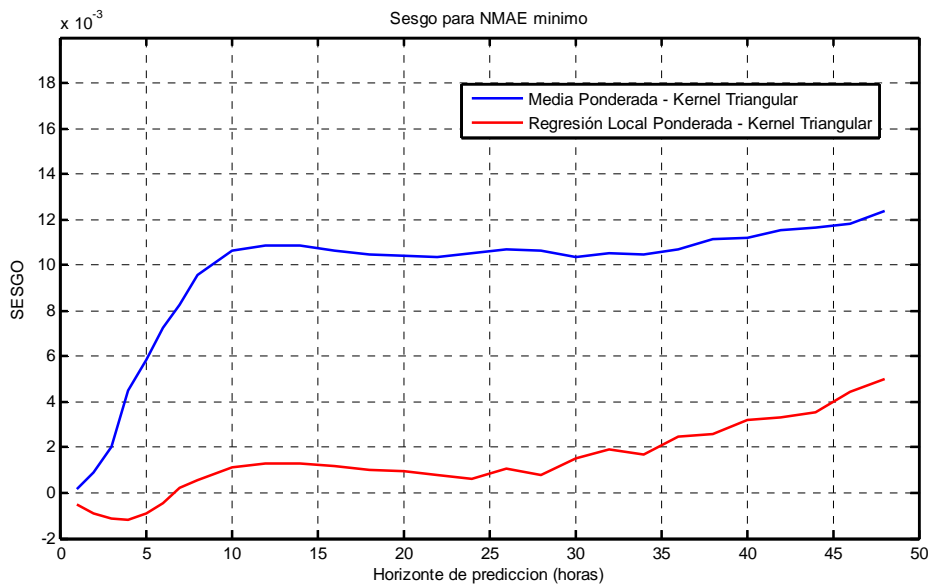


Figura 5.19: Comparación de sesgo de los modelos de predicción de Media Ponderada y Regresión Local Ponderada calculados para NMAE mínimo

Métodos de Predicción de la Generación Agregada de Energía Eólica

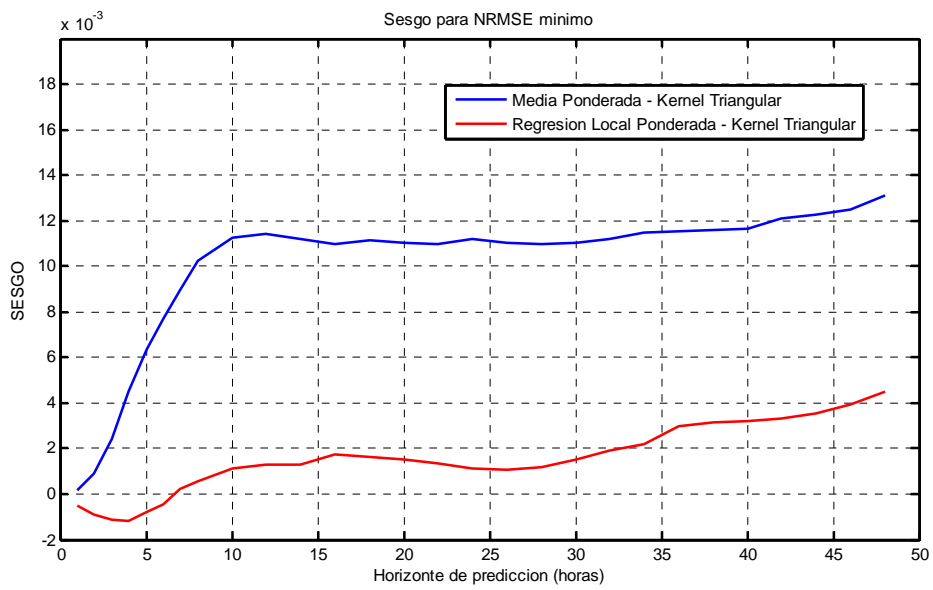


Figura 5.20: Comparación de sesgo de los modelos de predicción de Media Ponderada y Regresión Local Ponderada calculados para NRMSE mínimo

5.3 Análisis previo de las medidas de distancia propuestas

Para analizar las diferentes medidas de distancia propuestas se va a utilizar la serie de datos disponible para el año 2007, que cuenta con medidas de potencia eólica media horaria total peninsular (8760 datos), y las correspondientes velocidades de viento en 83 coordenadas seleccionadas en la región.

Previamente, se analiza la serie de potencias totales medidas durante el año 2007. En la Figura 5.21 se representa el histograma de la serie de datos, y en la Tabla 5.2 se detallan algunos parámetros que orientan el resto del análisis.

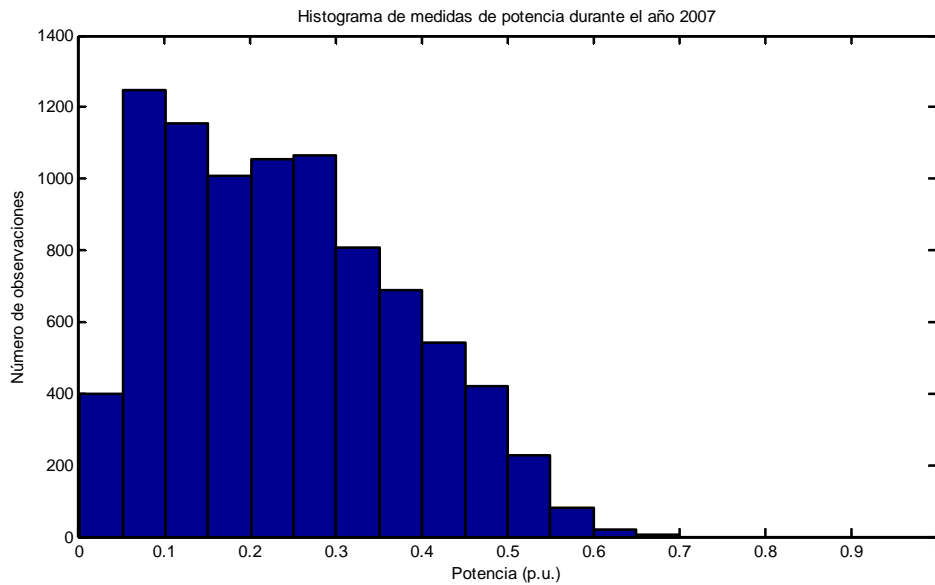


Figura 5.21: Histograma de medidas de potencia total peninsular registrada durante el año 2007

Tabla 5.2: Parámetros estadísticos básicos de la serie de medidas de potencia media horaria peninsular durante el año 2007

Media	Mediana	Mínimo	Máximo	Desviación típica	Varianza
0.244	0.2278	0.0028	0.6997	0.1364	0.0186

Métodos de Predicción de la Generación Agregada de Energía Eólica

A continuación, en la Tabla 5.3 se calculan los percentiles del 10, 25, 50 75 y 90%, que representan el porcentaje de datos registrados con un valor inferior al valor del percentil calculado.

Tabla 5.3: Percentiles de medidas de potencia media horaria peninsular durante el año 2007

Percentil (%)	10	25	50	75	90
Valor (p.u.)	0.0736	0.1248	0.2278	0.3377	0.4373

Del histograma representado y los parámetros calculados se desprende que a nivel del sistema total peninsular no es esperable una potencia eólica superior al 70% de la potencia instalada, pues no se observan potencias agregadas en el sistema superiores a 0.7 p.u., estando la mayor concentración de potencias observadas, el 80% según los percentiles calculados, en el entorno desde 0.074 p.u. hasta 0.437 p.u.

Por otro lado, el valor medio de producción, de 0.244 p.u., indica que el factor de utilización medio, y por tanto la potencia media esperable en términos globales es algo inferior al 25% de la potencia instalada en el sistema, correspondiente en este caso a 2137 horas de uso anual equivalente.

5.3.1 Dependencia del nivel de potencia prevista

La distribución de las distancias calculadas entre el nuevo vector de predicciones de viento v_{t+h} y los vectores históricos almacenados v_i dependerá del nivel de potencia esperada para la nueva predicción de viento.

Para realizar el análisis de cómo influye el nivel de potencia eólica total de la región en la forma de la distribución de las distancias, se representarán estas distribuciones a cinco niveles de potencia distintos, especificados por los cinco percentiles calculados previamente.

Se considerará un valor de potencia muy bajo el correspondiente al percentil 10%, es decir, un nivel de potencia que será superado el 90% del tiempo, que en nuestro caso es de aproximadamente 0.074 p.u. Como valor de potencia bajo se considerará el percentil 25%, al que le corresponde una potencia de 0.125 p.u. Como valor de potencia medio se seleccionará un instante en el que la potencia sea de aproximadamente 0.24 p.u., correspondiente a la potencia media anual y al percentil 50% aproximadamente. Y como valores de potencia elevado y muy elevado se seleccionarán horas con valores de potencia total medida de aproximadamente 0.34 y 0.44 p.u., correspondientes a los percentiles 75% y 90% respectivamente.

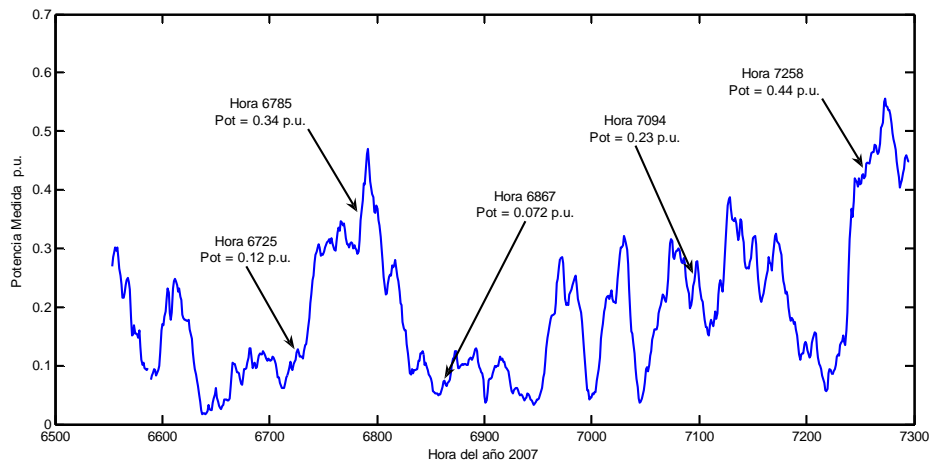


Figura 5.22: Selección de instantes a representar con varios niveles de potencia correspondientes a los percentiles 10, 25, 50, 75 y 90 %

Observando la serie de datos disponible, correspondiente al año 2007, se selecciona un periodo de datos cualquiera. En este caso se ha seleccionado el mes de octubre, al que corresponden desde la hora 6553 hasta la hora 7296 del total anual (Figura 5.22). Se selecciona el mes de octubre porque pertenece a una estación del año en la que se podrán encontrar fácilmente niveles de potencia bajos, medios y altos, debido a la variabilidad climatológica de la estación otoñal. Además, al ser octubre un mes cercano al final del año, se tendrá un mayor conjunto de datos históricos desde la primera hora del mes de enero.

A continuación, tal y como se muestra en la figura, se seleccionan algunos instantes en los que se haya registrado una medida de potencia muy baja, baja, media, elevada y muy elevada, de acuerdo a los valores de los percentiles calculados, para representar la distribución y dispersión de los valores de distancia calculados utilizando las distintas medidas de distancia propuestas.

Las horas seleccionadas del mes de octubre de 2007, y los percentiles y potencias eólicas totales medidas correspondientes se detallan en la Tabla 5.4.

Tabla 5.4: Datos seleccionados a varios niveles de potencia correspondientes a los percentiles 10, 25, 50, 75 y 90 %.

Percentil aproximado (%)	Hora del año	Fecha (aaaammddhh)	Potencia (p.u.)
10	6867	2007101402	0.0725
25	6725	2007100804	0.1231
50	7094	2007102313	0.2362
75	6785	2007101016	0.3483
90	7258	2007103009	0.4440

Métodos de Predicción de la Generación Agregada de Energía Eólica

Con el objetivo de conocer cómo son las distribuciones de distancias en función del nivel de potencia prevista, se representarán los histogramas y las dispersiones de puntos de las distancias entre el vector de viento previsto y los vectores históricos almacenados, $d(\mathbf{v}_t, \mathbf{v}_{t+h})$, para valores de potencia esperada muy bajo, bajo, medio, elevado y muy elevado, a un horizonte de predicción de 24 horas. De modo que, partiendo de un instante t situado 24 horas antes de las horas propuestas, se compara la predicción de viento para el instante $t+24$ con los vectores históricos de viento registrados hasta ese instante t .

Por simplicidad, únicamente se han representado las dispersiones e histogramas para los tipos de distancia absoluta y euclídea. Un análisis detallado de la influencia del tipo de distancia utilizado se hará más adelante.

En las representaciones de la Figura 5.23 se puede observar la forma que adquiere la dispersión de datos en función del nivel de potencia esperado. Así, para niveles de potencia bajos y muy bajos (gráficas superiores), se observa una gran concentración de puntos cercanos, a poca distancia del nuevo vector. Esto se manifiesta en un histograma muy concentrado en distancias bajas, de menos de 2 m/s, con poca concentración de datos a distancias elevadas que se traduce en las colas largas a la derecha del histograma (Figura 5.24).

A medida que el nivel de potencia esperado va aumentando (gráficas inferiores), se observa cómo la punta del histograma se va desplazando hacia la derecha, debido a que cada vez hay menor concentración de datos cercanos, puesto que hay menos datos observados con potencia más elevada, que equivaldría a valores elevados de velocidad de viento en la región.

De las dispersiones de datos se puede intuir el valor de potencia esperado, observando los valores de potencia que se registraron en instantes pasados y su distancia. La potencia prevista será una ponderación o combinación de los valores de potencia cuyos puntos en la dispersión se encuentran más próximos al eje de ordenadas, es decir, aquellos cuya distancia es menor.

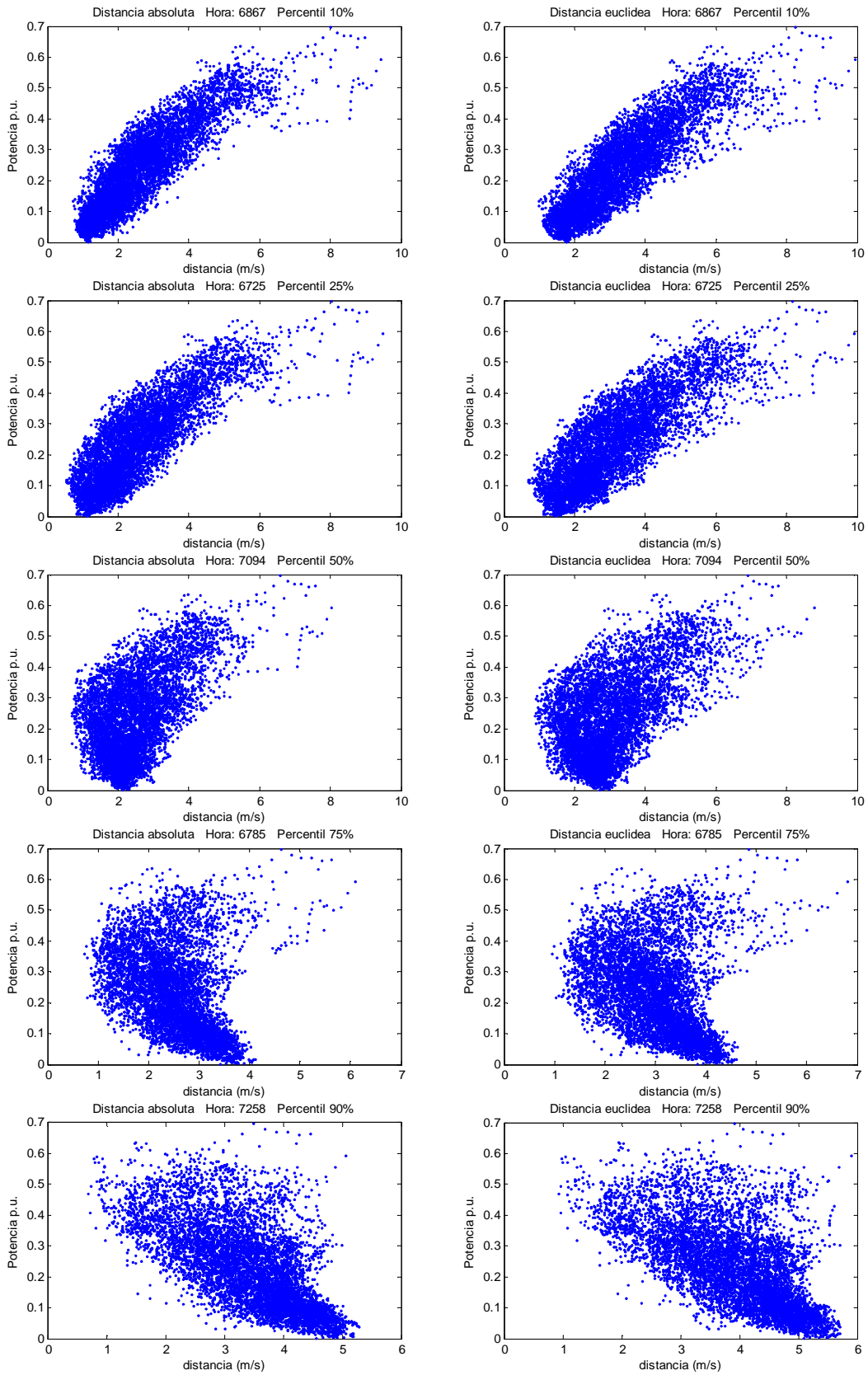


Figura 5.23: *Dispersión de distancias para potencia esperada muy baja, baja, media, alta y muy alta*

Métodos de Predicción de la Generación Agregada de Energía Eólica

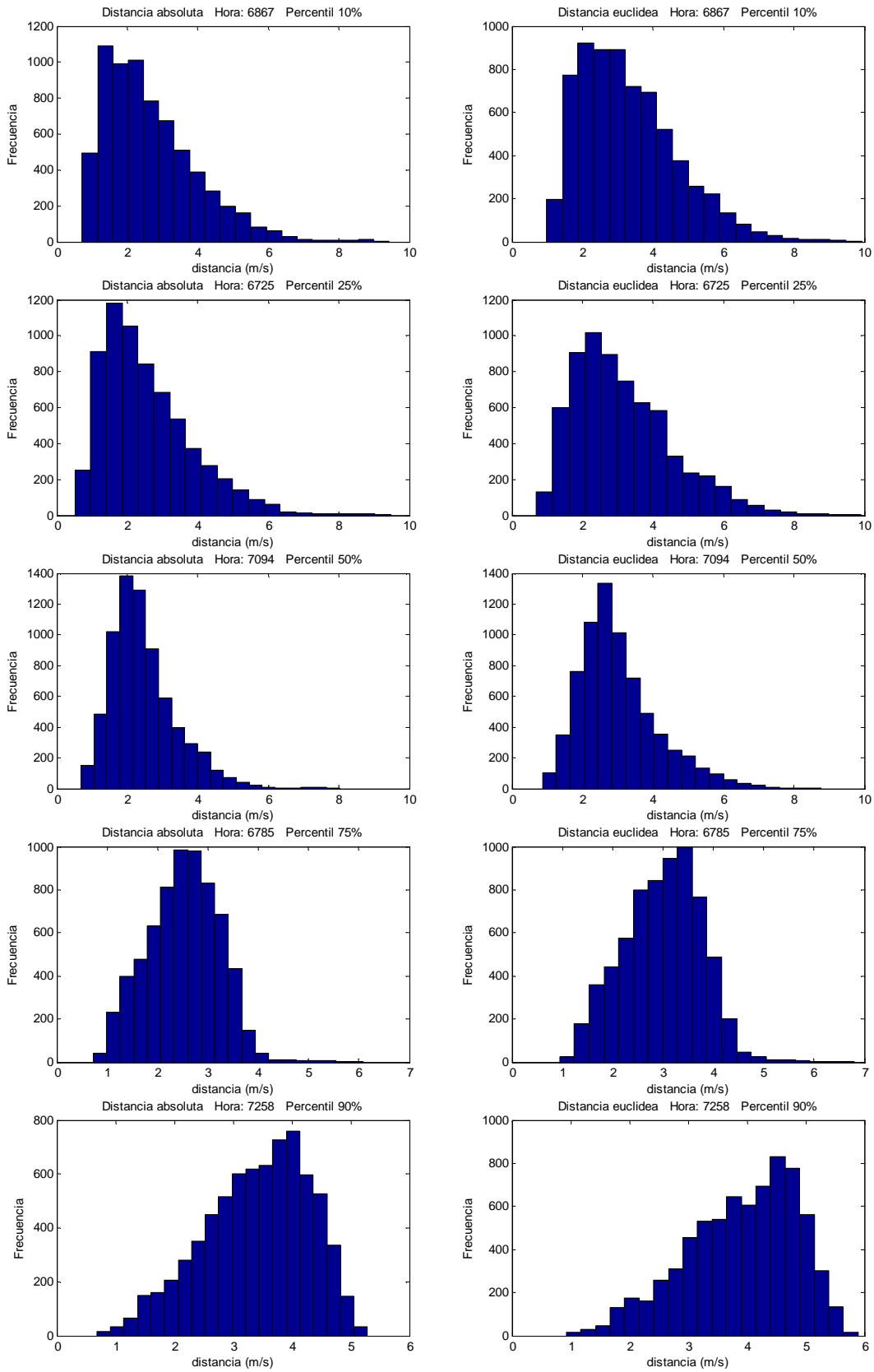


Figura 5.24: Histograma de distancias para potencia esperada muy baja, baja, media, alta y muy alta

5.3.2 Dependencia del tipo de medida de distancia elegido

A continuación se desean comparar las dispersiones e histogramas resultantes de calcular los distintos tipos de distancia propuestos: Distancia de velocidad media, absoluta, cúbica, euclídea, absoluta estandarizada con la media, absoluta estandarizada con la desviación típica, cúbica estandarizada con la desviación típica, euclídea estandarizada con la varianza y Mahalanobis.

Para hacer la comparación, se eligen los instantes en los que la potencia tuvo unos valores bajo, medio y elevado. En este caso, se han elegido horas en las que la potencia registrada fue de 0.123, 0.236 y 0.348 p.u., correspondientes a las horas 6725, 7094 y 6785 del año respectivamente (Tabla 5.4). El horizonte de la predicción del vector de viento para calcular la distancia $d(\mathbf{v}_i, \mathbf{v}_{t+h})$ se ha fijado de nuevo en $h=24$ horas.

En las figuras de las páginas siguientes se representan las dispersiones e histogramas de las distancias calculadas a los tres niveles de potencia elegidos. Las dos gráficas de la fila superior representan la dispersión e histograma para el nivel de potencia bajo, la segunda fila para el nivel de potencia medio, y las gráficas inferiores son las obtenidas para el nivel de potencia elevado.

Distancia de velocidad media

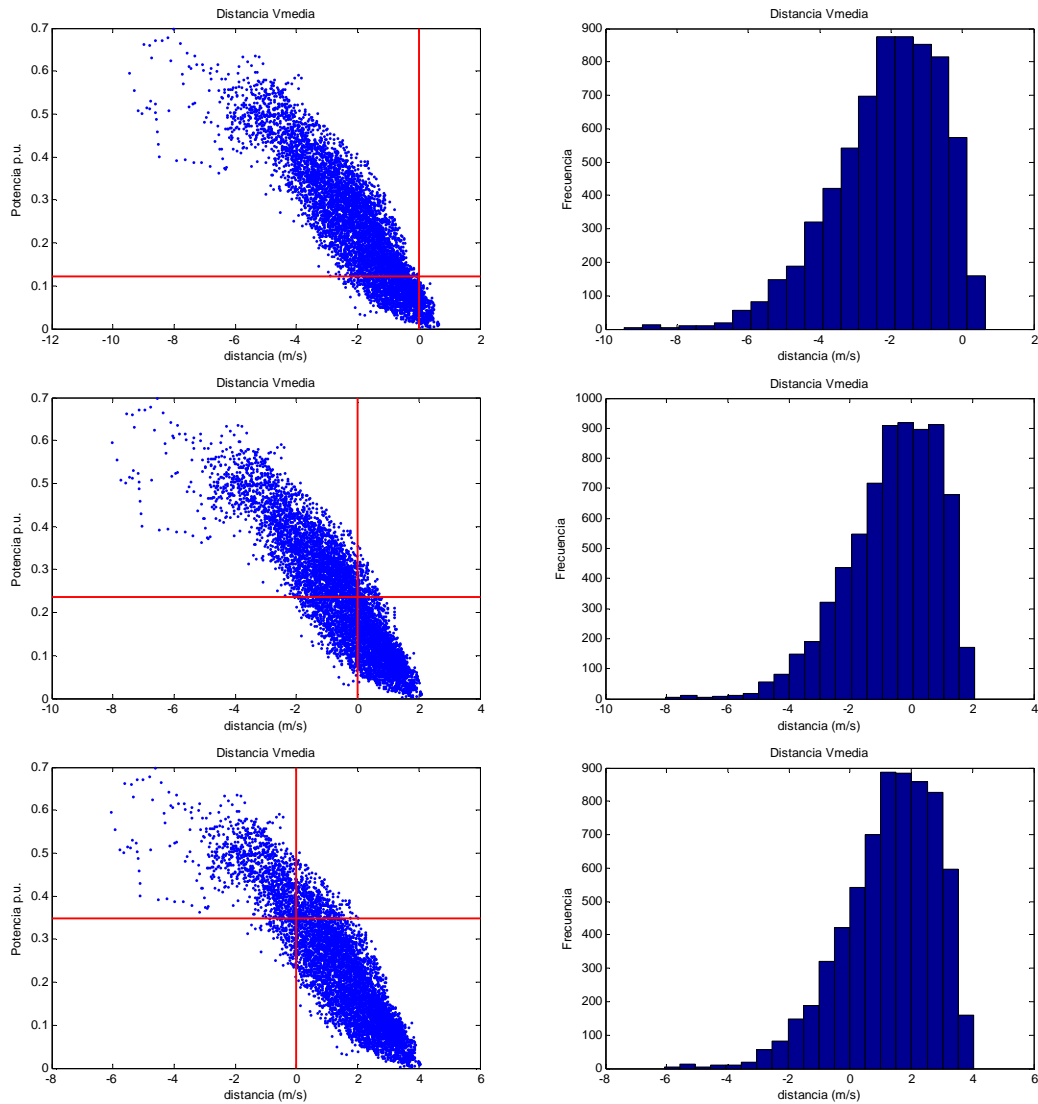


Figura 5.25: Dispersiones e histogramas para distancias de velocidad media

Si el modo de comparación de vectores es la velocidad media, según la ecuación (4.10), se observa que el paso de la nube de puntos por distancia $d = 0$ se encuentra en el entorno de la potencia esperada, aunque con un rango muy elevado de valores de potencia a poca distancia (Figura 5.25). Así, por ejemplo, a distancia cero encontramos valores entre 0 y 0.1 p.u. cuando la potencia esperada es de 0.12, como es el caso de potencia baja representado, o un rango de valores entre 0.05 y 0.34 p.u. cuando la potencia esperada es de 0.24 p.u., correspondiente al caso de potencia media.

El gran rango de valores que adquieren los datos de potencia que tuvieron una velocidad media similar a la del nuevo vector, y que por tanto se encuentran a

distancia muy reducida, ya nos indica que este tipo de distancia no dará buenos resultados, ya que no tiene suficiente capacidad de selección o de asignación de importancia relativa a los datos históricos observados.

Una característica particular de la medida de distancia de velocidad media es que, excepto por el nuevo punto que se añade cada vez que se tenga un nuevo dato histórico, la forma de la dispersión de puntos será siempre la misma, al igual que el histograma, independientemente del nivel de potencia esperable para el nuevo vector de viento. Solamente se desplazarán los puntos a la izquierda o derecha sobre el eje de abscisas donde se representan las distancias.

Distancia absoluta

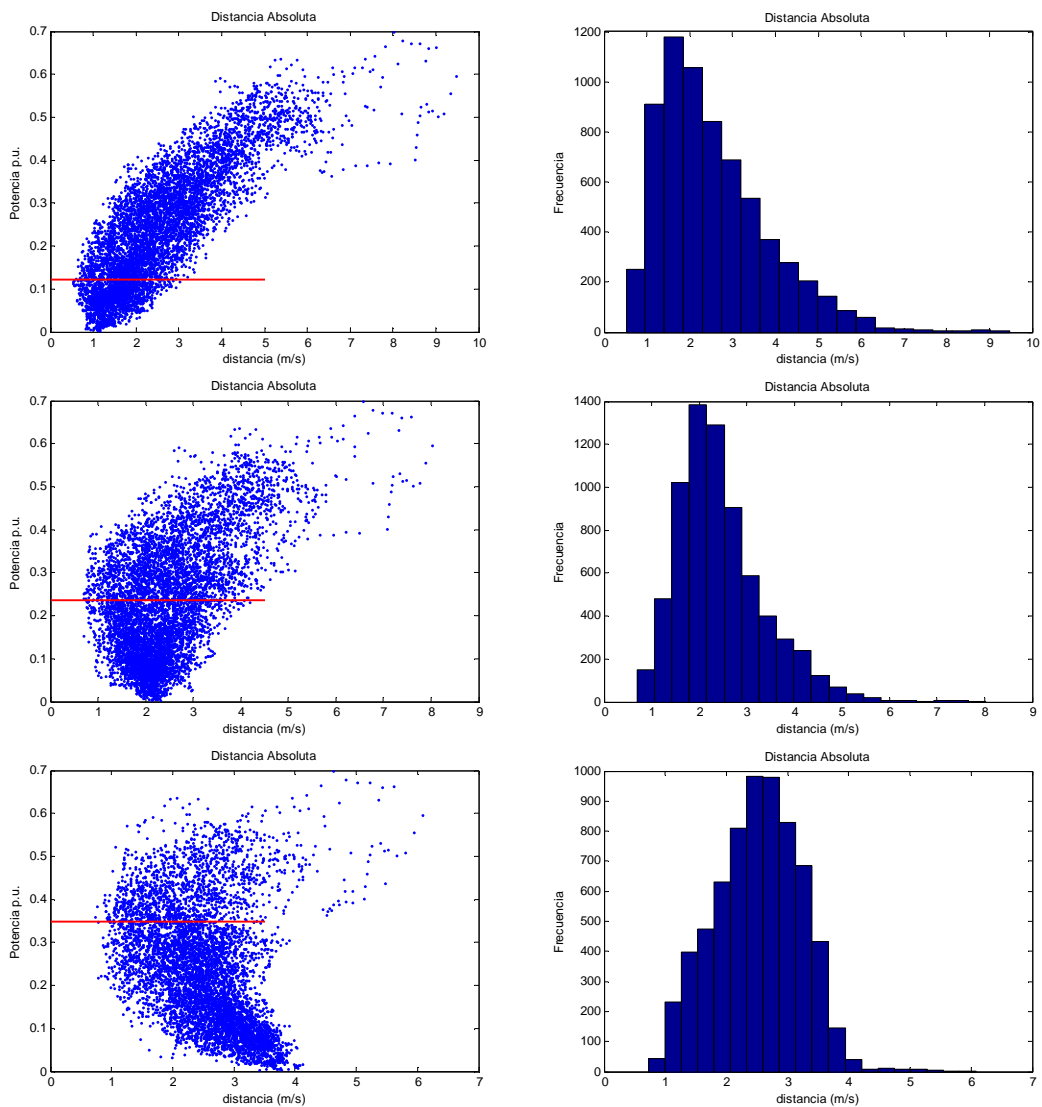


Figura 5.26: Dispersiones e histogramas para distancia absoluta

Distancia absoluta estandarizada con la media

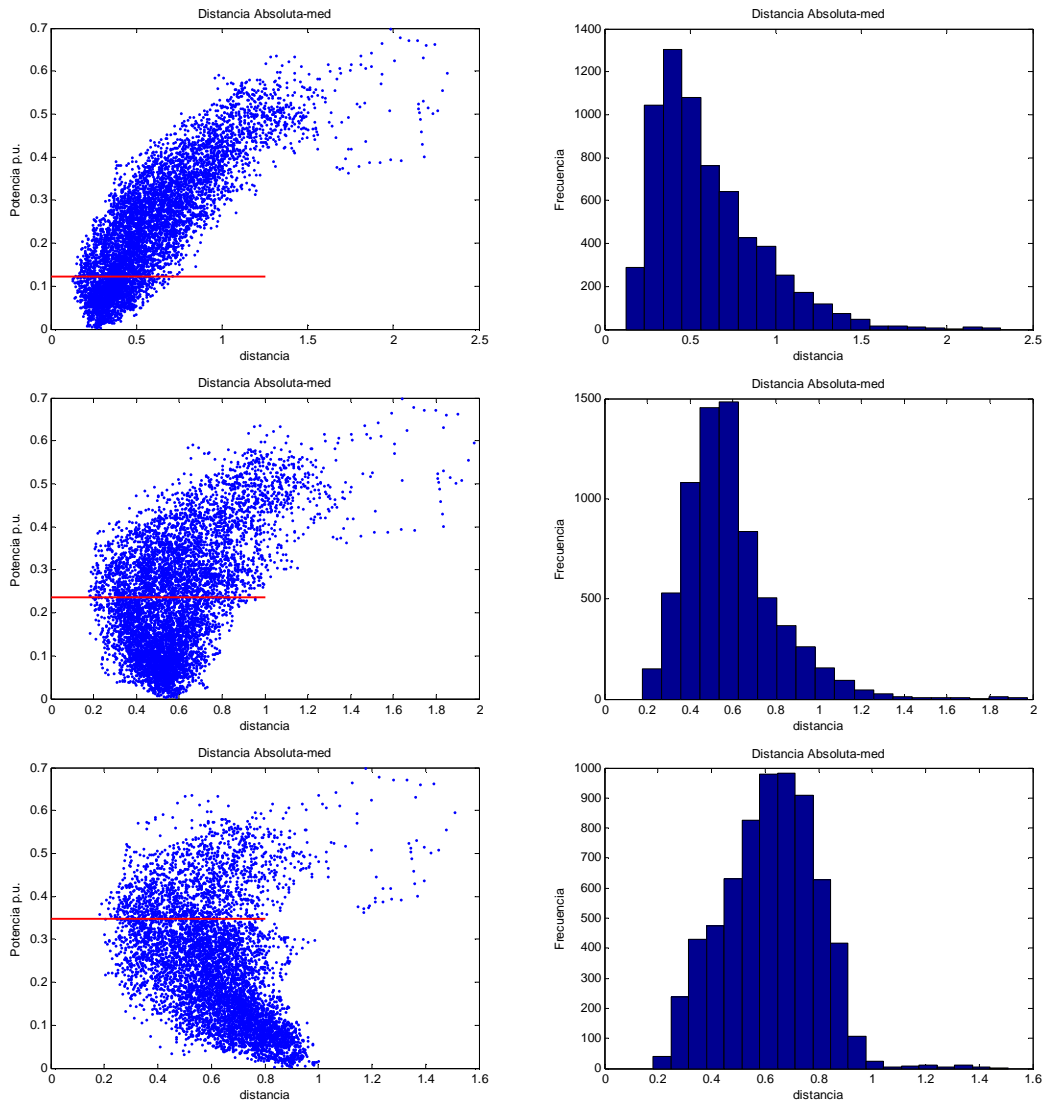


Figura 5.27: Dispersiones e histogramas para distancia absoluta estandarizada con la media

Distancia absoluta estandarizada con la desviación típica

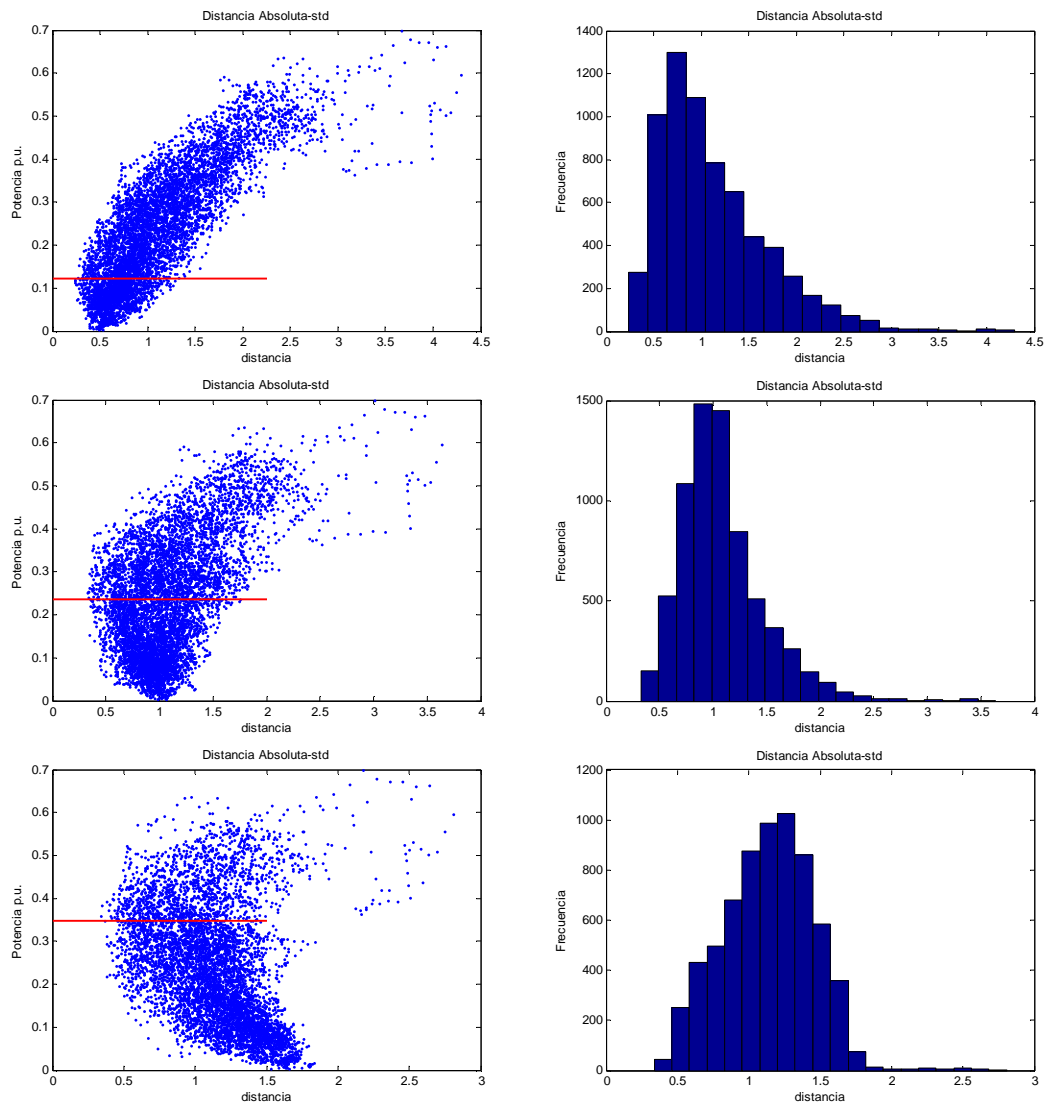


Figura 5.28: Dispersiones e histogramas para distancia absoluta estandarizada con la desviación típica

En la representación de las dispersiones e histogramas correspondientes a distancia absoluta, ecuación (4.13) y Figura 5.26, absoluta estandarizada con la media, ecuación (4.16) y Figura 5.27, y absoluta estandarizada con la desviación típica, ecuación (4.18) y Figura 5.28, la forma de las dispersiones e histogramas depende del nivel de potencia esperada, observándose una gran concentración de datos a baja distancia para potencias bajas, y desplazándose dicha concentración de datos hacia distancias más elevadas conforme aumenta la potencia esperada, encontrándonos en estos casos con poca cantidad de datos a baja distancia.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Se puede apreciar en las figuras que los puntos a menor distancia de las dispersiones tienden a valores de potencia en torno a los esperados, de 0.12, 0.24 y 0.35 p.u. para potencias esperadas baja, media y elevada respectivamente.

Por otro lado, apenas se puede apreciar visualmente la diferencia en la forma de la dispersión o del histograma para un mismo nivel de potencia dependiendo de si se utiliza la distancia absoluta, o si ésta se estandariza con la media o la desviación típica. La principal diferencia fácilmente observable se encuentra en el rango de valores que pueden adquirir las distancias calculadas, tal y como se puede apreciar en las escalas de los ejes, donde encontramos valores de hasta unos 9 m/s para la distancia absoluta, mientras que las medidas de distancia estandarizadas con la media y la desviación típica tienen valores hasta rangos de unos 2 ó 4 m/s respectivamente.

Distancia euclídea

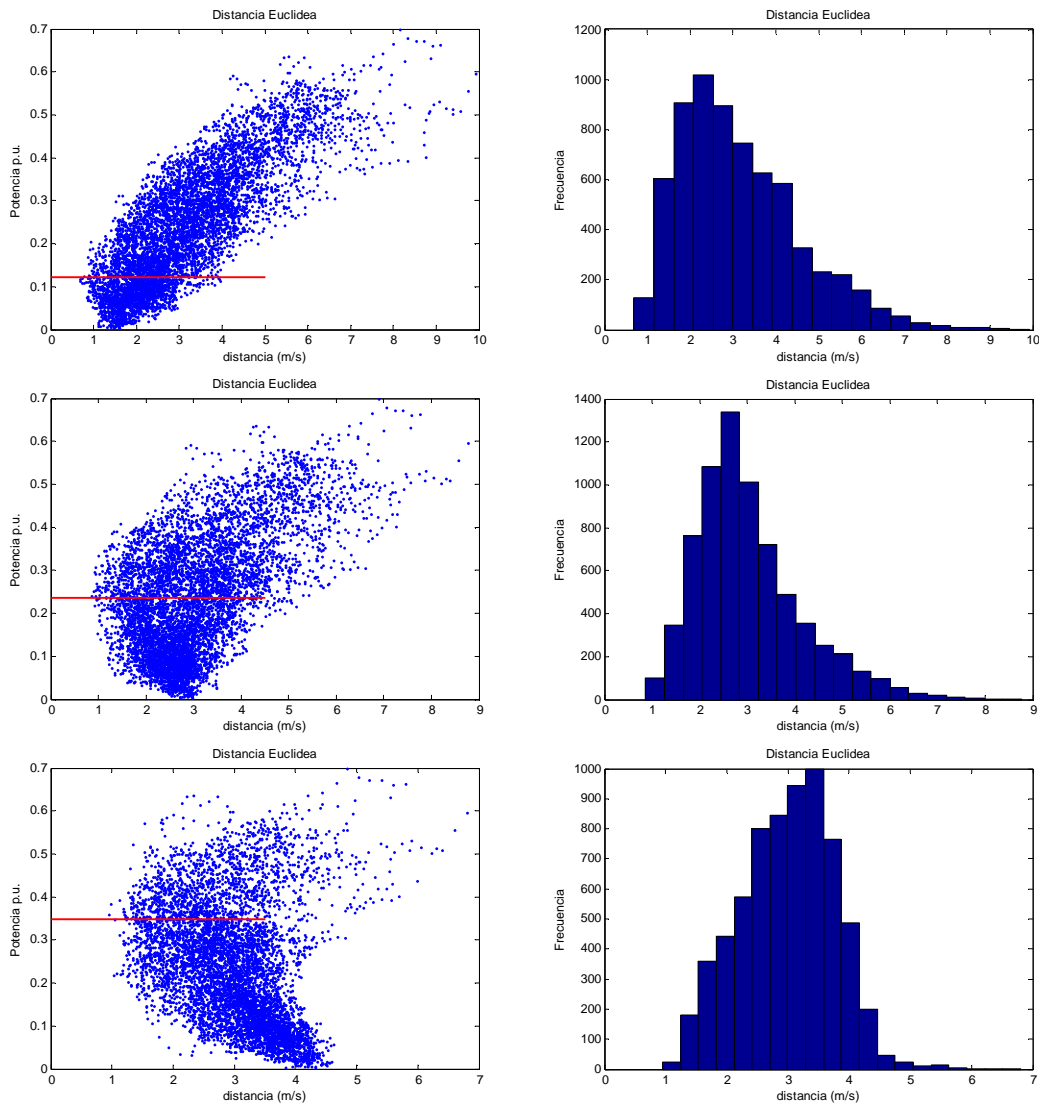


Figura 5.29: Dispersiones e histogramas para distancia euclídea

Distancia euclídea estandarizada con la varianza

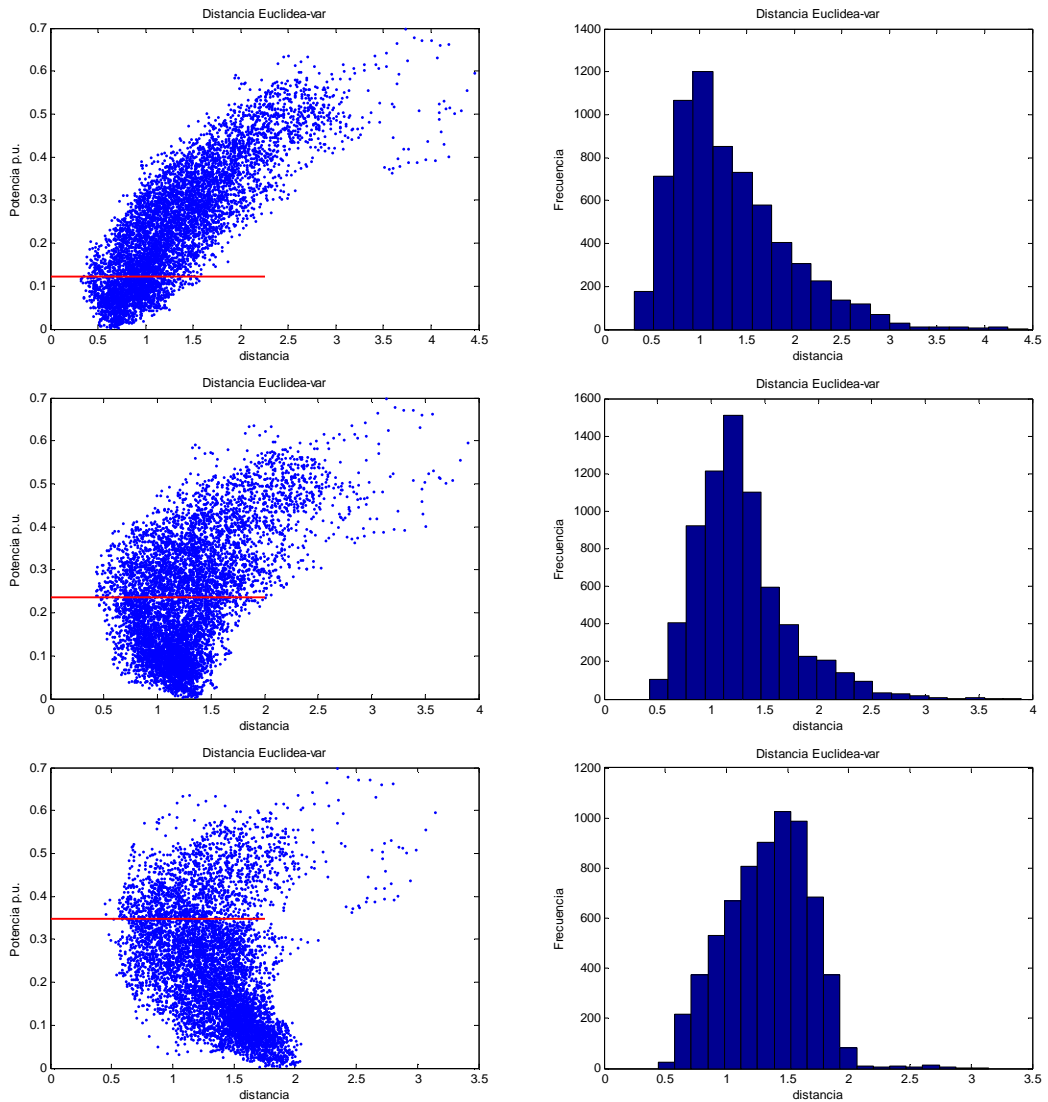


Figura 5.30: Dispersiones e histogramas para distancia euclídea estandarizada con la varianza

Si nos fijamos en la Figura 5.29 y Figura 5.30, que representan las dispersiones e histogramas obtenidos con la medida de distancia euclídea (4.14) y euclídea estandarizada con la varianza (4.20), y se compara con lo que se obtuvo calculando distancias absolutas, se observa cierta diferencia entre las distancias euclídeas y absolutas, que se pone de manifiesto con un desplazamiento hacia la derecha de la moda o intervalo con más frecuencia en el histograma representado con la familia de distancias euclídeas para todos los niveles de potencia esperados, ya sea bajo, medio o elevado.

Esto se explica porque el cálculo de distancia euclídea penaliza más que la distancia absoluta a los vectores con algunas coordenadas con grandes diferencias de velocidad

Métodos de Predicción de la Generación Agregada de Energía Eólica

entre el nuevo vector de viento y el vector histórico comparado. Así, en el histograma con distancias euclídeas, se habrán alejado aquellos vectores históricos que contengan alguna coordenada con una velocidad de viento muy distinta de la predicción en la misma coordenada del vector nuevo, habiendo por tanto menos concentración de elementos en la zona de distancias bajas.

Por otra parte, de nuevo se puede apreciar que la principal diferencia claramente visible entre el cálculo de distancia euclídea y euclídea estandarizada con la varianza se encuentra en el rango de valores que adquieren dichas distancias, y que se puede comprobar en los valores del eje horizontal donde se representan las distancias, alcanzándose distancias de hasta 8 m/s si se usa la distancia euclídea, y de unos 4 m/s si la estandarizamos con la varianza.

Distancia cúbica

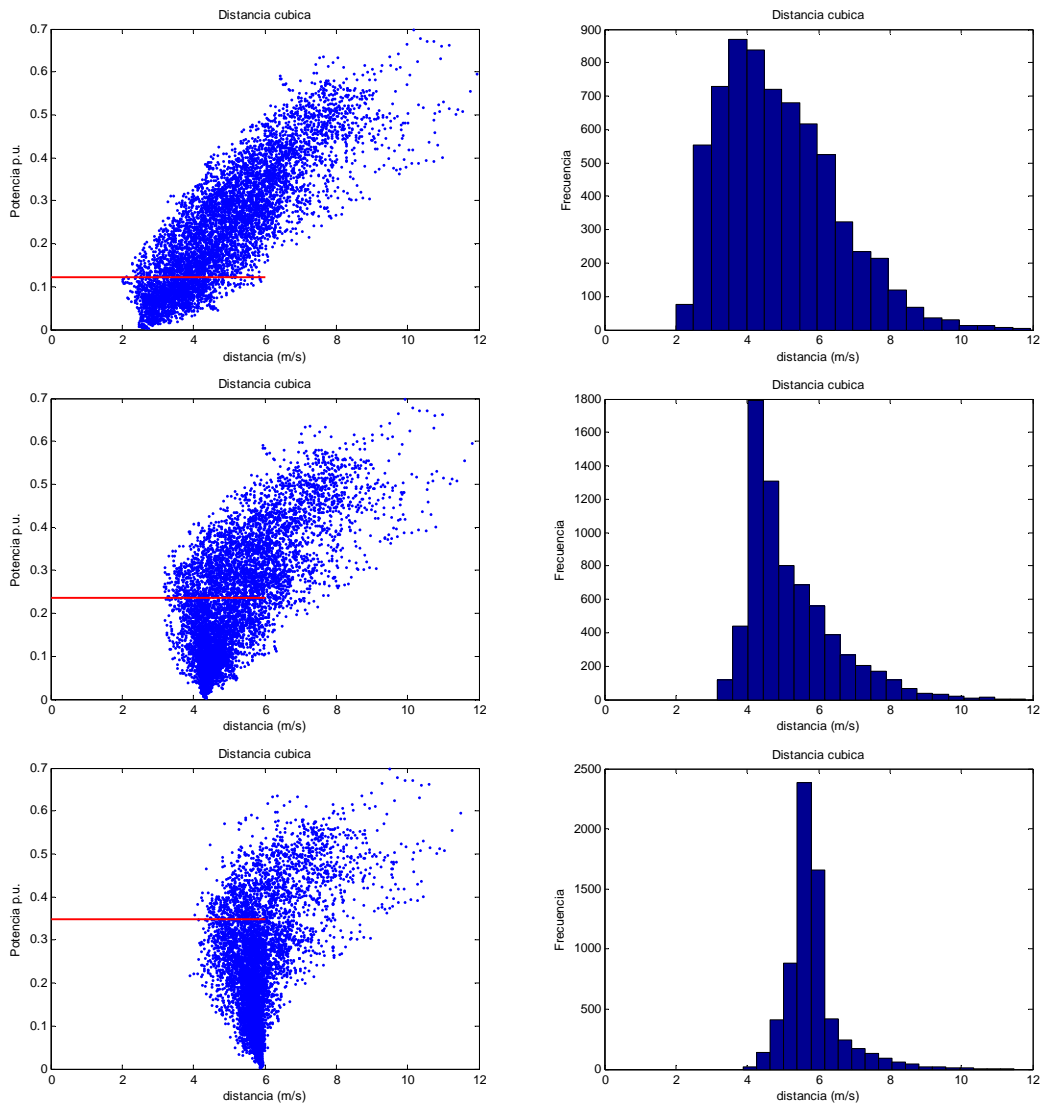


Figura 5.31: Dispersiones e histogramas para distancia cúbica

Distancia cúbica estandarizada con la desviación típica

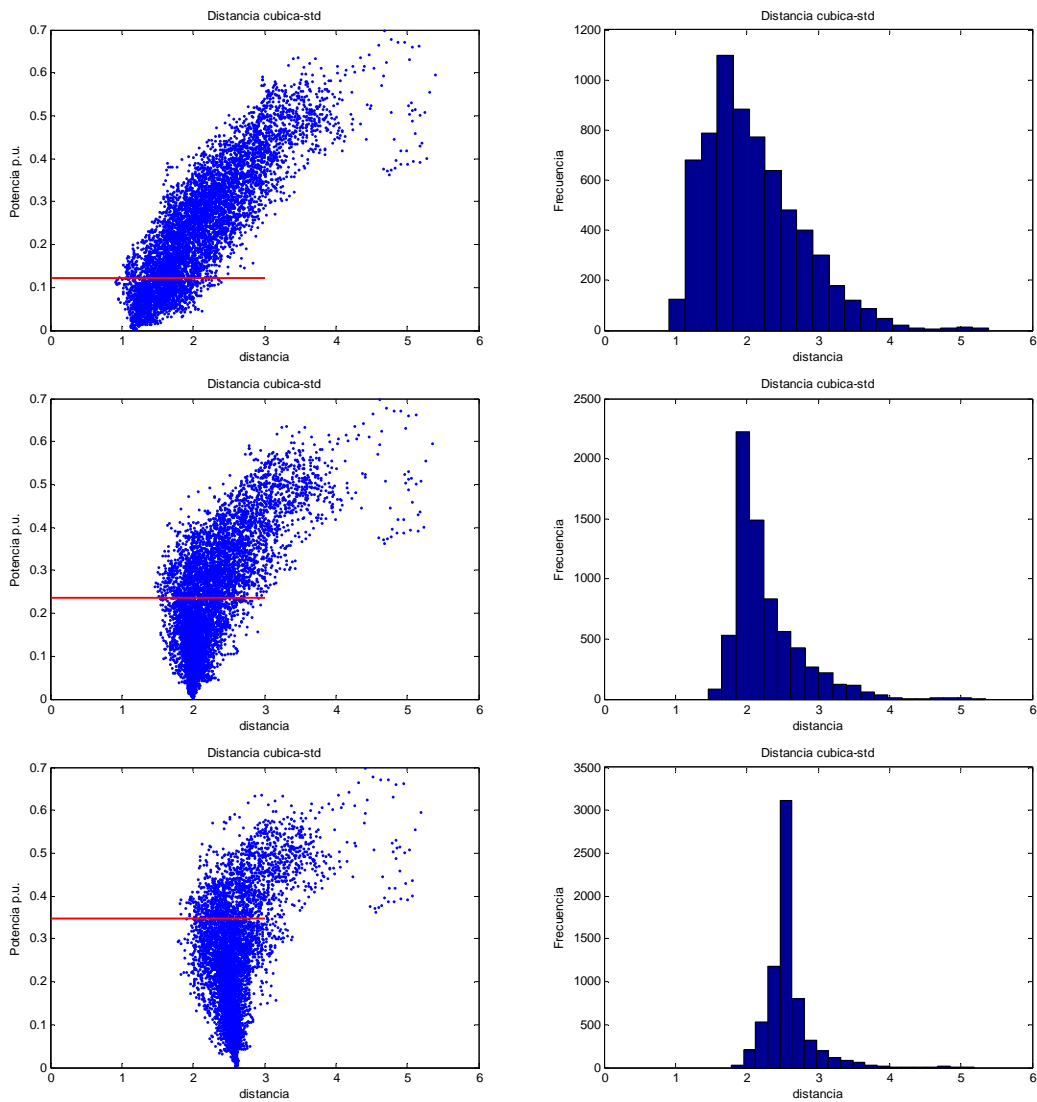


Figura 5.32: Dispersiones e histogramas para distancia cúbica estandarizada con la desviación típica

Tanto la distancia cúbica, ecuación (4.15) y Figura 5.31, como la cúbica estandarizada, ecuación (4.21) y Figura 5.32, presentan dispersiones e histogramas similares. Se aprecia en cambio que la cúbica es una medida de distancia aún más selectiva que las distancias absoluta o euclídea, ya que la concentración de datos a baja distancia es menor, pues los histogramas presentan menor cantidad de datos en las primeras clases. Esta mayor selectividad se debe a que la distancia cúbica penaliza en mayor medida aquellos datos con valores de velocidad de viento muy distintos en alguna de las coordenadas.

Al igual que con otros tipos de distancias, la estandarización se puede apreciar en el cambio de escala en el eje que mide las distancias, pasando de rangos de valores de unos 12 m/s a rangos de 5 m/s si se estandarizan las distancias.

Distancia de Mahalanobis

En último lugar, se representan las dispersiones para potencias esperadas baja, media y elevada usando la distancia de Mahalanobis, ecuación (4.22). En esta ocasión, y al contrario de lo que ocurriría con las distancias absoluta, euclídea o cúbica, la dispersión de puntos no apunta claramente a ninguna potencia concreta en la zona de distancias bajas. Se observan así, por ejemplo, valores históricos cercanos que varían desde 0.05 p.u. hasta 0.45 p.u. cuando se espera una potencia media de 0.24 p.u. Debido a esto no se espera un buen funcionamiento del modelo de predicción que utilice distancia Mahalanobis, aunque esto último se comprobará más adelante.

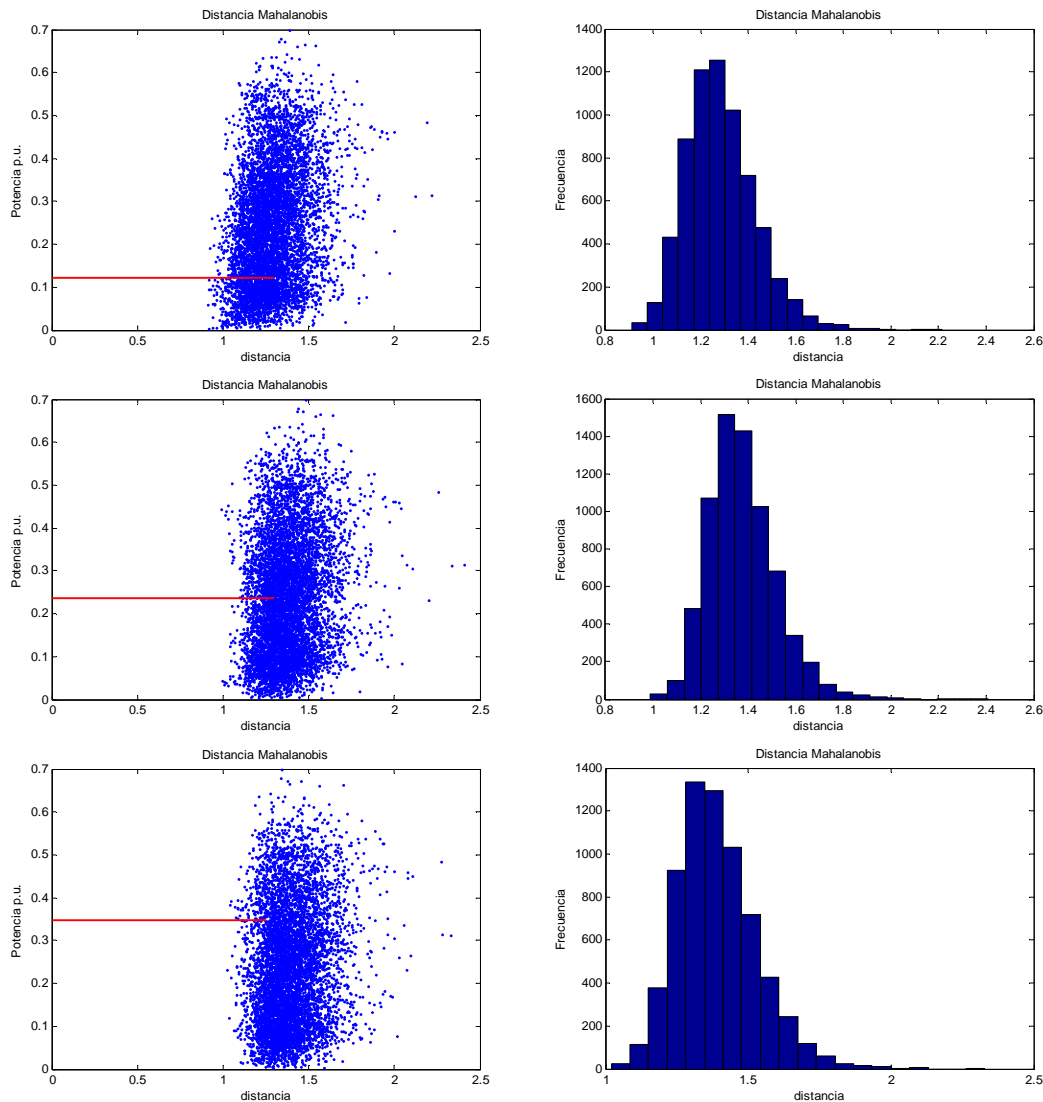


Figura 5.33: Dispersiones e histogramas para distancia de Mahalanobis

5.4 Evaluación de los métodos de selección

En esta sección se realizará una evaluación de los métodos que se han propuesto para seleccionar los datos históricos cuyos vectores de velocidad de viento son más cercanos al nuevo vector de predicción de viento para las coordenadas de la región.

Para la evaluación se han utilizado los datos de viento y de potencia total eólica horaria en la península durante el año 2007, lo que resulta en un total de 8760 datos. Se ha establecido como periodo histórico inicial de datos los primeros seis meses, de enero a junio de 2007. El periodo de prueba comprende la segunda mitad del año, de julio a diciembre de 2007.

Dado que el objetivo en esta sección es la evaluación de los métodos de selección propuestos, se realiza un análisis de todos los métodos de selección, aunque por simplicidad este análisis se ha realizado únicamente usando las medidas de distancia absoluta y euclídea. El análisis del error para los distintos tipos de medida de distancia propuestos se realizará con posterioridad.

Utilizando los modelos locales de Media Ponderada y de Regresión Local Ponderada, se representan las curvas de error para los horizontes $h = 8, 12, 18, 24, 36$ y 48 horas, en función del valor del parámetro que decide el número de datos seleccionados dependiendo del método de selección elegido.

5.4.1 Datos a una distancia inferior a un valor umbral fijo

Con este método se selecciona un número variable de elementos cercanos que estén situados a una distancia inferior a un valor umbral fijo de distancia, definido por un ancho de banda. Usando un modelo local de Media Ponderada, ecuación (3.3), se representan las curvas de error para los horizontes 8, 12, 18, 24, 36 y 48 horas, para un rango de valores de distancia umbral entre 0.1 y 2 m/s, (Figura 5.34).

Hay que tener en cuenta que, para evitar la posibilidad de no existencia de datos por debajo de una distancia dada, en el caso de que no se encuentre ningún dato bajo el valor de distancia umbral seleccionada, se ha escogido el dato más cercano encontrado, para evitar no tener ningún dato para calcular la potencia. Este caso equivaldría a usar el método de selección de M vectores más cercanos, siendo $M = 1$.

Esto último se puede apreciar en las representaciones del error NMAE y NRMSE, puesto que el valor del error se mantiene constante para los valores de distancia umbral hasta 0.4 m/s y 0.6 m/s si se usan distancia absoluta y euclídea respectivamente. Esto nos indica que todas las distancias absoluta y euclídea calculadas fueron superiores a 0.4 y 0.6 m/s respectivamente para los horizontes representados.

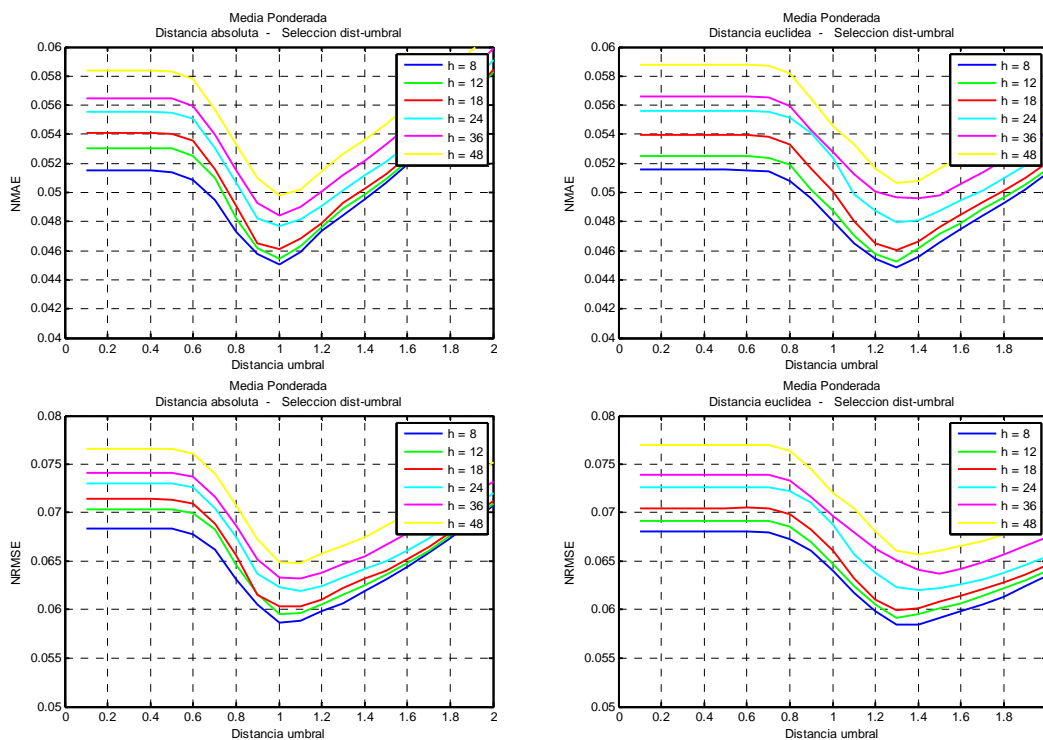


Figura 5.34: Errores NMAE y NRMSE para el método de selección de distancia umbral con modelo de Media Ponderada

Con un modelo de Media Ponderada, la distancia umbral óptima si se utiliza el método de cálculo de distancia absoluta se sitúa entre 1 y 1.1 m/s, mientras que si se usa la distancia euclídea el valor óptimo de ancho de banda se sitúa entre 1.3 y 1.5 m/s para todos los horizontes representados.

Además, se aprecia gran dependencia del resultado de los errores en función del ancho de banda de selección, aumentando drásticamente el error en cuanto nos alejemos ligeramente de los valores óptimos del ancho de banda.

Si se utiliza un modelo de Regresión Local Ponderada, (Figura 5.35), se observa que el valor mínimo del ancho de banda para que se puedan tener predicciones es sensiblemente superior al necesario para un modelo de Media Ponderada. Con anchos de banda inferiores a 2.2 o 2.5 m/s para las distancias absoluta y euclídea respectivamente, no se pudieron calcular predicciones por falta de datos. Esto es debido a que para poder resolver las ecuaciones de la regresión, y calcular todos los parámetros β en (3.11), es necesario disponer de un número suficiente de datos históricos, ya que la dimensión de nuestro problema es bastante elevada, pues contamos con 83 coordenadas de predicción de viento.

El valor óptimo de ancho de banda en este caso se sitúa entre los 2.4 y 2.8 m/s si se usa la distancia absoluta, y entre 3 y 3.5 m/s si se utiliza la distancia euclídea, dependiendo del horizonte de predicción.

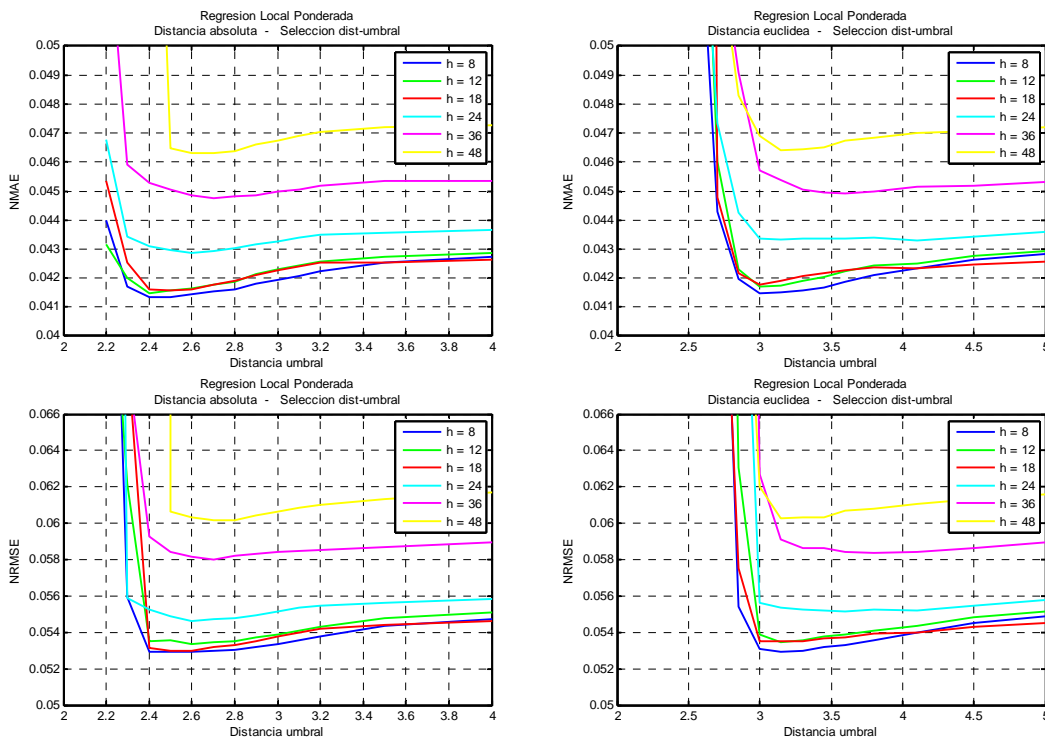


Figura 5.35: Errores NMAE y NRMSE para el método de selección de distancia umbral con modelo de Regresión Local Ponderada

5.4.2 M vecinos más cercanos

A continuación se presentan las curvas de los errores NMAE y NRMSE calculados en función del número M de datos seleccionados, utilizando el método de selección de los M vectores más próximos para los métodos de medida de distancia absoluta y euclídea.

Utilizando un modelo de Media Ponderada, se representan las curvas para los horizontes 8, 12, 18, 24, 36 y 48 horas, con un rango del parámetro M entre 1 y 100 datos seleccionados (Figura 5.36).

Se puede comprobar la bondad del efecto del suavizado, ya que se obtiene menor error al hacer la media de varios elementos cercanos al nuevo vector de predicción de viento que si se utiliza como predicción únicamente la potencia eólica del elemento más cercano, tal y como puede observarse en las curvas de error, que presentan un error máximo cuando $M=1$ para todos los horizontes representados. Con esto queda patente la utilidad del suavizado con puntos cercanos para el modelado en el problema de predicción.

Por otro lado, se puede apreciar que cuando se utiliza un modelo de Media Ponderada el valor óptimo de M se sitúa entre los 20 y 50 datos más cercanos, zona donde las curvas que representan el error NMAE y NRMSE en función de M encuentran su mínimo.

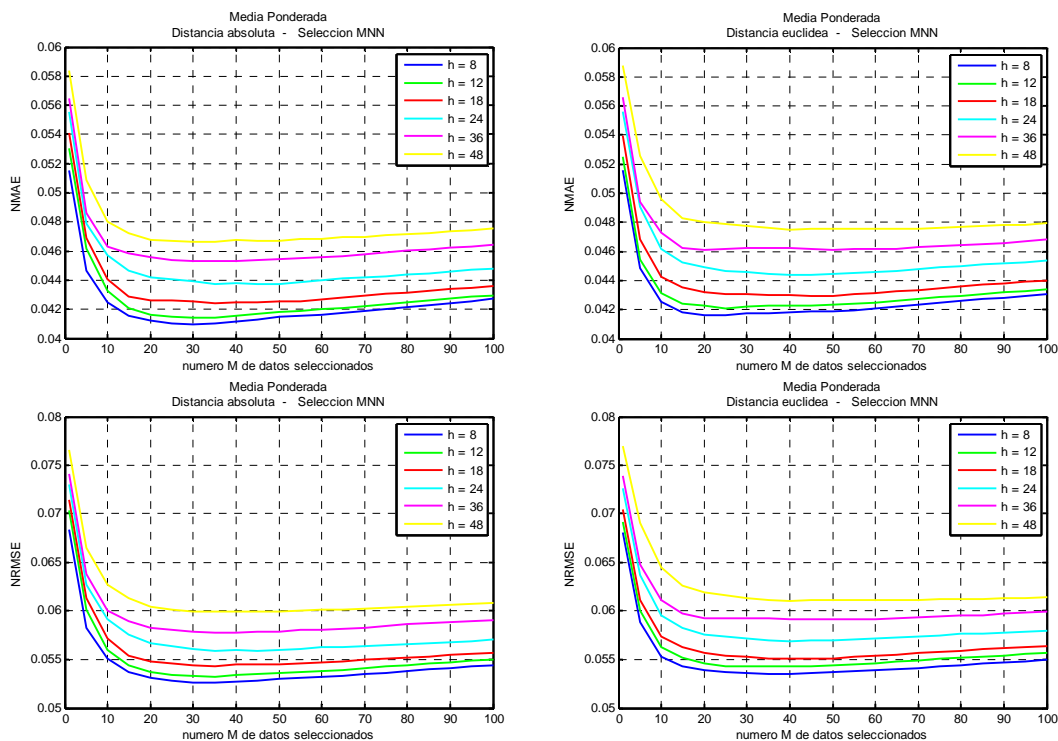


Figura 5.36: Errores NMAE y NRMSE para el método de selección de M vecinos más cercanos con modelo de Media Ponderada

Además, cabe destacar que los menores valores de error obtenidos para todos los horizontes utilizando este método de selección son claramente inferiores a los que se obtuvieron utilizando el método de selección de ancho de banda fijo, aunque esto se verá con más detalle en el resumen final de esta sección.

Por último, el error mínimo para los horizontes representados es ligeramente inferior utilizando la distancia absoluta que utilizando distancia euclídea, aunque esta diferenciación se estudiará con más detalle en la siguiente sección, en la que se compararán los distintos procedimientos de cálculo de distancias.

Si se representan los errores obtenidos utilizando un modelo de Regresión Local Ponderada se obtiene la Figura 5.37, donde se aprecia que el número de datos necesario para poder resolver el problema de regresión se incrementa considerablemente. Así, con un número de datos inferior a 400, se tuvieron problemas para poder calcular la inversa en la ecuación de regresión (3.11). El valor adecuado para obtener los mínimos errores se sitúa en torno a 2000 o 2500 datos, dependiendo del tipo de distancia utilizado y del horizonte de predicción, aunque en general se aprecia que a medida que aumenta el horizonte de predicción, aumenta el número de datos necesarios para minimizar el error. Esto se debe a que conforme aumenta el horizonte las predicciones de viento son menos precisas, existiendo gran variabilidad entre los valores de potencia cercanos encontrados, por lo que se hace necesario un mayor suavizado, y por tanto mayor número de datos, para minimizar en lo posible el error de predicción.

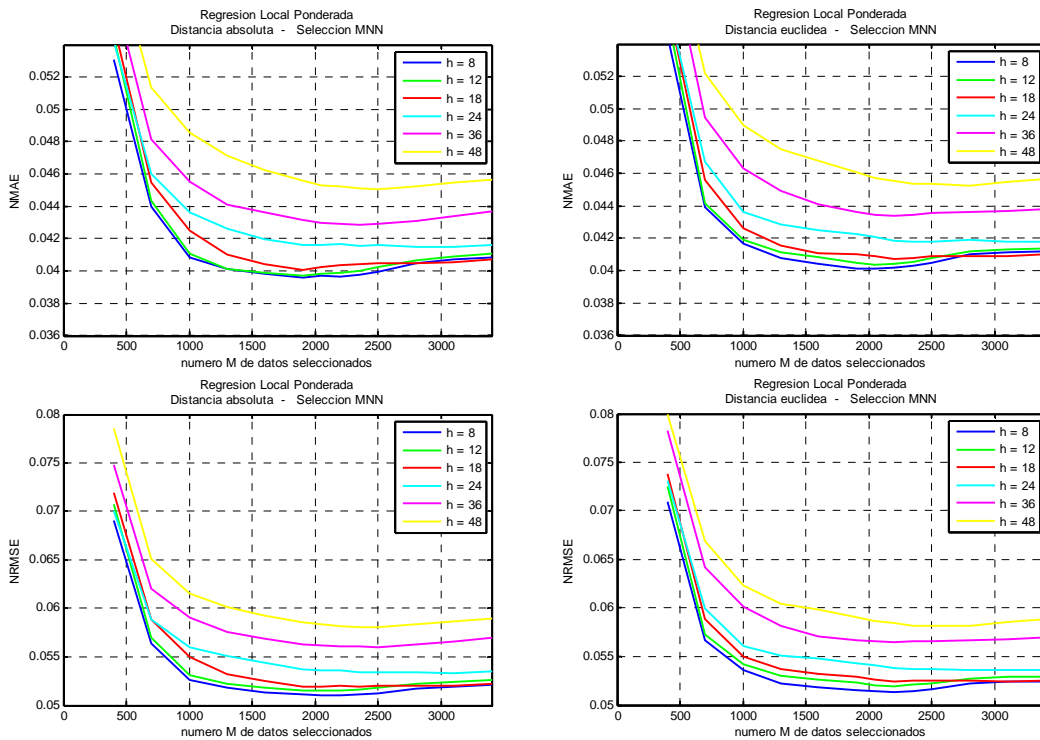


Figura 5.37: Errores NMAE y NRMSE para el método de selección de M vecinos más cercanos con modelo de Regresión Local Ponderada

5.4.3 M como un porcentaje fijo de datos más cercanos del total de datos disponibles

En la Figura 5.38 se representan las curvas de error calculadas utilizando el método de selección de un porcentaje fijo de los datos históricos, para unos valores de porcentaje p entre 0.1% y 2%. Se presentan las curvas de dependencia de NMAE y NRMSE, en función del parámetro p , para los horizontes de predicción 8, 12, 24, 36 y 48 horas, utilizando un modelo de Media Ponderada para calcular la predicción.

Los mínimos errores obtenidos con este método son muy similares a los obtenidos utilizando el método de selección de los M vecinos más cercanos. En este caso, el valor óptimo del parámetro p está en el entorno de 0.3 al 0.7 %, dependiendo del horizonte de predicción y del método de cálculo de distancia.

Al igual que en el caso de selección de M vecinos más cercanos, los resultados obtenidos usando la distancia absoluta son ligeramente mejores que los obtenidos usando distancia euclídea para los horizontes analizados.

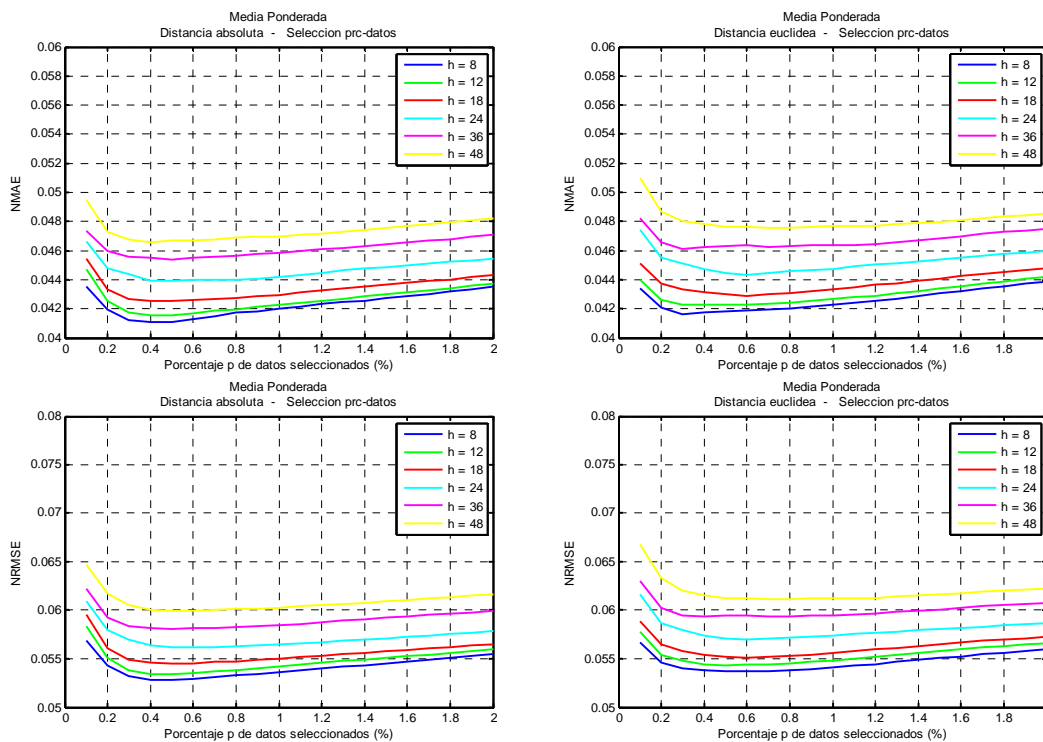


Figura 5.38: Errores NMAE y NRMSE para el método de selección de porcentaje de datos más cercanos con modelo de Media Ponderada

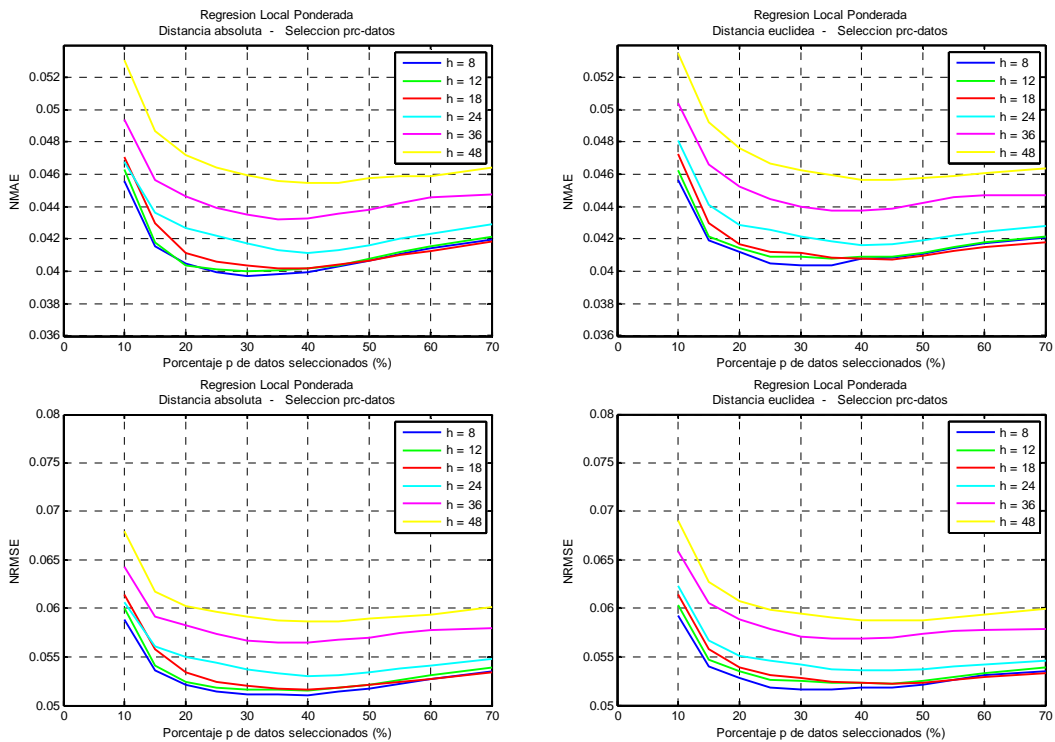


Figura 5.39: Errores NMAE y NRMSE para el método de selección de porcentaje de datos más cercanos con modelo de Regresión Local Ponderada

Si se representan los errores obtenidos mediante un modelo de Regresión Local Ponderada para calcular la predicción (Figura 5.39), se obtienen de nuevo conclusiones similares. Por un lado, se aprecia que es necesario un buen número de datos para poder calcular predicciones, debido al elevado número de dimensiones del problema a resolver. Así, con porcentajes inferiores al 10 % de datos se tuvieron problemas para invertir la matriz en la ecuación de regresión.

Los valores óptimos de datos se sitúan en torno al 35 a 45 %, dependiendo del horizonte de predicción y del tipo de distancia utilizado, con los que de nuevo se obtienen unos errores mínimos similares a los obtenidos con el método de selección de M vecinos más cercanos.

Al igual que en el caso de selección de M vecinos más cercanos, con un modelo de Regresión Local Ponderada los resultados obtenidos usando la distancia absoluta también son ligeramente mejores que los obtenidos usando distancia euclídea para los horizontes analizados.

5.4.4 Datos a una distancia inferior a un porcentaje fijo del rango de distancias

En la Figura 5.40 se representan los errores que se obtienen utilizando un método de cálculo de predicciones de Media Ponderada con el método de selección de un valor de distancia umbral definido por un porcentaje del rango de distancias calculadas. El parámetro p de este método define el porcentaje del rango de distancias del que se seleccionarán los elementos cercanos. Se han representado las curvas de error NMAE y NRMSE para un conjunto de valores del parámetro p entre el 1% y 10% del rango de distancias.

Los resultados obtenidos indican que el porcentaje óptimo del rango total de distancias se encuentra entre el 3% y 6%, dependiendo de si se utiliza la distancia absoluta o euclídea para realizar las predicciones utilizando un modelo de Media Ponderada.

Aunque el error obtenido es inferior al método de selección en que se definía la distancia umbral como un valor fijo, los mejores valores de error son mayores a los errores obtenidos por el método de selección de un porcentaje de datos o de M vecinos más cercanos.

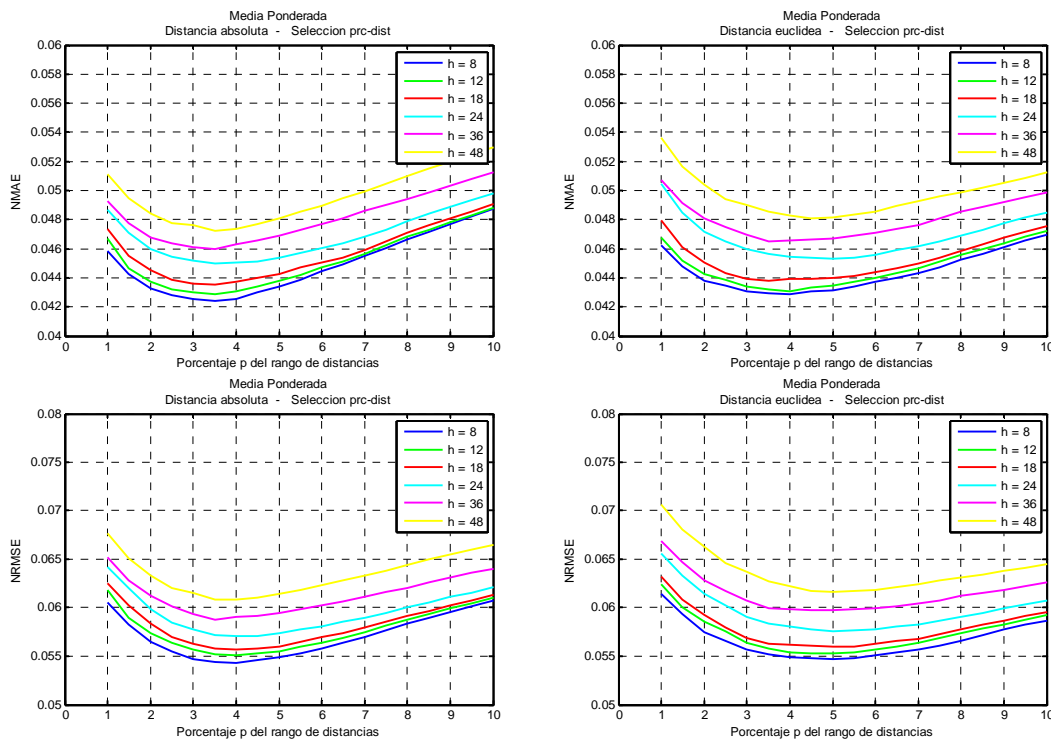


Figura 5.40: Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias con modelo de Media Ponderada

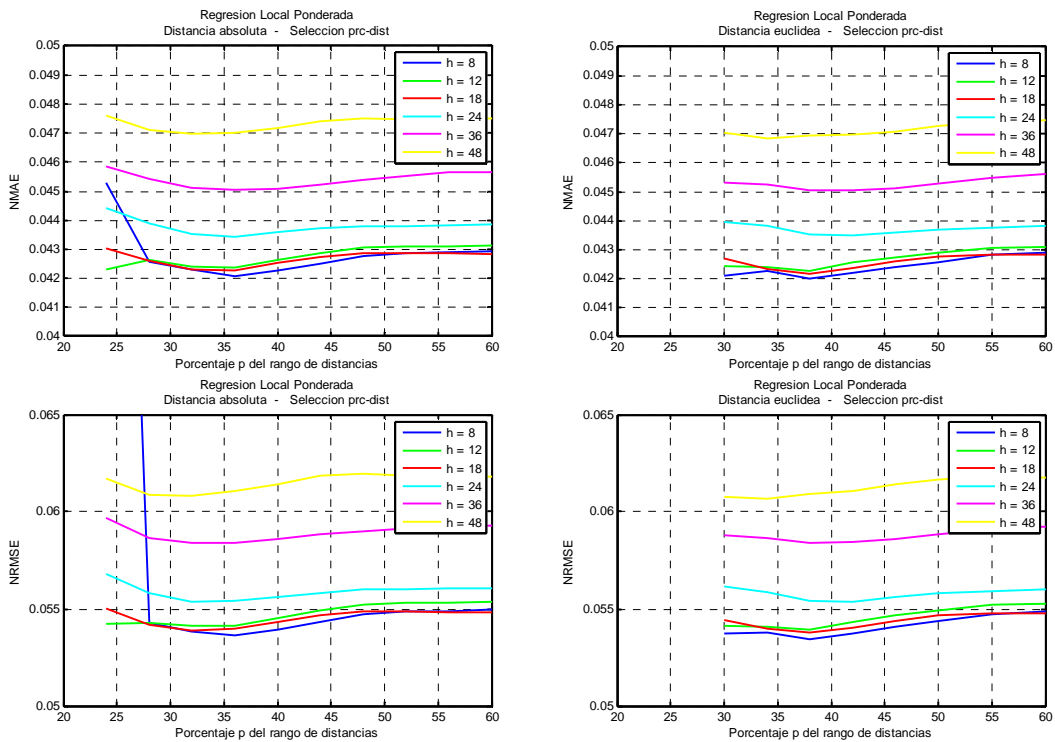


Figura 5.41: Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias con modelo de Regresión Local Ponderada

Los errores obtenidos dependiendo del porcentaje del rango de distancias seleccionado se muestran en las curvas de la Figura 5.41, con un modelo de predicción basado en una Regresión Local Ponderada. En ellas se observa que es necesario aproximadamente entre un 30 y un 40 % de ancho de banda de selección respecto de todo el rango de distancias observado, no pudiéndose calcular las predicciones con porcentajes de distancia inferiores a un 20% por insuficiencia de datos. En cualquier caso, el error mínimo obtenido es también, al igual que ocurría al utilizar un modelo de Media Ponderada, sensiblemente superior al obtenido con otros métodos de selección.

5.4.5 Datos a una distancia inferior a un porcentaje fijo del rango de distancias hasta la mediana

A continuación se representan los errores del método alternativo, en el que el rango de distancias se define hasta la mediana de los datos para evitar vectores de viento a distancias atípicamente elevadas que dan lugar a rangos de distancia muy amplios.

Con un modelo de predicción de Media Ponderada se han representado las curvas de error para un conjunto de valores del parámetro p entre el 2% y el 40% del rango de distancias hasta la mediana, obteniéndose unos resultados que indican que el porcentaje óptimo se encuentra entre el 15% y el 20% del rango de distancias hasta la mediana, dependiendo del tipo de error representado y del tipo de medida de distancia calculada, (Figura 5.42).

Se observa que, para todos los horizontes representados, el método de selección que solo usa los datos hasta la mediana tiene errores menores que el que usa todo el rango de distancias. Los valores mínimos de error calculados por este método son muy similares a los mínimos errores que se obtuvieron utilizando el método de selección de M vectores más cercanos o de un porcentaje de datos históricos.

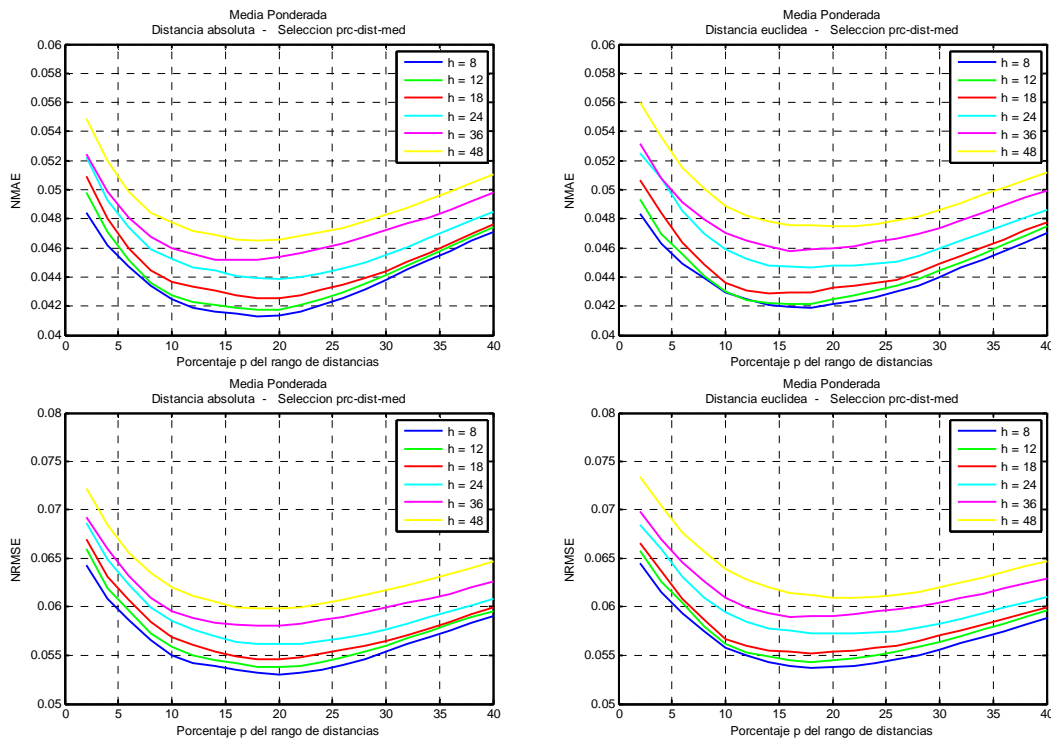


Figura 5.42: Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias hasta la mediana con modelo de Media Ponderada

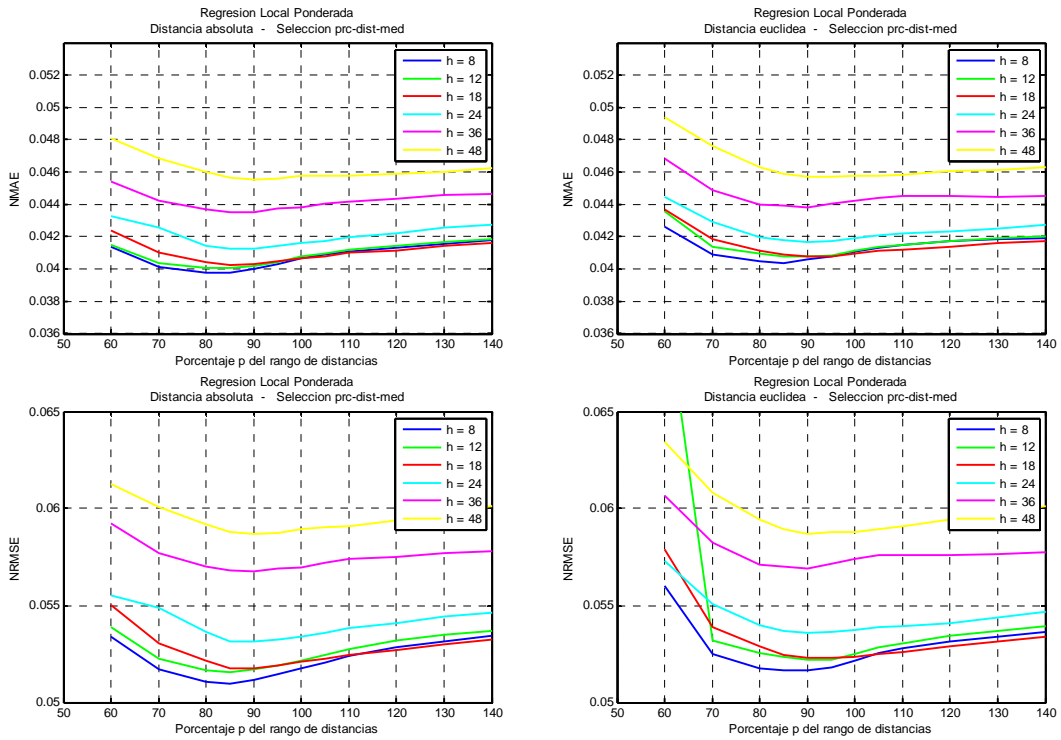


Figura 5.43: Errores NMAE y NRMSE para el método de selección de umbral de distancia variable dependiente del rango de distancias hasta la mediana con modelo de Regresión Local Ponderada

Utilizando un modelo de Regresión Local Ponderada, los errores también disminuyen sensiblemente al seleccionar un ancho de banda relativo a la distancia de la mediana, frente a los obtenidos utilizando todo el rango de distancias calculadas. En este caso, los valores óptimos de selección se sitúan entre el 80 y 90 % de distancia hasta la mediana, (Figura 5.43).

5.4.6 Método basado en la tangente a los datos ordenados por distancia

Por último se representan los resultados obtenidos utilizando el método propuesto en la sección 4.5.6 como método de selección, para un rango del parámetro de porcentaje de pendiente p entre el 10% y el 100% de la pendiente de referencia definida por la recta que une el origen de coordenadas y la mediana, usando un modelo de Media Ponderada (Figura 5.44).

Los valores óptimos del parámetro p están en el entorno del 20% al 30%, dependiendo del horizonte de predicción y de si la distancia utilizada es la absoluta o euclídea. El error mínimo observado para este método es similar al que se obtuvo utilizando los otros métodos que dieron mejores resultados.

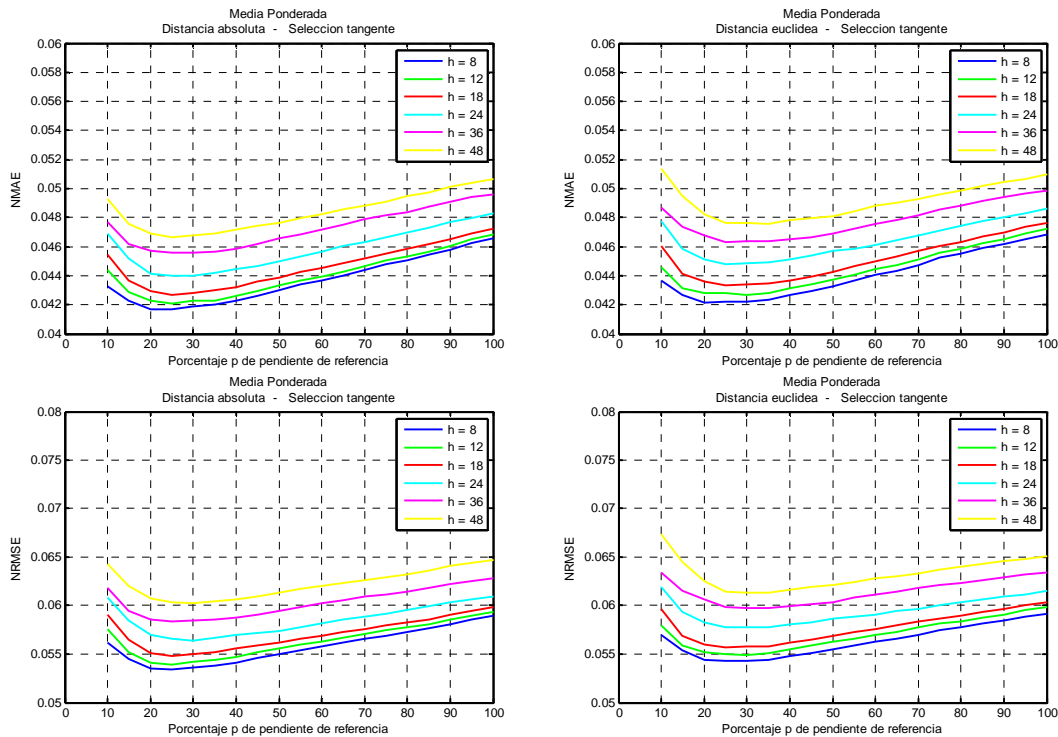


Figura 5.44: Errores NMAE y NRMSE para el método de selección basado en la tangente con modelo de Media Ponderada

Los errores obtenidos en función del porcentaje de pendiente cuando se emplea un modelo de Regresión Local Ponderada se representan en la Figura 5.45. En este caso, la pendiente óptima está entre 3 y 4 veces la pendiente de referencia, de modo que el número de datos seleccionados es mucho mayor que cuando se utiliza la Media Ponderada para calcular la predicción.

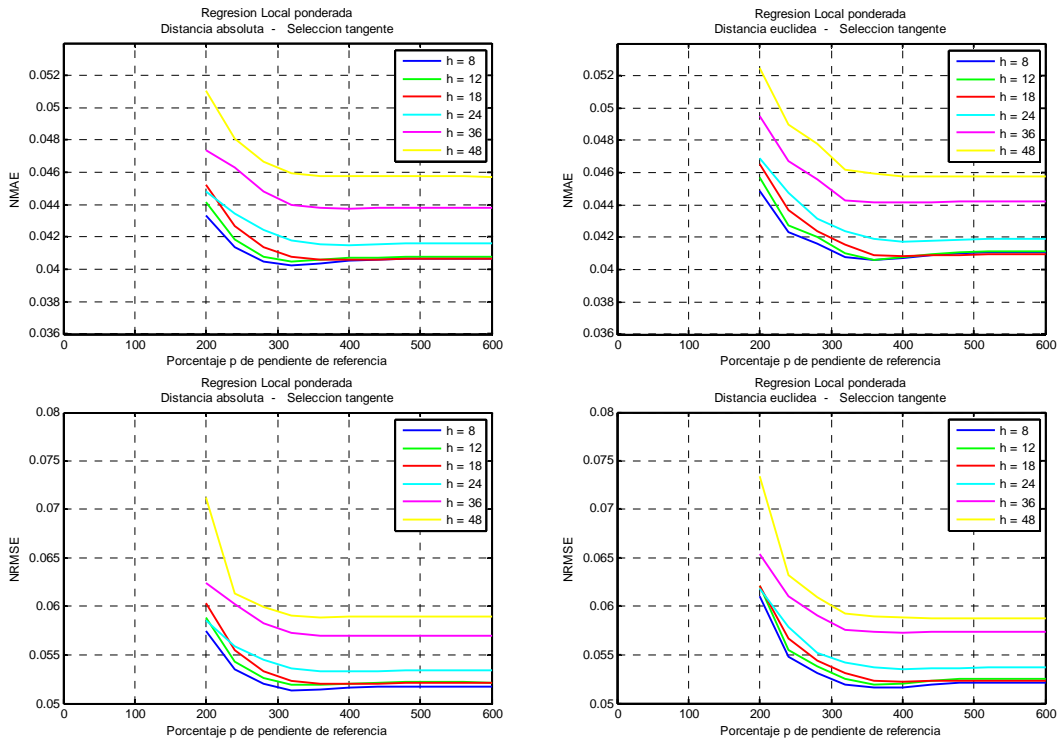


Figura 5.45: Errores NMAE y NRMSE para el método de selección basado en la tangente con modelo de Regresión Local Ponderada

5.4.7 Resumen de la evaluación de métodos de selección

A modo de resumen, se representarán a continuación los menores errores que fueron obtenidos con cada uno de los métodos de selección evaluados, usando en cada uno de ellos, para cada horizonte, el valor óptimo de los parámetros de selección de los que dependen. Se representan los errores de las predicciones calculadas entre los horizontes 5 a 48 horas, usando como medidas de distancia la distancia absoluta y distancia euclídea.

En primer lugar se representan los errores obtenidos utilizando como método de predicción un modelo local de Media Ponderada (Figura 5.46). Se puede apreciar que tanto comparando los errores NMAE como los errores NRMSE, el método de selección que proporciona peores resultados es el de ancho de banda fijo, seguido de la selección por medio de ancho de banda variable determinado por un porcentaje del rango de distancias. Mientras, los métodos de selección con mejores resultados son el de M vecinos más cercanos, el porcentaje de datos más cercanos, datos bajo una distancia umbral determinada por un porcentaje de distancias hasta la mediana, y en menor medida el método basado en la tangente.

Métodos de Predicción de la Generación Agregada de Energía Eólica

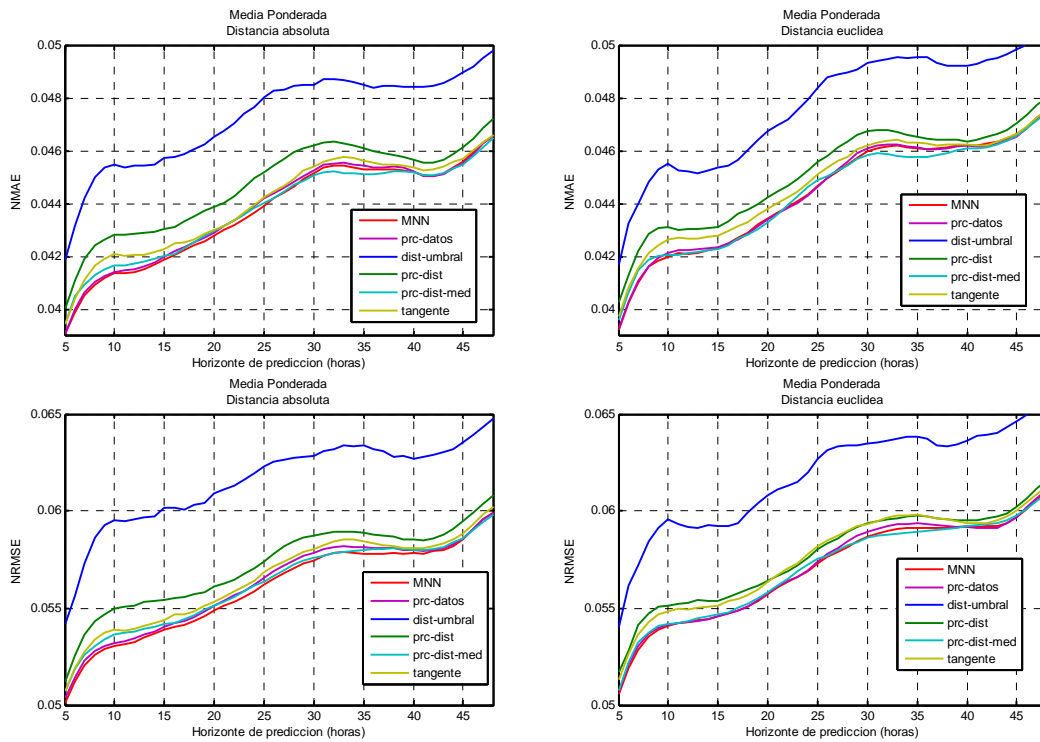


Figura 5.46: Comparación de errores NMAE y NRMSE para todos los métodos de selección con un modelo de Media Ponderada

Si se representan los errores de las predicciones calculadas usando un modelo de Regresión Local Ponderada (Figura 5.47), que han sido calculados para los horizontes $h= 4, 8, 12, 18, 24, 36$ y 48 horas, de nuevo se aprecia que los métodos de selección que proporcionan mejores resultados son los de M vecinos más cercanos, el porcentaje de datos más cercanos, y el ancho de banda variable determinado por un porcentaje fijo del rango de distancias hasta la mediana.

Evaluación de los Métodos de Suavizado Propuestos

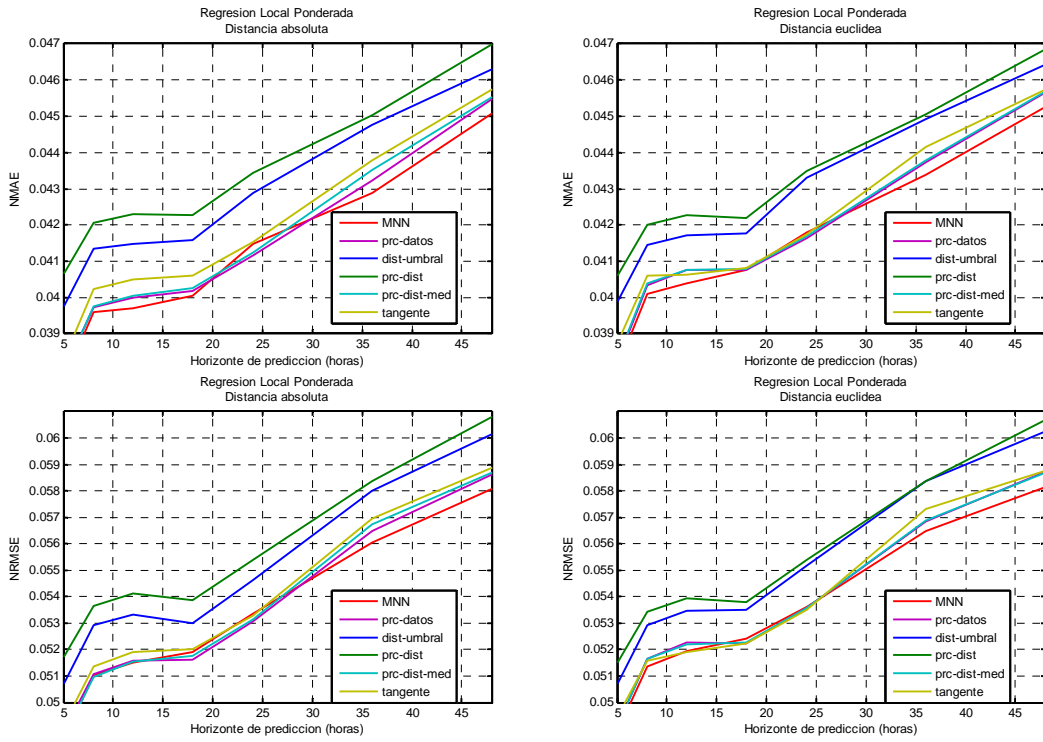


Figura 5.47: Comparación de errores NMAE y NRMSE para todos los métodos de selección con un modelo de Regresión Local Ponderada

5.5 Evaluación de las medidas de distancia

Se realizará a continuación una evaluación de las distintas medidas propuestas para el cálculo de las distancias entre el nuevo vector v_{t+h} de predicciones de viento en las coordenadas y los vectores de viento históricos v_i almacenados.

Los datos utilizados para la evaluación han sido de nuevo los datos de viento y potencia media horaria medida durante el año 2007, estableciendo como periodo inicial histórico los primeros seis meses del año, y como periodo de cálculo de predicciones la segunda mitad del año.

Por simplicidad, el análisis se ha realizado únicamente para dos de los métodos de selección de datos cercanos que mostraron un buen funcionamiento en el análisis anterior. Así, se han elegido los métodos de selección de un porcentaje fijo de datos y el método de selección de datos a una distancia inferior a un porcentaje fijo del rango de distancias hasta la mediana.

La evaluación se ha realizado calculando el NMAE y NRMSE de las predicciones calculadas. Los horizontes de predicción evaluados comprenden desde 1 a 48 horas en el caso predicción con modelo de Media Ponderada, y los horizontes $h = 4, 8, 12, 18, 24, 36$ y 48 horas para las predicciones calculadas con un modelo de Regresión Local Ponderada.

En la Figura 5.48 se representan los gráficos de error NMAE y NRMSE en función del horizonte de predicción, para cada uno de los métodos de cálculo de distancia usando una Media Ponderada y un porcentaje de datos como método de selección. Para cada horizonte de predicción se ha utilizado el valor óptimo del parámetro de porcentaje p que resulta en un menor error de predicción.

En estas curvas iniciales, en las que se ha utilizado el método de selección un porcentaje fijo de datos cercanos, ya puede concluirse que los métodos de cálculo de distancia de velocidad media, cúbica, cúbica estandarizada y Mahalanobis son claramente peores que el resto de métodos de cálculo de distancias propuestos.

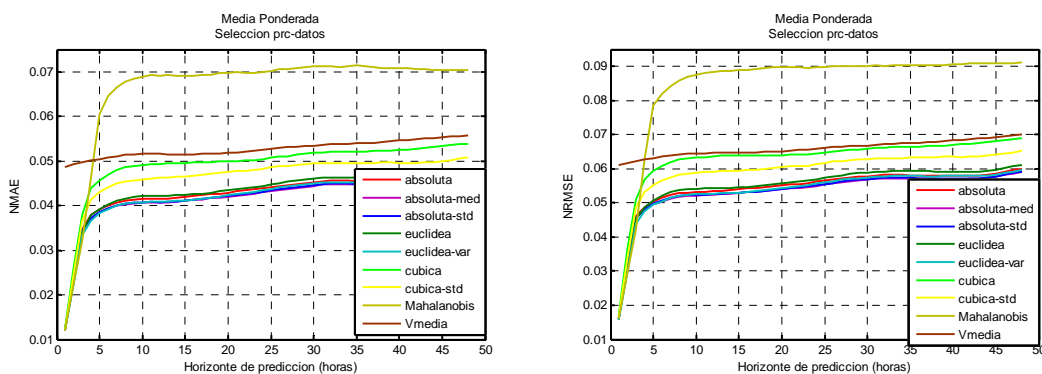


Figura 5.48: Evaluación inicial de las medidas de distancia

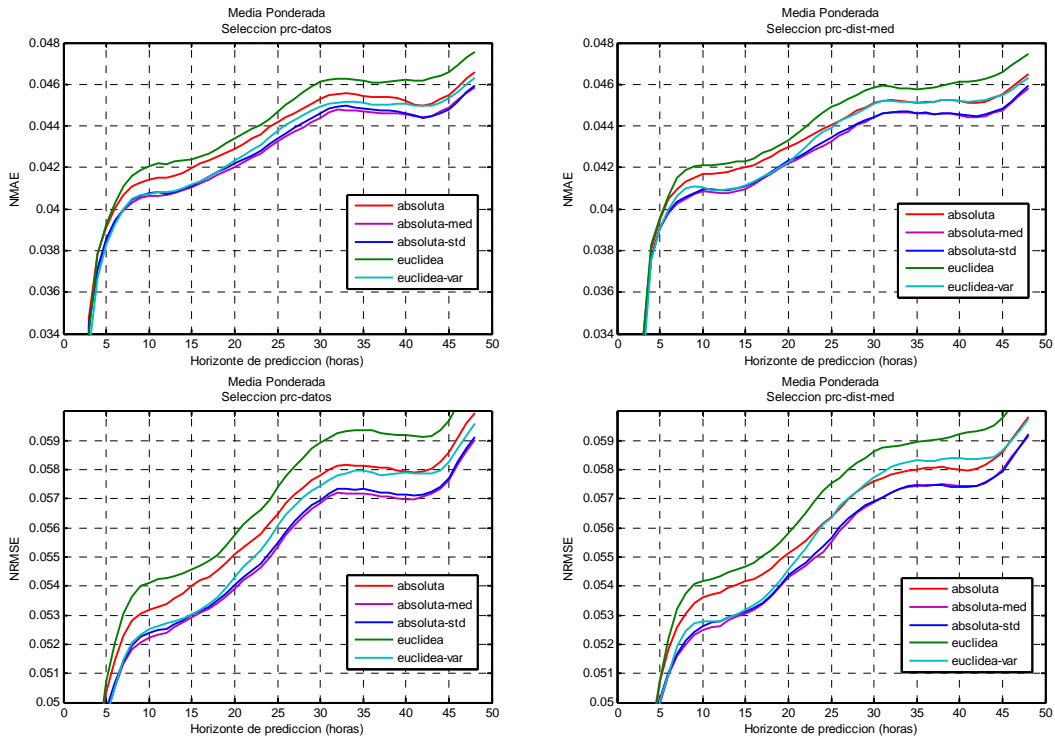


Figura 5.49: Evaluación inicial de las medidas de distancia para selección de porcentaje de datos y umbral como porcentaje de distancia hasta la mediana, con modelo de Media Ponderada

En la Figura 5.49 se han representado con mayor detalle los errores NMAE y NRMSE obtenidos con los métodos de cálculo de distancia basados en distancia absoluta y en distancia euclídea, tanto si se utiliza el método de selección de un porcentaje fijo de datos, como el método de selección de datos bajo un porcentaje fijo del rango de distancias hasta la mediana.

También se representan en la Figura 5.50 los errores de predicción utilizando un modelo de Regresión Local Ponderada, para las medidas de distancia basadas en distancia absoluta y distancia euclídea.

Métodos de Predicción de la Generación Agregada de Energía Eólica

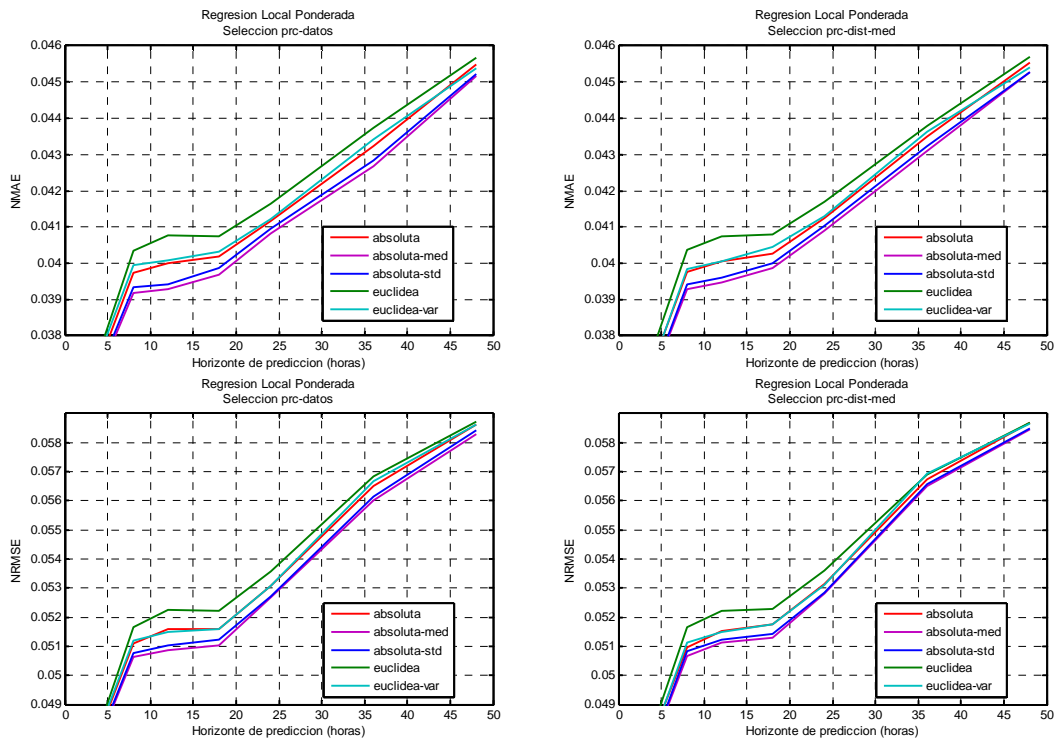


Figura 5.50: Evaluación inicial de las medidas de distancia para selección de porcentaje de datos y umbral como porcentaje de distancia hasta la mediana, con modelo de Regresión Local Ponderada

De entre los dos métodos de cálculo basados en la distancia euclídea, se observa un mejor comportamiento, para todos los horizontes, de la distancia euclídea estandarizada por la varianza, frente a la distancia euclídea sin estandarizar, independientemente de si se utiliza un modelo de Media Ponderada o de Regresión Local Ponderada, o de si la selección se hace con un porcentaje de datos o con un ancho de banda como porcentaje de distancia hasta la mediana.

De igual modo, de entre los métodos de cálculo de distancia basados en la distancia absoluta, los errores obtenidos con las distancias absolutas estandarizadas por la media y la desviación típica son menores, para todos los horizontes, que el error correspondiente a la distancia absoluta sin estandarizar, independientemente del modelo utilizado.

El mejor funcionamiento de los métodos de cálculo de distancia con estandarización ocurre tanto para el error NMAE como NRMSE, y también para los dos métodos de selección de datos representados. Por tanto, se podrían descartar los métodos de cálculo de distancia absoluta y euclídea sin estandarizar y continuar el análisis únicamente con los métodos de cálculo de distancia estandarizados.

Por otro lado, se puede apreciar que para los dos métodos de cálculo de distancia absoluta estandarizados, uno con la media y el otro con la desviación típica, las curvas de error son casi coincidentes. Esto podría indicar que las distancias calculadas por uno y otro método son casi equivalentes, como veremos a continuación:

La distancia absoluta estandarizada con la media viene dada por:

$$d_{i,t+h} = \frac{1}{K} \sum_{k=1}^K \left| \frac{v_{k,i} - v_{k,t+h}}{\bar{v}_k} \right| \quad (5.1)$$

Mientras, la distancia absoluta estandarizada con la desviación típica se calculará como:

$$d_{i,t+h} = \frac{1}{K} \sum_{k=1}^K \left| \frac{v_{k,i} - v_{k,t+h}}{s_k} \right| \quad (5.2)$$

donde:

\bar{v}_k = media de la velocidad de viento histórico registrado en la coordenada k

$$\bar{v}_k = \frac{1}{n} \sum_{i=1}^n v_{k,i} \quad (5.3)$$

s_k = desviación típica de la velocidad de viento en la coordenada k

$$s_k^2 = \frac{1}{n-1} \sum_{i=1}^n (v_{k,i} - \bar{v}_k)^2 \quad (5.4)$$

Así, para cada una de las coordenadas k donde se tengan predicciones de viento, la relación existente entre estandarizar con la media o estandarizar con la desviación típica vendrá dada por el Coeficiente de Variación, definido para cada coordenada k como la relación entre la desviación típica de las velocidades de viento y el valor absoluto de la media de velocidad de viento en la coordenada.

$$CV = \frac{s_k}{|\bar{v}_k|} \quad (5.5)$$

De este modo, los dos métodos de cálculo de distancia estandarizados serán equivalentes si el Coeficiente de Variación es semejante en todas las coordenadas $k=1, \dots, K$, donde se tienen predicciones de viento.

Para comprobar esto último se ha calculado el coeficiente de variación en las 83 coordenadas para las que se tienen datos de viento durante todo el año 2007.

Métodos de Predicción de la Generación Agregada de Energía Eólica

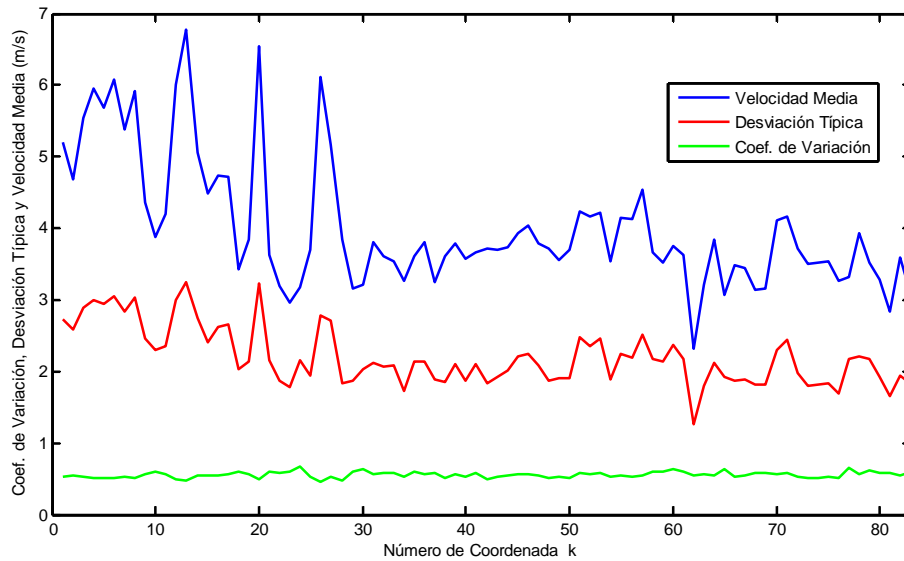


Figura 5.51: Coeficiente de Variación del viento en las coordenadas

En la Figura 5.51 se representa la velocidad media, desviación típica, y coeficiente de variación calculados según (5.3), (5.4) y (5.5). Se puede observar que para las coordenadas en las que la velocidad media es elevada, la desviación típica también resulta elevada, y viceversa. De modo que el coeficiente de variación CV resulta en un valor casi constante para todas las coordenadas.

Esto indica que resulta prácticamente equivalente estandarizar la distancia absoluta con la desviación típica o con la velocidad media en las coordenadas, de modo que se puede descartar una de las dos estandarizaciones.

Por otro lado, tal y como se muestra en las gráficas de la Figura 5.52, donde se ha representado el NMAE y NRMSE obtenidos con modelos de Media Ponderada y Regresión Local Ponderada, para el método de selección de un porcentaje de datos y para el método de selección de un porcentaje del rango de distancias hasta la mediana respectivamente, el método de distancia absoluta estandarizado con la media presenta para todos los horizontes unos resultados muy similares aunque algo mejores que el método de cálculo de distancia absoluta estandarizada con la desviación típica. Por tanto, se seleccionará como medida de distancia para el resto del análisis la distancia absoluta estandarizada con la media.

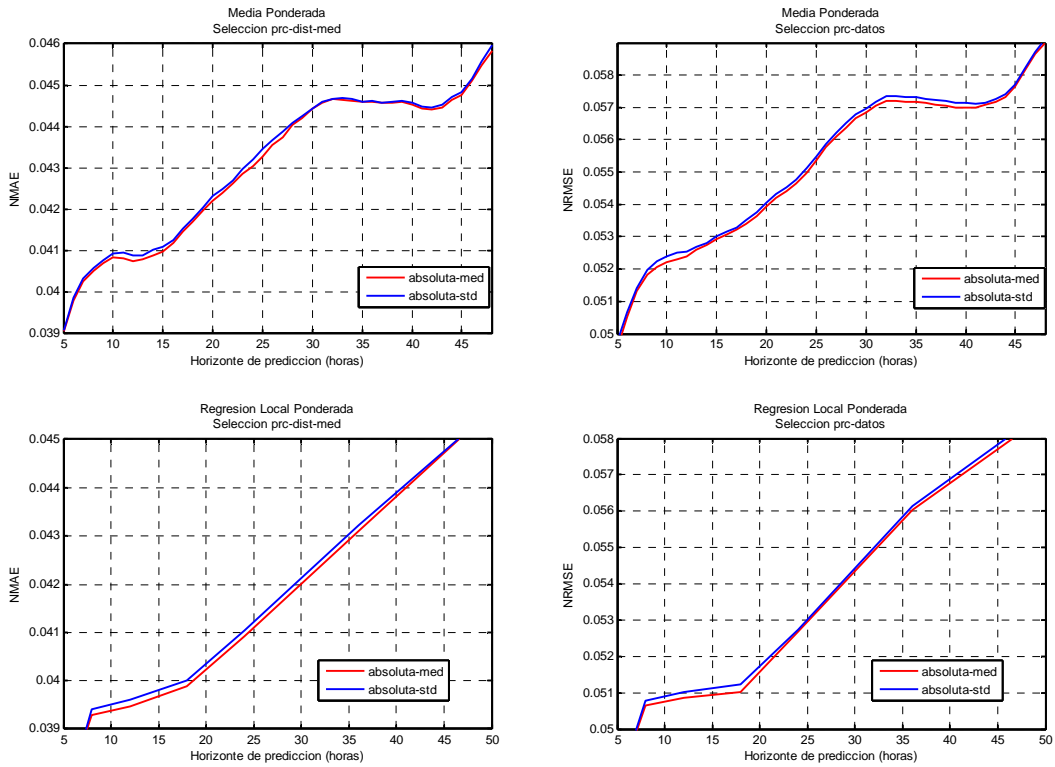


Figura 5.52: Comparación de errores para distancia absoluta estandarizada con la media y estandarizada con la varianza, con modelo de Regresión Local Ponderada

5.5.1 Resumen de la evaluación de medidas de distancia

Como conclusión, de entre todos los métodos de cálculo de distancias propuestos, se han descartado las medidas de distancia de Mahalanobis, cúbica y de velocidad media, pues presentaban unos errores mucho mayores que el resto de métodos propuestos. Además, de los métodos basados en distancias absoluta y euclídea, se ha observado mejor comportamiento con el uso de estandarización, descartando los métodos no estandarizados. Por último, se ha llegado a la conclusión de que las dos distancias absolutas estandarizadas son equivalentes, descartando una de ellas, la estandarizada con la desviación típica, por presentar unos resultados algo peores que la estandarizada con la media.

Por tanto, los métodos de cálculo de distancias que, estando basados en procedimientos distintos, han presentado mejores resultados y con los que se continuará el estudio son la distancia absoluta estandarizada con la media y la distancia euclídea estandarizada con la varianza.

5.6 Evaluación de las funciones de ponderación por distancia

Tras la selección de los elementos más cercanos al nuevo vector de velocidades de viento en base a alguna medida de distancias, la predicción de potencia para la región será una combinación de las potencias medidas en los instantes correspondientes a los vectores de viento más cercanos seleccionados. Esta predicción se podrá obtener con un modelo de Media Ponderada, según (3.3), o con un modelo de Regresión Local Ponderada, resolviendo la ecuación (3.11) de regresión.

Esta combinación tendrá en cuenta unos coeficientes o pesos, que dependerán de la distancia entre el nuevo vector de predicción de viento y cada uno de los vectores seleccionados.

En la sección 4.6 se propusieron dos métodos de ponderación por distancia, uno inversamente proporcional a la distancia, definido por (4.28), y otro de decaimiento exponencial con la distancia, definido por (4.29). Estas funciones de ponderación dependen de un parámetro α , que determina la severidad con que se tiene en cuenta de distancia de los datos seleccionados.

En la Figura 5.53 se muestra la influencia del parámetro α sobre el NMAE, para los horizontes $h = 8, 12, 18, 24, 36$ y 48 horas, utilizando los dos métodos de ponderación por distancia propuestos. Como medida de distancia entre vectores se ha utilizado la distancia absoluta estandarizada con la media y como método de selección un porcentaje fijo de datos, con $p = 0.8\%$ cuando el modelo de predicción es de Media Ponderada, y $p = 50\%$ cuando se calculan las predicciones con un modelo de Regresión Local Ponderada.

Como se puede observar, en los casos representados se consigue una disminución del error de predicción si se hace una ponderación por distancia de los datos con un valor adecuado de α . El caso de $\alpha = 0$ sería el equivalente a no hacer ponderación alguna, dando el mismo peso a todos los elementos seleccionados.

En los casos de predicciones con un modelo de Media Ponderada, representados en las gráficas superiores de la figura, los valores de α que consiguen minimizar el error NMAE se sitúan alrededor de 30 si se utiliza una ponderación de decaimiento exponencial, y cercanos a 6 si se utiliza una ponderación inversa a la distancia.

Cuando las predicciones se calculan con un modelo de Regresión Local Ponderada, representado en las gráficas inferiores de la figura, los menores errores se obtienen con valores de α entre 8 y 10 si se utiliza una función de decaimiento exponencial con la distancia, y valores entre 3 y 4 si se hace una ponderación inversa a la distancia.

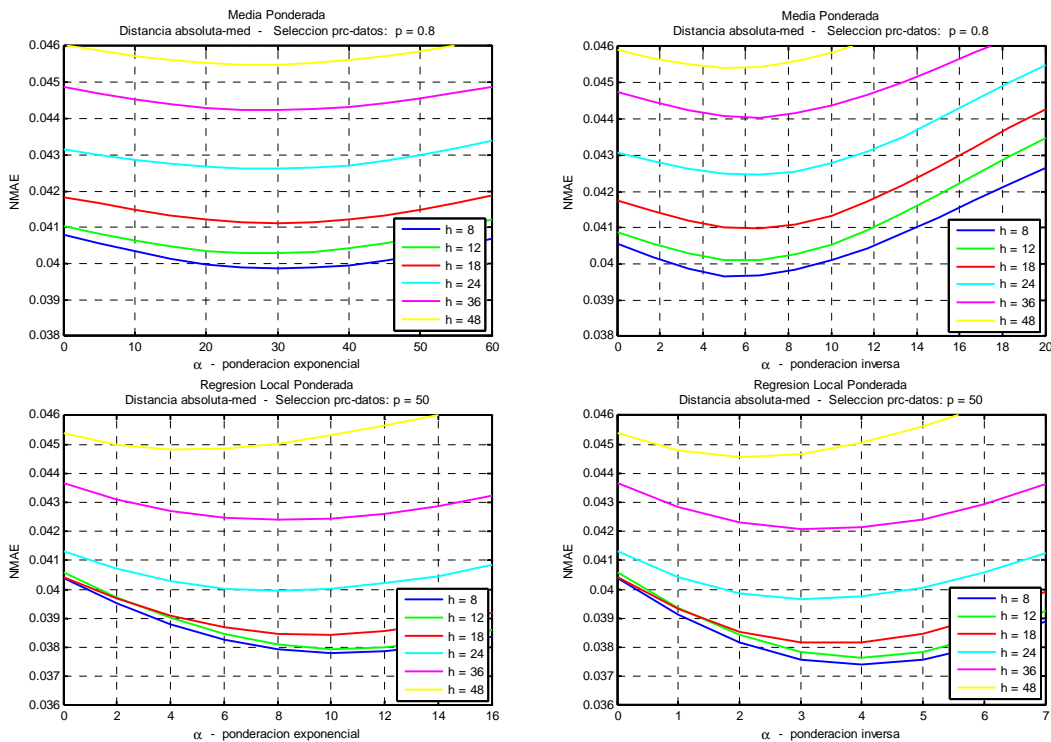


Figura 5.53: Errores NMAE para las funciones de ponderación inversa y de decaimiento exponencial con la distancia, en función del parámetro α

En cualquier caso, el valor de α que minimiza el error depende del horizonte de predicción, teniendo valores ligeramente inferiores a medida que aumenta el horizonte. Esto se debe a que para horizontes elevados las predicciones de velocidad de viento son menos precisas, y por tanto la predicción de potencia requerirá un mayor suavizado, equivalente a utilizar más datos en la ponderación y por tanto a que la función de ponderación sea menos severa con la distancia.

Aunque en los casos representados se ha fijado el porcentaje de datos seleccionados a ponderar en $p = 0.8\%$ y $p = 50\%$ cuando se utiliza un modelo de Media Ponderada y Regresión Local Ponderada respectivamente, se debe tener en cuenta que los valores adecuados del parámetro α de la función de ponderación dependerán del porcentaje de datos seleccionados.

En la superficie de la Figura 5.54 se ha representado el error NMAE para un horizonte de predicción $h = 24$ horas, en función del porcentaje p de datos seleccionados y del parámetro α para una ponderación de decaimiento exponencial con la distancia, utilizando un modelo de Media Ponderada. Se puede observar que existe una combinación óptima de los parámetros que minimiza el error NMAE de predicción, en este caso con p alrededor del 1% de los datos y α en torno a 30.

Métodos de Predicción de la Generación Agregada de Energía Eólica

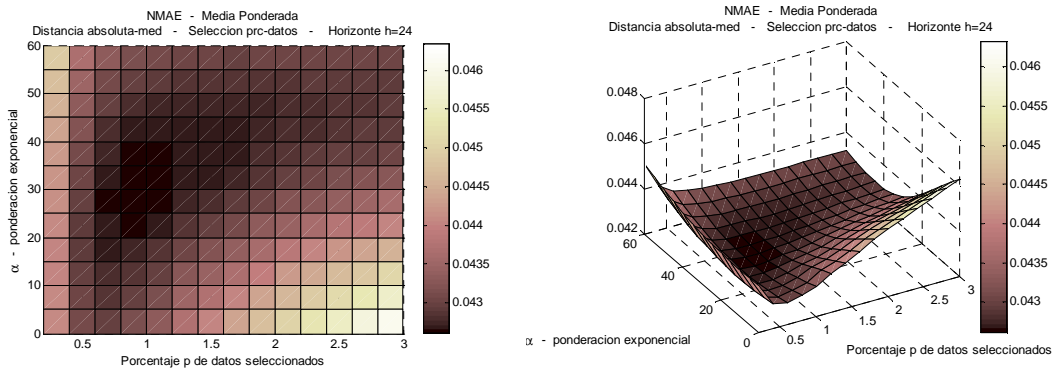


Figura 5.54: Error NMAE para la función de ponderación de decaimiento exponencial con la distancia, en función de los parámetros p y α , con modelo de Media Ponderada

Del mismo modo, en la superficie de la Figura 5.55 se ha representado el error NMAE para un horizonte $h = 24$ horas en función de los parámetros p y α para una ponderación inversa a la distancia, utilizando un modelo de predicción de Media Ponderada. En este caso la combinación óptima de los parámetros que minimiza el error de predicción se encuentra con p entre el 1% y 2% de los datos y α entre 7 y 10.

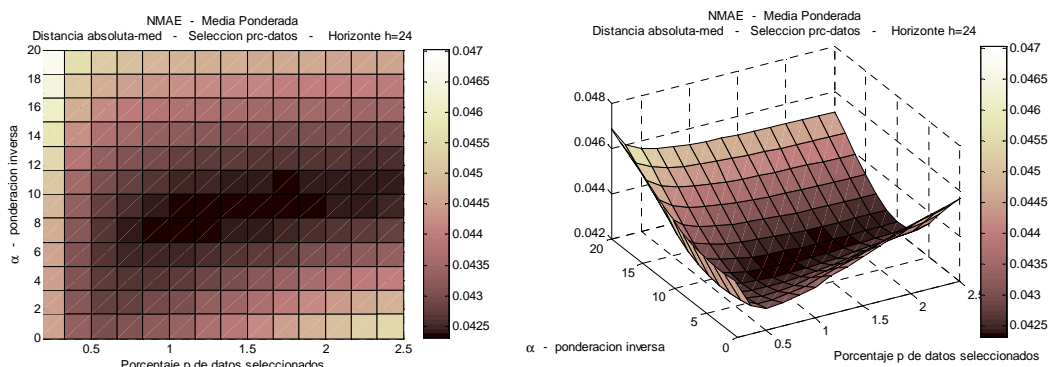


Figura 5.55: Error NMAE para la función de ponderación inversa con la distancia, en función de los parámetros p y α , con modelo de Media Ponderada

De un modo similar, se representan en las siguientes figuras las superficies de error en función de los parámetros p y α para una ponderación con decaimiento exponencial (Figura 5.56) e inversa a la distancia (Figura 5.57), utilizando un modelo de Regresión Local Ponderada para calcular las predicciones.

En ambos casos se observa también una combinación óptima de parámetros que minimiza el error, que se sitúa en porcentajes de datos de aproximadamente un 45 o 55%, y valores de α entre 8 y 10 para funciones de ponderación con decaimiento exponencial, y entre 3 y 4 para funciones de ponderación inversa a la distancia.

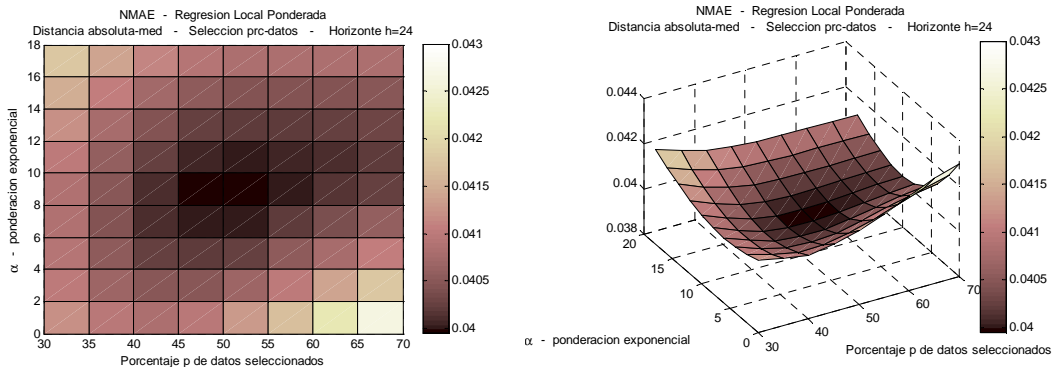


Figura 5.56: Error NMAE para la función de ponderación de decaimiento exponencial con la distancia, en función de los parámetros p y α , con modelo de Regresión Local Ponderada

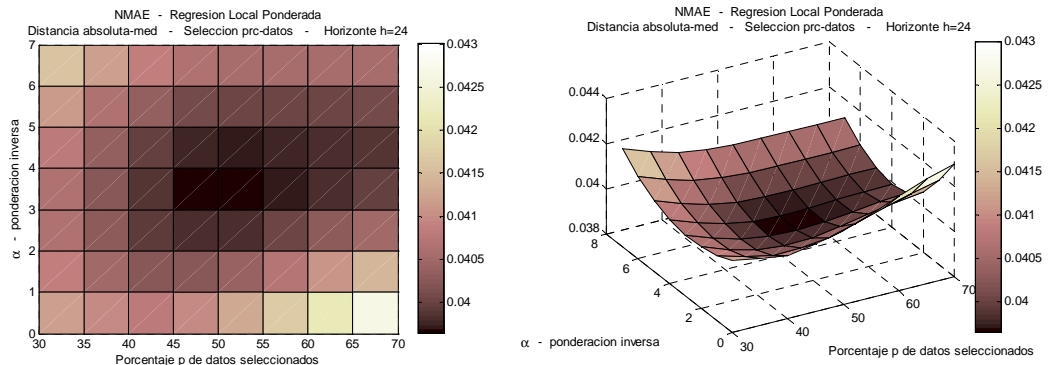


Figura 5.57: Error NMAE para la función de ponderación inversa con la distancia, en función de los parámetros p y α , con modelo de Regresión Local Ponderada

5.6.1 Comparación de las funciones de ponderación propuestas

A continuación se realizará una evaluación y comparación de los dos métodos propuestos de ponderación por distancias para todos los horizontes. En la Figura 5.58 se representa el error NMAE para los métodos de medida de distancia absoluta estandarizada con la media y euclídea estandarizada con la varianza, utilizando como método de selección un porcentaje p de datos. Se ha seleccionado, para cada horizonte, la mejor combinación del parámetro α de la función de ponderación y del parámetro de porcentaje p de datos seleccionados que minimiza el error de predicción (Figura 5.59), utilizando un modelo de Media Ponderada para calcular las predicciones.

Métodos de Predicción de la Generación Agregada de Energía Eólica

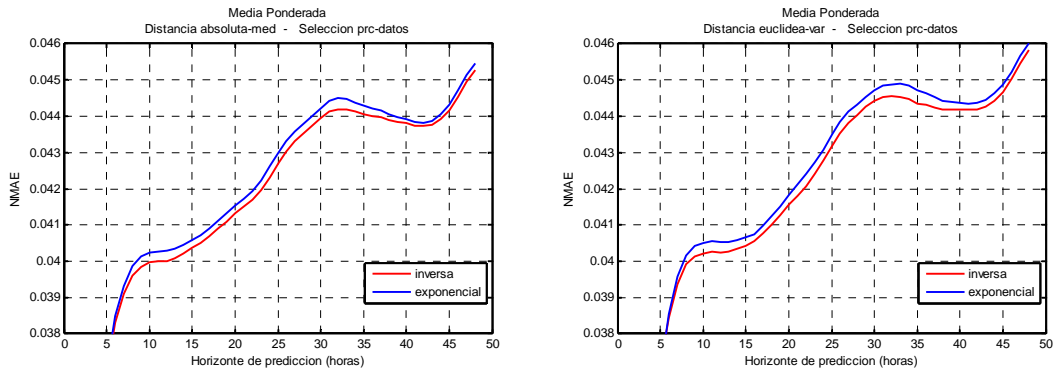


Figura 5.58: Comparación de errores para ponderación inversa y decaimiento exponencial con la distancia, con método de selección de un porcentaje de datos y modelo de Media Ponderada

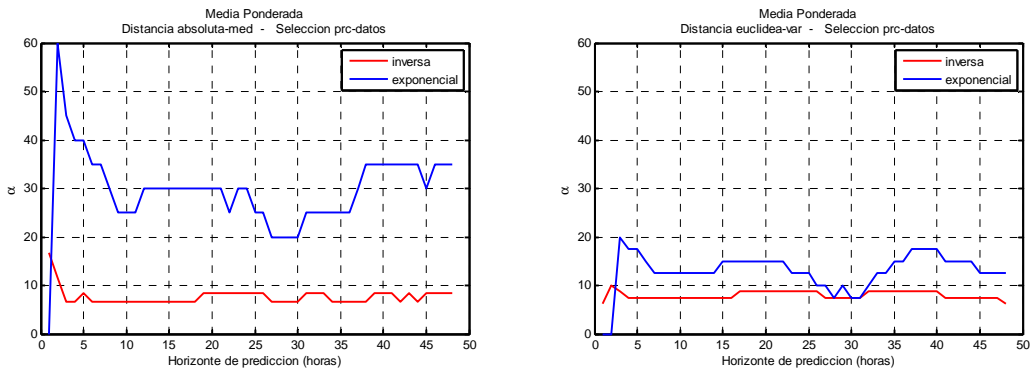


Figura 5.59: Parámetros α con los que se obtiene el menor error NMAE, con método de selección de un porcentaje de datos y modelo de Media Ponderada

También se representa en la Figura 5.60 el error NMAE, utilizando en esta ocasión como método de selección un porcentaje fijo del rango de distancias hasta la mediana, seleccionando en cada horizonte la mejor combinación de valores de los parámetros p y α (Figura 5.61), y utilizando también un modelo de predicción de Media Ponderada.

Evaluación de los Métodos de Suavizado Propuestos

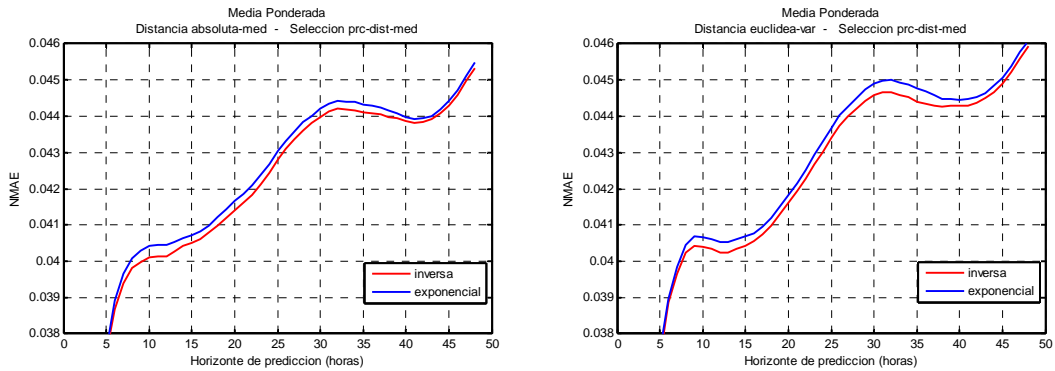


Figura 5.60: Comparación de errores para ponderación inversa y decaimiento exponencial con la distancia, con método de selección de un porcentaje de distancia hasta la mediana y modelo de media Ponderada

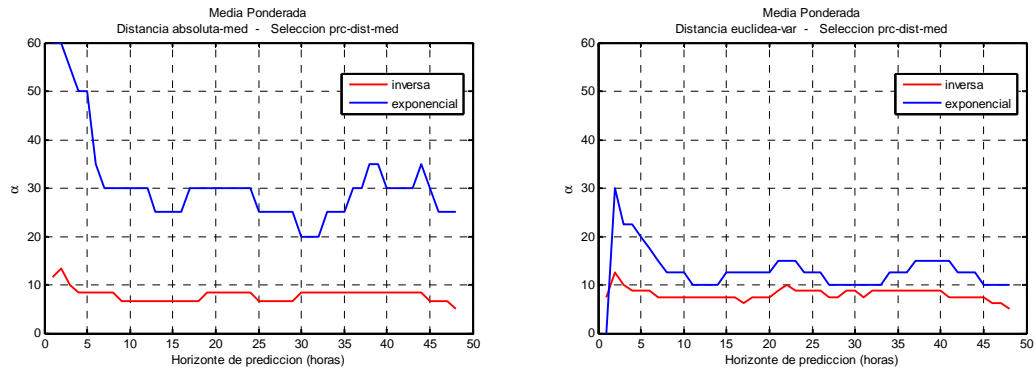


Figura 5.61: Parámetros α con los que se obtiene el menor error NMAE, con método de selección de un porcentaje de distancia hasta la mediana y modelo de Media Ponderada

Se puede observar en la Figura 5.58 y Figura 5.60 que el método de ponderación que da un peso inversamente proporcional a la distancia entre vectores presenta, en todos los casos y para todos los horizontes de predicción, unos resultados mejores que el método de ponderación de decaimiento exponencial con la distancia.

Esto también se cumple si se utiliza un modelo de Regresión Local Ponderada para calcular la predicción, como puede comprobarse en la Figura 5.62, donde se han representado los errores NMAE y NRMSE en función del horizonte de predicción, para un modelo de predicción que utiliza como medida de distancia la distancia absoluta estandarizada con la media, y un método de selección de porcentaje de datos. Los parámetros p y α que proporcionan estos errores mínimos se muestran en la Figura 5.63. En ambos casos, la función de ponderación inversa proporciona mejores resultados que la función con decaimiento exponencial a la distancia.

Métodos de Predicción de la Generación Agregada de Energía Eólica

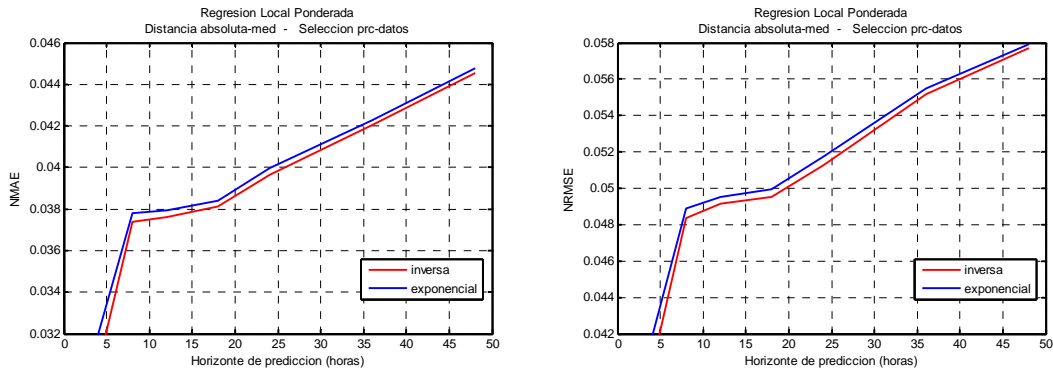


Figura 5.62: Comparación de errores para ponderación inversa y decaimiento exponencial con la distancia, con método de selección de un porcentaje de datos y modelo de Regresión Local Ponderada

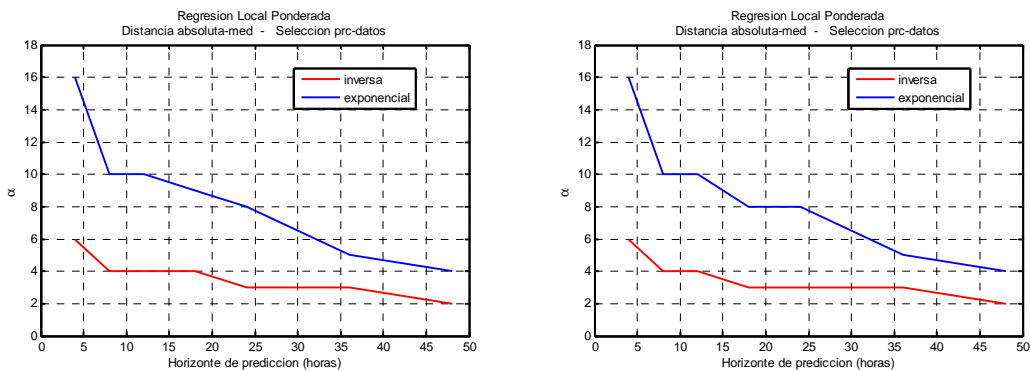


Figura 5.63: Parámetros α con los que se obtiene el menor error NMAE, con método de selección de un porcentaje de datos y modelo de Regresión Local Ponderada

En último lugar, también puede apreciarse que los valores del parámetro α que minimizan el error son inferiores conforme aumenta el horizonte de predicción. Esto de nuevo se debe a que las predicciones de viento son menos precisas a medida que aumenta el horizonte de predicción, de modo que se hace necesario un mayor suavizado, y por tanto una función de ponderación menos severa, para horizontes de predicción más elevados.

5.6.2 Corrección del parámetro α en función de la distribución de distancias

Para tener en cuenta la forma de las distribuciones de distancia calculadas y por tanto la mayor o menor concentración de datos cercanos, se hará una corrección del parámetro α en la función de ponderación, tal y como se describe en (4.30) y (4.31). La corrección se hará dividiendo, en cada instante, los valores calculados de distancias entre un valor μ representativo del orden de magnitud de dichas distancias calculadas, y por tanto de la concentración de datos cercanos.

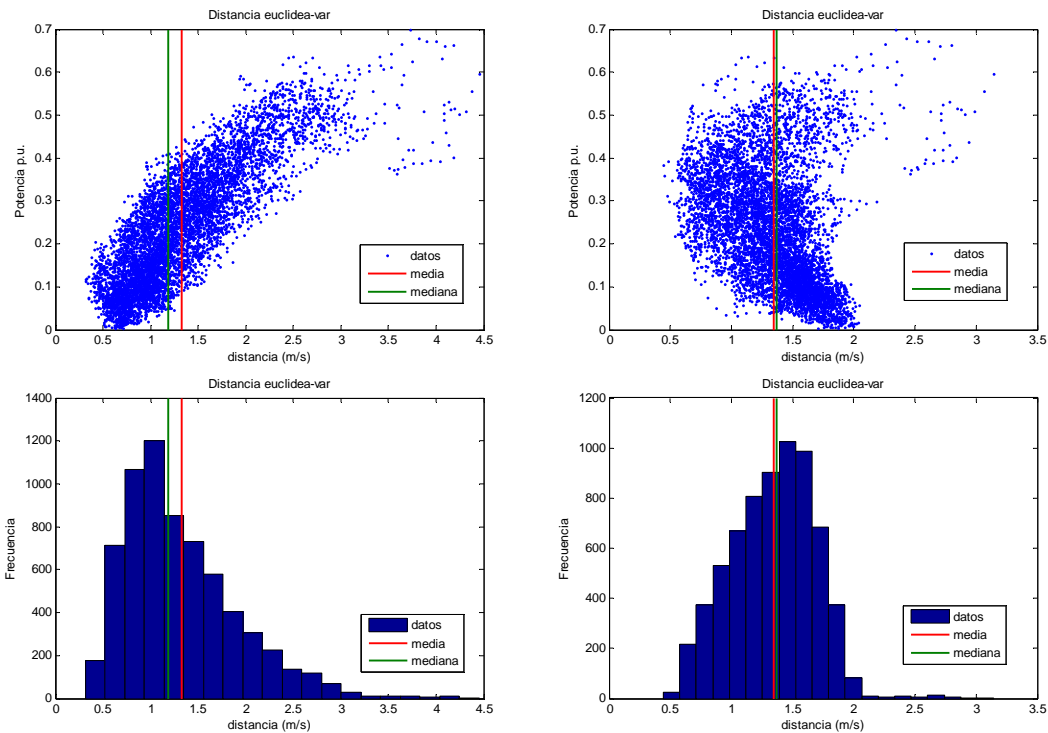


Figura 5.64: Dispersiones e histogramas, media y mediana de las distancias para situaciones de baja y alta potencia prevista.

Se probarán dos tipos de corrección: estandarización con la mediana y estandarización con la media de la distribución de distancias calculada en cada instante. Así, si se representan las distribuciones de distancias calculadas para un mismo tipo de distancia y un mismo horizonte, pero con vectores de viento v_{t+h} distintos, el valor de la media y mediana de las distancias dependerá de la forma de la distribución.

En el caso en que el valor de μ sea pequeño, que corresponde a instantes como los de las dispersiones representadas en la parte izquierda en la Figura 5.64, donde la distribución presenta una gran concentración de elementos a poca distancia, el cociente α/μ será elevado, de modo que el decaimiento exponencial será más severo, coherentemente con la gran concentración de elementos cercanos a ponderar.

Mientras, en el caso de que el valor de μ sea más elevado, que corresponde a instantes en los que la distribución presenta pocos elementos a una distancia cercana, como se observa en la parte derecha de la figura anterior, el cociente α/μ será algo más pequeño, de modo que el decaimiento exponencial será menos acentuado y se tendrán por tanto en cuenta suficientes datos históricos.

En la Figura 5.65 se representan las curvas de error NMAE resultantes de las predicciones calculadas con un modelo local de Media Ponderada y una ponderación con decaimiento exponencial por distancias, realizando la estandarización del parámetro α con la media y con la mediana (4.31). También se muestran en la Figura 5.66 los valores del parámetro α que han dado lugar a las mejores predicciones.

Métodos de Predicción de la Generación Agregada de Energía Eólica

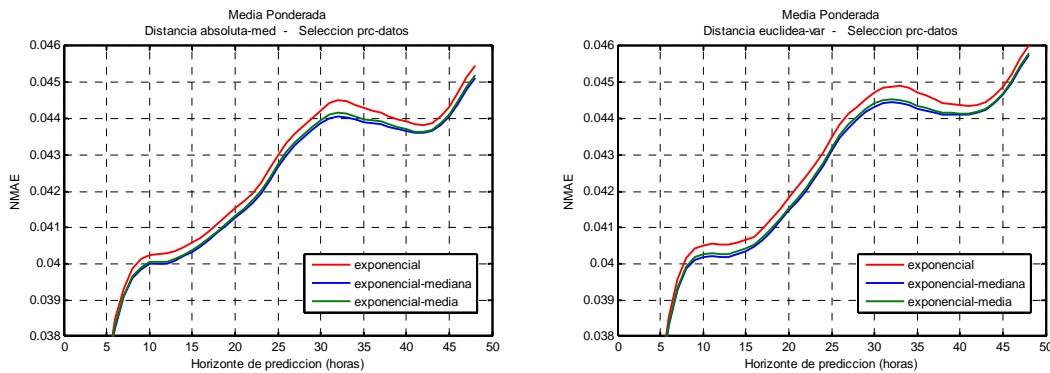


Figura 5.65: Comparación de errores para ponderación con decaimientos exponencial sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada

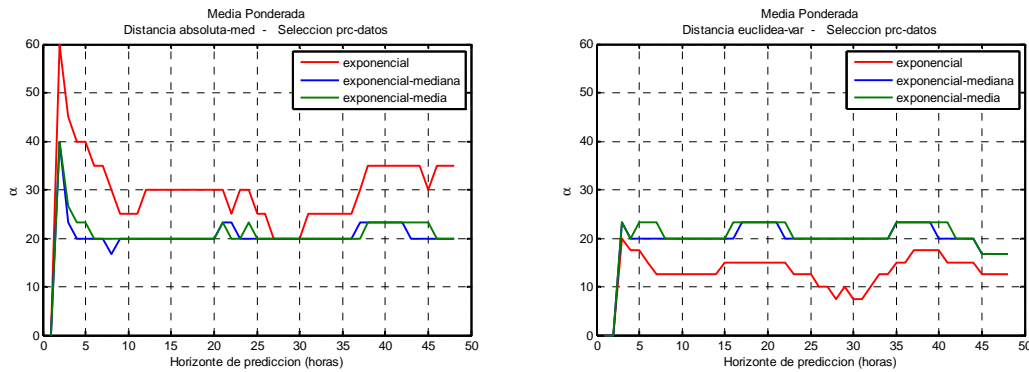


Figura 5.66: Parámetros α con los que se obtiene el menor error NMAE, para ponderación con decaimiento exponencial sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada

Se puede comprobar que los errores de predicción NMAE disminuyen significativamente si se realiza la corrección del parámetro α en cada instante utilizando la media o la mediana, para todos los horizontes, independientemente del tipo de distancia calculado.

Esta mejoría de las predicciones cuando se utiliza la estandarización con la media o la mediana se debe a que, al estandarizar en cada instante, se obtiene un valor añadido, que es el hecho de tener en cuenta la forma de las distribuciones de distancia.

Por último, se puede apreciar una leve mejoría en el error si se utiliza la mediana de la distribución de distancias para estandarizar en cada instante, frente al uso de la media. Esta mejoría se debe a que las distribuciones de distancias calculadas resultan muy asimétricas y con colas muy pesadas, como puede observarse en los histogramas representados en la Figura 5.64. Por ello, la media no es un buen indicador de la localización del centro de la distribución, pues al ser el centro de gravedad se ve muy influenciada por las observaciones que están muy alejadas. La estandarización con la mediana resulta más robusta que la estandarización con la media en estos casos.

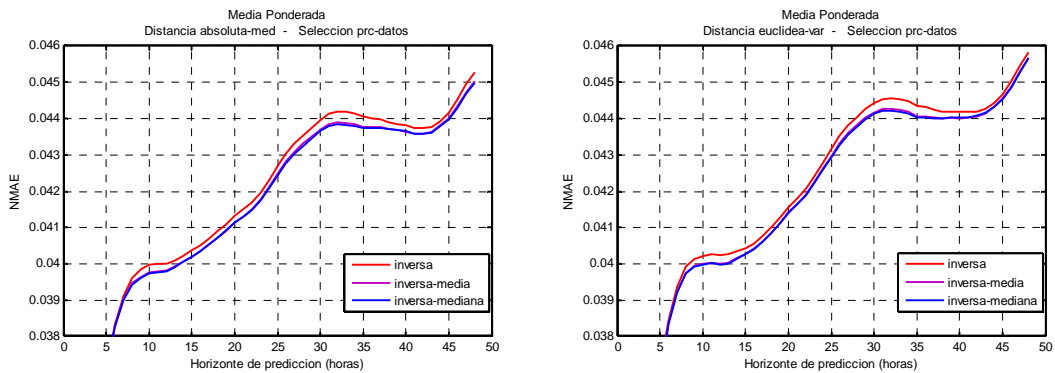


Figura 5.67: Comparación de errores para ponderación inversa sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada

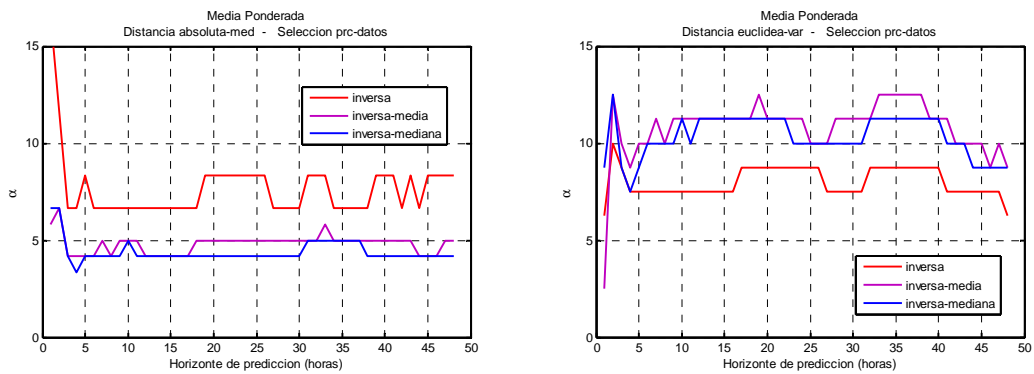


Figura 5.68 : Parámetros α con los que se obtiene el menor error NMAE, para ponderación inversa sin estandarizar y estandarizado con la media y mediana de distancias, con modelo de Media Ponderada

La ventaja observada con esta corrección del parámetro α también se puede aplicar sobre la función de ponderación inversa (4.30). De este modo, el decaimiento será más o menos severo dependiendo de si hay muchos o pocos datos a poca distancia. En la Figura 5.67 y la Figura 5.68 se representan el error NMAE y los valores del parámetro α con los que se obtienen los menores errores, utilizando la media y la mediana como corrector del parámetro α , utilizando un modelo de Media Ponderada con una función de ponderación inversa a la distancia.

De nuevo se aprecia una disminución del error de predicción si se afecta el parámetro α con la media o con la mediana. Además, en este caso también se observa que el uso de la mediana proporciona unos resultados ligeramente mejores que el uso de la media.

Las mejoras observadas con la corrección del parámetro de ponderación en función de la forma de la distribución de distancias mediante el uso de la mediana como estandarizador también se ha comprobado utilizando un modelo de Regresión Local Ponderada.

Métodos de Predicción de la Generación Agregada de Energía Eólica

En la Figura 5.69 se muestra el error NRMSE calculado para los horizontes $h = 4, 8, 12, 18, 24, 36$ y 48 horas utilizando una Regresión Local Ponderada como modelo de predicción, con medida de distancia absoluta estandarizada con la mediana, un método de selección de porcentaje de datos, y corrigiendo el parámetro α con la mediana de la distribución de distancias calculada en cada instante. También se han representado los valores del parámetro α con los que se han obtenido los menores errores para cada horizonte, (Figura 5.70).

Tanto si se utiliza una función de ponderación de decaimiento exponencial como si se pondera inversamente a la distancia, la corrección del parámetro de ponderación con la mediana μ de la distribución de distancias en cada instante consigue disminuir el error de predicción. Los valores que minimizan el error de nuevo dependen del horizonte de predicción, teniendo una ponderación más suave, es decir, con valores de α más pequeños, conforme aumenta el horizonte de predicción.

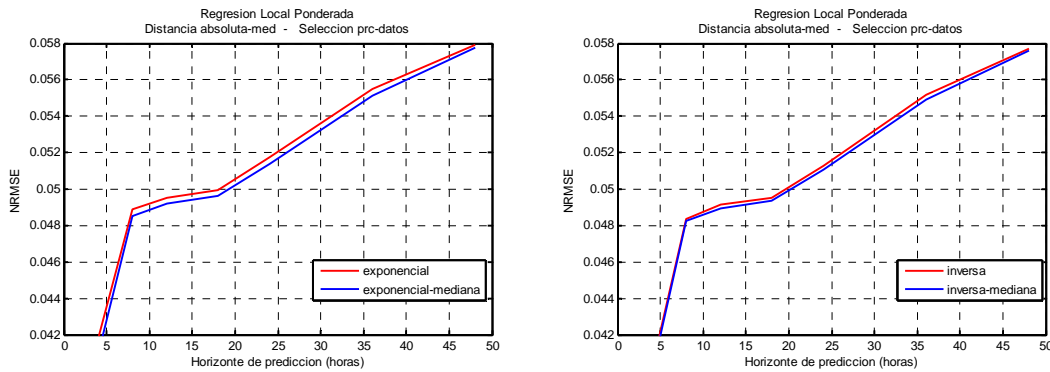


Figura 5.69: Comparación de errores para ponderación inversa sin estandarizar y estandarizada con la mediana de distancias, con modelo de Regresión Local Ponderada

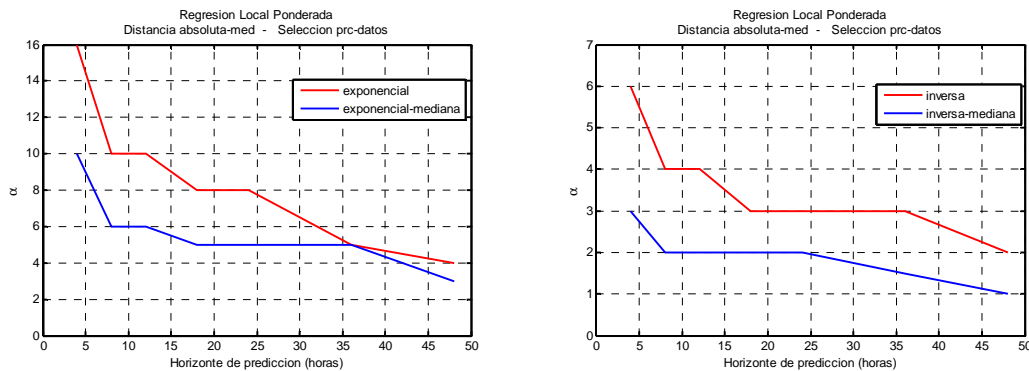


Figura 5.70: Parámetros α con los que se obtiene el menor error NRMSE, para ponderación inversa sin estandarizar y estandarizada con la mediana de distancias, con modelo de Regresión Local Ponderada

5.6.3 Comparación de las funciones de ponderación con el parámetro α corregido

Se concluirá este estudio comparativo de las funciones de ponderación por distancias comparando el método de ponderación inversa a la distancia con el parámetro α estandarizado por la mediana, y el método de ponderación de decaimiento exponencial con el parámetro α también estandarizado por la mediana, que quedaron definidos por las ecuaciones (4.30) y (4.31) respectivamente.

En la Figura 5.71 se dibujan algunos de los errores NMAE y NRMSE obtenidos con un modelo de Media Ponderada, usando ambos métodos de ponderación, utilizando la distancia absoluta estandarizada con la media y la distancia euclídea estandarizada con la varianza, tanto con el método de selección de un porcentaje fijo de datos, como seleccionando los datos por debajo de un porcentaje de distancia hasta la mediana.

Como se puede comprobar en las curvas de error representadas, en todos los casos el método de ponderación inversa con el parámetro α corregido por la mediana tiene unos errores inferiores a los errores del método de ponderación exponencial con el parámetro α corregido con la mediana.

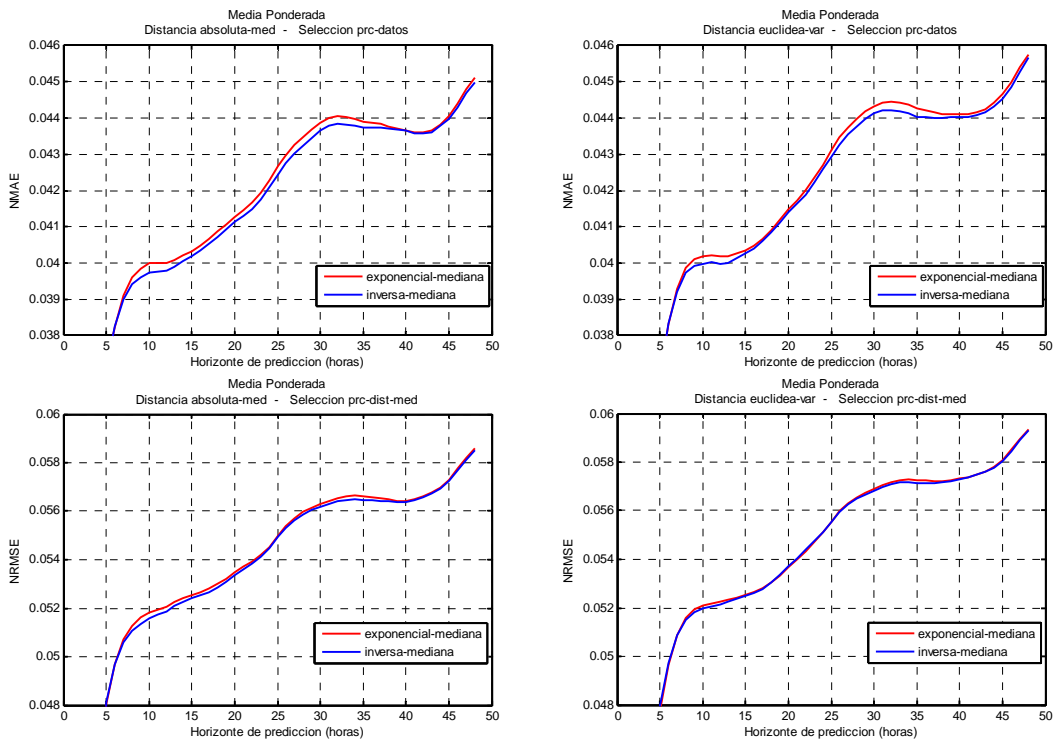


Figura 5.71: Comparación de errores NMAE y NRMSE entre ponderación inversa estandarizada con la mediana y ponderación de decaimiento exponencial estandarizado con la mediana, modelo de Media Ponderada

Métodos de Predicción de la Generación Agregada de Energía Eólica

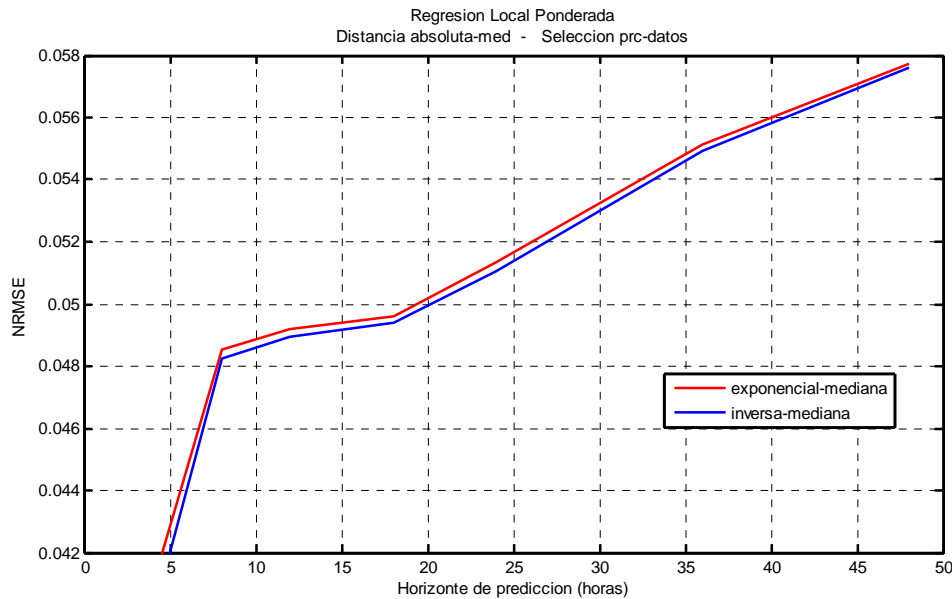


Figura 5.72: Comparación de errores NMAE y NRMSE entre ponderación inversa estandarizada con la mediana y ponderación de decaimiento exponencial estandarizado con la mediana, modelo de Regresión Local Ponderada

Esto mismo ocurre si el modelo de predicción está basado en una Regresión Local Ponderada, tal y como puede comprobarse en la Figura 5.72, en la que se ha representado el error NRMSE obtenido utilizando ambas funciones de ponderación por distancia, aplicadas a la medida de distancia absoluta estandarizada con la media, y un método de selección de un porcentaje de datos.

Por tanto, se puede concluir que el método de ponderación inversa a la distancia con el parámetro α corregido con la mediana en cada instante es el más indicado para realizar la ponderación por distancias, y será el utilizado en el resto del análisis.

5.7 Comparación entre las funciones Kernel y el método de suavizado propuesto

A continuación se hará una evaluación comparativa entre el uso de funciones Kernel en el cálculo de los coeficientes de ponderación ω_m para (3.3) y (3.11), y el uso del método propuesto, en el que se tratan de forma independiente la medida de distancia, el proceso de selección y la función de ponderación.

Para ambos casos, se representará la opción que ha proporcionado mejores resultados en el periodo de ensayo, que en el caso de suavizado por Kernel resultó ser la función Kernel triangular, y en el método de suavizado propuesto corresponde a la combinación de medida de distancia absoluta estandarizada con la media con un método de selección de un porcentaje de datos y una función de ponderación inversa a

la distancia con corrección del parámetro de ponderación en función de la mediana de la distribución de distancias.

En la Figura 5.73 se representa el NMAE y en la Figura 5.74 se muestra el NRMSE para ambas opciones de ponderación, tanto si el modelo de predicción es de Media Ponderada como si se trata de un modelo de Regresión Local Ponderada.

En cualquier caso, se puede apreciar que los errores obtenidos son inferiores con el método propuesto de utilizar de forma independiente la medida de distancias, la selección de datos y la función de ponderación que más se adecue a nuestro problema.

Además, también se puede concluir que proporciona mejores resultados un modelo de Regresión Local Ponderada que un modelo de Media Ponderada, si bien en todos los procesos de evaluación realizados se ha comprobado que todas las características y conclusiones que puede proporcionar un modelo de Media Ponderada son aplicables al modelo de predicción de Regresión Local Ponderada.

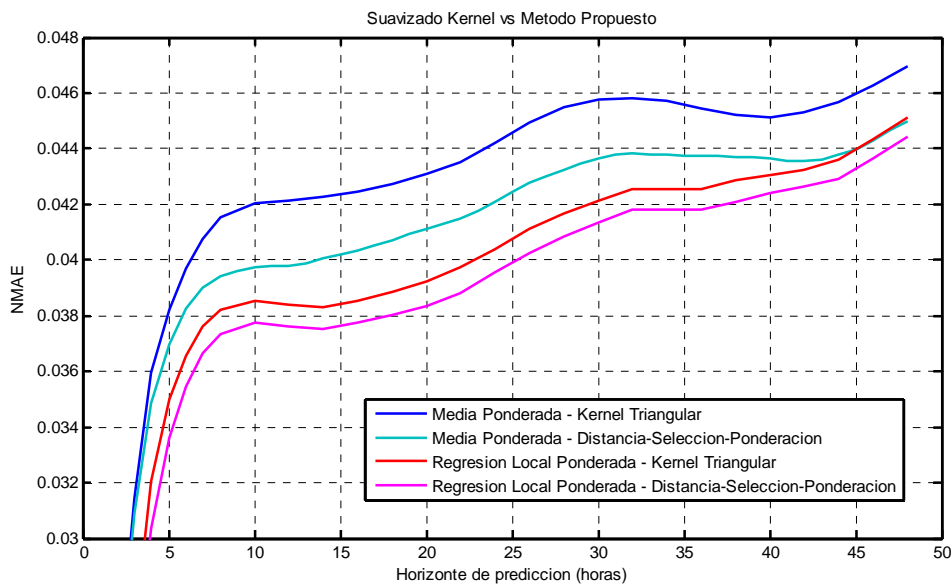


Figura 5.73: Comparación de error NMAE entre función de suavizado Kernel y método de ponderación propuesto.

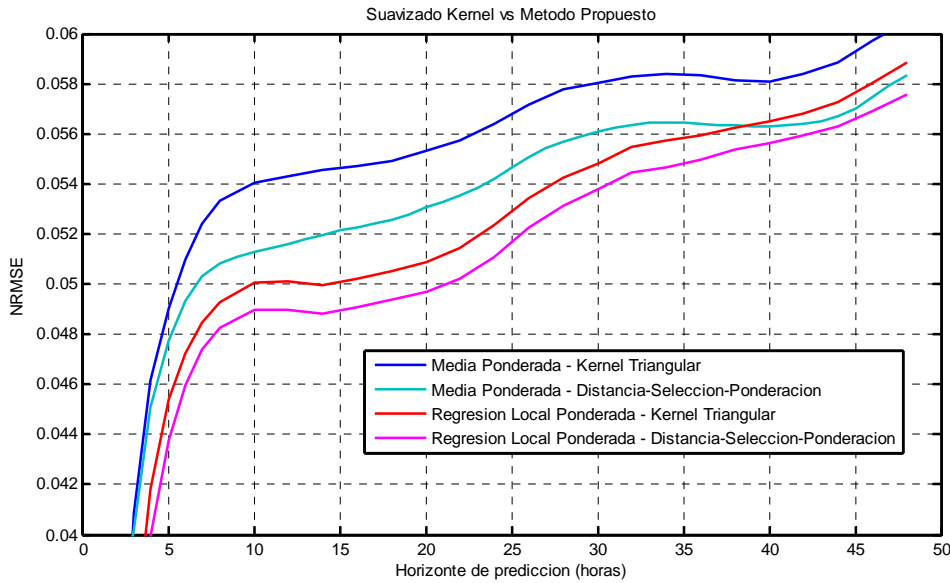


Figura 5.74: Comparación de error NRMSE entre función de suavizado Kernel y método de ponderación propuesto.

5.8 Evaluación de la ponderación por antigüedad

En el momento de hacer la combinación de medidas de potencia correspondientes a los vectores más cercanos al nuevo vector de viento previsto, el método agregado propuesto tendrá en cuenta la antigüedad de los datos históricos observados.

Por tanto, se deberán tener en cuenta una serie de coeficientes que dependerán de la antigüedad de los datos seleccionados. Así, de todos los datos seleccionados se dará menor peso a los que tienen una antigüedad mayor, y mayor peso a aquellos que ocurrieron en un instante más reciente. Para tener en cuenta la antigüedad de los datos seleccionados se propuso el coeficiente de ponderación definido por la ecuación (4.33).

El factor de olvido λ tomará valores entre 0 y 1 y se ajustará de modo que se minimice el error de predicción. El factor de olvido λ está relacionado con la memoria asintótica γ según la ecuación (4.39). Para comprobar el error de predicción dependiendo del valor del parámetro λ se irá variando progresivamente el valor en horas de la memoria asintótica γ hasta encontrar los valores que minimizan el error.

Para conocer el rango de valores óptimos del factor de olvido λ que proporcionan los mejores resultados, se han utilizado de nuevo los datos de potencia medida y predicciones de viento del año 2007, utilizando los primeros seis meses como conjunto de datos históricos, y la segunda mitad del año como periodo de ajuste de parámetros.

En primer lugar, se han ajustado los valores del parámetro α , que define los coeficientes de ponderación por distancia, y del parámetro p , que define el porcentaje de datos históricos cercanos utilizados en la ponderación. Posteriormente se ajustan los valores de λ para minimizar el error en cada horizonte de predicción.

En la Figura 5.75 se muestran los valores de α y p que minimizaron los errores NMAE y NRMSE utilizando como medida de distancia la distancia absoluta estandarizada con la media y seleccionando un porcentaje de los datos históricos como elementos cercanos seleccionados. Se ha utilizado la función de ponderación inversa a la distancia con el parámetro α corregido con la mediana μ para determinar los valores óptimos de α .

Como se puede observar, cuando se usa un modelo local de Media Ponderada para calcular la predicción (gráficas superiores), el porcentaje óptimo de datos a utilizar en la ponderación ronda del 1 al 1.5%, y los valores de α que minimizan los errores NMAE y NRMSE se sitúan en torno a 4, con pequeñas variaciones salvo para horizontes muy bajos, que tienen un comportamiento especial y adquieren un valor más elevado de α , tal y como se explica más adelante.

Cuando las predicciones se calculan con un modelo de Regresión Local Ponderada (gráficas inferiores), el porcentaje adecuado de datos a usar en la ponderación oscila entre el 40 y el 60%, y el parámetro α que minimiza los errores varía entre 1 y 2.5, siendo decreciente conforme aumenta el horizonte de predicción.

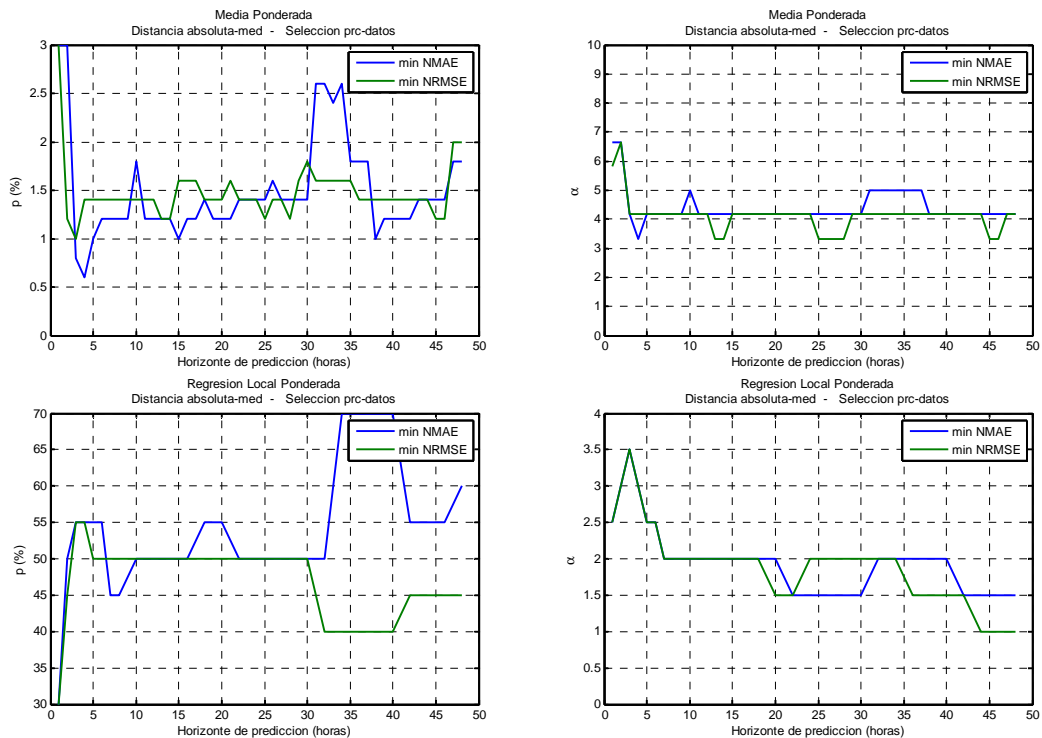


Figura 5.75: Valores de parámetros p y α que minimizan los errores NMAE y NRMSE, con modelos de Media ponderada y Regresión Local Ponderada

Métodos de Predicción de la Generación Agregada de Energía Eólica

Una vez conocida la mejor combinación de valores de α y p que proporciona el menor error de predicción en cada horizonte, se incluyen los coeficientes de ponderación por antigüedad manteniendo constantes los valores de α y p calculados anteriormente, y se comprueban los valores de λ que minimizan el error.

En la Figura 5.76 se comprueba que el valor del factor de olvido λ que minimiza el error NMAE se sitúa entre 0.9996 o 0.9997 para horizontes de 5 horas, lo que implica una memoria asintótica γ de unas 3000 horas, y crece paulatinamente con el horizonte de predicción hasta valores de λ superiores a 0.9999, con una memoria asintótica asociada superior a 10.000 horas, equivalente a más de un año de datos.

Para horizontes de predicción muy bajos se llega al extremo de que el factor de olvido baje hasta $\lambda=0$, como ocurre para horizontes inferiores a 3 horas cuando se utiliza una Media Ponderada para calcular la predicción, lo que implica una memoria asintótica de 1 hora y que, por tanto, es equivalente a utilizar un modelo de persistencia. Usando el modelo de Regresión Local Ponderada no se puede comprobar este fenómeno, ya que si se disminuye demasiado la memoria asintótica el número equivalente de datos seleccionados disminuye considerablemente, de modo que no habría suficientes datos para calcular la predicción, al tener un problema con muchas dimensiones y no poder invertir la matriz en la ecuación de regresión (3.11).

La tendencia a la persistencia en los primeros horizontes de predicción está relacionada con el hecho de que, cuando no se utiliza ponderación por antigüedad, los valores que se obtienen del parámetro α de ponderación por distancias que minimizan el error para horizontes inferiores a 3 horas sean muy superiores a los que se obtienen para horizontes mayores (Figura 5.75). Para horizontes de predicción muy cortos, el vector más cercano en distancia suele ser el último registrado, correspondiente al instante t . Un parámetro α elevado en la ponderación por distancia será más selectivo, y por tanto tenderá a tomar solo en cuenta el elemento más cercano en distancia, que para horizontes de predicción muy bajos coincidirá con el último vector observado, y por tanto se asemejará a usar el método de persistencia.

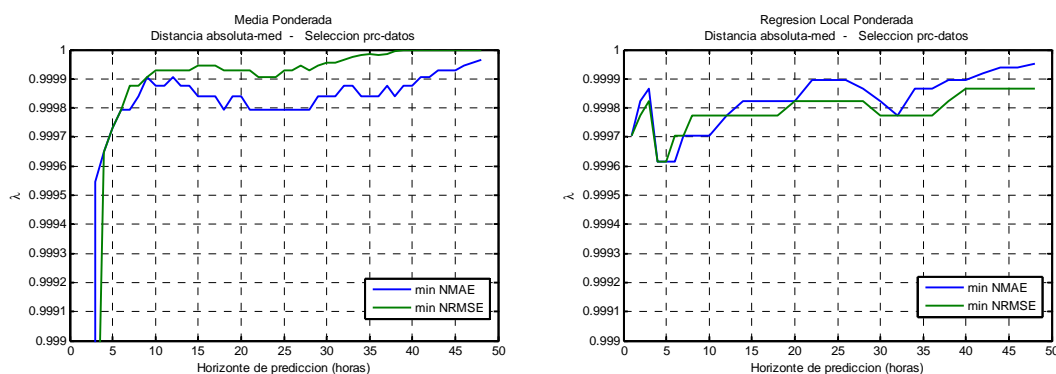


Figura 5.76: Valores del parámetro λ que minimizan los errores NMAE y NRMSE, con modelos de Media ponderada y Regresión Local Ponderada

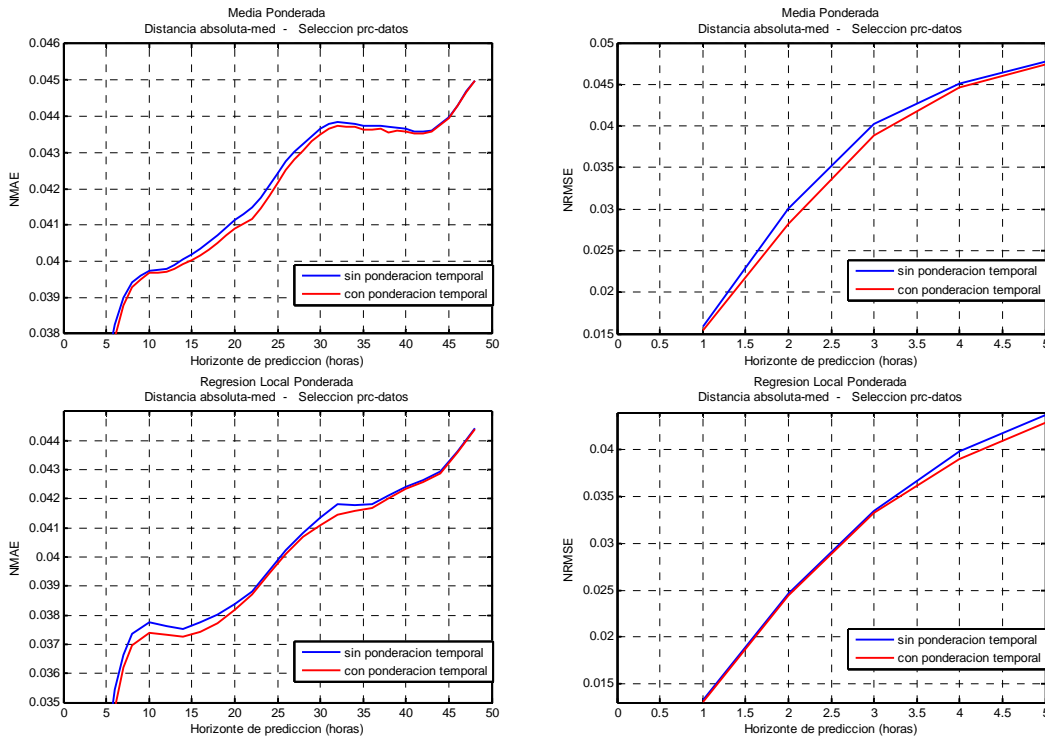


Figura 5.77: Comparación de los errores de predicción sin ponderación por antigüedad y con ponderación por antigüedad

Para evaluar la influencia de incluir una función de ponderación por antigüedad, en las gráficas de la Figura 5.77 se muestran las curvas de error de predicción NMAE y NRMSE para horizontes superiores e inferiores a 5 horas respectivamente. También se muestra en la Figura 5.78 la disminución porcentual, o coeficiente de mejora según (3.18), de dichos errores en función del horizonte de predicción. Como puede apreciarse, la disminución del error al incluir los coeficientes de ponderación por antigüedad es mucho más acusada para horizontes bajos cuando se usa un modelo de Media Ponderada, donde se llega a disminuir el error más de un 5% si se utilizan coeficientes de ponderación temporal. Cuando el modelado es de Regresión Local Ponderada, la ponderación temporal no permite ser tan severa, pues al tener pocos datos existen problemas para invertir la matriz de la ecuación de regresión (3.11).

En todo caso, el coeficiente de ponderación por antigüedad adquiere mayor importancia en horizontes de predicción relativamente bajos, ya que la potencia esperada para la región a corto plazo estará muy influenciada por los últimos valores de potencia registrados en la región. En cambio, conforme aumenta el horizonte de predicción la mejora obtenida disminuye paulatinamente

Métodos de Predicción de la Generación Agregada de Energía Eólica

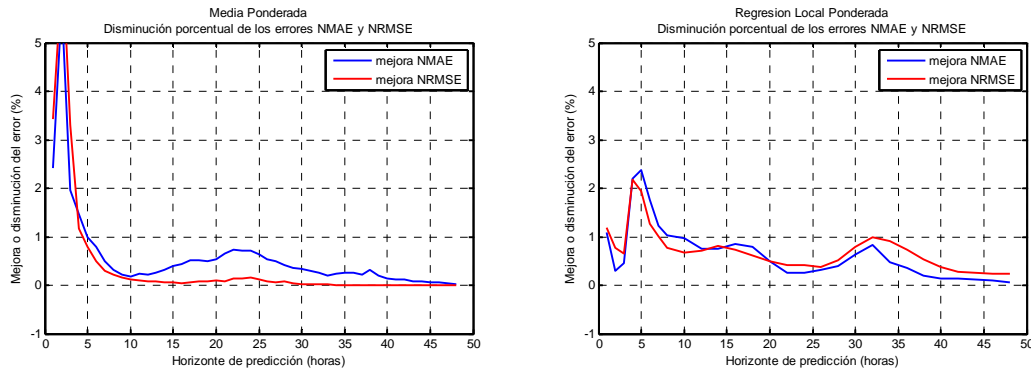


Figura 5.78: Coeficiente de mejora o disminución de los errores NMAE y NRMSE con el uso de ponderación por antigüedad de los datos

5.9 Resumen general de la evaluación

En este capítulo se han evaluado los métodos de suavizado con funciones Kernel y el procedimiento de suavizado propuesto. Tras las evaluaciones realizadas, se ha podido concluir que se obtienen mejores predicciones utilizando el procedimiento de suavizado propuesto, en el que se tratan de forma independiente la medida de distancias, la selección de datos y la función de ponderación que más se adecúe a nuestro problema.

Se ha comprobado que la distancia absoluta es la mejor opción de las alternativas propuestas en cuanto a medida de distancias se refiere. También se ha comprobado que el uso de factores de escala, con distancias estandarizadas, mejora los resultados, de modo que la medida de distancia que ha ofrecido mejores resultados ha sido la distancia absoluta estandarizada con la media.

En cuanto al método de selección de datos cercanos, se ha comprobado que la selección de un porcentaje p fijo de datos es la mejor opción de entre todas las propuestas, aparte de ser probablemente la más intuitiva y más sencilla de implementar, lo que se convierte en una ventaja computacional para el algoritmo final.

La función de ponderación por distancia que mejores resultados ha proporcionado ha sido la ponderación inversa con la distancia, en la que el parámetro a ajustar es α . Además, se ha comprobado que se pueden obtener mejoras si se tiene en cuenta la forma de la distribución de distancias, mediante la corrección del parámetro α con la mediana μ de las distancias calculadas en cada instante.

Por último, se ha comprobado que el uso de una función de ponderación por antigüedad de las observaciones, con un factor de olvido λ , también puede mejorar las predicciones, especialmente para horizontes de predicción bajos, en los que la potencia eólica agregada vendrá muy influenciada por las observaciones recientes.

6. Estimación Adaptativa o Evolutiva de Hiperparámetros

Resumen

Frente a la posibilidad de estimar y fijar los valores de los parámetros del modelo de predicción con un periodo de entrenamiento, en este capítulo se propone un método de estimación adaptativa de los hiperparámetros p , α y λ , que definen el número de datos seleccionados y la ponderación de estos datos en función de su distancia y antigüedad.

En un entorno cambiante, como lo es la situación meteorológica o el número y características de los parques instalados en una región, el ajuste automático de los hiperparámetros puede mejorar las predicciones respecto a una estimación global de parámetros en base a un único periodo de entrenamiento. Además, la estimación adaptativa posibilita el funcionamiento del modelo de predicción en un entorno real sin necesidad de supervisión continua.

6.1 Introducción

Como se ha explicado en las secciones anteriores, una vez establecida una medida de distancias entre vectores de viento y un método de selección de datos cercanos, la predicción de potencia eólica será igual a una combinación de los valores de potencia correspondientes a los elementos seleccionados de entre los registrados en el periodo de datos históricos, según un modelo local de Media Ponderada (3.3), o de Regresión Local Ponderada (3.11).

El número de datos históricos seleccionados en cada caso dependerá de un parámetro de selección p , y el modo en que se ponderarán los datos seleccionados vendrá determinado por los parámetros α y λ , asociados a los coeficientes de ponderación por distancia y antigüedad respectivamente.

Debido a que la relación existente entre estos parámetros y la función objetivo a minimizar, ya sea el error NMAE o NRMSE, es altamente no lineal, se debe emplear un método de entrenamiento para ajustar los valores óptimos de estos parámetros.

Se puede hacer un ajuste global, en el que los parámetros del modelo se fijan globalmente mediante un proceso de optimización que habitualmente minimizará el error de predicción sobre los datos de entrenamiento. Así, un valor adecuado para los parámetros se puede determinar utilizando un periodo de ensayo razonablemente amplio, probando unos rangos de valores de los parámetros p , α y λ , y estableciendo finalmente unos valores fijos para los parámetros con los que se minimice el error de predicción durante el periodo de ensayo. De este modo, se tendría un modelo global de predicción en el que los parámetros que proporcionan el menor error se suponen constantes e iguales para todo el espacio de entrada.

También se podría tratar de hacer un ajuste local, en el que los parámetros sean ajustados localmente a cada zona del espacio de entrada. De este modo, se utilizarían unos valores u otros de los parámetros dependiendo de los datos de entrada, o vector de viento previsto en nuestro caso. El problema es que cuando el espacio de entrada es muy grande, como ocurre en nuestro caso dado que tenemos muchas coordenadas de predicción de viento como entrada, sería necesaria una cantidad enorme de datos de entrenamiento, lo que hace inviable un ajuste local de parámetros para nuestro problema de predicción.

En nuestro caso lo que buscaremos será minimizar el error cuadrático medio de forma global durante un periodo de ensayo, y posteriormente trataremos de mantener mínimo ese error de forma adaptativa con el transcurso del tiempo, para lo cual iremos variando paulatinamente los parámetros del modelo.

Métodos de Predicción de la Generación Agregada de Energía Eólica

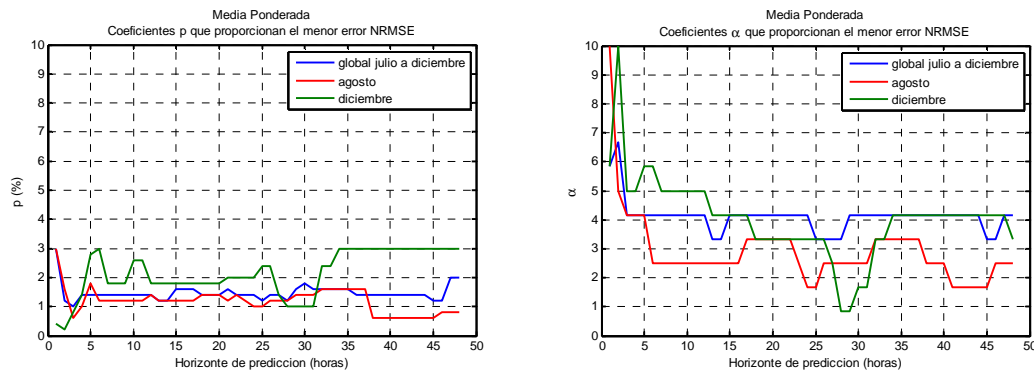


Figura 6.1: Parámetros p y α que proporcionan el menor error NRMSE en distintos periodos

Si se dibuja el valor de los parámetros que hubiesen derivado en un menor error durante un periodo de tiempo, se puede observar que el valor de los parámetros que minimiza el error evoluciona con el tiempo. En la Figura 6.1 se representa el valor de los parámetros α y p que dan lugar a los menores errores NRMSE de predicción del modelo local de Media Ponderada para los horizontes $h=1$ hasta $h=48$ horas durante la segunda mitad completa del año 2007. También se representan los valores de α y p que proporcionaron el menor NRMSE para los meses individuales de agosto y diciembre.

Se puede apreciar que los mejores valores de los parámetros son distintos para el mes de agosto que para el mes de diciembre, y a su vez distintos de los que minimizan el error de los seis meses de ensayo de forma conjunta.

Esto nos lleva a considerar que los valores más adecuados de los parámetros del modelo de predicción agregado no tienen por qué ser algo global o estático, sino que sus valores pueden ir variando a lo largo del tiempo. Por tanto, sería esperable que un método de estimación que permita cierto grado de libertad en la elección de los parámetros p , α y λ para calcular las predicciones proporcione mejores resultados que si se establecen unos valores fijos de los parámetros.

Por otro lado, la potencia eólica generada en una región depende de un modo no lineal de multitud de variables meteorológicas y parámetros de funcionamiento de los parques eólicos. Por elaborada que pueda ser la solución propuesta para la estimación de la potencia eólica, siempre será un modelo muy simple para un problema extremadamente complejo. Por tanto, los parámetros que se puedan estimar en nuestro modelo durante un periodo de ensayo no tendrán por qué ser válidos indefinidamente en un entorno cambiante como lo son las condiciones meteorológicas o las características de los parques eólicos en la región.

Gracias a que la evolución meteorológica a nivel regional es en general relativamente lenta, y debido a que las características de los parques instalados en una región suficientemente grande no cambiarán bruscamente, permitiremos que los parámetros se ajusten por sí solos paulatinamente, de modo que conforme vayan evolucionando

las condiciones, el modelo se vaya ajustando a las nuevas condiciones. Esta adaptabilidad temporal será equivalente a optimizar localmente los parámetros.

Además, lo que se busca en un modelo de predicción es que pueda estar funcionando sin supervisión durante largos periodos, de modo que lo más conveniente es que el modelo recalcule de un modo autónomo todos los parámetros necesarios y se adapte a las situaciones cambiantes. Pero este ajuste no puede hacerse de modo individual, sino que será necesario ajustar la combinación de todos los parámetros de modo simultáneo, ya que si un parámetro varía, habitualmente el valor óptimo de los otros parámetros también variará.

Para este fin, lo que se propone en este caso es una estimación adaptativa o evolutiva de los parámetros p , α y λ del modelo de predicción de forma conjunta, de modo que, cada vez que se reciba una medida de potencia eólica, se calculen los errores cometidos por un conjunto de predicciones alternativas que hayan utilizado diferentes combinaciones de estos parámetros, y se vayan adaptando para los siguientes cálculos aquellos valores de parámetros que estén proporcionando mejores resultados.

Por otra parte, el hecho de calcular varias predicciones alternativas nos permite hacer una combinación de predicciones o “ensemble”, solo que en este caso las alternativas no serán distintas predicciones meteorológicas de entrada, sino distintas combinaciones de valores de los parámetros. Como ejemplo, en [115] se sugiere calcular predicciones con modelos que utilicen en la selección diferentes cantidades de vecinos cercanos. Con esto se podrían mejorar las predicciones, al tener una colección de predicciones que estarán mejor o peor adaptadas dependiendo del punto de entrada.

6.2 Vectores de parámetros y cubo de predicciones alternativas

El objetivo de esta estimación adaptativa es buscar, en cada momento, el valor más adecuado de cada uno de los tres parámetros p , α y λ , que consiga minimizar los errores de predicción. Esto se hará calculando un conjunto de predicciones alternativas que utilicen valores de parámetros en el entorno de los valores que han proporcionado mejores resultados hasta ese instante.

Así, cuando en el instante t se calculen las predicciones alternativas de horizonte h , para el instante $t+h$, una de las predicciones será la que utilice los valores de los parámetros que minimizaron el error de predicción de horizonte h hasta el instante t , y el resto de predicciones alternativas serán calculadas aplicando pequeñas variaciones a dichos parámetros.

Para este fin, se definen unos vectores de parámetros $\mathbf{p}_{t,h}$, $\boldsymbol{\alpha}_{t,h}$, y $\boldsymbol{\lambda}_{t,h}$, centrados en los valores que han proporcionado los mejores resultados en las predicciones de horizonte h hasta el instante t , valores a los que llamaremos $\mathbf{p}_{t,h}^{(0)}$, $\boldsymbol{\alpha}_{t,h}^{(0)}$ y $\boldsymbol{\lambda}_{t,h}^{(0)}$.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Los vectores de parámetros serán de la forma:

$$\mathbf{p}_{t,h} = \left[\dots, p_{t,h}^{(-2)}, p_{t,h}^{(-1)}, p_{t,h}^{(0)}, p_{t,h}^{(+1)}, p_{t,h}^{(+2)}, \dots \right] \quad (6.1)$$

$$\mathbf{\alpha}_{t,h} = \left[\dots, \alpha_{t,h}^{(-2)}, \alpha_{t,h}^{(-1)}, \alpha_{t,h}^{(0)}, \alpha_{t,h}^{(+1)}, \alpha_{t,h}^{(+2)}, \dots \right] \quad (6.2)$$

$$\mathbf{\lambda}_{t,h} = \left[\dots, \lambda_{t,h}^{(-2)}, \lambda_{t,h}^{(-1)}, \lambda_{t,h}^{(0)}, \lambda_{t,h}^{(+1)}, \lambda_{t,h}^{(+2)}, \dots \right] \quad (6.3)$$

donde:

$\mathbf{p}_{t,h}$ = vector de parámetros p , centrado en el valor $p_{t,h}^{(0)}$ que proporcionó el menor error de predicción para el horizonte h hasta el instante t

$\mathbf{\alpha}_{t,h}$ = vector de parámetros α , centrado en el valor $\alpha_{t,h}^{(0)}$ que proporcionó el menor error de predicción para el horizonte h hasta el instante t

$\mathbf{\lambda}_{t,h}$ = vector de parámetros λ , centrado en el valor $\lambda_{t,h}^{(0)}$ que proporcionó el menor error de predicción para el horizonte h hasta el instante t

Los vectores de parámetros tendrán un número fijo e impar de elementos, n_{z_p} , n_{z_α} y n_{z_λ} , ocupando la posición central de cada vector el valor del parámetro que minimizó el error en la predicción de horizonte h durante los últimos cálculos.

Cada uno de los valores de los elementos de los vectores de parámetros será calculado con pequeños incrementos o decrementos porcentuales sobre el valor central, que proporcionó los menores errores en la predicción de horizonte h hasta el instante t . Esto sería similar al procedimiento de cálculo de predicciones ensembles meteorológicas, según el cual se introducen pequeñas variaciones sobre los valores de las condiciones iniciales y de contorno para tener varias predicciones meteorológicas alternativas.

Se definen así unos coeficientes de incremento r_p , r_α , y r_γ que definen la relación entre los valores de dos elementos consecutivos de un vector de parámetros. Por tanto, los coeficientes de incremento vendrán dados por:

$$r_p = \frac{p_{t,h}^{(z_p+1)}}{p_{t,h}^{(z_p)}} \quad (6.4)$$

$$r_\alpha = \frac{\alpha_{t,h}^{(z_\alpha+1)}}{\alpha_{t,h}^{(z_\alpha)}} \quad (6.5)$$

Hay que tener en cuenta que si lo que se desea es aumentar o disminuir en un cierto porcentaje el número equivalente de datos antiguos seleccionados para la ponderación, no habrá que aumentar o disminuir el factor de olvido λ , sino que el valor que deberá aumentar y disminuir será la memoria asintótica γ de la ecuación (4.39). Se tiene en este caso por tanto:

$$r_\gamma = \frac{\gamma_{t,h}^{(z_\lambda+1)}}{\gamma_{t,h}^{(z_\lambda)}} \quad (6.6)$$

Los coeficientes de incremento tendrán un valor $r_{param} > 1$, que indicarán en qué proporción aumentan los valores de los parámetros dentro del vector. Así, por ejemplo, un valor de $r_p = 1.1$ indicaría un incremento de un 10% en el número de datos cercanos seleccionados entre un elemento y el siguiente elemento del vector de parámetros de selección $\mathbf{p}_{t,h}$.

De este modo, el valor de cualquier elemento del vector de parámetros, localizado en la posición z_{param} , vendrá dado por el valor del parámetro que proporcionó los menores errores de predicción hasta ese instante, multiplicado por el coeficiente de incremento r_{param} elevado al valor de z_{param} , siendo $\mathbf{z}_{param} = [\dots, -2, -1, 0, +1, +2, \dots]$ un vector centrado en cero de $n z_{param}$ elementos y que define cuántas alternativas o posiciones en incremento y decremento se contemplarán para un parámetro determinado.

Para los parámetros p y α , asociados respectivamente al número de datos cercanos seleccionados y al decaimiento en el coeficiente de ponderación con la distancia, los valores de los parámetros en los vectores $\mathbf{p}_{t,h}$ y $\boldsymbol{\alpha}_{t,h}$ serán por tanto:

$$\mathbf{p}_{t,h}^{(z_p)} = \mathbf{p}_{t,h}^{(0)} \cdot r_p^{z_p} \quad (6.7)$$

$$\boldsymbol{\alpha}_{t,h}^{(z_\alpha)} = \boldsymbol{\alpha}_{t,h}^{(0)} \cdot r_\alpha^{z_\alpha} \quad (6.8)$$

Obteniéndose un vector de parámetros en el que cada elemento se diferencia del siguiente en un incremento porcentual, definido por r_p o r_α , del número p de elementos seleccionados o del exponente α en el coeficiente de decaimiento con la distancia.

Los valores alternativos de la memoria asintótica se calcularán de forma similar, con incrementos y decrementos porcentuales sobre el valor central, de la forma:

$$\gamma_{t,h}^{(z_\gamma)} = \gamma_{t,h}^{(0)} \cdot r_\gamma^{z_\lambda} \quad (6.9)$$

Métodos de Predicción de la Generación Agregada de Energía Eólica

De la ecuación (4.39), que relaciona la memoria asintótica γ con el factor de olvido λ , los valores alternativos del factor de olvido que contendrá el vector $\lambda_{t,h}$ y que serán utilizados en la ponderación por antigüedad vendrán dados por la relación entre la memoria asintótica y el factor de olvido mediante:

$$\lambda_{t,h}^{(z\lambda)} = 1 - \frac{1}{\gamma_{t,h}^{(z\lambda)}} \quad (6.10)$$

y usando la ecuación (6.9) se obtienen los valores alternativos del factor de olvido como:

$$\lambda_{t,h}^{(z\lambda)} = 1 - \frac{1}{\frac{1}{1 - \lambda_{t,h}^{(0)}} \cdot (r_\gamma)^{z\lambda}} \quad (6.11)$$

Dado que se tienen tres vectores de parámetros, se tendrá un conjunto tridimensional de combinaciones de parámetros y por tanto de predicciones alternativas. Así, si se tienen unos vectores $\mathbf{p}_{t,h}$, $\mathbf{\alpha}_{t,h}$ y $\lambda_{t,h}$, con nz_p , nz_α , y nz_λ elementos cada uno, se tendrán un total de $Z = nz_p \cdot nz_\alpha \cdot nz_\lambda$ predicciones alternativas para el instante $t+h$, organizadas en lo que llamaremos un “cubo de predicciones”, centrado en la predicción que ha sido calculada con la combinación de parámetros que obtuvo los menores errores de predicción durante los cálculos previos al instante t .

6.3 Error de las predicciones alternativas. Cubo de errores

El objetivo final del procedimiento propuesto en este capítulo es adaptar de modo continuo los valores de los parámetros del modelo de predicción agregado, aumentando o reduciendo su valor de modo que la predicción final de potencia tenga el mínimo error. Se irán desplazando así los vectores de parámetros, manteniendo fijo el número de elementos en cada uno de los vectores, hacia valores más elevados o más bajos de los parámetros, en busca del valor de cada parámetro que proporciona los mejores resultados en cada momento.

El modo en que se decidirá si hay que aumentar o disminuir los valores de los parámetros será calculando los errores de predicción y desplazando los vectores hacia valores de parámetros donde se obtenga el menor error de predicción. Así, cada vez que se reciba el valor de potencia agregada medida en la región en un instante t , se calculará el error cometido por cada una de las predicciones alternativas del “cubo de predicciones de horizonte h ”, que fueron calculadas en el instante $t-h$. Se obtendrá de este modo un “cubo de errores” para las predicciones de horizonte h .

A la hora de decidir en qué dirección se desean desplazar los vectores de parámetros se debe decidir previamente cuál es el error que se desea minimizar. Si lo que se

desease fuera minimizar en lo posible el error absoluto de predicción en término medio, se debería calcular el error de predicción NMAE para cada una de las combinaciones de parámetros.

Sin embargo, para la operación de un sistema eléctrico con alta penetración eólica, no es tan importante el error de predicción en término medio a lo largo de un periodo, y cobran mayor importancia los valores extremos en los desvíos de la predicción eólica, que además coincidirán generalmente con valores extremos de potencia eólica registrada y por tanto con los instantes más críticos para la operación del sistema eléctrico. El error cuadrático medio penaliza en mayor medida los valores extremos del error de predicción, de modo que será este tipo de error, el error cuadrático medio o NMSE, el que se tratará de minimizar.

6.4 Estimación adaptativa del NMSE mediante EWMA

El error NMSE de cada predicción alternativa, fruto de una combinación dada de parámetros p , α , y λ , se calculará de modo recursivo, utilizando una media móvil ponderada exponencialmente o EWMA (Exponentially Weighted Moving Average). De este modo, se hace innecesario almacenar los datos históricos de error para calcular el error medio de un periodo, actualizándose el valor medio de error cada vez que se tiene un nuevo dato de potencia.

Cada vez que en un instante t llega un nuevo dato de potencia medida en la región, el error cometido por una predicción z del total de Z predicciones alternativas de horizonte h que fue calculada en el instante $t-h$ por una combinación dada de parámetros $p_{t-h,h}^{(z)}$, $\alpha_{t-h,h}^{(z)}$ y $\lambda_{t-h,h}^{(z)}$, será:

$$e_{t|t-h}^{(z)} = P_t - \hat{P}_{t|t-h}^{(z)} \quad (6.12)$$

donde:

$e_{t|t-h}^{(z)}$ = Error de predicción de horizonte h de la predicción alternativa z

P_t = Potencia eólica registrada en la región en el instante t

$\hat{P}_{t|t-h}^{(z)}$ = Predicción alternativa z de horizonte h calculada en el instante $t-h$

El error cuadrático medio normalizado se actualizará de forma adaptativa por medio de una media móvil ponderada exponencialmente, así el nuevo NMSE en el instante t se calculará del siguiente modo:

$$NMSE_{t|t-h}^{(z)} = \frac{1}{N_{z,t}} \cdot \Phi_{t|t-h}^{(z)} \quad (6.13)$$

Métodos de Predicción de la Generación Agregada de Energía Eólica

donde:

$NMSE_{t|t-h}^{(z)}$ = Estimación en el instante t del error cuadrático medio normalizado de las predicciones alternativas z , de horizonte h , calculadas con la combinación de parámetros $p_{t-h,h}^{(z)}$, $\alpha_{t-h,h}^{(z)}$, $\lambda_{t-h,h}^{(z)}$.

con

$$\Phi_{t|t-h}^{(z)} = \left(e_{t|t-h}^{(z)} \right)^2 + \lambda_{MSE} \cdot \Phi_{t-1|t-h-1}^{(z)} \quad (6.14)$$

y con

$$N_{z,t} = 1 + \lambda_{MSE} \cdot N_{z,t-1} \quad (6.15)$$

que se inicializa con los valores $N_{z,0}=0$ y $\Phi_0^{(z)}=0$

donde λ_{MSE} es el factor de olvido en la estimación del NMSE. El factor de olvido tomará valores en el intervalo $0 < \lambda_{MSE} < 1$, de modo que un valor pequeño de λ_{MSE} implicaría una menor influencia de los valores más antiguos de error en la estimación, mientras que un valor próximo a 1 implicaría mayor lastre de los errores que ocurrieron hace mucho tiempo.

La memoria asintótica en la estimación adaptativa del error, relacionada con el número de muestras equivalentes $N_{z,t}$ en (6.15) cuando t sea suficientemente elevado, será el resultado de la serie:

$$N_{z,t} = 1 + \lambda_{MSE} + \lambda_{MSE}^2 + \dots + \lambda_{MSE}^t = \frac{1}{1 - \lambda_{MSE}} \quad (6.16)$$

El valor adecuado del factor de olvido λ_{MSE} , y por tanto del número de muestras equivalentes $N_{z,t}$, será un valor a determinar utilizando para ello un periodo de ensayo. La memoria asintótica tendrá que ser lo suficientemente pequeña como para detectar con suficiente rapidez los cambios en los errores cometidos por cada predicción alternativa, y lo suficientemente elevada como para que no haya una gran variabilidad en los valores de error calculados.

Los valores de error medio se calcularán recursivamente para cada una de las $z=1, \dots, Z$ predicciones alternativas del cubo de predicciones, y para cada uno de los horizontes de predicción.

6.5 Combinación de predicciones alternativas del cubo de predicciones

Una vez calculada cada una de las Z predicciones alternativas, fruto de las $Z = nz_p \cdot nz_\alpha \cdot nz_\lambda$ combinaciones posibles de parámetros p , α y λ , la predicción final del modelo agregado para cada horizonte h será una combinación de las predicciones alternativas de potencia calculadas. La combinación de predicciones proporcionará mayor robustez al modelo frente a la posibilidad de dar como predicción final la correspondiente a la predicción alternativa con menor error. Esta combinación se hará teniendo en cuenta los errores cometidos por cada predicción alternativa, estimados mediante EWMA, (6.13) a (6.15), por medio de unos coeficientes de ponderación.

Si se tienen Z predicciones alternativas para el horizonte h , la predicción combinada final será entonces:

$$\hat{p}_{t+h|t}^{(C)} = \sum_{z=1}^Z w_{z,t}(h) \cdot \hat{p}_{t+h|t}^{(z)} \quad (6.17)$$

donde:

$\hat{p}_{t+h|t}^{(C)}$ = Predicción final, fruto de la combinación de predicciones alternativas, calculada en el instante t para el instante $t+h$

$w_{z,t}(h)$ = Coeficiente de ponderación de la predicción alternativa z de horizonte h , calculado en t

$\hat{p}_{t+h|t}^{(z)}$ = Predicción alternativa z para el instante $t+h$ calculada en t , fruto de una combinación dada de parámetros

Los coeficientes de ponderación de las predicciones alternativas, $w_{z,t}(h)$, se calcularán teniendo en cuenta el error cuadrático medio NMSE cometido por las distintas predicciones alternativas, según un método de promedio óptimo mediante:

$$w_{z,t}(h) = \frac{1}{\frac{NMSE_{t|t-h}^{(z)}}{\sum_{j=1}^Z \frac{1}{NMSE_{t|t-h}^{(j)}}}} \quad (6.18)$$

De modo que los coeficientes de ponderación serán inversamente proporcionales al error cuadrático medio. Se debe destacar que, al calcular recursivamente el NMSE, también serán recalculados en cada instante t los coeficientes de ponderación de cada predicción alternativa para cada horizonte h , con lo que los coeficientes $w_{z,t}(h)$ irán variando con el paso del tiempo.

6.6 Desplazamiento de los vectores de parámetros

Calculado el cubo de errores de las predicciones alternativas, se procurará que los valores de los parámetros queden centrados en el entorno de los valores que han proporcionado los menores errores. Así, cada periodo de tiempo definido por el número de muestras equivalentes $N_{z,t}$, y que depende del factor de olvido λ_{MSE} en la estimación del error según (6.16), se facilitará un desplazamiento a cada uno de los vectores de parámetros $\mathbf{p}_{t,h}$, $\mathbf{\alpha}_{t,h}$, y $\mathbf{\lambda}_{t,h}$ en el sentido en que se hayan obtenido los mejores resultados.

Pero este ajuste no puede hacerse de modo individual, sino que será necesario ajustar la combinación de todos los parámetros de modo simultáneo, ya que si un parámetro varía, habitualmente el valor óptimo de los otros parámetros también variará.

Para decidir en qué sentido se deben desplazar los valores de los parámetros, creciente o decreciente, se puede calcular el error medio de las predicciones alternativas que fueron calculadas con cada uno de los valores de un parámetro dado al ser utilizado con todas las combinaciones posibles del resto de parámetros. Obviando en la notación la referencia al instante t y al horizonte de predicción h , los errores medios \overline{NMSE} se calculan para cada valor de los parámetros por medio de:

$$\overline{NMSE}_{p(z_p)} = \frac{1}{nz_\alpha \cdot nz_\lambda} \sum_{z_\alpha} \sum_{z_\lambda} NMSE^{(z_p, z_\alpha, z_\lambda)} \quad (6.19)$$

$$\overline{NMSE}_{\alpha(z_\alpha)} = \frac{1}{nz_p \cdot nz_\lambda} \sum_{z_p} \sum_{z_\lambda} NMSE^{(z_p, z_\alpha, z_\lambda)} \quad (6.20)$$

$$\overline{NMSE}_{\lambda(z_\lambda)} = \frac{1}{nz_p \cdot nz_\alpha} \sum_{z_p} \sum_{z_\alpha} NMSE^{(z_p, z_\alpha, z_\lambda)} \quad (6.21)$$

donde:

$\overline{NMSE}_{p(z_p)}$ = Media de los errores NMSE de horizonte h calculados en el instante t de las predicciones alternativas fruto de combinar todos los valores de los vectores de parámetros $\mathbf{\alpha}_{t,h}$ y $\mathbf{\lambda}_{t,h}$ con el valor de parámetro p situado en la posición z_p del vector de parámetros $\mathbf{p}_{t,h}$.

$\overline{NMSE}_{\alpha(z_\alpha)}$ = Media de los errores NMSE de horizonte h calculados en el instante t de las predicciones alternativas fruto de combinar todos los valores de los vectores de parámetros $\mathbf{p}_{t,h}$ y $\mathbf{\lambda}_{t,h}$ con el valor de parámetro α situado en la posición z_α del vector de parámetros $\mathbf{\alpha}_{t,h}$.

$\overline{NMSE}_{\lambda(z_\lambda)}$ = Media de los errores NMSE de horizonte h calculados en el instante t de las predicciones alternativas fruto de combinar todos los valores de los vectores de parámetros $\mathbf{p}_{t,h}$ y $\mathbf{\alpha}_{t,h}$ con el valor de parámetro λ situado en la posición z_λ del vector de parámetros $\mathbf{\lambda}_{t,h}$.

$NMSE^{(z_p, z_\alpha, z_\lambda)}$ = Error MSE de horizonte h calculado en el instante t de la predicción alternativa calculada con los valores situados en las posiciones z_p , z_α y z_λ de los vectores de parámetros $\mathbf{p}_{t,h}$, $\boldsymbol{\alpha}_{t,h}$, y $\boldsymbol{\lambda}_{t,h}$.

El nuevo valor central del vector de parámetros seleccionado para el siguiente desplazamiento de vectores sería aquel que ha proporcionado el menor error en término medio hasta el instante t :

$$param_{t+1,h}^{(0)} = param_{t,h}^{(sel)} \mid \overline{NMSE}_{param^{(sel)}} = \min\left(\overline{NMSE}_{param^{(z_{param})}}\right) \quad (6.22)$$

donde “sel” se refiere al valor seleccionado, y “param” se refiere indistintamente a cualquiera de los parámetros p , α o λ

Sin embargo, valores anormalmente elevados de error de alguna combinación concreta de parámetros podrían afectar al cálculo de la media de error $\overline{NMSE}_{param^{(z_{param})}}$ que se obtiene con un valor de parámetro concreto, aumentando el valor de la media de error y ocultando por tanto la posición en la que se podría encontrar el mínimo error.

Dado que nuestra búsqueda se centra en encontrar la combinación de parámetros que proporciona menores errores, o lo que es lo mismo, mayores coeficientes de ponderación $w_{z,t}$ en (6.17), lo que se hará será desplazar los vectores hacia valores para los que la media de coeficientes de ponderación de predicciones alternativas sea más elevada. Obviando por simplicidad la referencia al instante t y al horizonte h , estas medias de coeficientes serán:

$$\bar{w}_{p^{(z_p)}} = \frac{1}{nz_\alpha \cdot nz_\lambda} \sum_{z_\alpha} \sum_{z_\lambda} w^{(z_p, z_\alpha, z_\lambda)} \quad (6.23)$$

$$\bar{w}_{\alpha^{(z_\alpha)}} = \frac{1}{nz_p \cdot nz_\lambda} \sum_{z_p} \sum_{z_\lambda} w^{(z_p, z_\alpha, z_\lambda)} \quad (6.24)$$

$$\bar{w}_{\lambda^{(z_\lambda)}} = \frac{1}{nz_p \cdot nz_\alpha} \sum_{z_p} \sum_{z_\alpha} w^{(z_p, z_\alpha, z_\lambda)} \quad (6.25)$$

donde:

$\bar{w}_{p^{(z_p)}}$ = Media de los coeficientes de ponderación de las predicciones alternativas fruto de combinar todos los valores de los vectores de parámetros $\boldsymbol{\alpha}_{t,h}$ y $\boldsymbol{\lambda}_{t,h}$ con el valor de parámetro p situado en la posición z_p del vector de parámetros $\mathbf{p}_{t,h}$.

$\bar{w}_{\alpha^{(z_\alpha)}}$ = Media de los coeficientes de ponderación de las predicciones alternativas fruto de combinar todos los valores de los vectores de parámetros $\mathbf{p}_{t,h}$ y $\boldsymbol{\lambda}_{t,h}$ con el valor de parámetro α situado en la posición z_α del vector de parámetros $\boldsymbol{\alpha}_{t,h}$.

Métodos de Predicción de la Generación Agregada de Energía Eólica

$\bar{w}_{\lambda(z_{\lambda})}$ = Media de los coeficientes de ponderación de las predicciones alternativas fruto de combinar todos los valores de los vectores de parámetros $\mathbf{p}_{t,h}$ y $\mathbf{a}_{t,h}$ con el valor de parámetro λ situado en la posición z_{λ} del vector de parámetros $\lambda_{t,h}$.

$w^{(z_p, z_{\alpha}, z_{\lambda})}$ = Coeficiente de la predicción alternativa calculada con los valores situados en las posiciones z_p , z_{α} y z_{λ} de los vectores de parámetros $\mathbf{p}_{t,h}$, $\mathbf{a}_{t,h}$, y $\lambda_{t,h}$.

Este método es menos sensible a la posibilidad de que alguna combinación dada de parámetros que pudiera tener un error de predicción anormalmente elevado dificultase la identificación de otra combinación de parámetros con error de predicción bajo y, por tanto, mayor coeficiente de ponderación asignado en el reparto.

Calculadas las medias de los coeficientes de ponderación para todos los valores de un vector de parámetros, si el valor máximo de la media de coeficientes no corresponde al valor del parámetro en la posición central del vector, se desplazará el vector de parámetros en sentido creciente o decreciente, con el objetivo de mantenerse centrado en el valor del parámetro con que se obtienen los mayores valores medios de los coeficientes de ponderación.

El nuevo valor central del vector de parámetros seleccionado para el siguiente desplazamiento de vectores será aquel que ha obtenido el mayor coeficiente de ponderación en término medio hasta el instante t :

$$param_{t+1,h}^{(0)} = param_{t,h}^{(sel)} \mid \bar{w}_{param^{(sel)}} = \max\left(\bar{w}_{param^{(z_{param})}}\right) \quad (6.26)$$

Finalmente, se recalculan los nuevos valores de los vectores de parámetros según las ecuaciones (6.7), (6.8) y (6.11).

Desplazamiento del cubo de errores

Los valores de error almacenados en el cubo de errores de predicciones alternativas también son desplazados en la misma dirección y sentido que los vectores de parámetros, asumiendo que el error medio inicial de las celdas vacías correspondientes a los nuevos valores de parámetros será igual al error medio de las celdas más cercanas. Por este motivo debe transcurrir un tiempo suficiente para que esas estimaciones iniciales de error de los nuevos parámetros se recalculen con datos de error reales.

La frecuencia con que se lleva a cabo este movimiento y recálculo de vectores de parámetros depende del factor de olvido en la estimación del NMSE, λ_{MSE} , produciéndose una comprobación y movimiento de los vectores con una periodicidad igual a la memoria asintótica o número de muestras $N_{z,t}$ equivalentes de la media móvil que se ha empleado para estimar el error cuadrático medio de las predicciones alternativas. De este modo se asegura que el valor del NMSE estimado sea representativo del error de las predicciones calculadas en el último periodo, desde la anterior ocasión en que se produjo un desplazamiento de vectores.

El periodo de desplazamiento de vectores de parámetros se divide en tres sub-periodos de la misma duración ($\frac{N_{z,t}}{3}$), tras cada uno de los cuales se desplaza y recalcula uno de los tres vectores de parámetros del modelo agregado. Así el vector $p_{t,h}$ tiene su desplazamiento tras el primer sub-periodo, $\alpha_{t,h}$ tras el segundo sub-periodo y $\lambda_{t,h}$ tras el tercero.

Al hacer el desplazamiento de cada vector de parámetros en un sub-periodo diferente, y por tanto el desplazamiento del cubo de errores, se consigue asegurar que los errores correspondientes a cada predicción alternativa hayan sido calculados con un número suficiente de datos, y por tanto se correspondan con cada valor de los vectores de parámetros.

6.7 Evaluación de resultados obtenidos

Para comprobar el funcionamiento del algoritmo propuesto de estimación adaptativa de parámetros para el modelo de predicción regional, se utilizan los datos disponibles de viento y potencia eólica registrados durante la segunda mitad de 2007.

Con los datos correspondientes a los meses de julio, agosto y septiembre, se han calculado los valores de parámetros que dan lugar a los menores errores de predicción NRMSE de forma global. Estos valores se representan en la Figura 6.2, tanto para un modelo de predicción de Media Ponderada como para un modelo de Regresión Local Ponderada.

Utilizando un modelo de Media Ponderada, los valores más adecuados del parámetro p se sitúan entre el 1.5% y el 3%. Para el parámetro α en la función de ponderación por distancia los valores que proporcionan el menor error se sitúan entre 3 y 4, mientras que los mejores valores del parámetro λ en la función de ponderación por antigüedad se sitúan entre 0.9995 y 1 dependiendo del horizonte de predicción.

El menor NRMSE si se usa un modelo de Regresión Local Ponderada se obtiene con p entre un 50 y 70% de los datos, α entre 1 y 3 dependiendo del horizonte de predicción, adquiriendo valores mayores cuanto menor es el horizonte de predicción, y λ se sitúa entre 0.9991 y 0.9994 para horizontes de predicción entre 4 y 48 horas. Para horizontes de predicción inferiores el factor de olvido debe tener menos fuerza para evitar una carencia de datos en el cálculo de la regresión, lo que se traduce en unos valores de λ mayores.

Métodos de Predicción de la Generación Agregada de Energía Eólica

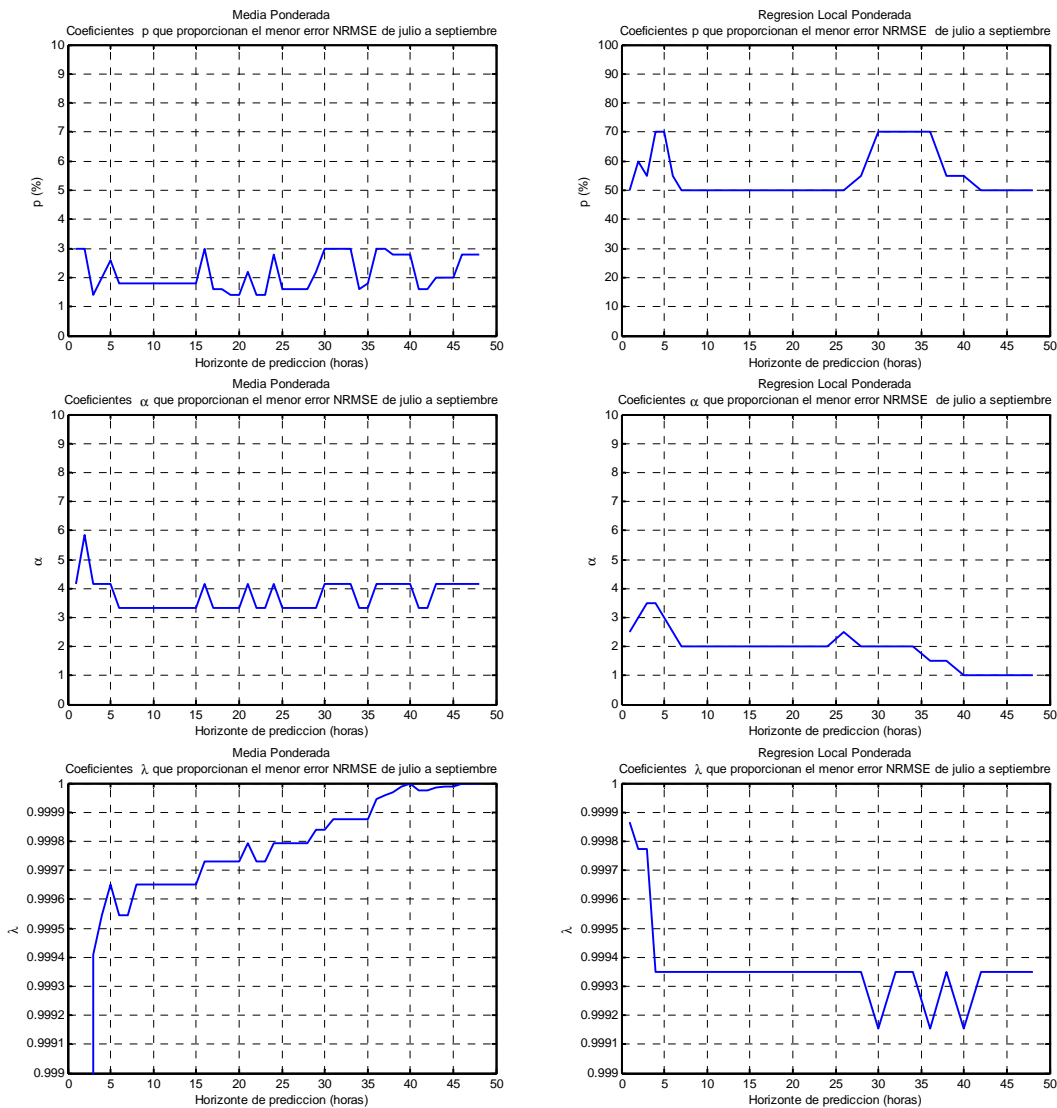


Figura 6.2: Parámetros p , α y λ que proporcionan el menor error NRMSE en el periodo de julio a septiembre de 2007

Todos estos valores de parámetros son relativamente similares a los que se obtuvieron para el período completo de seis meses de ensayo en capítulos anteriores. Establecidos los valores de parámetros que han proporcionado los mejores resultados durante el periodo de ensayo de tres meses, se han calculado los errores del modelo de predicción sin estimación evolutiva de parámetros, durante los meses de octubre a diciembre, utilizando los mejores valores de parámetros p , α y λ constantes e iguales a los del periodo de ensayo de julio a septiembre.

Por otro lado, se han calculado los errores de un modelo de predicción con el procedimiento de estimación adaptativa de parámetros presentado. Como valores centrales iniciales de los vectores de parámetros para la primera hora del mes de

octubre se han utilizado los valores de p , α y λ que minimizaron el error NRMSE durante el periodo de julio a septiembre.

El tamaño de los vectores de parámetros para el modelo de predicción de Media Ponderada se ha establecido en cinco elementos para cada vector, con lo que se tiene un total de $Z = 5 \cdot 5 \cdot 5 = 125$ predicciones alternativas en el cubo de predicciones basadas en un modelo de Media Ponderada. Para un modelo de Regresión Local Ponderada se establece el tamaño de estos vectores en solo tres elementos, debido a que este tipo de modelo requiere mucha más carga de cálculo, teniendo en este caso un total de $Z = 3 \cdot 3 \cdot 3 = 27$ predicciones alternativas en el cubo de predicciones basadas en un modelo de Regresión Local Ponderada.

Los coeficientes de incremento se han establecido en $r_p=1.1$, $r_\alpha=1.1$ y $r_\gamma=1.3$, de modo que el valor de los coeficientes de ponderación con la distancia y el número de datos seleccionados variarán en un 10% entre cada elemento de los vectores de parámetros p y α , y la memoria asintótica γ se verá incrementada o decrementada en un 30% en cada uno de los elementos del vector de parámetros λ .

Finalmente, se han probado diferentes factores de olvido λ_{MSE} para la estimación del NMSE en el cubo de errores, y que es utilizado para el cálculo de coeficientes $w_{z,t}$ de ponderación de predicciones alternativas. Se tienen así diferentes valores en el periodo de actualización o número de muestras equivalentes $N_{z,t}$. Los valores estudiados han variado desde las 72 horas hasta un total de 720 horas, equivalentes a entre 3 días y un mes de datos respectivamente.

En la Figura 6.3 y Figura 6.4 siguientes se muestran los errores NRMSE en función del horizonte de predicción utilizando algunos valores de λ_{MSE} . Se representan los errores para valores equivalentes $N_{z,t}$ de 72, 168 y 240 horas, utilizando modelos de predicción de Media Ponderada y Regresión Local Ponderada respectivamente.

Se puede apreciar cómo el error de predicción con estimación evolutiva de parámetros disminuye respecto del error sin estimación evolutiva. Esta disminución se produce para casi todos los horizontes cuando se utiliza una Regresión Local Ponderada, mientras que para un modelo de Media Ponderada se observa mejora sobre todo para horizontes de predicción medios entre 5 y 20 horas, aunque también se observa una leve disminución del error para horizontes más elevados.

También se pueden apreciar ciertos saltos muy acentuados en el error para los horizontes de 36 y 40 horas cuando se usa un periodo de actualización de parámetros de tan solo 72 horas con una Regresión Local Ponderada. Esto nos puede estar indicando que se puede producir cierta inestabilidad si el periodo de actualización de parámetros es demasiado corto, por lo que aunque se observen menores errores para algunos horizontes con este periodo de actualización de parámetros, resultaría más recomendable utilizar un periodo $N_{z,t}$ más conservador.

Métodos de Predicción de la Generación Agregada de Energía Eólica

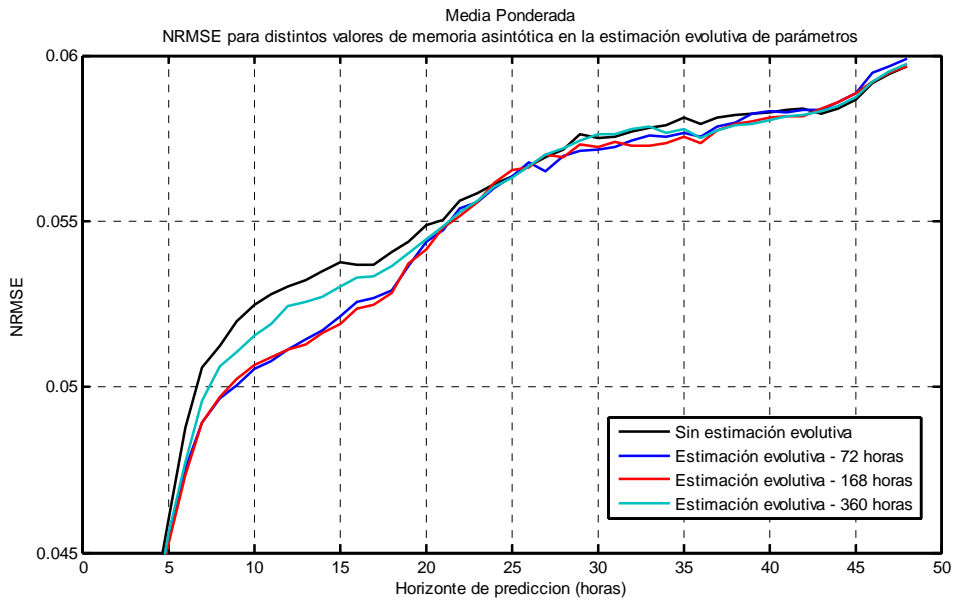


Figura 6.3: NRMSE para distintos valores de memoria asintótica en la estimación evolutiva de parámetros, con modelo de Media Ponderada

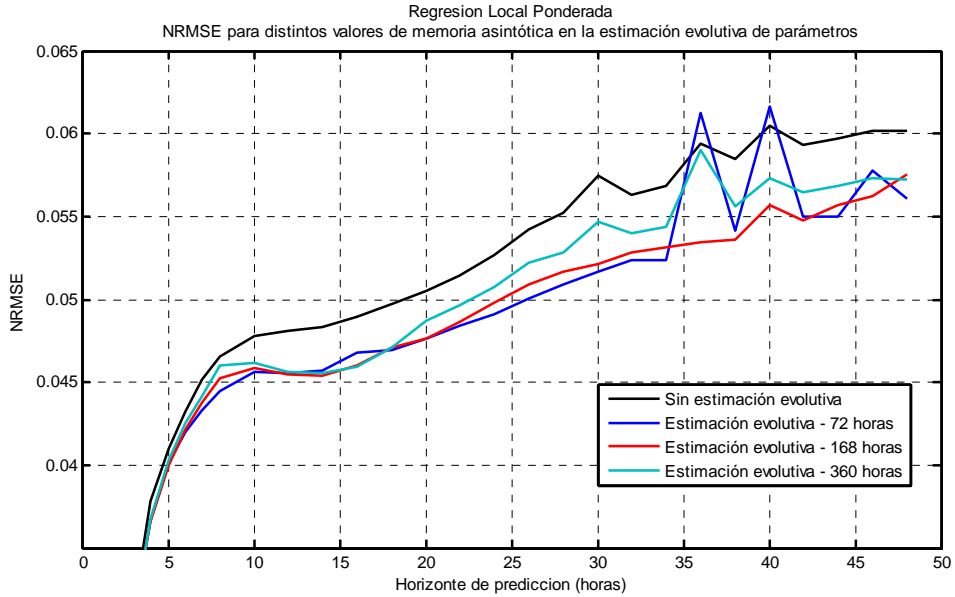


Figura 6.4: NRMSE para distintos valores de memoria asintótica en la estimación evolutiva de parámetros, con modelo de Regresión Local Ponderada

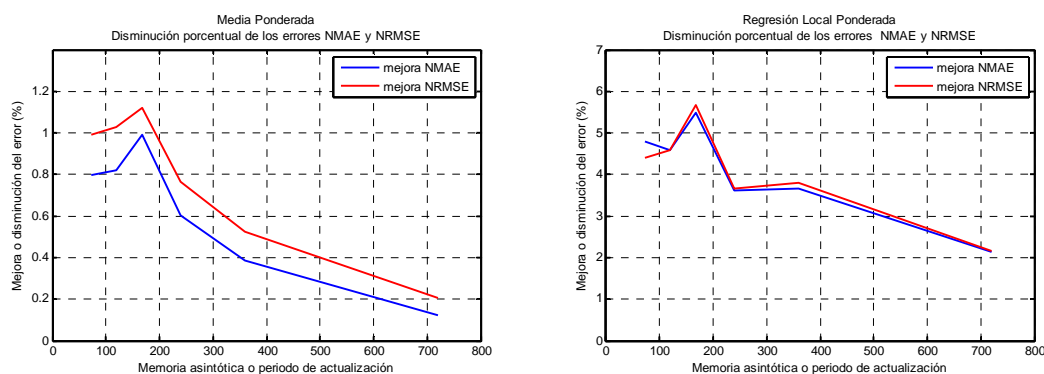


Figura 6.5: Disminución de los errores NMAE y NRMSE medios en función del periodo de actualización de los vectores de parámetros y el cubo de errores

Para evaluar qué factor de olvido o número de muestras equivalentes es el más adecuado, se calcula el valor medio del error de predicción en todos los horizontes, en función del número de muestras equivalentes o periodo de actualización $N_{z,t}$ de los vectores de parámetros.

En la Figura 6.5 se representa, en función del periodo de actualización, la mejora de los errores medios NMAE y NRMSE, para un modelo de Media Ponderada y un modelo de Regresión Local Ponderada. Se han empleado valores de $N_{z,t} = 72, 120, 168, 240, 360$ y 720 horas, que equivalen a periodos de actualización de 3, 5, 7, 10, 15 y 30 días respectivamente.

Como se puede comprobar, en el periodo estudiado tanto el error NMAE como NRMSE mejoran o disminuyen en mayor medida usando un periodo de actualización $N_{z,t} = 168$ horas, equivalente a una semana de datos o un factor de olvido $\lambda_{MSE} = 0.994$. Además, a medida que se aumenta la memoria asintótica o número de muestras equivalentes, la mejora del método de predicción con estimación evolutiva de parámetros va disminuyendo, al hacerse la adaptación muy lenta y convertirse casi en un método no evolutivo para valores muy elevados del periodo de actualización.

En cualquier caso, la elección del periodo de actualización debe ser un compromiso entre la velocidad de adaptación del modelo de predicción y la variabilidad que puede provocar un factor de olvido demasiado bajo. Un periodo de actualización de una semana se puede considerar razonable, pues es de un orden de magnitud similar al periodo con el que pueden producirse cambios significativos en la meteorología.

Para finalizar, se representa la evolución de los valores centrales de los vectores de parámetros, $p_{t,h}^{(0)}$, $\alpha_{t,h}^{(0)}$ y $\lambda_{t,h}^{(0)}$, que han sido utilizados en los modelos de predicción de Media Ponderada y de Regresión Local Ponderada, para los horizontes $h=12, 24$ y 36 horas, durante las horas o instantes t comprendidos entre la primera hora del mes de octubre (hora 6529 del año) y la última hora del mes de diciembre de 2007 (hora 8760 del año).

Métodos de Predicción de la Generación Agregada de Energía Eólica

En la Figura 6.6 se aprecia la evolución de los valores centrales de los vectores de parámetros, que se desplazan, cada 168 horas en este caso, hacia valores con los que se consigue reducir el error de predicción. En todos los casos se puede comprobar que los parámetros parten del valor inicial con el que tuvieron el menor error NRMSE durante el periodo de entrenamiento entre julio y septiembre de 2007, y posteriormente se van desplazando cambiando de valor cada 168 horas. El movimiento es parecido para los tres horizontes de predicción representados, de manera que las subidas y bajadas en el valor de los parámetros se realizan de forma bastante acompasada para todos los horizontes.

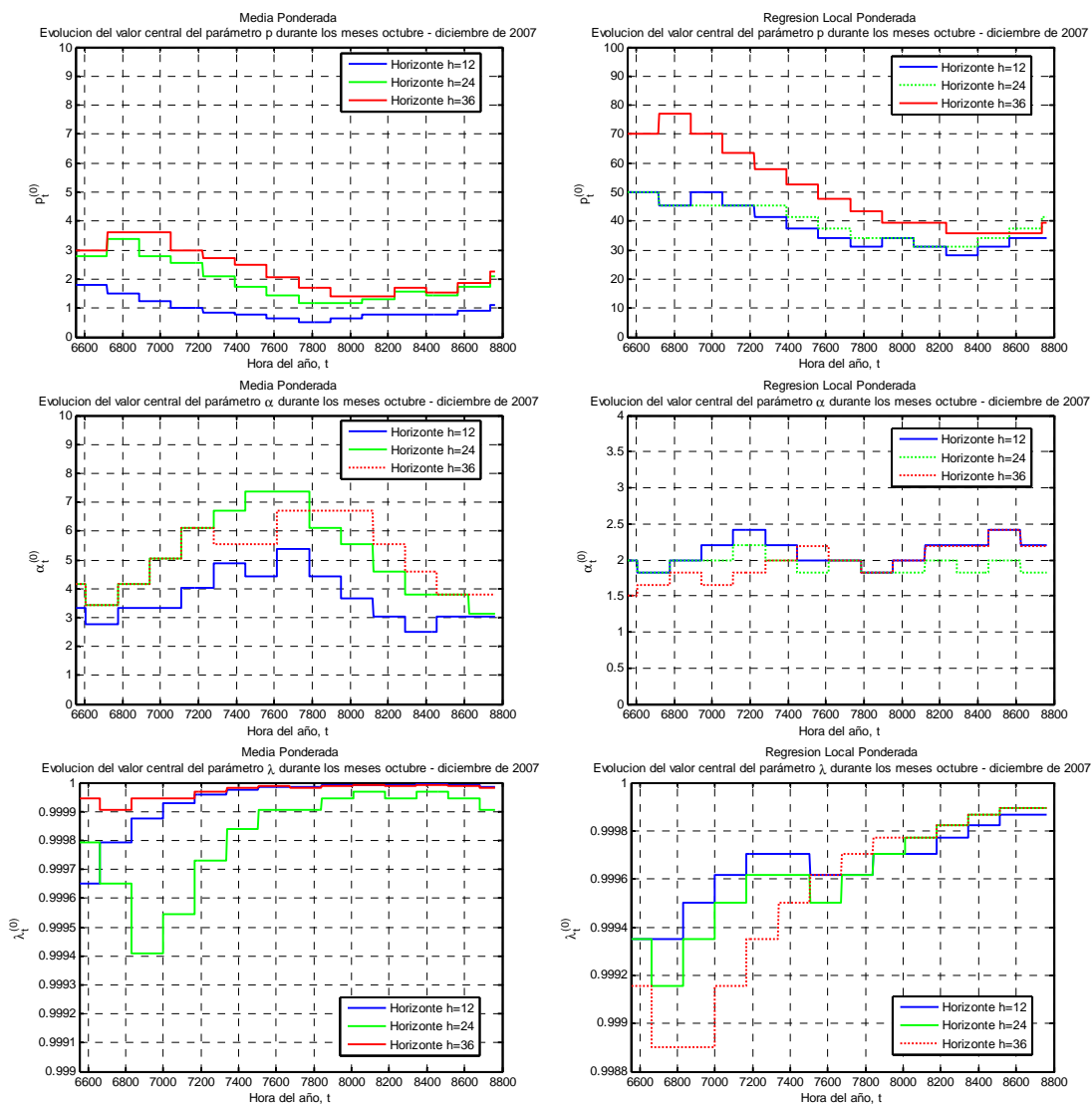


Figura 6.6: Evolución del valor central de los vectores de parámetros p , α y λ durante el periodo de octubre a diciembre de 2007

7. Combinación Adaptativa de Predicciones

Resumen

En este capítulo se propone combinar de forma adaptativa las predicciones obtenidas con las dos alternativas de modelado local: la Media Ponderada, y la Regresión Local Ponderada. De esta forma, se aprovecharán las ventajas de ambos modelos en la adaptación al problema de predicción, que depende de la complejidad local del mismo en cada instante.

El proceso de combinación utilizado es una combinación en dos etapas, que se beneficia de las ventajas que pueden ofrecer los dos enfoques básicos de combinación: el promedio óptimo y la combinación de mínimo riesgo.

Este procedimiento de combinación adaptativa también se puede utilizar para combinar predicciones de potencia que hayan sido calculadas a partir de predicciones de viento de distintas agencias meteorológicas. Por último, se comprobará si la combinación se puede aplicar a las distintas opciones posibles en el suavizado, como son las medidas de distancias, los métodos de selección de datos y las funciones de ponderación.

7.1 Introducción

En el ámbito de la energía eólica es bastante común que se tengan distintas predicciones de viento como datos de entrada al modelo de predicción de potencia, obtenidas de la ejecución de diferentes modelos meteorológicos. Además, estos datos meteorológicos pueden ser tratados con modelos distintos para convertirlos en predicciones de potencia eólica. En estos casos, en los que se puede disponer de varias predicciones de potencia, se deseará obtener una única predicción final que englobe las ventajas de todas las predicciones alternativas disponibles.

Dado que lo que se desea al final del proceso es una única predicción eólica, se tratará de hacer una combinación de las predicciones alternativas que fueron obtenidas de modos o fuentes diferentes. El objetivo último del proceso de combinación será que la predicción final resultante sea mejor o al menos igual de precisa que la mejor de las predicciones que se combinan.

La precisión relativa de los diferentes modelos a combinar puede variar a lo largo del tiempo o en función del horizonte de predicción, ya sea porque las predicciones meteorológicas de entrada sean diferentes, o porque varíe la precisión de las distintas aproximaciones para obtener las predicciones de potencia, dependiendo de su capacidad para capturar en cada momento la relación no lineal existente entre el viento y la potencia eólica generada. Por tanto, el peso o valor que se le deberá dar a cada predicción alternativa en la combinación para obtener la mejor predicción también será variable con el tiempo.

Por otro lado, para que un procedimiento de combinación de predicciones sea útil en nuestro contexto, es necesario que sea flexible y adaptativo, es decir, debe ser capaz de aprender y ajustarse automáticamente en cada momento, de modo que no sea necesaria la intervención de un usuario experto. Esto es necesario debido a que normalmente un algoritmo de predicción de potencia eólica estará funcionando en tiempo real, tratando datos de entrada y proporcionando informes de salida sin supervisión constante.

7.2 Modelos de combinación

Existen distintos métodos para combinar predicciones, y aquí se describirá un procedimiento de combinación en dos etapas que tratará de sacar partida a las ventajas de los diferentes métodos de combinación. En una primera etapa se utilizarán dos métodos alternativos de combinación de predicciones, el promedio óptimo y la combinación de mínimo riesgo, basados en propiedades diferentes y que presentan ciertas ventajas e inconvenientes cada uno. En una segunda etapa se utilizarán las dos combinaciones obtenidas en la primera etapa, combinándolas de nuevo para formar la predicción final, siendo este segundo paso, por tanto, una combinación de las combinaciones de la primera etapa.

Métodos de Predicción de la Generación Agregada de Energía Eólica

Este procedimiento es el mismo que utiliza la herramienta Sipleólico para la combinación de modelos de predicción de potencia en parques eólicos [61], y también se aplicó en el proyecto ANEMOS para combinar las predicciones generadas por los distintos modelizadores participantes y obtener la predicción combinada final [63].

Independientemente del método de combinación empleado en cada etapa, la formulación general de un modelo de combinación en el que se pretenden combinar Z predicciones alternativas será de la forma (6.17), que se recuerda a continuación:

$$\hat{P}_{t+h|t}^{(C)} = \sum_{z=1}^Z w_{z,t}(h) \cdot \hat{P}_{t+h|t}^{(z)} \quad (7.1)$$

donde:

$\hat{P}_{t+h|t}^{(C)}$ = Predicción final, fruto de la combinación de predicciones alternativas, calculada en el instante t para el instante $t+h$

$w_{z,t}(h)$ = Coeficiente de ponderación de la predicción alternativa z de horizonte h , calculado en t

$\hat{P}_{t+h|t}^{(z)}$ = Predicción alternativa z para el instante $t+h$, calculada en t

Los parámetros de interés cuando se combinan varias predicciones son los pesos o coeficientes de ponderación $w_{z,t}(h)$, que indicarán la importancia relativa de cada predicción individual a combinar.

Para que exista cierta adaptabilidad en la combinación, estos coeficientes de ponderación deberán ir variando con el tiempo, en función de la precisión relativa de cada predicción alternativa, y su valor deberá ser recalculado en cada periodo t en función de la información que se va acumulando sobre el error cometido por los predictores alternativos.

Los distintos métodos de combinación de predicciones se basan en diferentes esquemas de cálculo de la evolución de estos coeficientes. En la literatura se pueden encontrar dos enfoques o aproximaciones básicas para realizar combinación de predicciones, en función del objetivo principal de la combinación, que son la combinación de promedio óptimo y la combinación de riesgo mínimo [61],[119].

7.2.1 Combinación de promedio óptimo

La primera clase de métodos de combinación es la que se conoce como método de promedio óptimo. Según este enfoque, el objetivo es conseguir la mejor combinación lineal de un conjunto de predicciones, con el fin de mejorar la predicción final. Para realizar esta combinación óptima se suelen usar métodos basados en regresión, enfocados a minimizar la varianza residual de la combinación lineal.

Si se considera el error cuadrático medio de la predicción combinada como función de pérdidas a minimizar, y si se tratan los errores de predicción de las alternativas como si fueran independientes, los coeficientes de combinación dependerán únicamente de la precisión relativa de los modelos de predicción a los que representan [120]. Además, teniendo en cuenta que dado que las predicciones alternativas contribuirán en cierta fracción de la predicción final, la suma de todos los pesos debe ser uno, con lo que los coeficientes de combinación pueden ser calculados de la forma (6.18), que se repite a continuación:

$$w_{z,t}(h) = \frac{\frac{1}{NMSE_{t|t-h}^{(z)}}}{\sum_{j=1}^Z \frac{1}{NMSE_{t|t-h}^{(j)}}} \quad (7.2)$$

donde:

$NMSE_{t|t-h}^{(z)}$ = Estimación en el instante t del error cuadrático medio normalizado de las predicciones de horizonte h del modelo de predicción z

Puesto que se desea que los coeficientes de combinación sean variables en el tiempo, es necesario un método de estimación adaptativa de estos coeficientes a partir de los datos disponibles, y por tanto será necesaria una estimación adaptativa de los errores de cada predictor alternativo.

La estimación de los errores $NMSE_{t|t-h}^{(z)}$ puede ser obtenida mediante un procedimiento de media móvil ponderada exponencialmente o EWMA, similar al descrito en la sección 6.4. Así, cada vez que en un instante t llegue un nuevo dato de potencia medida en la región, P_t , el nuevo error en el instante t para cada predicción alternativa se calculará siguiendo el procedimiento de las ecuaciones (6.12) a (6.15).

De nuevo, la estimación mediante EWMA dependerá de un factor de olvido λ_{MSE} , con una memoria asintótica asociada, o número de muestras equivalentes $N_{z,t}$. Es importante destacar que la elección del factor de olvido es la pieza clave de la adaptación. A medida que λ_{MSE} sea más pequeño, menor será la influencia de los datos antiguos en la estimación, pero mayor puede ser la varianza del error. Así, la elección del factor de olvido es un compromiso entre la habilidad de percibir los cambios y la necesidad de reducir la varianza del error de predicción.

En cada periodo de cálculo se deberá actualizar por tanto el valor del error cuadrático medio $NMSE_{t|t-h}^{(z)}$, correspondiente a cada predictor alternativo z , y una vez calculado el error, la combinación de predicciones de promedio óptimo sería la resultante de aplicar la ecuación (7.1) con los coeficientes calculados en (7.2).

Idealmente, la combinación lineal siguiendo un método de promedio óptimo mejoraría las predicciones individuales. En la práctica, no está claro si se llega o no a la combinación ideal. Esta mejora sólo tendrá lugar en determinadas condiciones,

siendo posible que la predicción combinada obtenida pudiera ser peor que alguna de las predicciones originales individuales, especialmente cuando se trate de un problema no estacionario o con pocos datos. En estos casos, sería mejor solución hacer una selección de la mejor predicción en vez de una combinación de predicciones, aunque con el inconveniente de que no se puede conocer a priori cual será el mejor predictor.

7.2.2 Combinación de mínimo riesgo

La segunda clase de métodos de combinación es el llamado método de mínimo riesgo. En este caso se busca una combinación que funcione tan bien como el mejor de los predictores individuales. Esta segunda clase puede ser interpretada como un procedimiento de selección dinámica de modelos, en el que la combinación tiende a dar todo el peso al mejor modelo de predicción, cuando parezca claro que alguno de los predictores es el mejor. Por tanto, en este caso no se busca mejorar las predicciones alternativas.

Pero la simple selección del mejor predictor no suele ser la mejor opción, ya que los posibles cambios en la precisión de las alternativas harían que se estuviese cambiando continuamente de modelo seleccionado, y esa variabilidad perjudicaría a la predicción final. Para evitar esta variabilidad, la mejor opción será de nuevo una ponderación o promediado de los predictores alternativos, de modo que se añada cierto suavizado al proceso de selección, pero con la tendencia de otorgar cada vez mayor peso al modelo que esté mostrando mejores resultados. Así, si un modelo es el mejor de forma indefinida, se le acabará asignando todo el peso y en la práctica se tendrá el equivalente a una selección pura.

Se pueden encontrar en la literatura diferentes procedimientos de combinación orientados a la selección. En [119] se propone el método AFTER (Aggregated Forecasts Through Exponential Re-weighting), en el que los coeficientes de la combinación se calculan en cada instante de la forma:

$$w_{z,t}(h) = \frac{\frac{1}{\sqrt{NMSE_{t|t-h}^{(z)}}} \times e^{-\left(\frac{(e_{t|t-h}^{(z)})^2}{2 \cdot NMSE_{t|t-h}^{(z)}}\right)} \times w_{z,t-1}(h)}{\sum_{i=1}^Z \left(\frac{1}{\sqrt{NMSE_{t|t-h}^{(i)}}} \times e^{-\left(\frac{(e_{t|t-h}^{(i)})^2}{2 \cdot NMSE_{t|t-h}^{(i)}}\right)} \times w_{i,t-1}(h) \right)} \quad (7.3)$$

donde $e_{t|t-h}^{(z)} = P_t - \hat{P}_{t|t-h}^{(z)}$, es el último error de predicción observado.

Según la expresión (7.3), en los coeficientes $w_{z,t}(h)$ se está teniendo en cuenta el error cuadrático medio $NMSE_{t|t-h}^{(z)}$, de forma que se da mayor peso a aquellos predictores con menor error cuadrático medio.

Al mismo tiempo se está teniendo en cuenta, en el término exponencial de la expresión, el último error cometido $e_{t|t-h}^{(z)}$ estandarizado con el error cuadrático medio, de forma que si el último error observado es sensiblemente superior al valor habitual esperado se tiende a reducir su peso asignado, pues esto indicaría un deterioro de ese predictor. Así pues, el coeficiente asignado a cada modelo es doblemente ponderado (Re-weighted), por un lado a través de su error cuadrático medio, y por otro lado con el último error a través del término exponencial.

Además, todo queda multiplicado por el coeficiente calculado en el periodo anterior $w_{z,t-1}(h)$, de modo que si de forma sistemática un predictor es mejor o peor que el resto, su peso relativo asignado irá creciendo o decreciendo en cada periodo gracias a la multiplicación acumulativa de coeficientes, con lo que se conseguiría el equivalente a una selección.

Sin embargo, el inconveniente de este método de cálculo de los coeficientes de combinación radica en el hecho de que se vayan multiplicando los pesos de diferentes periodos sin tener en cuenta la antigüedad de los mismos. De este modo, el coeficiente $w_{z,t}(h)$ está siendo afectado en la misma medida por el último error de predicción que por el anterior coeficiente $w_{z,t-1}(h)$, que a su vez estará siendo afectado en la misma medida por el error cometido en ese instante $t-1$ y por el anterior coeficiente $w_{z,t-2}(h)$, y así sucesivamente.

A medida que transcurra el tiempo, los errores de predicción más antiguos serán igual de importantes que el último error de predicción en la determinación del coeficiente $w_{z,t}(h)$. Así, incluso podría darse el caso de que en un determinado momento un predictor tuviese un error excesivamente grande, con lo que su coeficiente $w_{z,t}(h)$ se hiciera casi a cero, y desde ese momento ya quedaría atrapado en un valor nulo, pues en cada periodo futuro estaría siendo afectado el peso nulo de la ocasión en que su error fue muy elevado.

En un entorno variable como el de predicción eólica que nos ocupa, sería deseable que la influencia de los errores de predicción se adapte dinámicamente, de modo que la precisión más reciente de los modelos de predicción tenga más peso en el cálculo de los coeficientes que las precisiones obtenidas en periodos muy antiguos.

Para conseguir esta adaptabilidad, en [61] se propone una modificación del método AFTER. Esta modificación se conoce como método AEC (Adaptive Exponential Combination), que permite la evolución temporal de los coeficientes por medio de un factor de olvido.

Métodos de Predicción de la Generación Agregada de Energía Eólica

La ecuación del método AFTER puede ser escrita de la forma:

$$w_{z,t}(h) = \frac{A_{z,t}}{\sum_{i=1}^Z A_{i,t}} \quad (7.4)$$

En la que cada término $A_{z,t}$ es el producto de la nueva información de error, y el término calculado en el periodo anterior, es decir:

$$A_{z,t} = I_{z,t}A_{z,t-1} = I_{z,t}I_{z,t-1}A_{z,t-2} = \prod_{i=1}^t I_{z,i} \quad (7.5)$$

Siendo $I_{z,t}$ la nueva información de error de la forma:

$$I_{z,t} = \frac{1}{\sqrt{NMSE_{t|t-h}^{(z)}}} \times e^{-\left(\frac{(e_{t|t-h}^{(z)})^2}{2 \cdot NMSE_{t|t-h}^{(z)}}\right)} \quad (7.6)$$

En la ecuación (7.5) se puede apreciar que la información antigua es tan importante como la nueva en el cálculo del nuevo coeficiente de combinación, lo que implica una falta de adaptabilidad en el proceso de actualización de los coeficientes, tal y como se ha justificado.

Para introducir adaptabilidad en la estimación se reemplaza el término $A_{z,t}$ de la ecuación (7.5) por un nuevo término $B_{z,t}$, y en el cálculo de los nuevos términos $B_{z,t}$ se introduce un nuevo factor de olvido λ_{AEC} de la forma:

$$B_{z,t} = I_{z,t}B_{z,t-1}^{\lambda_{AEC}} = I_{z,t}I_{z,t-1}^{\lambda_{AEC}}B_{z,t-2}^{\lambda_{AEC} \cdot \lambda_{AEC}} = \prod_{i=1}^t I_{z,i}^{\lambda_{AEC}^{t-i}} \quad (7.7)$$

donde cada término $B_{z,t-i}$ de periodos anteriores se eleva a un exponente $0 < \lambda_{AEC} < 1$. Se tendrá así un olvido de la información sobre el error que tuvo lugar en periodos antiguos, gracias a que su aportación estará elevada a un exponente, que para la información del instante i será λ_{AEC}^{t-i} . Este exponente tenderá a cero conforme pase el tiempo, de modo que los aportes muy antiguos, al estar elevados a un número próximo a cero, tenderán a convertirse en una multiplicación por uno en el cálculo del nuevo coeficiente.

Por tanto, y como resumen, para poder aplicar el método AEC de combinación de riesgo mínimo se debe actualizar en cada periodo de cálculo el valor del error cuadrático medio $NMSE_{t|t-h}^{(z)}$ correspondiente a cada predictor alternativo z . Una vez actualizado este error medio, y con el valor del último error cometido $e_{t|t-h}^{(z)}$, se calculará la nueva aportación $I_{z,t}$ del error, y con ella se actualizan los términos $B_{z,t}$. Finalmente, los coeficientes de la combinación de predicciones alternativas según el método AEC de riesgo mínimo se calcularán de la forma:

$$w_{z,t}(h) = \frac{B_{z,t}}{\sum_{i=1}^Z B_{i,t}} \quad (7.8)$$

El inconveniente de la combinación según el método de mínimo riesgo en un entorno no estacionario es que el mejor modelo de predicción puede cambiar a lo largo del tiempo. El problema de combinación se convierte así en un problema de eficiencia en rastrear el mejor predictor en cada momento, con el inconveniente de que su funcionamiento será muy pobre cuando el mejor predictor individual esté continuamente cambiando, pues la variabilidad en la selección se traducirá en variabilidad en las predicciones. En estos casos, sería mejor una combinación de promedio óptimo en vez de una combinación de mínimo riesgo.

7.3 Procedimiento de combinación en dos etapas

En la práctica, se trabaja con el inconveniente de no conocer de antemano si existe un modelo de predicción mejor que el resto, o si será mejor usar un método de combinación de promedio óptimo para mejorar las predicciones originales o una combinación de riesgo mínimo orientada a seleccionar el mejor predictor. Por eso, con el objetivo de beneficiarse de las características de ambos esquemas, se pueden aplicar los dos enfoques para combinar las predicciones de energía eólica en un procedimiento de dos etapas [61].

En la primera etapa se utilizan los dos tipos de combinaciones propuestas, basadas en métodos de propiedades distintas, generándose dos predicciones combinadas alternativas. El primer método consiste en promediar las predicciones iniciales para tratar de mejorarlas, asignando por tanto un peso a cada predicción utilizando el método de promedio óptimo. El segundo método utilizado en esta primera etapa tiene como objetivo seleccionar o rastrear el mejor predictor en cada momento, basándose en la minimización del riesgo según el método AEC expuesto, sin tratar de mejorar a las predicciones alternativas.

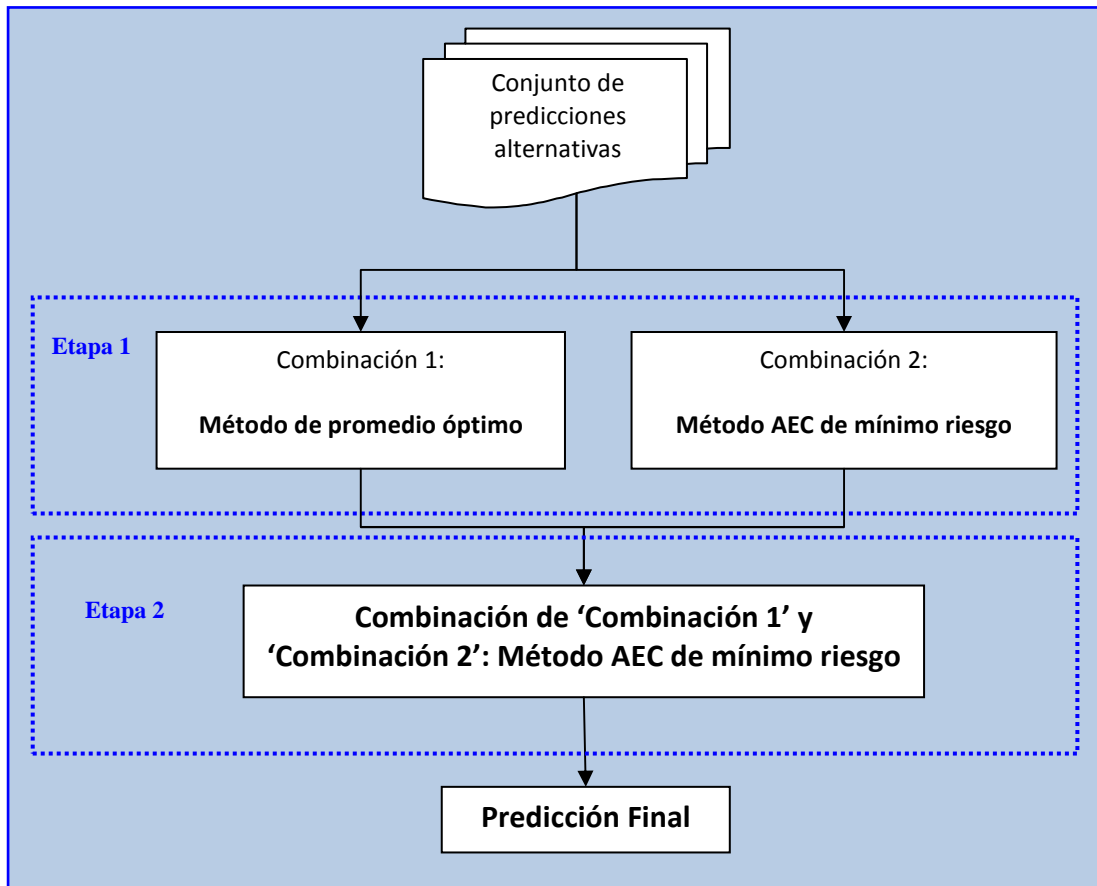


Figura 7.1: Procedimiento de combinación en dos etapas

Las dos predicciones combinadas alternativas, fruto de los dos enfoques de combinación descritos, se tratan de nuevo en un problema de combinación para obtener la predicción combinada final (Figura 7.1). En esta segunda etapa de la combinación se aplica únicamente el método AEC de riesgo mínimo a las dos combinaciones obtenidas en el primer paso. Así, en esta etapa no se espera mejorar las predicciones combinadas alternativas de la primera etapa, sino que el objetivo es asegurarse de que la predicción final sea tan buena como la mejor de las dos predicciones combinadas previas.

La ventaja principal de la combinación en dos etapas es que no es necesario conocer de antemano si existe un modelo de predicción claramente mejor que el resto, o si funcionará mejor un método de combinación de promedio óptimo o de mínimo riesgo, pues la segunda etapa de combinación descartará aquellos procedimientos de combinación que resulten ineficientes.

7.4 Evaluación de resultados obtenidos

7.4.1 Combinación de modelos de Media Ponderada y Regresión Local Ponderada

Aunque las pruebas realizadas en los capítulos anteriores con los datos de potencia eólica y viento previsto durante la segunda mitad de 2007 parecen indicar que el mejor modelo de suavizado para calcular las predicciones de potencia eólica es un modelo de regresión polinómica local ponderada, en la literatura no se encuentra una evidencia clara de que elegir un método de Regresión Local Ponderada sea mejor o peor que usar un modelo de Media Ponderada para calcular predicciones, especialmente en problemas tan complejos y con tantas dimensiones como el que nos ocupa [97].

Dependiendo de cómo sea la dispersión de datos a modelar localmente funcionará mejor un modelo u otro. Cuando la superficie a modelar sea más o menos plana o regular, un modelo de promedio local funcionará bien, mientras que para superficies con mayor curvatura funcionará mucho mejor un método de regresión polinómica local, aunque este último necesitará muchos más datos de entrenamiento, pues deberá estimar una cantidad mayor de parámetros.

Por tanto, dependiendo de las variables de entrada en cada momento es posible que funcione mejor un modelo de Media Ponderada en vez de un modelo de Regresión Local Ponderada. La elección de un modelo u otro se convierte en una competición entre varianza y sesgo, ya que modelos de mayor orden disminuyen el sesgo, pero puede aumentar la varianza de sus predicciones al tener que estimar más parámetros con el mismo número de datos.

En [103] se propone combinar predicciones calculadas con polinomios de distinto orden en función de la localización del punto de entrada para calcular la predicción final. Esta idea está orientada a determinar los coeficientes de ponderación que se aplicarán a cada modelo dependiendo de la localización del punto de entrada para el que se busca predicción, definido en nuestro caso por el vector de viento previsto. Si bien este procedimiento puede ser factible cuando el número de variables de entrada es reducido, puede ser irrealizable en un problema en el que el número de variables de entrada es muy elevado, ya que se necesitarían enormes cantidades de datos para poder barrer completamente el espacio de entrada y determinar cuáles serían los parámetros óptimos de combinación en cada punto de entrada.

Dado que la evolución del estado de la atmósfera suele ser relativamente lenta, se puede asumir que las variaciones en las velocidades de viento no serán demasiado grandes en todas las coordenadas de la región a la vez, de modo que la localización de los puntos a predecir no variará muy bruscamente con el tiempo. Gracias a esto último, en nuestro caso podremos estimar la combinación óptima de las predicciones de forma evolutiva y adaptativa con el tiempo. En vez de calcular los parámetros de la combinación de predicciones para cada localización del espacio de entrada,

Métodos de Predicción de la Generación Agregada de Energía Eólica

utilizaremos un parámetro de combinación para cada predicción alternativa que no dependerá del punto de entrada, pero que será variable con el tiempo, de modo que su valor se irá adaptando a las variaciones relativamente lentas de la velocidad de viento en las coordenadas, y por tanto a variaciones lentas en la localización del punto dentro del espacio de entrada.

Se aplicará el procedimiento de combinación en dos etapas descrito, con el objetivo de obtener una única predicción que mejore o al menos iguale a la mejor de entre las dos predicciones calculadas mediante el modelo de Media Ponderada y el modelo de Regresión Local Ponderada.

Utilizando los datos de potencia, viento, y las predicciones calculadas para la segunda mitad del año 2007 con los modelos ajustados a los valores de hiperparámetros p , α y λ que proporcionaron el menor error de predicción en cada caso, se ha calculado una nueva predicción combinada.

En la Figura 7.2 se muestran los errores NRMSE obtenidos con las combinaciones de promedio óptimo (MSE-1), que utiliza el NMSE para calcular los coeficientes, y mínimo riesgo (AEC-1) de la primera etapa. También se muestra el resultado de la combinación final (AEC-2) de la segunda etapa. Las curvas que se muestran son las que se obtienen con un factor de olvido $\lambda_{MSE} = \lambda_{AEC} = 0.9972$, que corresponden a una memoria asintótica de 360 horas, valor con el que se obtuvieron los mejores resultados.

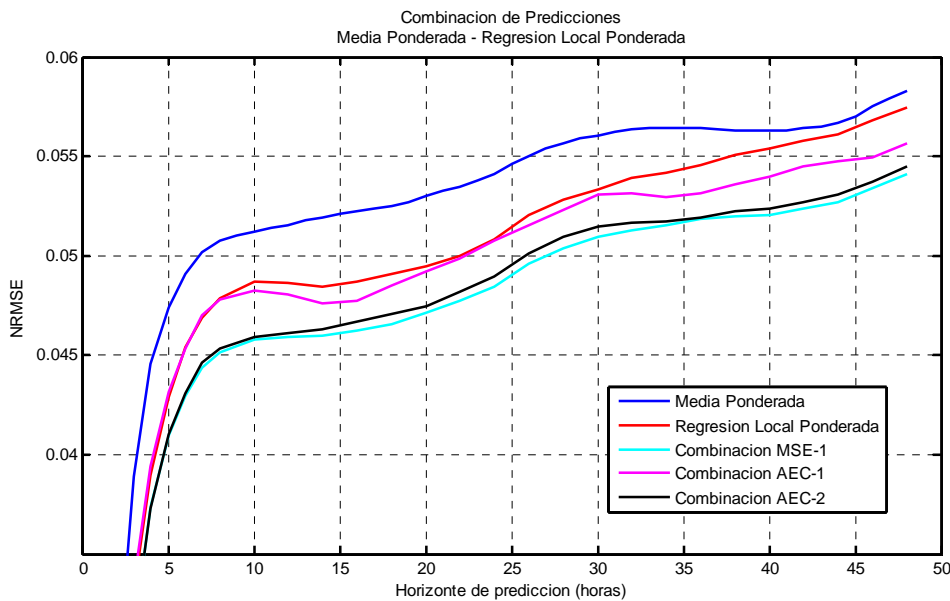


Figura 7.2: Combinación en dos etapas de predicciones basadas en modelos de Media Ponderada y Regresión Local Ponderada

Se puede comprobar que la predicción combinada de promedio óptimo de la primera etapa (MSE-1) mejora notablemente el resultado de las predicciones originales de los modelos de Media Ponderada y Regresión Local Ponderada, mientras que la predicción combinada de la primera etapa (AEC-1) apenas mejora la predicción de Regresión Local Ponderada. Sin embargo, en la segunda etapa, la combinación AEC-2 consigue seleccionar entre las predicciones de las dos combinaciones previas, MSE-1 y AEC-1, aproximándose a la mejor de ambas, es decir, a la combinación de promedio óptimo MSE-1, que obtuvo el menor error.

En la Figura 7.3 se muestra la influencia de la elección de los factores de olvido λ_{MSE} y λ_{AEC} , o más concretamente de la memoria asintótica equivalente. Se representa la mejora porcentual, en términos de NRMSE medio de todos los horizontes, de la predicción combinada de la segunda etapa, AEC-2 respecto de la mejor de las predicciones originales, en este caso la correspondiente al modelo de Regresión Local Ponderada.

Si bien el valor máximo de mejora se obtiene para una memoria asintótica de 360 horas, con una mejora del 5%, se puede observar que el porcentaje de mejora se mantiene prácticamente en el mismo valor, entre el 4.9 y 5%, para un rango muy amplio de valores de memoria asintótica, lo que indica que la elección del factor de olvido no es excesivamente crucial. En la mayoría de las aplicaciones de estimación adaptativa se suele usar un factor de olvido constante en el entorno entre $0.950 < \lambda < 0.999$. Aquí se utilizará un factor de olvido $\lambda_{MSE} = \lambda_{AEC} = 0.9972$, equivalente a 360 horas de memoria asintótica, para el resto de procedimientos de combinación.

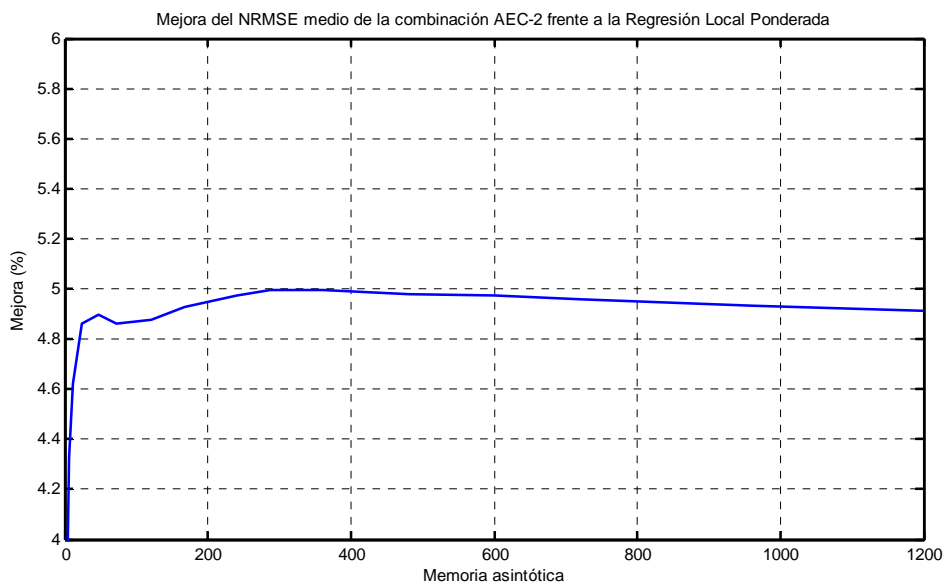


Figura 7.3: Coeficiente de mejora o disminución del error NRMSE en función de la memoria asintótica del modelo de combinación

7.4.2 Combinación de predicciones obtenidas a partir de distintas fuentes de viento

Los modelos de predicción meteorológicos resuelven las ecuaciones de la atmósfera sobre una malla de puntos en el espacio a partir de una serie de condiciones iniciales y de contorno estimadas a partir de las observaciones en estaciones de medida u otros medios de observación. Con estas condiciones iniciales se calcula la evolución de la atmósfera mediante un modelo numérico de predicción, y se obtiene la evolución que tomarán los valores de las variables de interés, que en el caso de generación eólica suelen ser la velocidad y dirección del viento, y en menor medida la temperatura o presión del aire.

La evolución prevista de la atmósfera puede ser muy variable, ya que depende en gran medida del modelo numérico de predicción meteorológica utilizado, de las condiciones iniciales y de contorno, o de la resolución de la malla de puntos que se use para resolver el problema. Por ello, las predicciones de viento que se utilizan en el modelo de predicción de potencia eólica pueden ser bastante distintas dependiendo de la agencia de previsión meteorológica que las proporcione.

Cuando se tienen varias predicciones de potencia obtenidas de fuentes distintas, como pueden ser distintas fuentes de predicción meteorológica, también puede resultar muy atractivo combinar las predicciones alternativas de potencia, para aprovechar la información contenida en cada una de ellas, y conseguir así una predicción de potencia más precisa.

Se ha comprobado este tipo de combinación utilizando una segunda agencia de predicción meteorológica (Agencia B). Estos datos de viento consisten en la velocidad del viento previsto durante el mismo periodo que la Agencia A usada con anterioridad en esta tesis, con predicciones en 79 coordenadas en este caso. Estas predicciones de viento en las coordenadas se actualizan cada 6 horas y tienen una resolución horaria.

Con esta nueva fuente de predicción de viento en las coordenadas de la región se han calculado las predicciones de potencia agregada obtenidas con un modelo de Media Ponderada y con un modelo de Regresión Local Ponderada, utilizando en ambos casos la misma medida de distancia, método de selección de datos, y funciones de ponderación que fueron utilizadas con la primera fuente de predicción de viento, aunque con los parámetros de selección y ponderación ajustados para obtener los mejores resultados con esta segunda fuente.

Las dos predicciones de potencia obtenidas con un modelo de Media Ponderada y de Regresión Local Ponderada también son combinadas para esta segunda agencia meteorológica siguiendo un proceso de combinación en dos etapas (Figura 7.4). Como se puede comprobar, se obtiene también una mejoría en la predicción combinada final (AEC-2), con respecto a las dos predicciones originales.

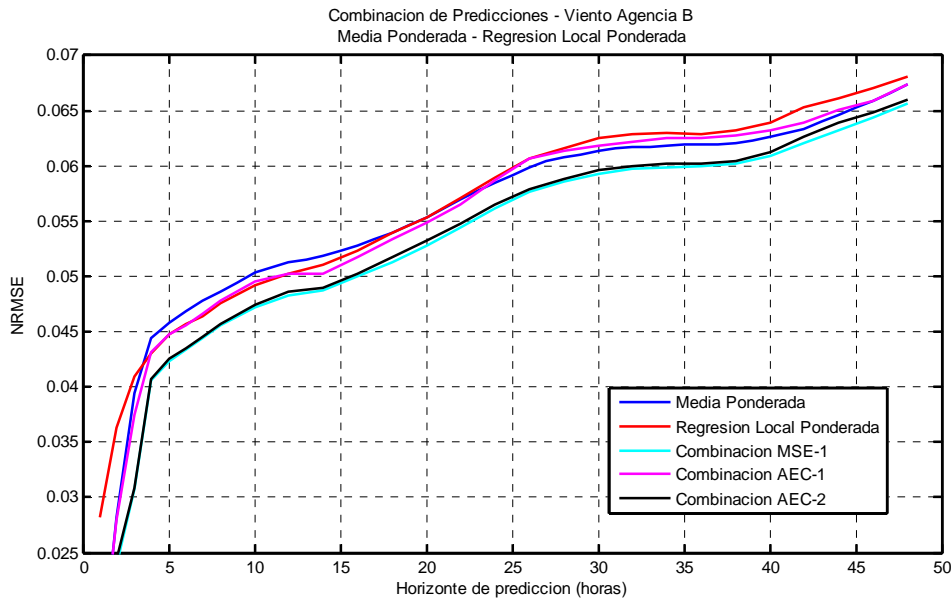


Figura 7.4: Combinación en dos etapas de predicciones basadas en modelos de Media Ponderada y Regresión Local Ponderada, obtenidas con una segunda fuente de viento

Una vez obtenidas las predicciones finales de cada una de las dos agencias (Agencia A y Agencia B), éstas se combinan en un nuevo proceso de combinación en dos etapas (Figura 7.5). Como puede observarse, la predicción de potencia obtenida con el viento de la Agencia A presenta en general un menor error que la predicción de potencia obtenida con el viento de la nueva Agencia B para todos los horizontes de predicción.

En la primera etapa se obtiene la predicción combinada de promedio óptimo (MSE-1), que mejora en gran medida a las dos predicciones originales, y la predicción combinada de mínimo riesgo (AEC-1), que no consigue mejorar a la mejor de las predicciones originales, especialmente para horizontes superiores a 22 horas. En la segunda etapa, de nuevo un procedimiento de combinación por el método del mínimo riesgo selecciona entre las dos predicciones combinadas de la primera etapa, obteniendo la predicción final combinada (AEC-2), que resulta casi tan precisa como la mejor de las dos predicciones combinadas anteriores.

La mejora porcentual media obtenida gracias a la adición de la nueva fuente de previsión de viento, en términos de error NRMSE, se sitúa en un 8.6% sobre el error obtenido al utilizar únicamente el viento de la agencia A, a pesar de que aparentemente la nueva agencia B de viento no parecía aportar demasiado, pues el error de predicción de potencia obtenido con sus predicciones de viento es superior al de la primera agencia en todos los horizontes.

Métodos de Predicción de la Generación Agregada de Energía Eólica

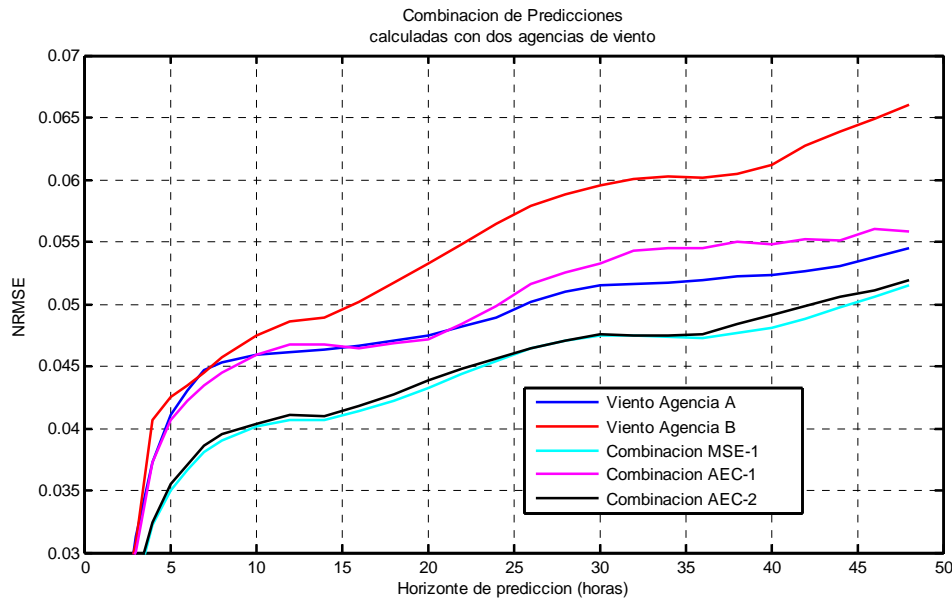


Figura 7.5: Combinación en dos etapas de predicciones obtenidas a partir de distintas fuentes de viento

A pesar de que las dos agencias proporcionan predicciones de viento para casi las mismas coordenadas geográficas de la región, cada una de ellas utiliza un modelo numérico de predicción meteorológica distinto, con resoluciones y condiciones iniciales diferentes, de modo que la evolución de las predicciones de viento será distinta para cada una de ellas. Dependiendo del estado y la evolución de la atmósfera, en unas ocasiones el modelo meteorológico de una agencia proporcionará mejores resultados que la otra, y viceversa, con lo que la bondad de las predicciones de potencia calculadas también irá variando entre una agencia y otra.

Los resultados obtenidos muestran la idoneidad de utilizar distintas fuentes de datos para resolver el mismo problema de predicción, ya que cuando estas fuentes son independientes y proporcionan errores razonablemente similares, la combinación de ambas puede dar lugar a predicciones mucho más precisas.

7.4.3 Análisis de otras posibilidades de combinación

Tras observar los buenos resultados derivados de la combinación de predicciones obtenidas con diferentes fuentes de viento, o de combinar modelos de predicción de distinto orden como la Media Ponderada y la Regresión Local Ponderada, cabe la posibilidad de plantearse combinar otras alternativas, como puedan ser distintas medidas de distancia, diferentes métodos de selección de datos, o utilizar las funciones Kernel para combinarlas con el método propuesto de tratar de forma independiente la medida de distancia, la selección y la ponderación.

Para analizar y en su caso aceptar o descartar estas posibilidades, se han comprobado los resultados de combinar en cada caso los dos procedimientos con los que se obtuvieron los mejores resultados.

Combinación de medidas de distancia

En el caso de las medidas de distancias, se ha probado a combinar las predicciones que se obtienen utilizando las medidas de distancia absoluta estandarizada con la media, y la distancia euclídea estandarizada con la desviación típica, utilizando en ambos casos un modelo de Media Ponderada y un método de selección de porcentaje de datos.

Ajustando los parámetros p , α y λ a los valores con los que se obtienen los mejores resultados en cada caso, se obtienen los errores NRMSE que se representan en la Figura 7.6, para cada tipo de distancia a combinar. Por medio de un procedimiento de combinación en dos etapas como el propuesto, se calcula la combinación de ambas predicciones, obteniendo la combinación final AEC-2 cuyo error también se representa en la figura.

Se puede observar que la predicción combinada se ajusta a la mejor de las dos predicciones originales, obtenida con la distancia absoluta estandarizada con la media, pero sin conseguir mejorarla, salvo tal vez para algunos horizontes en los que la mejora es casi despreciable, con lo que la combinación de medidas de distancia parece no merecer la pena.

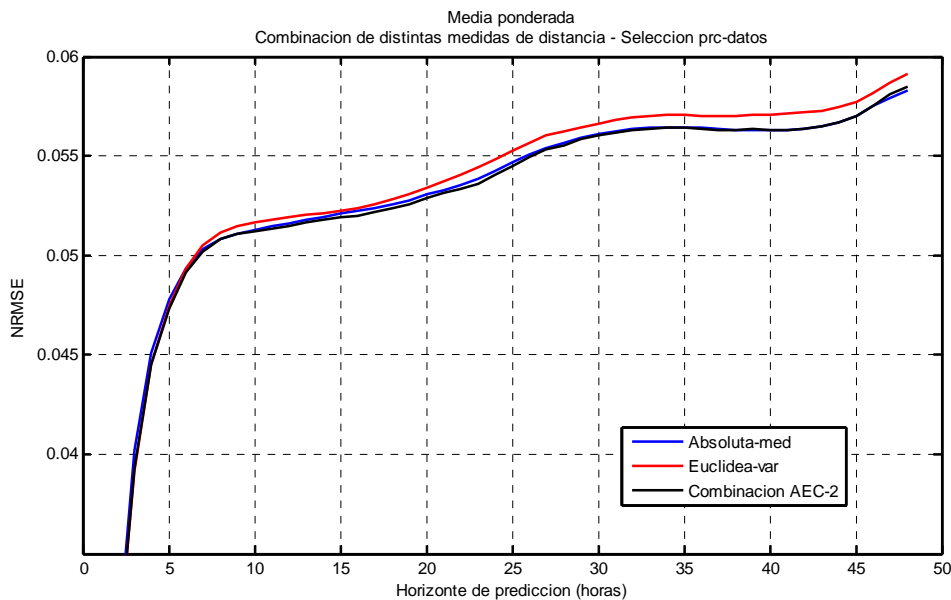


Figura 7.6: Combinación de predicciones obtenidas a partir de distintas medidas de distancia entre vectores de viento

Combinación de métodos de selección

De un modo similar se han combinado las predicciones obtenidas utilizando los dos métodos de selección que habían proporcionado mejores resultados en las pruebas correspondientes. Utilizando como medida de distancia la distancia absoluta estandarizada con la media y un modelo de predicción de Media Ponderada, se han calculado las predicciones de potencia con los métodos de selección de un porcentaje de datos, y de un porcentaje del rango de distancias hasta la mediana.

En la Figura 7.7 se muestran los errores NRMSE que se obtienen con los dos procedimientos de selección comparados, y el error de la predicción final AEC-2 en la segunda etapa de combinación.

De nuevo, se puede comprobar que la predicción combinada no es capaz de mejorar a las predicciones originales, sino que sencillamente se ajusta a la mejor predicción de ambas, la que utiliza un porcentaje de datos como método de selección, con lo que tampoco parece buena idea tratar de combinar predicciones utilizando más de un método de selección.

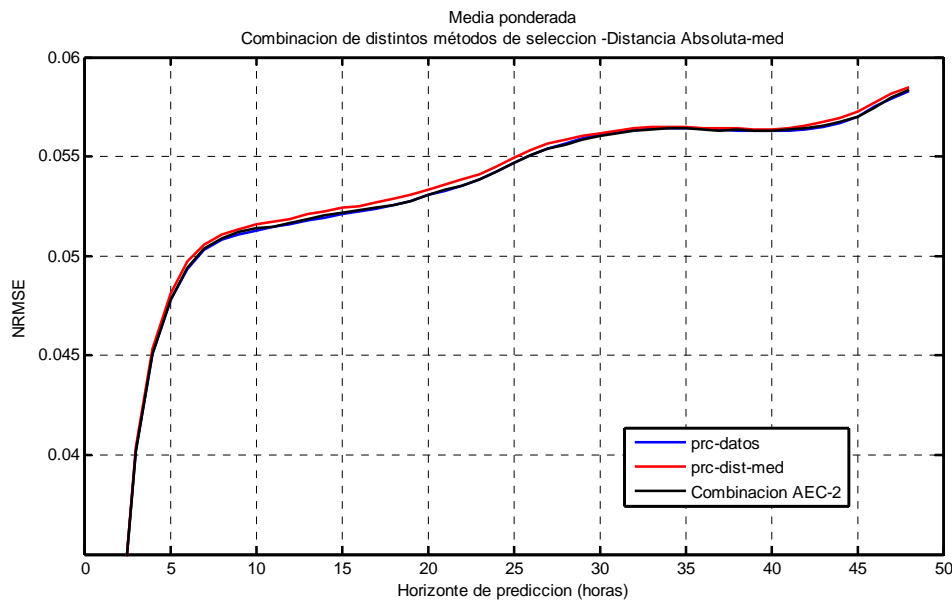


Figura 7.7: Combinación de predicciones obtenidas a partir de distintos métodos de selección de datos cercanos

Combinación de métodos de suavizado

Por último, se ha tratado la posibilidad de combinar predicciones de modelos que utilicen diferentes procedimientos o métodos de suavizado. En este caso, se ha tratado de combinar las predicciones obtenidas con funciones Kernel de suavizado y las predicciones obtenidas con el método de suavizado propuesto en esta tesis, en el que se trata de forma independiente la medida de distancias, la selección de datos en función de su distancia, y la función de ponderación de datos seleccionados.

En la figura siguiente se muestran los errores de predicción NRMSE obtenidos al utilizar el método de suavizado propuesto en esta tesis, y los errores obtenidos al utilizar una función Kernel triangular, que fue la que proporcionó mejores resultados de todas las funciones Kernel comprobadas. En ambos casos, se ha hecho una combinación previa de modelos de Media Ponderada y Regresión Local Ponderada con el fin de tener la mejor predicción final en cada caso.

De nuevo, siguiendo un procedimiento de combinación en dos etapas se obtiene una predicción combinada AEC-2 final, que no mejora la mejor de las predicciones originales, correspondiente al método de suavizado propuesto, con lo que se comprueba que no ha sido posible mejorar las predicciones mezclando diferentes métodos de suavizado.

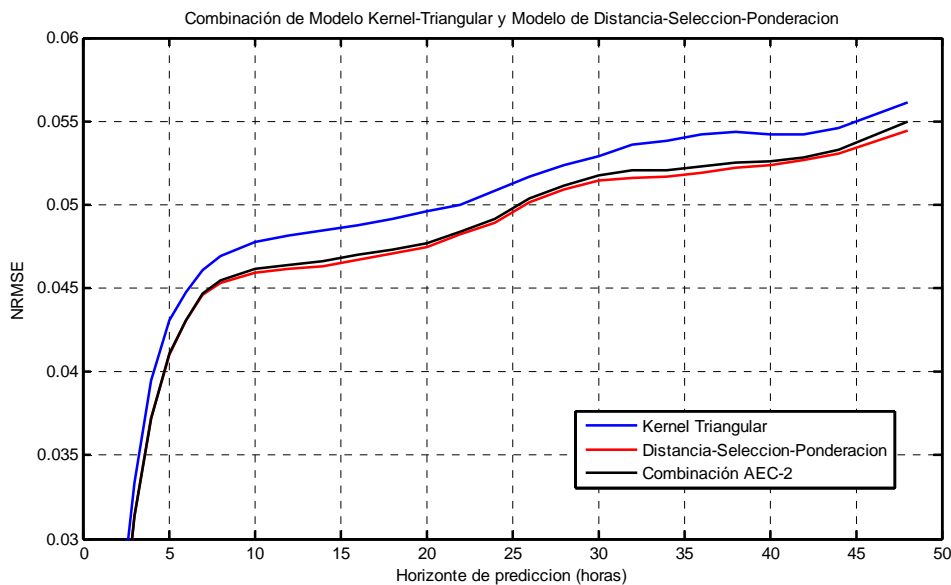


Figura 7.8: Combinación de predicciones obtenidas a partir de modelos que utilizan distintos métodos de suavizado

8. Evaluación del Modelo Propuesto de Predicción de Potencia Eólica Agregada

Resumen

En este capítulo se hace una evaluación de los resultados obtenidos con el modelo de predicción de potencia agregada propuesto en esta tesis, comparando sus predicciones con las obtenidas por el modelo de predicción Sipleólico, en funcionamiento en REE, y con las predicciones de otro modelo de predicción agregada.

8.1 Introducción

Para finalizar el estudio realizado y comprobar el funcionamiento del modelo de predicción regional propuesto en esta tesis, se compararán las predicciones generadas por nuestro modelo con las predicciones generadas por otros modelos de predicción.

La comparación se hará utilizando datos de potencia eólica y viento correspondientes a un periodo distinto al que ha sido utilizado en los capítulos precedentes, de manera que se compruebe el funcionamiento del modelo fuera de la muestra de pruebas y entrenamiento.

Los datos de potencia utilizados en el periodo de evaluación corresponden a la producción eólica total en el sistema eléctrico peninsular español, obtenidos como la integral horaria de la energía eólica registrada durante las 8761 horas del año 2008. Como datos históricos iniciales se utilizarán los datos correspondientes al año 2007, utilizados en los capítulos precedentes. Estos datos han sido facilitados por Red Eléctrica de España.

Se utilizarán las predicciones de viento de dos agencias meteorológicas (Agencia A y Agencia B), con el fin de comprobar la utilidad de combinar las predicciones de potencia calculadas a partir de distintas fuentes de predicción de viento. Ambas agencias proporcionan predicciones de viento con un horizonte superior a 48 horas, horizonte máximo para el que se hará la evaluación. Independientemente de la agencia de viento, dado que no se tienen medidas de viento en las coordenadas, en los datos del periodo histórico se considerarán como medidas de viento en las coordenadas las predicciones de menor horizonte que se tengan para cada hora.

Los datos de viento de la Agencia A consisten en la velocidad de viento prevista durante los años 2007 y 2008 en 83 coordenadas seleccionadas repartidas por el territorio peninsular, cercanas a los emplazamientos de los parques eólicos. Estas predicciones de viento en las coordenadas se actualizan cada 12 horas y tienen una resolución de 6 horas. Para obtener la predicción de viento a escala horaria se ha realizado una interpolación lineal de las predicciones originales de 6 horas de resolución.

Los datos de viento de la Agencia B consisten en la velocidad de viento prevista durante los años 2007 y 2008 en 79 coordenadas seleccionadas repartidas por el territorio peninsular, cercanas a los emplazamientos de los parques eólicos. Estas predicciones de viento en las coordenadas se actualizan cada 6 horas y tienen una resolución horaria.

La evaluación se realiza calculando el NRMSE de las predicciones de potencia en función del horizonte de predicción durante el periodo comprendido entre la primera hora del año 2008 y la última hora del año 2008 (8761 horas). Se considera además que, cada vez que el modelo calcula las predicciones de potencia en una hora determinada, la medida de potencia y las predicciones de viento en las coordenadas

Métodos de Predicción de la Generación Agregada de Energía Eólica

para esa hora pasan a formar parte del registro histórico almacenado para ser utilizadas en las horas siguientes como datos de entrenamiento. Así, en la primera hora del año 2008 se tendrá un histórico de 8760 horas, correspondientes al histórico previo del año 2007, en la segunda hora se tendrá un registro histórico 8761 horas, y así sucesivamente hasta que en la última hora del año 2008 se tenga un registro histórico previo almacenado de $8760+8760 = 17520$ horas.

Las predicciones de potencia total peninsular calculadas por el modelo de predicción agregada propuesto en esta tesis se comparan con las predicciones de potencia total peninsular que generó la herramienta de predicción Sipreólico durante las mismas horas del año 2008. Se debe tener en cuenta que Sipreólico calcula las predicciones de potencia total peninsular sumando las predicciones de potencia de cada uno de los parques eólicos instalados en la península. Se trata por tanto de una configuración en cascada, frente a la configuración directa del modelo de predicción propuesto en esta tesis.

También se incluye en la comparación la predicción de potencia de otro modelo de predicción regional, que cada 12 horas proporciona predicciones de potencia eólica total peninsular con una resolución de 1 hora y con un horizonte de predicción superior a las 48 horas que son objeto de este estudio.

8.2 Modelo de predicción utilizado

El modelo de predicción propuesto en esta tesis consiste en la combinación adaptativa de las predicciones obtenidas con los modelos locales de predicción de Media Ponderada (3.3) y Regresión Local Ponderada (3.11).

Las funciones de distancia y funciones de suavizado que utilizará el modelo de predicción durante el periodo de 2008 de evaluación final serán las mismas que proporcionaron los mejores resultados durante el periodo de 2007, con el que se han ido evaluando las diferentes opciones en el desarrollo de esta tesis.

Por tanto, la medida de distancias utilizada para medir la similitud entre vectores de predicción de viento ha sido la distancia absoluta estandarizada con la media, según la ecuación (4.16). El método de selección de datos cercanos es el descrito en la sección 4.5.3 como un porcentaje p de datos cercanos. La función de ponderación por distancia es la (4.30), correspondiente a una ponderación inversa a la distancia con el parámetro α corregido con la mediana de las distancias calculadas en cada instante. Por último, la función de ponderación por la antigüedad de los datos es la (4.33), dependiente de un factor de olvido λ .

En todos los casos, los parámetros correspondientes p , α y λ serán estimados de forma adaptativa mediante el procedimiento descrito en el capítulo 6, partiendo de unos valores iniciales estimados mediante entrenamiento con los datos del segundo semestre de 2007.

8.2.1 Comprobación de la estimación adaptativa de hiperparámetros

En primer lugar se comprobará el funcionamiento de la estimación evolutiva o adaptativa de los hiperparámetros de los modelos de predicción. Para ello, se calculan las predicciones obtenidas durante el año 2008 partiendo de los parámetros ajustados a cada modelo de forma global con los datos del segundo semestre del año 2007.

Se han calculado las predicciones obtenidas por los modelos de Media Ponderada y de Regresión Local Ponderada, tanto manteniendo fijos los parámetros estimados durante el entrenamiento como haciendo una estimación adaptativa de los mismos.

Para las dos agencias de viento A y B, el tamaño de los vectores de parámetros para los modelos de Media Ponderada se ha establecido en cinco elementos para cada vector. Para el modelo de Regresión Local Ponderada se establece el tamaño de estos vectores en solo tres elementos, debido a que este tipo de modelo requiere mucha más carga de cálculo. Los coeficientes de incremento que definen la relación entre los valores de dos elementos consecutivos en los vectores de parámetros se han establecido en $r_p=1.1$, $r_\alpha=1.1$ y $r_\gamma=1.3$. El periodo de actualización para la estimación adaptativa se ha fijado en $N_{z,t} = 168$ horas, equivalente a un factor de olvido $\lambda_{MSE} = 0.994$, que fue el que obtuvo mejores resultados en la evaluación correspondiente del capítulo 6.

En la Figura 8.1 se muestran los NRMSE de las predicciones obtenidas utilizando como entrada al modelo las predicciones de viento de la Agencia A. Se representan los errores de los modelos de Media Ponderada y Regresión Local Ponderada, tanto utilizando unos parámetros fijos como estimándolos de forma adaptativa.

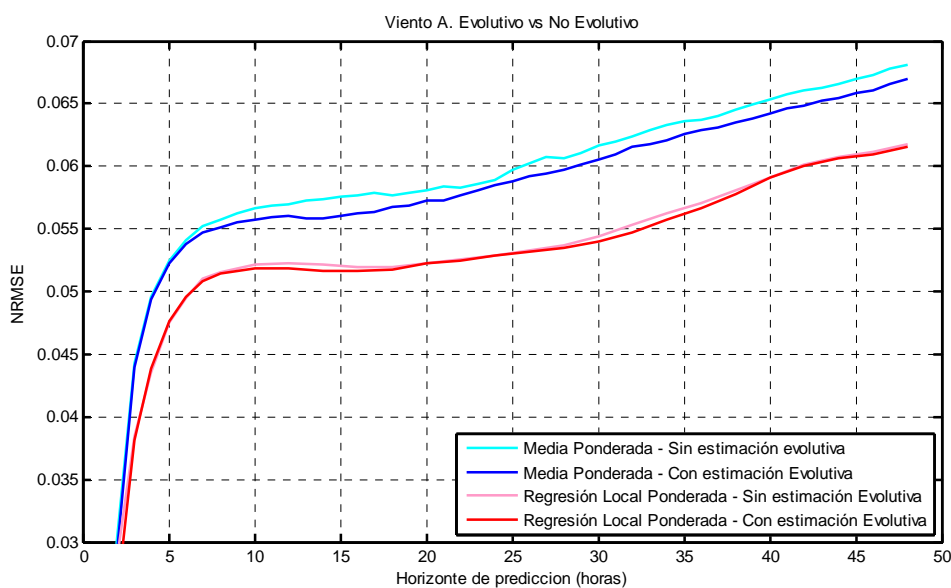


Figura 8.1: Error de las predicciones obtenidas con la agencia de Viento A, con modelos de Media Ponderada y Regresión Local Ponderada, y estimaciones evolutiva y no evolutiva de hiperparámetros

Métodos de Predicción de la Generación Agregada de Energía Eólica

Se puede comprobar que tanto si se utiliza un modelo local de Media Ponderada como una Regresión Local Ponderada, las predicciones obtenidas con una estimación adaptativa de parámetros son ligeramente mejores a las obtenidas manteniendo fijos los parámetros de los modelos de predicción. Además, también se puede observar que los modelos basados en la Regresión Local Ponderada tienen un error inferior que los modelos de Media Ponderada.

En la Figura 8.2 se representa el NRMSE de las predicciones obtenidas utilizando como fuente de predicción de viento a la Agencia B. En este caso también se han representado los errores correspondientes a los modelos de Media Ponderada y de Regresión Local Ponderada, tanto manteniendo fijos los parámetros calculados con los datos de entrenamiento, como permitiendo que el valor de dichos parámetros evolucione con el tiempo mediante el método de estimación adaptativa propuesto.

De nuevo se aprecia una mejoría de las predicciones si se estiman de forma adaptativa los parámetros de los modelos, frente a la opción de fijar los parámetros estimados de forma global con el periodo de entrenamiento.

Por tanto, se comprueba el adecuado funcionamiento del método de estimación adaptativa de hiperparámetros, pues en todos los casos representados las predicciones obtenidas con los modelos evolutivos mejoran a las predicciones obtenidas por los modelos que mantienen fijos los parámetros estimados con el periodo de entrenamiento.

Además, la estimación evolutiva aporta mayor robustez que la estimación global inicial frente a variabilidad en las condiciones futuras, tanto climáticas como de los parques. Esto permite a la herramienta de predicción funcionar en tiempo real y sin supervisión con mayor fiabilidad.

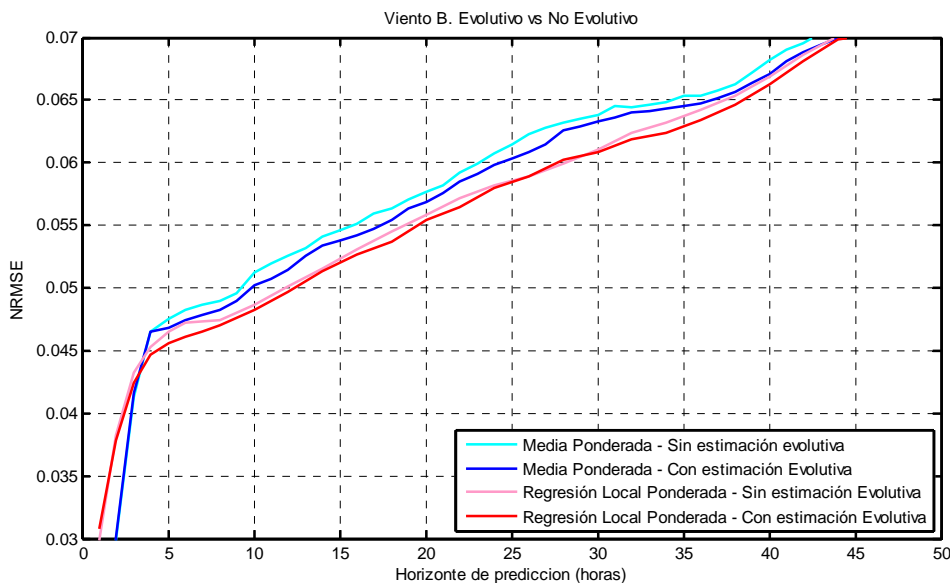


Figura 8.2: Error de las predicciones obtenidas con la agencia de Viento B, con modelos de Media Ponderada y Regresión Local Ponderada, y estimaciones evolutiva y no evolutiva de hiperparámetros

8.2.2 Comprobación de la combinación adaptativa de predicciones

Partiendo de las predicciones obtenidas anteriormente con los modelos que realizan la estimación adaptativa de los hiperparámetros propuesta en esta tesis, se procede a combinar estas predicciones para obtener unas predicciones finales mejoradas.

Se pueden combinar tanto las predicciones obtenidas con modelos de predicción de distinto orden (Media Ponderada y Regresión Local Ponderada), como las predicciones obtenidas con fuentes de predicción de viento diferentes (Agencia A y Agencia B). A continuación se hace una comprobación de los resultados obtenidos en ambos casos.

La combinación de predicciones se hace mediante un procedimiento de combinación en dos etapas como el descrito en el capítulo 7, utilizando el factor de olvido que obtuvo los mejores resultados en dicho capítulo con los datos de 2007, $\lambda_{MSE} = \lambda_{AEC} = 0.9972$.

Combinación de modelos de Media Ponderada y Regresión Local Ponderada

Como era esperable, independientemente de la agencia de viento utilizada, los modelos de Regresión Local Ponderada obtienen en general mejores resultados que los modelos de Media Ponderada. Los modelos de Regresión Local Ponderada son menos sesgados que los de Media Ponderada, de modo que en término medio tienen menor error. Sin embargo, en determinadas circunstancias los modelos de Regresión Local Ponderada aumentan la complejidad y la varianza de las estimaciones, de modo que la elección del orden del modelo se plantea como un compromiso entre reducir el sesgo o reducir la varianza.

Nuestro modelo de predicción final calcula las predicciones de potencia en base a combinar las predicciones obtenidas con ambas aproximaciones, para aprovechar las ventajas inherentes a cada una de ellas dependiendo de la complejidad o curvatura del problema a modelar en cada instante.

En la Figura 8.3 se representa el error NRMSE de las predicciones de potencia obtenidas con la agencia de viento A, tanto utilizando un modelo de Media Ponderada como de Regresión Local Ponderada, ambos con estimación evolutiva de hiperparámetros. También se representa el error de la predicción combinada, fruto de un proceso de combinación en dos etapas.

Se puede comprobar que la predicción combinada mejora las predicciones de los modelos de Media Ponderada y de Regresión Local Ponderada de forma individual para todos los horizontes.

Métodos de Predicción de la Generación Agregada de Energía Eólica

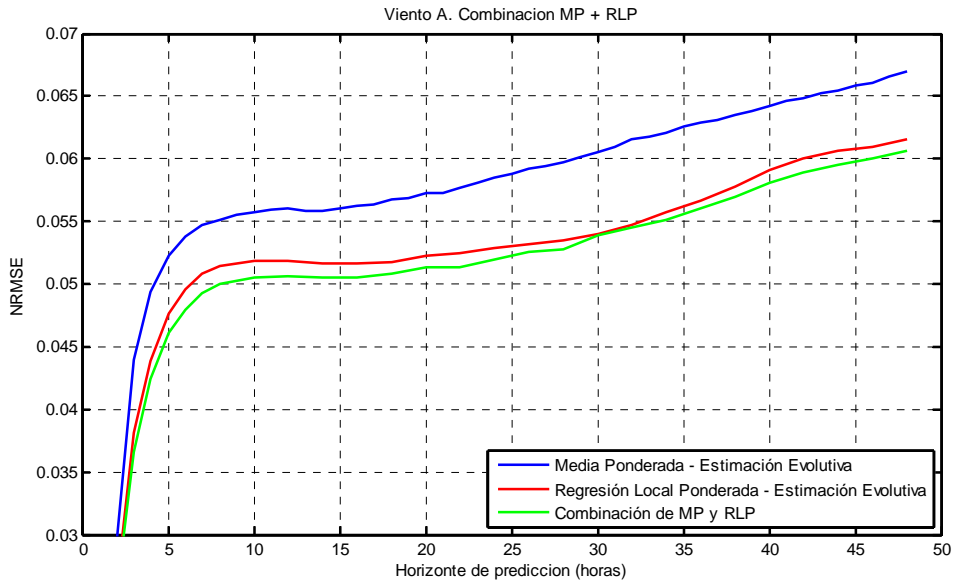


Figura 8.3: Error de las predicciones obtenidas con la agencia de Viento A, con modelos de Media Ponderada y Regresión Local Ponderada, y la combinación de ambos

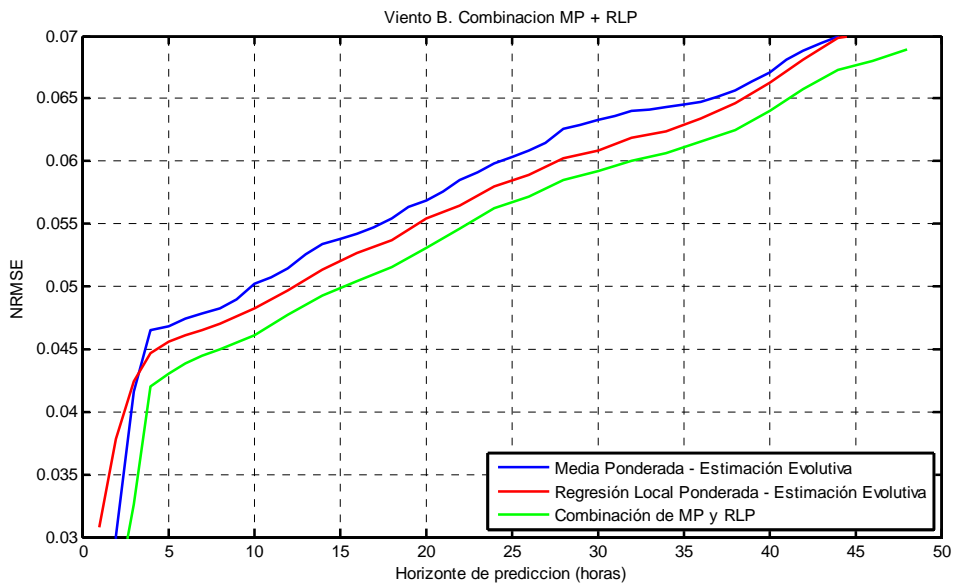


Figura 8.4: Error de las predicciones obtenidas con la agencia de Viento B, con modelos de Media Ponderada y Regresión Local Ponderada, y la combinación de ambos

Evaluación del Modelo Propuesto de Predicción de Potencia Eólica Agregada

Esta mejoría también se puede comprobar en la Figura 8.4, donde se muestra el NRMSE correspondiente a las predicciones generadas con modelos de Media Ponderada y Regresión Local Ponderada utilizando como fuente de viento las previsiones de la Agencia B. También en este caso se aprecia una disminución del error en la predicción combinada.

Se demuestra así la idoneidad de combinar las predicciones obtenidas con modelos de distinto orden, pues dependiendo del instante en el tiempo, y por consiguiente del vector de viento previsto, un modelo de Media Ponderada se puede adaptar al problema mejor que un modelo de Regresión Local Ponderada y viceversa. Estamos por tanto ante una competición entre eliminar el sesgo de los modelos de Media Ponderada o eliminar la posible varianza inherente a los modelos de Regresión Local Ponderada.

Combinación de predicciones obtenidas con distintas fuentes de predicción de viento

A continuación se comprueba el resultado de combinar las predicciones obtenidas utilizando como entrada a los modelos el viento previsto en la región por distintas agencias de predicción meteorológica.

En la Figura 8.5 se muestran los errores NRMSE de las predicciones finales de las agencias A y B, fruto de la combinación previa de los modelos de Media Ponderada y Regresión Local Ponderada en cada una de ellas, y el error de la predicción combinada final de ambas agencias.

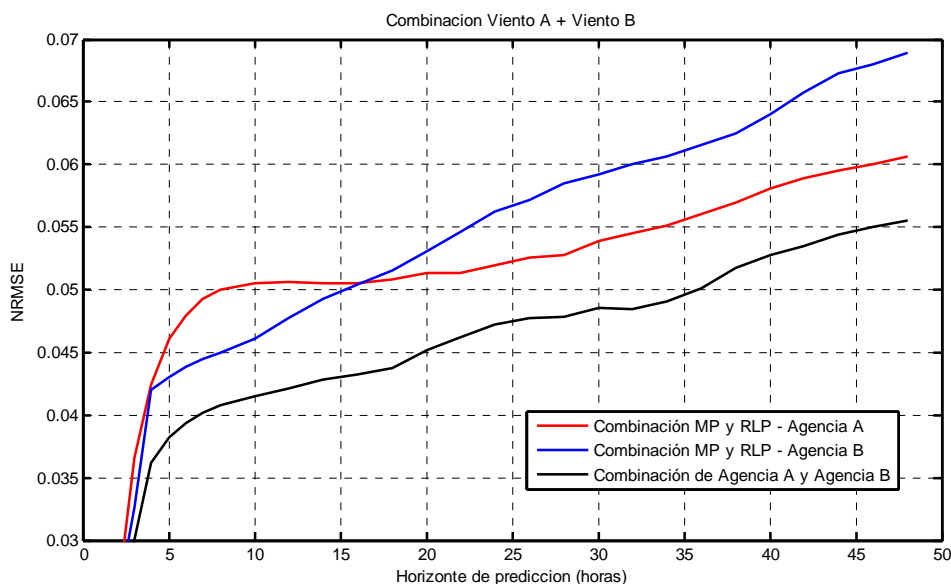


Figura 8.5: Error de las predicciones obtenidas con las agencias de Viento A y Viento B, y la combinación de ambas

En este caso se observa una gran mejoría al combinar las predicciones de potencia obtenidas a partir de las predicciones de viento de las dos agencias, ya que la predicción combinada es sensiblemente mejor que cualquiera de las agencias individuales para todos los horizontes de predicción. Esto pone de manifiesto la gran utilidad que tiene proporcionar a los modelos de predicción de potencia eólica las predicciones de viento de modelos atmosféricos distintos, ya que dependiendo de la situación meteorológica en cada instante unos modelos atmosféricos darán mejores predicciones de viento que otros y viceversa.

8.3 Comparación con otros modelos de predicción

Como evaluación final del modelo de predicción de potencia agregada propuesto en esta tesis, se hará una comparación de las predicciones finales obtenidas con el modelo de predicción propuesto, utilizando los datos del año 2008, con las predicciones generadas por la herramienta de predicción Sipreólico, que calcula predicciones agregadas para el total peninsular como la suma de las predicciones en los parques eólicos instalados. También se incluyen en la comparación las predicciones de potencia total peninsular proporcionadas por otra agencia de predicción de potencia agregada durante el mismo periodo.

La Figura 8.6 muestra los errores de predicción de los tres modelos de predicción comparados. Como se puede comprobar, el error de predicción del modelo agregado propuesto es sensiblemente inferior al error de predicción obtenido con la herramienta de predicción Sipreólico o el que se obtiene la otra agencia de predicción comparada.

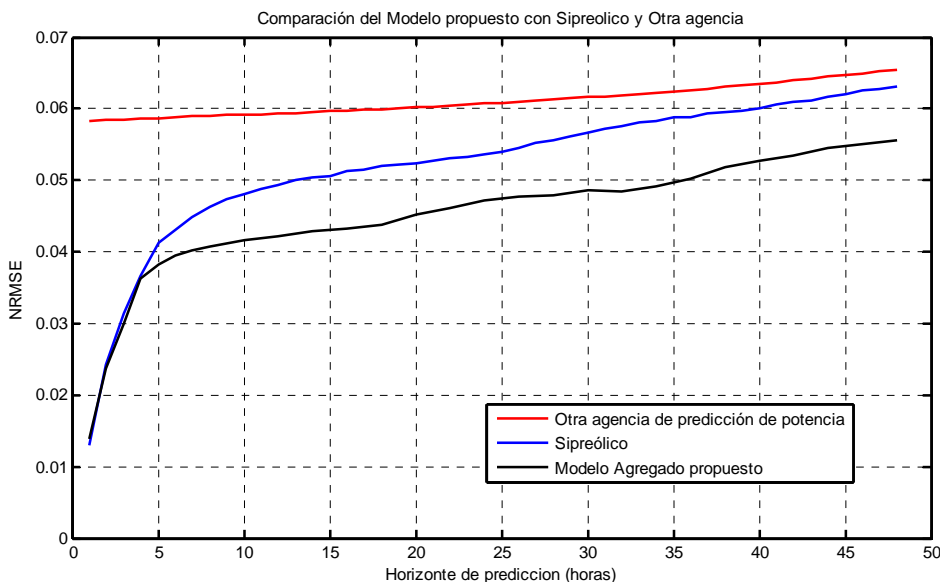


Figura 8.6: Comparación del modelo propuesto con Sipreólico y otra agencia de predicción de potencia agregada

Evaluación del Modelo Propuesto de Predicción de Potencia Eólica Agregada

Por otro lado, tanto Sipleólico como el modelo propuesto presentan errores de predicción similares para horizontes inferiores a 4 horas. Esto se debe a que ambos modelos utilizan cada hora los datos de potencia eólica registrada durante la hora anterior, de modo que las predicciones de las primeras horas a nivel agregado, muy condicionadas por la persistencia del viento, se muestran muy similares. Esto no se observa en las predicciones de la otra agencia, pues sus predicciones de potencia solo se actualizan cada 12 horas, lo que se ha presentado como una desventaja para sus predicciones a horizontes bajos.

Por último, en la Figura 8.7 se representa el coeficiente de mejora de la predicción del modelo agregado propuesto en esta tesis frente a la herramienta Sipleólico. Excepto para el horizonte de 1 hora, se puede comprobar que el modelo agregado mejora las predicciones de Sipleólico para todos los horizontes, llegando a valores de disminución del error que oscilan entre el 12 y el 16 % para horizontes superiores a 8 horas.

En cuanto al tiempo de cálculo empleado, el algoritmo de predicción agregada propuesto necesita menos de un segundo para calcular cada una de las predicciones para las 48 horas siguientes con un modelo de Media Ponderada. Las predicciones con un modelo de Regresión Local Ponderada tardan aproximadamente 15 segundos. Estos tiempos son los obtenidos cuando el periodo de datos históricos es de un año, es decir, 8760 pares de datos históricos (v_i, P_i) con los que comparar los vectores de predicción de viento. Conforme el periodo de datos históricos va aumentando el tiempo de cálculo de los modelos también aumenta.

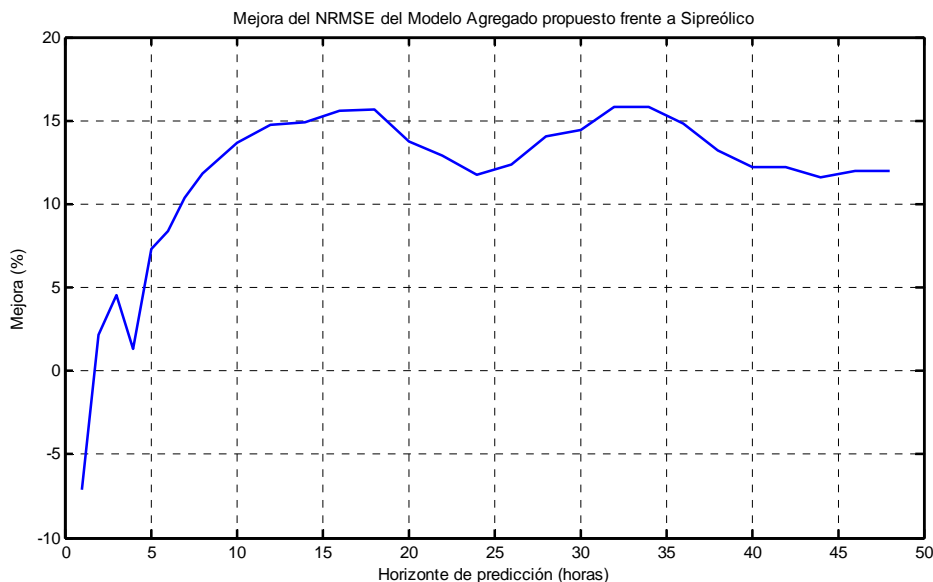


Figura 8.7: Mejora del modelo propuesto frente a Sipleólico

Métodos de Predicción de la Generación Agregada de Energía Eólica

Teniendo en cuenta que el modelo propuesto ha calculado las predicciones con modelos de Media Ponderada y Regresión Local Ponderada con datos de dos agencias de predicción de viento, el tiempo total de cálculo se sitúa en aproximadamente 32 segundos. El tiempo empleado en la combinación de predicciones es prácticamente despreciable.

Este tiempo de cálculo es muy inferior al tiempo de cálculo de Sipleólico, que siendo ejecutado con un ordenador de las mismas características ronda aproximadamente los 20 minutos, pues debe calcular las predicciones de potencia para cada uno de los parques eólicos instalados y posteriormente sumarlas. Por tanto, también se debe destacar la gran velocidad en el cálculo de predicciones con el modelo propuesto, que resulta unas 40 veces inferior al tiempo de cálculo habitual de la herramienta Sipleólico.

Si bien es cierto que un modelo agregado no ofrece predicciones para cada uno de los parques de forma individual, lo que supone un inconveniente para la identificación en detalle de los elementos de la red que se verán afectados por la generación eólica, la predicción agregada de potencia eólica se presenta como una herramienta de apoyo muy valiosa para mejorar los resultados de un modelo que calcule las predicciones de generación eólica parque a parque.

9. Modelos de Predicción Probabilista

Resumen

En este capítulo se propone un modelo de predicción probabilista condicionado al vector de viento previsto en la región, para lo cual se sirve de las medidas de distancia y funciones de ponderación por distancia propuestas en el capítulo 4. Para obtener la función de distribución del error asociado a la predicción de potencia se utilizan técnicas no paramétricas de estimación de densidades con funciones Kernel.

El modelo condicionado al vector de viento previsto se compara con otros modelos de predicción probabilista no condicionados, o condicionados al nivel de potencia prevista, evaluándolos con medidas de fiabilidad, agudeza y resolución de las predicciones probabilistas obtenidas.

9.1 Introducción

Habitualmente los modelos de predicción de energía eólica utilizan como datos de entrada predicciones meteorológicas, que se obtienen de la ejecución de modelos atmosféricos. Estos modelos proporcionan valores para la velocidad y dirección del viento, y en ocasiones otras variables como la temperatura, presión o humedad del aire, con cierta resolución hasta un horizonte temporal dado. A partir de estos valores de previsión meteorológica, se determina la potencia eléctrica que generarán los parques eólicos mediante modelos basados en métodos estadísticos o físicos.

Estos modelos atmosféricos que proporcionan los datos de entrada al modelo de predicción de potencia no están exentos de incertidumbre, ya sea por los propios datos iniciales de entrada al modelo atmosférico, las condiciones de contorno, o por la modelización física del terreno. Así, la predicción de potencia eólica se verá afectada por los posibles errores en la predicción de la velocidad del viento u otras variables explicativas utilizadas de los modelos atmosféricos, que se añadirán a la incertidumbre propia del modelo de predicción de potencia.

Los modelos de predicción deterministas proporcionan un único valor de predicción de energía eólica para cada horizonte de predicción, valor que necesariamente llevará asociada una incertidumbre debido a las limitaciones propias de los modelos de predicción. Para sacar mayor partido a los modelos de predicción, es necesario que éstos proporcionen predicciones probabilistas, que aportan información sobre la incertidumbre de las predicciones, y que resultan muy útiles como entrada para otros modelos o aplicaciones orientadas a la toma de decisiones bajo incertidumbre.

En el ámbito de la integración de la generación eólica en la red, entre las aplicaciones que utilizan las predicciones probabilistas están las encargadas de optimizar el nivel de reservas para compensar la variabilidad del viento. De hecho, la integración de la energía eólica en la red hace necesario que el operador del sistema eléctrico aumente la cantidad de reserva, precisamente debido a la incertidumbre asociada a la generación eólica, lo que incrementa los costes asociados a este tipo de energía renovable. Se ha demostrado también que el uso de predicciones probabilistas resulta muy útil para diseñar estrategias de oferta en los mercados de electricidad [72-74], o incluso en la coordinación de centrales hidráulicas con la generación eólica para facilitar su integración a la red [76],[80].

Recientemente han aparecido varios modelos de predicción probabilista en el ámbito de la generación eólica, debido al gran interés que despierta este tipo de predicción para ser utilizada en la toma de decisiones para la integración de energía eólica a gran escala u otras aplicaciones destinadas a estrategias de mercado.

Los modelos probabilísticos de predicción pueden proporcionar algunos percentiles o intervalos de predicción prefijados. Los percentiles indican la probabilidad con la que el dato real observado quedará por debajo de determinado valor. Así, por ejemplo una

Métodos de Predicción de la Generación Agregada de Energía Eólica

predicción de potencia de 0.4 p.u. en el percentil 85 indica que se estima un 85% de probabilidad de que la producción sea inferior al 40% de la potencia instalada.

Se pueden obtener intervalos de predicción como la diferencia entre dos percentiles. Los intervalos indican por tanto el rango de potencias dentro del que se espera encontrar la potencia observada con cierta probabilidad. Así, si por ejemplo el percentil de predicción del 15% es 0.1 p.u. y el percentil del 85% es de 0.4 p.u., se obtendrá un intervalo centrado de predicción que indica que la potencia observada estará entre 0.1 y 0.4 p.u. con un 70% de probabilidad.

En algunas ocasiones, como la del ejemplo anterior, los modelos proporcionan intervalos de predicción que indican los límites inferior y superior entre los cuales se espera que se encuentre el valor de generación real con determinada probabilidad. En otras ocasiones, el modelo puede estar preparado para proporcionar directamente la función de densidad de probabilidad o la función de distribución, funciones que proporcionan una estimación completa de la incertidumbre asociada a la predicción, pues a partir de ellas se puede obtener de forma sencilla cualquier percentil o intervalo de predicción.

En esta tesis se propone un método que proporciona la función de distribución $\hat{F}(x) = \int_{-\infty}^x \hat{f}_e(t) dt$, del error aleatorio e_{t+h} asociado a la predicción de potencia $\hat{P}_{t+h|t}$, en cada instante t y para cada horizonte de predicción $t+h$, basándose en técnicas no paramétricas de estimación de densidades con funciones Kernel.

Mediante el método propuesto se estimará la forma de la distribución del error en el entorno de la predicción puntual $\hat{P}_{t+h|t}$ que fue obtenida mediante el modelo de predicción de potencia agregada, de modo que la predicción probabilista de potencia se obtendrá superponiendo la incertidumbre del error a la predicción puntual de potencia. Obviando por simplicidad en la notación la dependencia del horizonte h y del instante t en el que se calcula la predicción, la función de densidad de probabilidad de potencia se obtendría entonces como:

$$\hat{f}_P = \hat{P} + \hat{f}_e \quad (9.1)$$

donde:

\hat{f}_P = función de densidad de potencia prevista

\hat{P} = predicción puntual de potencia

\hat{f}_e = función de densidad de error estimada

Los intervalos de predicción de potencia se obtendrán entonces como superposición de los intervalos de error, que se obtienen directamente de la función $\hat{F}_{t+h|t}$, sobre la potencia prevista.

9.2 Estimación de densidades con Kernel

La estimación de densidades con funciones Kernel es un método de estimación no paramétrico, es decir, no presupone que la distribución de datos, observaciones de error de predicción en nuestro caso, se ajuste a alguna familia de distribuciones prefijada y descrita por algunos parámetros, sino que estima la forma de la distribución directamente de los errores observados a lo largo del tiempo.

Hablamos de nuevo de un método de suavizado o aprendizaje basado en la memoria, con lo que la aproximación no paramétrica para estimar la función de densidad de los errores requiere un conjunto suficientemente grande de datos frente a los que serían necesarios si, con un modelo paramétrico, se estimasen los relativamente escasos parámetros que determinan una distribución prefijada. Pese a este posible inconveniente, la estimación no paramétrica presenta la gran ventaja de que no está limitada por la posibilidad de haber adoptado una hipótesis incorrecta en cuanto a la familia de la función de distribución de error.

Una forma sencilla de evaluar una función de densidad de forma no paramétrica podría basarse en la obtención del histograma. El método de obtención de un histograma consiste en dividir en C intervalos o clases, de una anchura δ_{dens} , el rango de valores posibles de una variable x , y contabilizar el número de observaciones n_x , del total de observaciones n , que están contenidas en cada uno de los intervalos. De esta forma se obtiene la frecuencia relativa de aparición de los datos dentro de cada intervalo c_j , con $c_j = [x_0 + (j-1) \cdot \delta_{dens}, x_0 + j \cdot \delta_{dens}]$, $j=1, \dots, C$, de la forma:

$$\hat{f}(x) = \frac{n_x}{n \cdot \delta_{dens}} \quad (9.2)$$

donde n_x es el número de observaciones que están dentro del intervalo que contiene a x .

Aunque el histograma es una herramienta muy útil para la representación de una función de densidad empírica, presenta ciertos inconvenientes, como que su forma depende del origen x_0 a partir del cual se definen los intervalos, del número C de intervalos o clases a representar, o que se asigna el mismo peso a cada dato recogido dentro de cada intervalo independiente de lo cerca que esté del centro del mismo. Como consecuencia el histograma es una función no continua, con saltos en los límites de cada intervalo que define sus clases.

Partiendo de la idea de construcción del histograma y teniendo en cuenta su problemática discontinuidad y dependencia de la elección de los intervalos, se puede intuir la necesidad del algún método de estimación de densidades basado en funciones de suavizado que evite estos inconvenientes. Una estimación de densidades por suavizado se aproximará más que un histograma a la función de densidad verdadera.

Para obtener el suavizado se busca una función que asigne un peso o contribución a cada observación e_i en función de su distancia al punto x para el que se quiere estimar

Métodos de Predicción de la Generación Agregada de Energía Eólica

la densidad. La estimación de la función de densidad en el entorno del punto x se calculará como la suma de todas estas contribuciones:

$$\hat{f}_e(x) = \frac{1}{n \cdot \delta_{dens}} \sum_{i=1}^n K(u_{dens}) \quad (9.3)$$

donde $K(u_{dens})$ es una función Kernel, dependiente de la variable u_{dens} que representa la distancia relativa entre el dato e_i observado, en nuestro caso una observación de error en el instante i , y el punto x para la que estima la función de densidad. Esta distancia relativa depende de un ancho de banda o parámetro de suavizado δ_{dens} .

$$u_{dens} = \frac{x - e_i}{\delta_{dens}} \quad (9.4)$$

Como función de suavizado se puede utilizar cualquiera de las funciones Kernel que ya se vieron para este cometido en la sección 4.1. Entre ellas podemos encontrar la uniforme, triangular, Epanechnikov, quartic, coseno o gaussiana entre otras, descritas en las ecuaciones (4.3) a (4.8).

Si bien en la literatura se suele encontrar que para la estimación de densidades por suavizado la elección de la función Kernel en (9.3) no es crucial, el ajuste del ancho de banda δ_{dens} sí suele ser importante [112]. De un modo similar al ancho de los intervalos en un histograma, el ancho de banda de la función Kernel controla el suavizado de la estimación, de forma que con un ancho de banda demasiado pequeño se obtendrá una función de densidad sobreajustada y con bastante ruido, mientras que un ancho de banda demasiado amplio dará lugar a un suavizado excesivo que podría pasar por alto la verdadera forma de la función de densidad.

La elección de un ancho de banda óptimo se puede llevar a cabo empíricamente, en base a calcular la función de densidad con un rango de valores de ancho de banda y comprobar cuál es el que mejor se ajusta a nuestro problema mediante la minimización de alguna función de pérdidas que mida la discrepancia entre la función estimada y lo que indican las observaciones reales.

Un modo de estimación del ancho de banda óptimo cuando la densidad esperada tenga forma aproximadamente gaussiana es el que establece la regla de Silverman, que proporciona una estimación de un ancho de banda adecuado a partir del número n de datos de la muestra y de la estimación de la desviación típica $\hat{\sigma}_{dens}$ o del Rango intercuartílico R_{75-25} [108].

El ancho de banda a utilizar para la estimación de la función de densidad se obtendría de esta forma como:

$$\hat{\delta}_{opt} = 1.06 \min \left\{ \hat{\sigma}_{dens}, \frac{R_{75-25}}{1.34} \right\} n^{-1/5} \quad (9.5)$$

También existe la posibilidad de variar este ancho de banda o parámetro de suavizado a lo largo del tiempo, de modo que se adapte a la distribución de datos en cada momento, con lo que se podrían obtener mejores estimaciones de la función de densidad.

9.3 Antigüedad de las observaciones de error

La forma de la distribución de los errores de predicción de potencia eólica puede variar a lo largo del tiempo, ya sea debido a variaciones estacionales o porque varíe la propia incertidumbre de las variables de entrada al modelo de predicción, o también por cambios en la precisión del modelo de predicción de potencia, lo que implicará variaciones en la anchura de la función de densidad del error.

Se puede tratar de aumentar la fiabilidad y adaptabilidad en la estimación de la función de densidad haciendo que las observaciones de error e_i a partir de las cuales se calcula la predicción probabilista también sean ponderadas con algún coeficiente en función de su antigüedad. Esta adaptabilidad puede ser beneficiosa ante los posibles cambios estacionales en la estabilidad atmosférica, de forma que se tenga en cuenta la posible variación a lo largo del año de la incertidumbre asociada a las predicciones de viento.

El coeficiente de ponderación por antigüedad de las observaciones de error propuesto es similar al utilizado para el modelo de predicción de potencia (4.33), obteniendo así el nuevo peso asignado a cada observación de error e_i , que será de la forma:

$$W\tau_i = \frac{\lambda_{prob}^{\tau_i}}{\sum_{j=1}^n (\lambda_{prob}^{\tau_j})} \quad (9.6)$$

donde:

$W\tau_i$ = Coeficiente de ponderación por antigüedad del error observado e_i

τ_i = antigüedad en horas, u otro intervalo de tiempo, del error observado e_i

λ_{prob} = factor de olvido de la predicción probabilista, es un parámetro a ajustar

Al introducir el factor de olvido, el peso asignado a cada observación histórica del error para calcular la función de densidad se calculará como el producto de dos pesos, el correspondiente a la función Kernel de suavizado para la estimación de la función de densidad de errores $K(u_{dens})$, y el peso correspondiente a la antigüedad de las observaciones de error:

$$\hat{f}_e(x) = \frac{1}{n \cdot \delta_{dens}} \sum_{i=1}^n W\tau_i \cdot K(u_{dens}) \quad (9.7)$$

9.4 Predicción probabilista condicionada al nivel de potencia prevista

La aproximación más simple para obtener la distribución de error consiste simplemente en procesar los errores históricos observados y evaluar su distribución de sin tener en cuenta otras variables que pudieran ser explicativas. De este modo se obtiene una estimación de incertidumbre en la que la posible variabilidad o desviación con respecto a las predicciones no tendrá en cuenta de forma explícita la incertidumbre asociada a cada condición meteorológica específica.

Un modo de proporcionar más información a la incertidumbre o distribución del error es hacerla dependiente de alguna variable explicativa, como por ejemplo el nivel de potencia prevista. Se pueden encontrar en la literatura varias referencias a la dependencia entre la incertidumbre y el nivel de potencia [86],[121].

Para niveles de generación bajos y elevados la distribución de errores de predicción es más estrecha que para niveles de generación media. Esto se debe a la fuerte pendiente que presenta la curva de potencia de los generadores eólicos para valores medios de viento, correspondientes a la parte central de la curva de potencia, con lo que pequeñas variaciones o errores de la velocidad prevista de viento implican grandes variaciones en la potencia finalmente generada frente a la potencia inicialmente prevista.

En la Figura 9.1 se representa la desviación típica del error de predicción en función del nivel de potencia prevista, para las predicciones agregadas de horizonte $h=24$ horas calculadas con los datos de potencia total peninsular durante el segundo semestre del año 2007. El rango de valores posibles de potencia agregada queda truncado en 0.7 p.u., pues no se observa en ningún instante una medida de potencia eólica total peninsular superior durante el periodo de ensayo seleccionado.

Si bien este es un caso diferente y mucho más complejo que el caso de un parque eólico o turbina independiente, pues al tratarse de potencia agregada en toda una región no se conoce a priori el nivel porcentual de potencia que se estará generando en cada parque con respecto a su potencia nominal, también se observa una variabilidad del error mucho más elevada en niveles de potencia medios, y menos variabilidad para niveles de potencia bajos y altos.

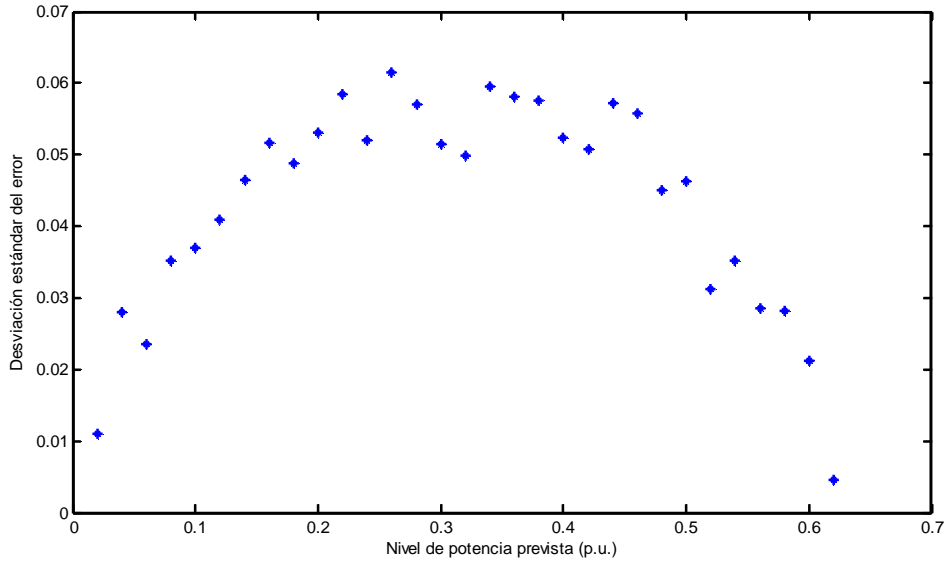


Figura 9.1: Desviación estándar del error de predicción en función del nivel de potencia

Para tener en cuenta el nivel de potencia se puede añadir un nuevo peso Wp_i que pondere cada dato de error histórico e_i observado en el instante i , en función de la diferencia entre la predicción de potencia agregada de horizonte h que fue prevista para el instante i , $\hat{P}_{i|i-h}$, y la nueva potencia prevista $\hat{P}_{t+h|t}$, obtenida con el modelo agregado de predicción en el instante t para el instante $t+h$.

Este peso se puede calcular de nuevo con una función Kernel de suavizado $K(u_p)$ como las descritas en (4.3) a (4.8), en la que $u_{p,i}$ dependa en este caso de la diferencia entre la predicción histórica $\hat{P}_{i|i-h}$ y la nueva predicción de potencia $\hat{P}_{t+h|t}$. De esta manera tienen más peso aquellas observaciones de error que ocurrieron en situaciones con un nivel de potencia prevista similar al actual. De algún modo, se podría decir que este procedimiento espera que la incertidumbre del error de predicción de potencia sea parecida a la incertidumbre que se obtuvo con niveles de potencia similares. Los nuevos pesos a aplicar a cada observación histórica de error e_i serán por tanto de la forma:

$$Wp_i = \frac{K(u_{p,i})}{\sum_{j=1}^n (K(u_{p,j}))} \quad (9.8)$$

con

$$u_{p,i} = \frac{\hat{P}_{t+h|t} - \hat{P}_{i|i-h}}{\delta_p} \quad (9.9)$$

Métodos de Predicción de la Generación Agregada de Energía Eólica

donde:

Wp_i = Nuevo peso, dependiente del nivel de potencia prevista, a aplicar al error observado e_i

$\hat{P}_{t+h|t}$ = Predicción puntual de horizonte h obtenida del modelo de predicción de potencia agregada

$\hat{P}_{i|i-h}$ = Potencia de horizonte h que fue prevista en el instante $i-h$ para el instante i

δ_p = ancho de banda de la función Kernel de ponderación por nivel de potencia

Finalmente, el peso asignado a cada observación histórica del error será ahora el producto de todos los pesos: el proporcionado por la función Kernel de suavizado para la estimación de la función de densidad de errores $K(u_{dens})$, el correspondiente a la antigüedad de las observaciones de error $W\tau_i$, y el proporcionado por la dependencia del nivel de potencia prevista, Wp_i . La nueva estimación de la función de densidad del error, condicionada al nivel de potencia se obtendrá como:

$$\hat{f}_e(x) = \frac{1}{n \cdot \delta_{dens}} \sum_{i=1}^n Wp_i \cdot W\tau_i \cdot K(u_{dens}) \quad (9.10)$$

9.5 Predicción probabilista condicionada al vector de viento previsto

Cuando la predicción de potencia se calcula para una turbina o un parque eólico individual, la variable de entrada en el modelo suele ser únicamente la previsión de velocidad media de viento en el emplazamiento. Sin embargo, cuando se calcula la predicción de potencia agregada en una región extensa utilizando un vector de predicciones de viento en varias coordenadas de la región, como es el caso tratado en esta tesis, aumenta la dimensión del espacio de entrada en el modelo. En el caso de predicción regional existirá por tanto un gran número de combinaciones de viento en las coordenadas, correspondientes a situaciones meteorológicas muy diferentes con las que se obtiene el mismo nivel de predicción de potencia.

Esto nos da la idea de que se puede hacer una discriminación más exhaustiva de las observaciones de error, en lugar de la discriminación por nivel de potencia expuesta en el apartado anterior. Esta discriminación más precisa para calcular la función de densidad de error esperada se puede conseguir si, en lugar de condicionar con el nivel de potencia prevista, se tiene en cuenta la similitud entre los vectores de predicción de viento.

Aparte, debido a que se dispone de las previsiones de viento en muchas coordenadas de la región, éstas ya serán bastante representativas de la estabilidad o inestabilidad de la situación meteorológica, al igual que lo serían si fuesen utilizadas como entrada en

un modelo de predicción atmosférica. Si bien el objetivo en este caso no es conocer cómo será la evolución de la atmósfera, sino evaluar su inestabilidad.

La idea que se propone es que si fueron muy grandes los errores de predicción de potencia observados cuando la previsión de viento en las coordenadas fue similar a la previsión de viento actual, entonces podemos concluir que la actual situación meteorológica prevista es muy inestable o difícil de predecir. Por el contrario, si cada vez que se registraron predicciones de viento en la región similares a la actual, el error fue pequeño, podremos concluir que la situación actual es fácil de predecir, y por tanto el error esperado también será menor.

Por tanto, otra forma de condicionar la función de densidad del error con variables explicativas es hacer que dependa de la previsión meteorológica, en cuyo caso las variables explicativas serán las predicciones de viento en las coordenadas de la región. En este caso, en lugar de buscar similitudes en el nivel de potencia, buscaremos similitudes entre las previsiones meteorológicas del pasado y la previsión meteorológica actual, mediante medidas de distancia entre los vectores de predicción de viento en las coordenadas.

Se puede medir la similitud entre situaciones meteorológicas mediante medidas de distancia entre vectores de predicción de viento de un modo similar al que se propuso en la sección 4.3. Utilizando las medidas de distancia se obtendrá un escalar $d_{i-h,t+h}=d(\mathbf{v}_{i,i-h}, \mathbf{v}_{t+h,t})$, que resume la similitud entre el nuevo viento $\mathbf{v}_{t+h,t}$ previsto en t para el instante $t+h$ y el viento $\mathbf{v}_{i,i-h}$ que fue previsto en el instante $i-h$ con horizonte h para el instante histórico i , en el que se registró un error e_i .

Con estas medidas de distancia se pueden añadir unos coeficientes de ponderación o pesos a las observaciones, calculados de nuevo mediante alguna función de suavizado que otorgue más peso a aquellos errores e_i que fueron observados en situaciones con una previsión meteorológica similar a la nueva situación de viento previsto. Así, los nuevos pesos a aplicar a cada observación histórica e_i serán de la forma:

$$Wd_i = \frac{\omega^d_{i-h}}{\sum_{j=1}^n \omega^d_{j-h}} \quad (9.11)$$

donde $\omega^d_{i-h} = \omega(d(\mathbf{v}_{i,i-h}, \mathbf{v}_{t+h|t}))$ puede ser alguna de las funciones de ponderación por distancia propuestas la sección 4.6 para el modelo de predicción de potencia. Utilizando la función de ponderación inversa a la distancia (4.28), con la que se obtuvieron los mejores resultados en el modelo de predicción de potencia, se tendría:

$$\omega^d_{i-h} = \frac{1}{(d_{i-h,t+h})^{\alpha_{prob}}} \quad (9.12)$$

Métodos de Predicción de la Generación Agregada de Energía Eólica

donde:

Wd_i = Peso a aplicar al error observado e_i , dependiente de la similitud entre situaciones meteorológicas, que será medida con distancias

$d_{i-h,t+h}$ = distancia entre los vectores de predicción de viento $\mathbf{v}_{i,i-h}$ y $\mathbf{v}_{t+h,t}$

α_{prob} = parámetro a optimizar

De nuevo, el peso asignado a cada observación histórica del error se calculará como el producto de los pesos correspondientes a la función Kernel de suavizado para la estimación de la función de densidad de errores $K(u_{dens})$, el correspondiente a la antigüedad de las observaciones de error $W\tau_i$, y el proporcionado por la similitud de la situación meteorológica prevista con otras situaciones meteorológicas previstas en el pasado Wd_i . La estimación de la función de densidad del error condicionada a la situación meteorológica en la región se obtiene entonces como:

$$\hat{f}_e(x) = \frac{1}{n \cdot \delta_{dens}} \sum_{i=1}^n Wd_i \cdot W\tau_i \cdot K(u_{dens}) \quad (9.13)$$

9.6 Métodos de evaluación de las predicciones probabilistas

9.6.1 Fiabilidad o calibración

El modelo de predicción probabilista propuesto, que estima la función de distribución del error $\hat{F}_{t+h|t}$, presenta la gran ventaja de que a partir de ella se puede obtener cualquier percentil o intervalo deseado. A partir de la función de distribución del error se pueden calcular los percentiles y los intervalos de predicción de potencia superponiendo los percentiles del error a la predicción puntual de potencia eólica $\hat{P}_{t+h|t}$.

La fiabilidad será un indicador de la habilidad del modelo de predicción probabilista para obtener dicha función \hat{F} , y por tanto de lo fiables que serán los percentiles e intervalos de predicción obtenidos a partir de ella. Cuanto menor sea la diferencia entre los percentiles o intervalos de predicción calculados y los observados a partir de los datos reales, más fiable será nuestra estimación.

La evaluación de las predicciones probabilistas debe estar basada exclusivamente en las distribuciones de error estimadas y los valores de error finalmente observados. Un modo muy extendido para medir la fiabilidad es utilizar el histograma de la PIT (Probability Integral Transform) [122],[123]. Este histograma se obtiene a base de hacer pasar en cada instante t las observaciones de error e_t a través de la función

$\hat{F}_{t|t-h}$ que estimó su distribución, con lo que se obtendrá una serie de valores comprendidos entre 0 y 1. Si el modelo de estimación de las funciones \hat{F} es correcto, los valores obtenidos se distribuirán como una uniforme $U(0,1)$.

Con estos valores se puede dibujar un histograma, del que se espera que tenga una distribución uniforme. El grado de ajuste del histograma PIT se puede evaluar comparando la frecuencia relativa y_i , $i=1, \dots, C$, observada en cada clase del histograma, con la frecuencia relativa esperada si hubiese tenido una distribución perfectamente uniforme, que será por tanto $1/C$ en todas las clases, siendo C el número de clases representadas en el histograma [124],[125].

Un modo de evaluar el ajuste del PIT es calcular la raíz del error cuadrático medio, $RMSE_{PIT}$, entre la altura de las clases del histograma PIT y la altura de una uniforme ideal. Habitualmente el número de clases en que se divide el histograma PIT para su evaluación suele ser de $C=20$. De este modo se puede evaluar con los percentiles habituales en escalones de un 5% de probabilidad.

$$RMSE_{PIT} = \sqrt{\frac{1}{C} \sum_{i=1}^C \left(y_i - \frac{1}{C} \right)^2} \quad (9.14)$$

También se puede evaluar esta fiabilidad representando la diferencia entre los intervalos de predicción calculados con las funciones \hat{F} y los intervalos finalmente observados [126]. Cuanto menor sea el valor absoluto del desvío de la cobertura nominal de los intervalos de predicción, mayor será la fiabilidad del modelo de estimación de la función de densidad.

Dependiendo del modelo de predicción probabilista y de los parámetros de suavizado ajustados en las funciones de ponderación para la estimación de la función de densidad, se pueden obtener intervalos de predicción demasiado anchos o demasiado estrechos. El objetivo será encontrar los parámetros que consiguen minimizar la discrepancia entre el histograma PIT obtenido con los datos observados y una distribución uniforme ideal, lo que indicaría que la función de densidad está bien estimada.

En la Figura 9.2 se muestra el resultado de los histogramas PIT calculados bajo distintas circunstancias, comparándolos con la que hubiese sido una estimación correcta de la función de densidad. Este ejemplo supone que la distribución real del error es una distribución normal $N(0,1)$ [127]. El primero es el único histograma que indicaría un ajuste correcto, con un PIT bastante homogéneo y que se aleja poco de la uniforme $U(0,1)$. El segundo histograma mostraría el efecto de haber obtenido una función de densidad demasiado estrecha respecto de la verdadera, con lo que resultaría en intervalos de predicción demasiado estrechos y por tanto el PIT tendría forma de U.

Métodos de Predicción de la Generación Agregada de Energía Eólica

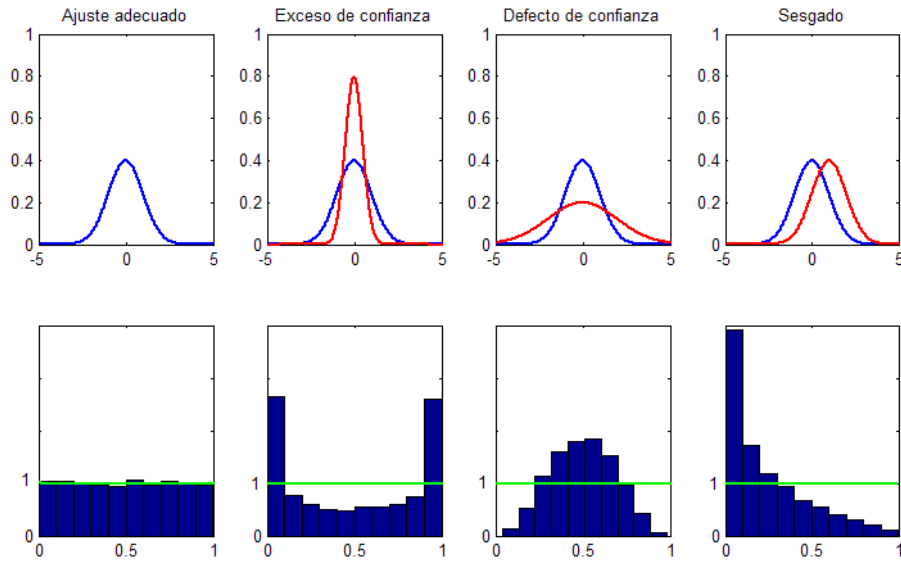


Figura 9.2: Posibles formas que puede adquirir el PIT

El tercer caso, con forma de joroba, correspondería a una estimación de la función de densidad demasiado ancha o suave, con lo que los intervalos de predicción serían muy gruesos, equivalente a una predicción probabilista demasiado conservadora. El cuarto caso correspondería al PIT obtenido cuando el modelo proporciona una estimación sesgada de la función de densidad. En este último caso el PIT adquiere una forma triangular, y en casos extremos puede llegar a tener forma de L o de J dependiendo del signo del sesgo.

9.6.2 Concentración o agudeza (sharpness)

La uniformidad del PIT es una condición necesaria pero no suficiente para que un modelo de predicción probabilista sea útil. La evaluación de fiabilidad solo demuestra que los intervalos calculados son probabilísticamente correctos en el sentido de que a largo plazo respetan la distribución de los errores observados. Pero aunque la comparación del PIT con la uniforme nos puede dar una idea de si el modelo de incertidumbre está bien calibrado, esto no quiere decir necesariamente que el modelo sea útil, ya que podría estar dando unos intervalos de predicción muy amplios.

Esto es lo que ocurriría si, por ejemplo, se tuviese un modelo que estimase una predicción probabilista de potencia basándose exclusivamente en los datos históricos de potencias medidas. Lo que se estaría representando en tal caso sería simplemente la distribución climática de potencias observadas, pero aunque estuviese bien calibrada y se obtuviese un PIT uniforme al estimar su distribución, no sería de gran utilidad.

Para que los intervalos de predicción sean útiles deben ser lo más estrechos posibles. Por tanto, a la evaluación de un modelo de predicción probabilista se le debe añadir

también una medida de su concentración o agudeza, referenciada en la bibliografía como “sharpness”, de manera que cuanto más concentradas sean las funciones de densidad, mejor será el modelo de predicción. La agudeza es por tanto una medida de la incertidumbre del modelo de predicción, de modo que estará limitada por el propio error medio del modelo de predicción. Un modelo de predicción probabilista por tanto será útil si proporciona unos intervalos de predicción estrechos y además esos intervalos se cumplen, lo que debe comprobarse previamente con la uniformidad del PIT.

La agudeza se puede medir como el tamaño medio de los intervalos de predicción obtenidos con el transcurso del tiempo para cada horizonte de predicción, a partir de las funciones de distribución $\hat{F}_{t|t-h}$ que se han ido estimando en cada instante para las predicciones de horizonte h [123],[126].

Si se define θ como la probabilidad de que el error observado e_t caiga dentro del intervalo centrado de amplitud $\psi_{t|t-h}^\theta$ que fue calculado a partir de la función de distribución $\hat{F}_{t|t-h}$, el intervalo $\psi_{t|t-h}^\theta$ estará delimitado por los percentiles $[\rho_{t|t-h}^{(1/2-\theta/2)}, \rho_{t|t-h}^{(1/2+\theta/2)}]$. Obviando en la notación la referencia al horizonte de predicción y el instante en que se calculó el intervalo, $(t/t-h)$, la agudeza de la función de densidad de probabilidad para el intervalo de cobertura nominal θ se puede calcular como:

$$\bar{\psi}^\theta = \frac{1}{n} \sum_{i=1}^n \psi_i^\theta = \frac{1}{n} \sum_{i=1}^n (\rho^{(1/2+\theta/2)} - \rho^{(1/2-\theta/2)}) \quad (9.15)$$

donde

$\bar{\psi}^\theta$ = agudeza o amplitud media de los intervalos de predicción de probabilidad θ

ψ_i^θ = amplitud del intervalo de probabilidad θ calculado con la función de distribución \hat{F}_i

$\rho^{(1/2-\theta/2)}$ = percentil o valor por encima del cual se materializará la observación con una probabilidad de $(1-(1/2 - \theta/2))$, correspondiente al límite inferior del intervalo de predicción

$\rho^{(1/2+\theta/2)}$ = percentil o valor por encima del cual se materializará la observación con una probabilidad de $(1-(1/2 + \theta/2))$, correspondiente al límite superior del intervalo de predicción

Para evaluar la agudeza de distintos modelos de predicción probabilista se suelen utilizar diagramas que representan la amplitud media de los intervalos, $\bar{\psi}^\theta$, en función de la cobertura nominal θ de los mismos, para un horizonte de predicción determinado. De entre los modelos comparados, y siempre que superen la calibración del histograma PIT, será mejor aquel que presente una amplitud media de los intervalos menor.

En cualquier caso, se debe tener en cuenta que la agudeza o concentración de la función de densidad del error es una cualidad limitada por la precisión de las predicciones de potencia. Es decir, no se podrían obtener distribuciones de error más estrechas en término medio que la magnitud del error propio asociado al algoritmo de predicción. Por tanto, esta cualidad tiene más que ver con la precisión del modelo de predicción de potencia que con el modelo de estimación de incertidumbre.

9.6.3 Resolución

El último criterio utilizado para evaluar el valor o utilidad de las predicciones probabilistas es la resolución. Independientemente de que varios procedimientos de estimación de la función de densidad estén bien calibrados, es decir, proporcionen un PIT uniforme, y de que sus estimaciones de intervalos sean igual de estrechas en término medio, cualidad que se mide con la agudeza, aún será más útil aquel procedimiento en el que la amplitud de los intervalos obtenidos se adecue a cada situación concreta de predicción.

Por tanto, a igualdad de amplitud media en el tamaño de los intervalos, tendrá mayor utilidad el modelo que presente mayor variabilidad en dicha amplitud, pues esto significa que la amplitud de los intervalos se va adaptando a la incertidumbre propia de cada nivel de predicción o de la situación meteorológica. En este último caso, dependiendo de si la situación meteorológica prevista es más o menos inestable, la amplitud de los intervalos de predicción será mayor o menor, con lo que el modelo estará aportando más información.

Esta variabilidad se puede evaluar visualmente mediante diagramas de cajas en los que se representen los rangos intercuartílicos de las amplitudes obtenidas [123],[128]. Otra forma de evaluar la variabilidad consiste en medir la desviación típica de las amplitudes a lo largo del tiempo [86], de la forma:

$$\sigma_{\theta} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\psi_i^{\theta} - \bar{\psi}^{\theta})^2} \quad (9.16)$$

Según el criterio de resolución, los modelos de predicción probabilista se evalúan utilizando diagramas que representan la desviación típica σ_{θ} de los intervalos ψ_i^{θ} , en función de la cobertura nominal θ de los mismos para un horizonte dado. En este caso, a igualdad de agudeza y fiabilidad, será mejor el modelo que presente una desviación típica mayor, pues indica que capta mejor la variabilidad en la incertidumbre de las predicciones de potencia.

La heterocedasticidad o variabilidad de la desviación típica de la distribución de errores, y por tanto en la amplitud de los intervalos de predicción, es en gran medida debida a la no linealidad que relaciona el viento y la potencia eólica generada, pues se observan diferentes variabilidades del error dependiendo del nivel de potencia.

Además, la incertidumbre de las predicciones de viento es muy dependiente de la situación meteorológica, presentándose en ocasiones situaciones de gran estabilidad en las que no se producen cambios bruscos en el viento, o situaciones muy inestables en las que la variación del viento puede ser muy rápida y por tanto resulta mucho más difícil acertar en su predicción, lo que implicaría la necesidad de intervalos de confianza más amplios.

9.7 Evaluación de resultados obtenidos

Se hace a continuación una evaluación de los modelos propuestos de estimación de la distribución del error para el cálculo de predicciones probabilistas de potencia eólica. Para ello, se utilizan los datos de error de las predicciones de potencia agregada de horizonte $h=24$ horas, calculadas con un modelo local de Media Ponderada durante el periodo comprendido entre los meses de julio y diciembre de 2007.

9.7.1 Estimación de funciones de densidad basadas únicamente en observaciones históricas del error

En primer lugar se hará una evaluación del método de estimación de funciones de densidad utilizando el suavizado Kernel, teniendo en cuenta que la bondad de esta estimación depende del ancho de banda δ_{dens} utilizado en la función Kernel. También se incluirá como parámetro el factor de olvido λ_{prob} , que define el peso asignado a cada observación histórica de error en función de su antigüedad, asignando más peso a las observaciones más recientes para estimar la función de densidad.

Elección del ancho de banda

Para intuir los valores aproximados de ancho de banda que darán buenos resultados, partimos del valor del error cuadrático medio de predicción, RMSE. Para horizontes de predicción medios, y dependiendo del modelo de predicción concreto utilizado, el error cuadrático medio obtenido por el modelo de predicción de potencia agregada varía entre valores de 0.05 y 0.055 p.u. Concretamente, para el horizonte seleccionado $h=24$ horas, y utilizando un modelo de Media Ponderada para calcular las predicciones, el error cuadrático medio ronda el valor $RMSE=0.055$ p.u., como se pudo comprobar en la Figura 7.2.

Si se supone que la dispersión de errores de predicción sigue una distribución aproximadamente normal o gaussiana $N(\mu, \sigma_{dens})$, y si se considera despreciable el sesgo del error ($\mu \sim 0$), el valor de la desviación típica σ_{dens} de la distribución gaussiana coincidirá con el valor del RMSE ($\sigma_{dens} \sim 0.055$ p.u.). El RMSE nos proporciona así una idea aproximada del grosor de la distribución de error supuestamente gaussiana.

Por otro lado, la práctica totalidad de los puntos observados en una distribución gaussiana se encuentran dentro del intervalo $[\mu - 3 \cdot \sigma_{dens}, \mu + 3 \cdot \sigma_{dens}]$, de modo que lo que

Métodos de Predicción de la Generación Agregada de Energía Eólica

se desea es dividir este intervalo, de amplitud $6 \cdot \sigma_{dens}$, en suficiente número de clases como para que la distribución quede bien representada. Si suponemos que una distribución puede quedar bien representada cuando se dibujan entre 10 clases como mínimo y 100 clases como máximo de su histograma, podemos concluir que los valores aproximados de ancho de banda del histograma se situarán entre $6 \cdot \sigma_{dens}/100$ y $6 \cdot \sigma_{dens}/10$, que para un valor de $\sigma_{dens}=0.055$ p.u. nos da unos anchos de banda a probar en un rango entre 0.0033 y 0.033.

En la Figura 9.3 se muestra el error cuadrático medio de un histograma PIT de 20 clases que evalúa el ajuste de la función de densidad, utilizando un Kernel triangular, en función del ancho de banda δ_{dens} utilizado, con una memoria asintótica de las observaciones de error fijada en 500 horas. Para un rango bastante amplio de valores de ancho de banda se obtiene un error en el PIT aproximadamente similar. Así, con anchos de banda con una diferencia de hasta un orden de magnitud, entre 0.003 y 0.03 p.u., se obtiene un $RMSE_{PIT}$ similar, en torno a 0.07 p.u.

También se ha comprobado el resultado si el ancho de banda se calcula según una regla de ancho de banda óptimo en función del número de datos y la variabilidad de los mismos (9.5), utilizando en este caso un Kernel gaussiano. En este caso, el $RMSE_{PIT}$ resulta en un valor de 0.067 p.u., muy similar a los obtenidos anteriormente con un Kernel triangular y anchos de banda entre 0.003 y 0.3 p.u.

Por tanto, si bien la elección del ancho de banda resulta importante, los resultados parecen indicar que hay un rango razonablemente amplio para su elección. De igual modo, la elección de la función Kernel tampoco parece especialmente relevante.

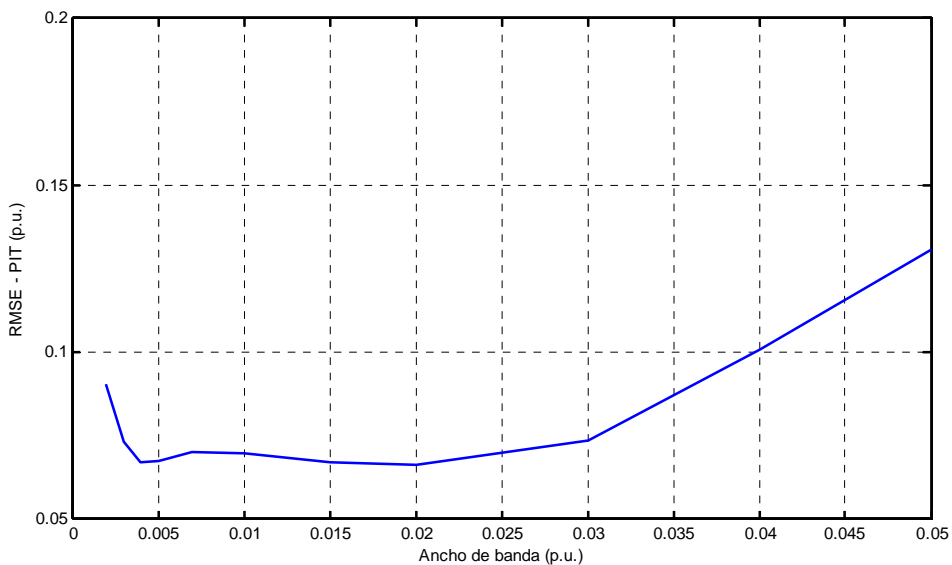


Figura 9.3: Ajuste del PIT de la estimación de la función de densidad en función del ancho de banda

La Figura 9.4 muestra los histogramas PIT correspondientes a algunos valores de ancho de banda. En la gráfica superior izquierda se muestra el PIT resultante cuando el ancho de banda es muy pequeño, en este caso de 0.002 p.u. La falta de suavizado provocada por un ancho de banda tan pequeño da lugar a un sobreajuste de la función de densidad y a que aparezcan algunos saltos en el PIT. En todo caso, dado que el histograma solo se representa con 20 clases, la altura de cada barra queda suavizada, y no se desvía demasiado de lo que sería una uniforme ideal.

El histograma de la parte superior derecha muestra el resultado de elegir un ancho de banda demasiado grueso, de 0.06 p.u. en este caso. El resultado es una forma de joroba en el PIT, lo que indica que la función de densidad estimada es demasiado gruesa con respecto a la función de densidad del error real, debido precisamente al excesivo tamaño del ancho de banda.

En los histogramas de la parte inferior izquierda e inferior derecha se muestran los PIT obtenidos utilizando un ancho de banda de 0.02 p.u., que proporcionó el menor $RMSE_{PIT}$, y un ancho de banda variable calculado en cada instante según una regla de ancho de banda óptimo respectivamente. En ambos casos se aprecia un histograma PIT bastante uniforme, lo que indica que el modelo de predicción probabilista está bien calibrado para los dos anchos de banda.

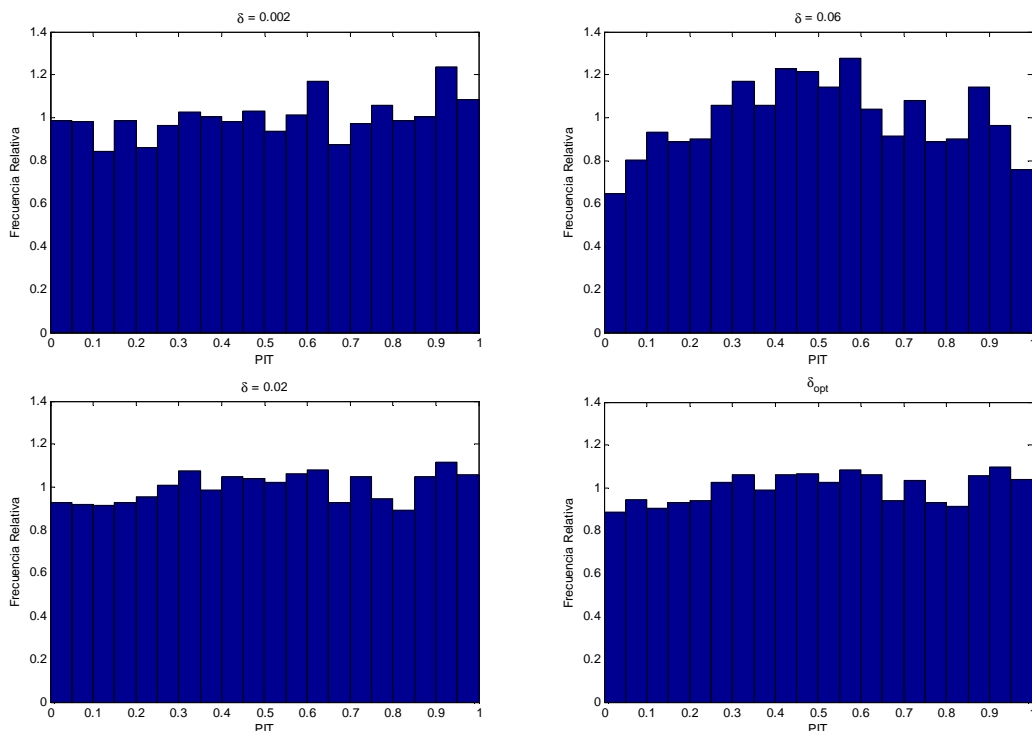


Figura 9.4: Histogramas PIT de la estimación de la función de densidad para varios anchos de banda

Métodos de Predicción de la Generación Agregada de Energía Eólica

Se debe destacar como conclusión en este caso que el ajuste de la función de densidad, pese a depender del ancho de banda δ_{dens} elegido, es poco sensible para un rango de valores bastante amplio. Si los percentiles de interés son limitados, como es el caso cuando los principales percentiles de interés son el 5%, 10%, ..., 95%, es decir, subdivisiones del 5% para poder obtener intervalos centrados en subdivisiones de un 10% de amplitud, la elección de un ancho de banda u otro no será tan relevante siempre y cuando no sea demasiado estrecho o demasiado grueso.

Memoria asintótica de las observaciones de error

A continuación se evalúa el efecto de la memoria asintótica o número de observaciones equivalentes para el cálculo de la función de densidad, relacionada con el factor de olvido λ_{prob} por la ecuación (4.39). Se hará variar esta memoria asintótica entre las 100 horas de antigüedad, correspondiente a cerca de 4 días de datos, hasta un valor de 2000 horas, equivalente a casi tres meses de observaciones. Se ha incluido también en caso de no tener en cuenta la antigüedad, según el cual todas las observaciones de error tendrían el mismo peso a la hora de determinar la función de densidad.

En la Figura 9.5 se aprecia que el ajuste del histograma PIT depende del factor de olvido, o en este caso de la memoria asintótica o número de observaciones equivalentes de error que se tengan en cuenta para modelar la función de densidad de error, que ha sido estimada usando una función Kernel triangular con un ancho de banda $\delta_{dens}=0.02$ p.u. Se obtiene un valor mínimo de error para una memoria asintótica de unas 500 horas, equivalente a unos 20 días, aunque se observan desvíos razonables, con errores inferiores a 0.1 p.u., con memorias entre 200 y 800 horas.

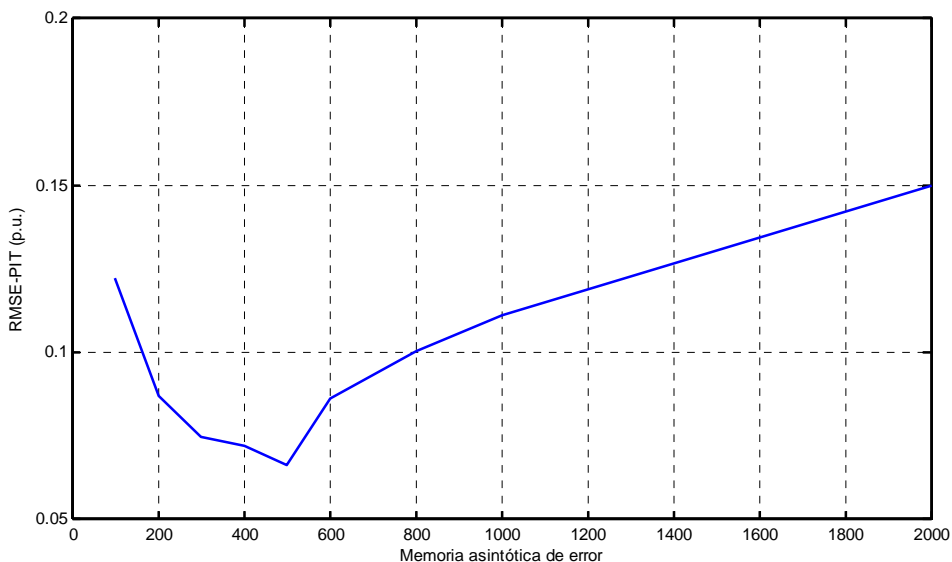


Figura 9.5: Ajuste del PIT de la estimación de la función de densidad en función de la memoria asintótica de las observaciones de error

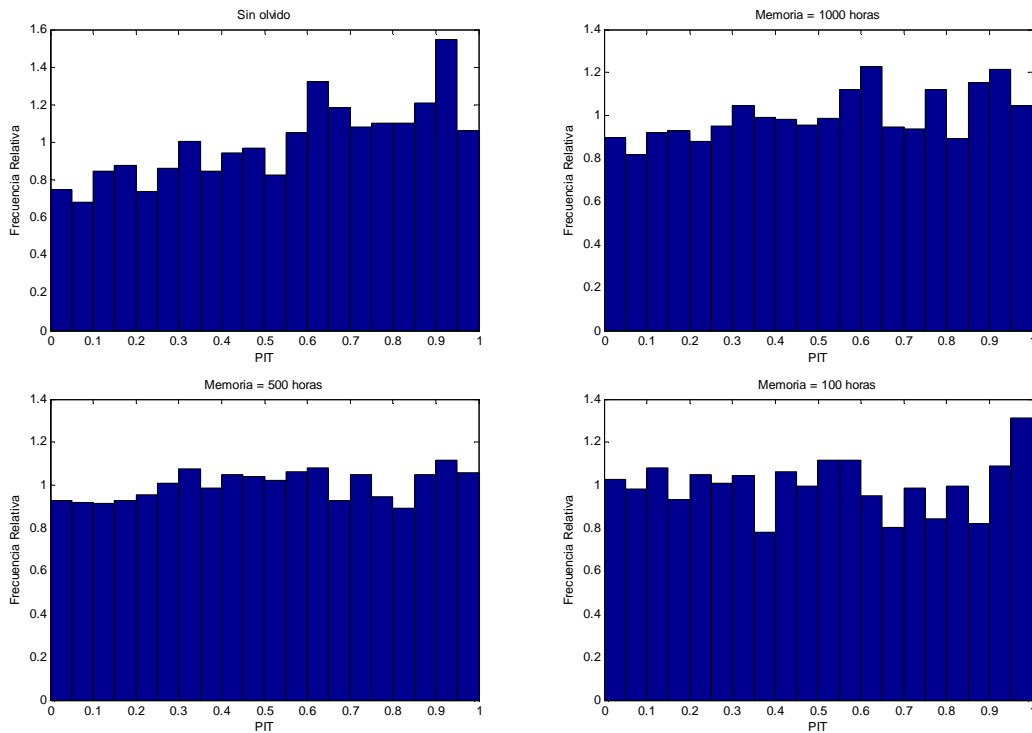


Figura 9.6: Histogramas PIT de la estimación de la función de densidad para varias memorias asintóticas de las observaciones de error

En la Figura 9.6 se muestran los histogramas PIT para algunos valores de memoria asintótica del error en el cálculo de la función de densidad. La imagen superior izquierda corresponde al caso de no tener en cuenta la antigüedad de los datos, que en este caso se ha traducido en un sesgo en la estimación del error, con lo que el histograma PIT tiene forma de rampa. Esto puede ser debido a la posible variación estacional de la distribución de error, lo que indica que resultará más ventajoso tener en cuenta cierto olvido en las observaciones de error para proporcionar adaptabilidad al modelo de predicción probabilista.

La imagen superior derecha representa el PIT para una memoria de 1000 horas, donde apenas se observa cierta desviación respecto de lo que sería una uniforme ideal. Para este factor de olvido el $RMSE_{PIT}$ resultante es de casi 0.12 p.u. (Figura 9.5). Esto parece indicar que aún se podrían tolerar errores del orden de 0.1 p.u. en el histograma PIT sin comprometer la fiabilidad del modelo.

El histograma en la parte inferior izquierda representa el caso en el que la memoria es de 500 horas, con el que se obtienen los menores desvíos en el PIT respecto de la uniforme, con un $RMSE_{PIT}$ de aproximadamente 0.07 p.u.

Mientras, el histograma en la parte inferior derecha muestra el caso en el que la memoria es demasiado reducida. En este caso se observan muchos altibajos en el histograma PIT, y un salto o pico en el extremo derecho, lo que indica que la cola de la función de densidad real es más pesada que la cola derecha de la función de

densidad estimada, con lo que varias observaciones quedaron fuera, a la derecha de la función de densidad estimada.

La influencia de la memoria de error da una idea de la existencia de variabilidad en la distribución de los errores a lo largo del tiempo, de modo que resultará más ventajoso estimar su función de densidad con una muestra relativamente reciente de datos, en vez de modelarla con el conjunto completo de observaciones históricas.

9.7.2 Predicción probabilista condicionada al nivel de potencia prevista

Una vez evaluado el método de suavizado por Kernel para estimar densidades de error basándose únicamente en las observaciones históricas del error, se procede a incluir nuevos factores que pueden influir en la distribución de los errores, y que ayudarán a hacer la estimación de incertidumbre condicionada a cada situación concreta de predicción, aumentando así su resolución.

A continuación se evalúa la influencia de incluir el nivel de potencia prevista como factor de ponderación en las observaciones históricas de error. Para ello, se calculan los $RMSE_{PIT}$ resultantes de comparar los histogramas PIT de 20 clases con una distribución uniforme ideal, en función del tamaño del ancho de banda δ_p utilizado en el Kernel de ponderación por nivel de potencia (9.9).

Debido a que existe la posibilidad de elegir entre un ancho de banda fijo o un ancho de banda variable, se comprueba la influencia de esta elección probando los resultados con ambas aproximaciones. En ambos casos, la función de suavizado utilizada ha sido un Kernel triangular.

Ancho de banda fijo

En una primera aproximación se ha utilizado un ancho de banda δ_p fijo. En la Figura 9.7 se observa la dependencia entre el tamaño de ancho de banda utilizado y la evaluación del histograma PIT. A medida que se reduce el ancho de banda, el Kernel se hace más selectivo en cuanto a la diferencia en nivel de potencia, de modo que el error del PIT disminuye, hasta el punto en el que el ancho de banda es tan pequeño que el número de observaciones utilizadas se reduce demasiado y la estimación de la función de densidad empeora drásticamente, con un $RMSE_{PIT}$ muy elevado.

En el caso representado la memoria asintótica para ponderar la antigüedad de los datos se ha fijado en 700 horas, lo que corresponde a casi un mes de observaciones, pues es la que ha proporcionado mejores resultados. Este leve incremento en el número de observaciones necesarias respecto del caso anterior basado solamente en las observaciones históricas del error, en el que no se hacía ponderación por nivel de potencia, se debe a que la inclusión de un nuevo factor de ponderación hace que disminuya el peso de las observaciones, con lo que equivalentemente se hace necesario un mayor número de datos para poder estimar correctamente la densidad.

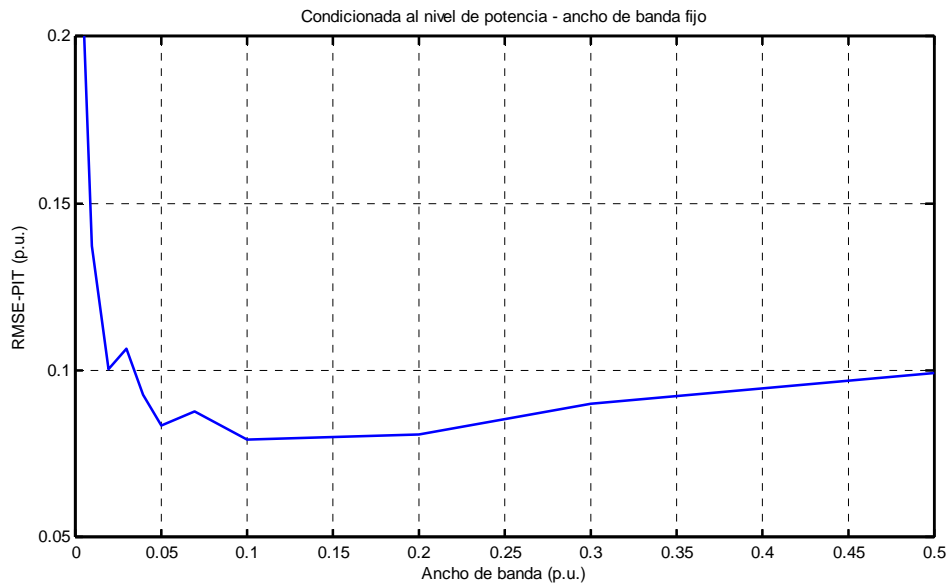


Figura 9.7: Ajuste del PIT de la estimación de la función de densidad condicionada al nivel de potencia prevista en función del ancho de banda

Los valores de error $RMSE_{PIT}$ observados para anchos de banda entre 0.1 p.u. y 0.2 p.u. son del orden de 0.08 p.u., razonablemente similares aunque algo superiores a los mínimos obtenidos en el caso anterior, en que no se tenía en cuenta la influencia del nivel de potencia. En cualquier caso, se pudo comprobar que con valores de $RMSE_{PIT}$ en el entorno de 0.1 p.u. aún se obtienen distribuciones PIT bastante uniformes, si bien la validez de los modelos de predicción probabilista se determinará más adelante, cuando se examinen los desvíos de los intervalos de predicción con respecto a su cobertura nominal.

Ancho de banda variable

Frente al uso de un ancho de banda δ_p fijo para el Kernel de ponderación por similitud en el nivel de potencia prevista, se puede utilizar un ancho de banda variable que establezca un porcentaje fijo de datos seleccionados. En este caso, el valor del ancho de banda se ajusta en cada instante de manera que el número total de datos seleccionados por el Kernel triangular coincida con el porcentaje de datos elegido.

En la Figura 9.8 se muestra la dependencia del ajuste del PIT frente al porcentaje de datos seleccionados que se encuentran cerca del nivel de previsto, y que define el grosor del ancho de banda. En este caso, la memoria equivalente se ha fijado en 1000 horas, pues es la que proporciona mejores resultados.

Métodos de Predicción de la Generación Agregada de Energía Eólica

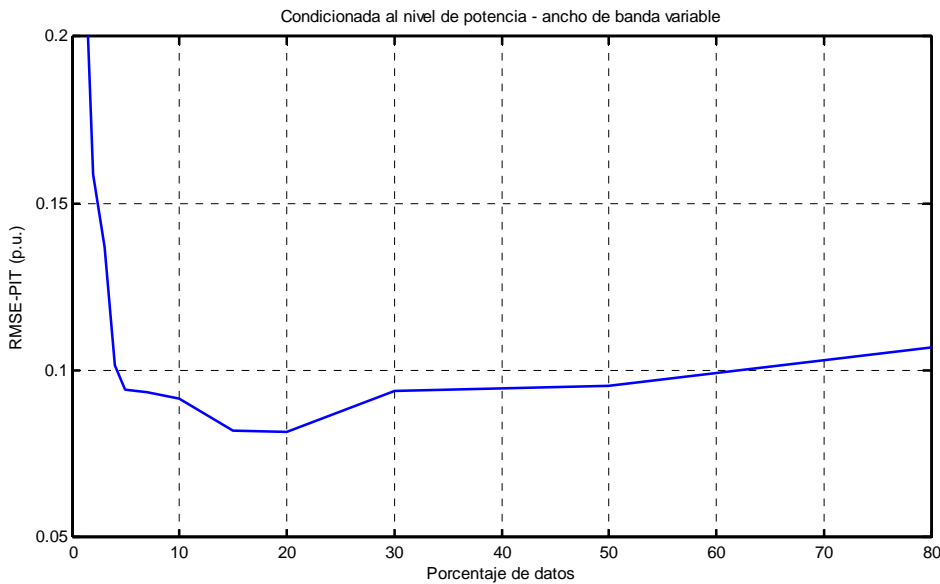


Figura 9.8: Ajuste del PIT de la estimación de la función de densidad condicionada al nivel de potencia prevista en función del porcentaje de datos seleccionados

De nuevo se observa un error mínimo en el ajuste del PIT para cierto rango de porcentaje de datos, entre el 15 y 20%. Para porcentajes de selección muy pequeños, inferiores a un 5%, aumenta drásticamente el $RMSE_{PIT}$, lo que indica de nuevo que una selección de observaciones demasiado escueta da lugar a una estimación muy pobre de la función de densidad por falta de datos.

9.7.3 Predicción probabilista condicionada al vector de viento previsto

Por último, se comprueba el funcionamiento del modelo de predicción probabilista condicionado a la distancia entre el nuevo vector de predicción de viento y los vectores históricos de viento previstos con el mismo horizonte para los periodos correspondientes a las observaciones históricas de error. Esta distancia dará una idea de la similitud entre la situación meteorológica actual prevista y las que fueron previstas en ocasiones anteriores.

En la Figura 9.9 se muestra la influencia del parámetro α_{prob} en la ecuación (9.12), correspondiente a una función de ponderación inversa con la distancia, frente al ajuste del histograma PIT. La medida de distancia utilizada para comparar los vectores de viento ha sido la distancia absoluta estandarizada con la media (4.16), pues fue la que mejores resultados proporcionó en el modelo de predicción de potencia. La memoria asintótica de las observaciones de error con la que se han obtenido los valores es de 700 horas.

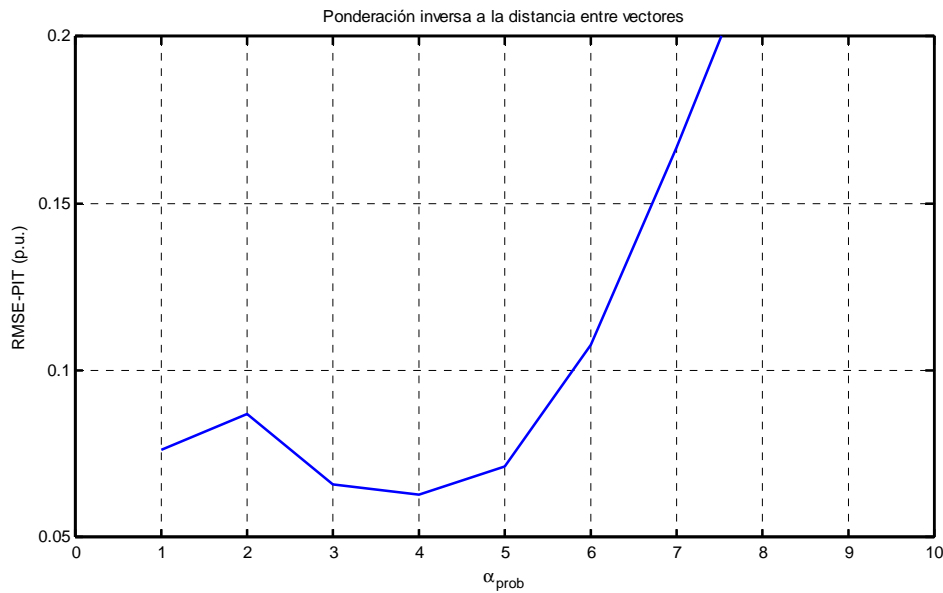


Figura 9.9: Ajuste del PIT de la estimación de la función de densidad condicionada al vector de viento previsto en función del parámetro α_{prob}

De nuevo se aprecia que hay cierto rango de valores del parámetro α_{prob} para los que el ajuste del PIT es adecuado, con valores $RMSE_{PIT}$ cercanos a 0.065 p.u., inferiores incluso los mejores que se obtuvieron en los casos anteriores. De igual modo, también se observa que si la selección es demasiado severa, correspondiente a valores de α_{prob} muy elevados, el ajuste se deteriora drásticamente por falta de observaciones para estimar la función de densidad, aumentando mucho el error $RMSE_{PIT}$.

9.7.4 Comparación de los modelos de predicción probabilista

A continuación se comparan los cuatro procedimientos alternativos para la predicción probabilista basada en estimación de densidades con funciones Kernel. En todos ellos se ha utilizado un factor de olvido de las observaciones de error λ_{prob} , y se diferencian en la posibilidad de uso de una función de ponderación para condicionar la predicción probabilista al nivel de potencia o la situación meteorológica. El modelo de predicción es de Media Ponderada y el horizonte de las predicciones es de 24 horas. Las alternativas son por tanto:

- Predicción probabilista sin condicionar
- Predicción probabilista condicionada al nivel de potencia prevista, con un ancho de banda fijo
- Predicción probabilista condicionada al nivel de potencia prevista, con un ancho de banda variable determinado por un porcentaje fijo de observaciones de potencia similar
- Predicción probabilista condicionada al vector de viento previsto

Fiabilidad o calibración. Uniformidad del PIT

Se representan en la Figura 9.10 los histogramas PIT correspondientes a los cuatro procedimientos alternativos, tomando los parámetros de ponderación que proporcionaron los mejores resultados en cada caso.

Todos los histogramas representados en la figura se aproximan razonablemente a una distribución uniforme, si bien es posible apreciar cierta mejoría con respecto al resto en el caso de condicionar la predicción probabilista mediante una ponderación inversa a la distancia entre vectores de viento previsto.

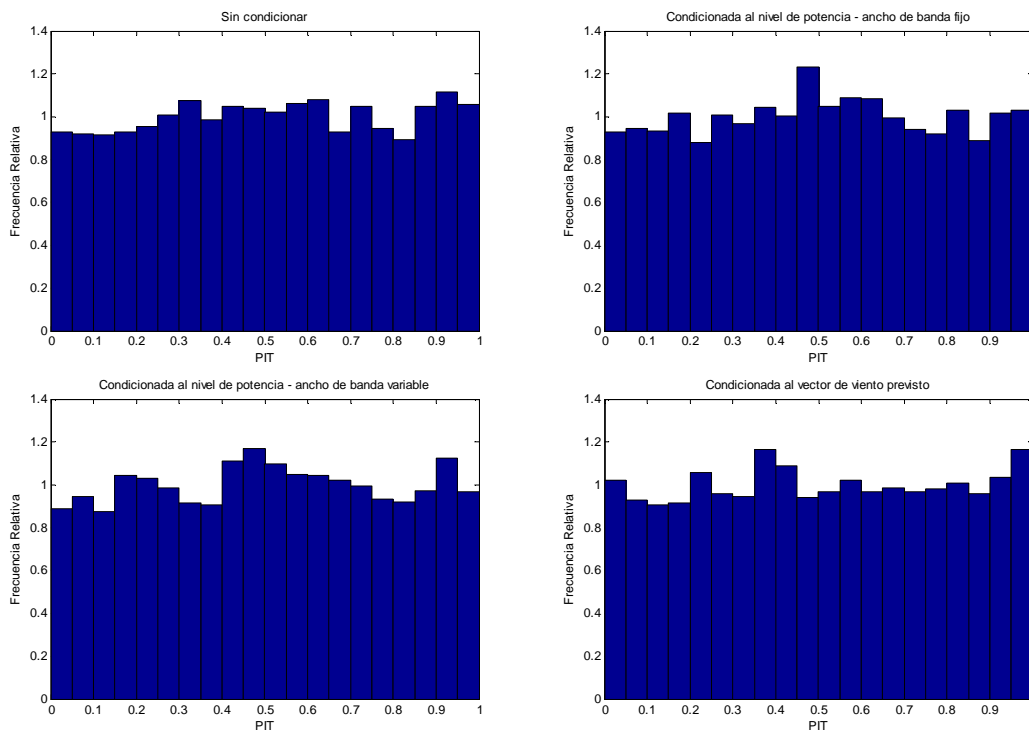


Figura 9.10: Histogramas PIT para los distintos métodos de estimación de la función de densidad expuestos

Agudeza

Una evaluación más exhaustiva de los procedimientos de predicción probabilista se puede obtener midiendo la agudeza y resolución de los intervalos de predicción que se obtienen. La Figura 9.11 muestra la agudeza o grosor medio de los intervalos de predicción en función de su cobertura. Se puede observar que, salvo leves diferencias, la agudeza resulta independiente del procedimiento de estimación de la densidad de error. Esto es debido a que la amplitud de la distribución de errores es una cualidad inherente al modelo de predicción de potencia, y no al modelo de estimación de errores.

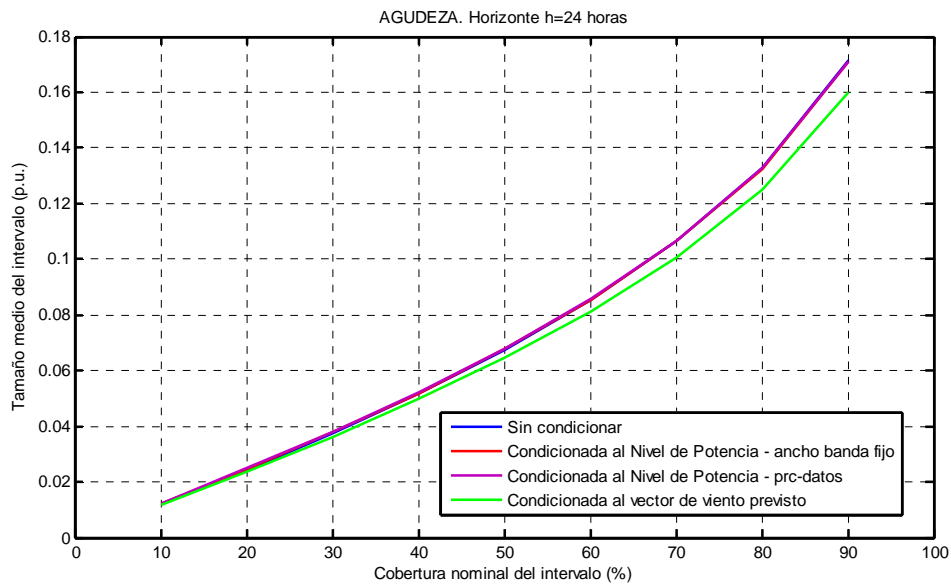


Figura 9.11: Tamaño medio de los intervalos de predicción en función de su cobertura

Se puede comprobar que los valores de agudeza en la Figura 9.11 se corresponden con el valor del error RMSE del modelo de predicción. Por ejemplo, vimos que para un horizonte de 24 horas, el error RMSE del modelo de predicción de Media Ponderada es de aproximadamente 0.055 p.u. (Figura 7.2). Si se supone que los errores siguen una distribución aproximadamente gaussiana no sesgada, el valor de la desviación típica será también de $\sigma_{dens} \sim 0.055$ p.u. Cuando la distribución de errores es gaussiana, aproximadamente el 70% de los datos se debe encontrar dentro del intervalo

$[-\sigma_{dens}, \sigma_{dens}]$, es decir, en un intervalo de tamaño $2 \cdot \sigma_{dens} = 0.11$ p.u. Se puede comprobar en la Figura 9.11 que efectivamente para un intervalo de cobertura 70% la agudeza o grosor del intervalo es de aproximadamente 0.11 p.u.

Resolución

La resolución se representa mediante la desviación estándar del tamaño de los intervalos de predicción en función de la cobertura nominal de los mismos. A igualdad de agudeza, un modelo con mayor resolución se estará adaptando mejor a las condiciones particulares de cada predicción, proporcionando intervalos de predicción con grosores más variables.

En la Figura 9.12 se representa la resolución de los diferentes procedimientos de estimación de densidad del error estudiados. En el caso de estimación teniendo en cuenta únicamente las observaciones de error y su antigüedad, la resolución es muy limitada con respecto al resto de casos, si bien ya es relativamente adaptativa gracias a la ponderación por antigüedad de los datos.

Métodos de Predicción de la Generación Agregada de Energía Eólica

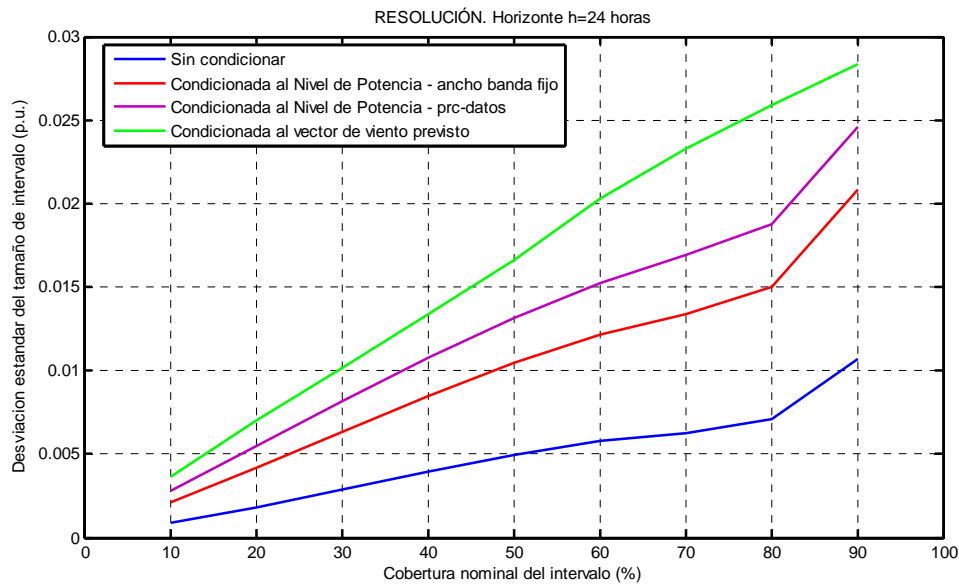


Figura 9.12: Desviación estándar de los intervalos de predicción en función de su cobertura

El hecho de incluir un coeficiente de ponderación que otorgue mayor peso a las observaciones de error que se aproximen a la situación actual hace que la resolución aumente. Así, si se tienen en cuenta los niveles de potencia prevista mediante un Kernel (9.8), la resolución casi se duplica con respecto al caso básico de no tener en cuenta niveles de potencia.

Además, se aprecia una mejora si se utiliza un ancho de banda variable con un porcentaje de datos para el Kernel, frente al uso de un ancho de banda δ_p fijo. Esto último es debido a que la concentración de observaciones también depende del nivel de potencia, encontrándose un gran número de observaciones a niveles de potencia bajos y medios, y pocas observaciones a niveles de potencia elevados o muy bajos, con lo que el uso de un ancho de banda variable permite adecuar la estimación de la función de densidad.

La mayor resolución se consigue si la predicción probabilista se condiciona al vector de viento previsto, teniendo en cuenta una función de ponderación en función de la distancia entre vectores de predicción de viento (9.12). Se observa así una variabilidad en la amplitud de los intervalos previstos aún mayor que la que proporciona un modelo condicionado al nivel de potencia prevista. Esto indica que el modelo de predicción probabilista condicionado a la distancia entre vectores de viento previsto es el que proporciona mayor valor o utilidad.

Desvío

Finalmente, en la Figura 9.13 se representa el desvío o error medio de la cobertura de los intervalos respecto de su cobertura nominal. En general se observan desvíos ligeramente negativos, que indican que la amplitud de los intervalos previstos es ligeramente superior a la amplitud real de las distribuciones de error. Es decir, en general los intervalos son demasiado conservadores y han caído más observaciones de las esperadas según su cobertura nominal.

De entre los cuatro modelos representados, el que en general presenta menores desvíos en valor absoluto ha sido el que condiciona la predicción probabilista a la distancia entre vectores de predicción de viento, mientras que el mayor desvío en términos absolutos corresponde a los modelos que ponderan las observaciones en función del nivel de potencia.

Estos desvíos resultan acordes con la forma de los histogramas PIT representados en la Figura 9.10, en los que se apreciaba mayor uniformidad para el caso de ponderación en función de la distancia entre vectores de predicción de viento. En cualquier caso, los desvíos se encuentran en el entorno del 2%, valor similar a los desvíos que proporcionan otros modelos de predicción probabilista que pueden encontrarse en la literatura [86],[126], donde se referencian desvíos entre el 2% y el 4%.

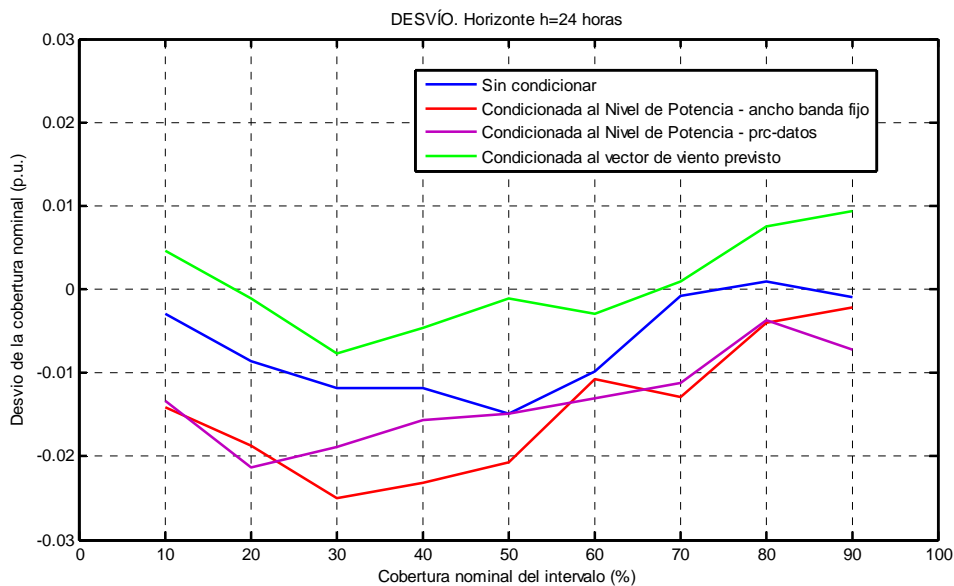


Figura 9.13: Desvío de la cobertura de los intervalos de predicción en función de su cobertura nominal

10. Conclusiones Generales

Resumen

En este capítulo se resumen las conclusiones generales obtenidas a lo largo del desarrollo de esta tesis, incluyendo una relación de las aportaciones originales realizadas. Finalmente se identifican y argumentan posibles perspectivas de investigación futura en base a este trabajo.

10.1 Conclusiones y aportaciones originales

Partiendo de la utilidad que presenta para el operador del sistema u otros agentes conocer con antelación suficiente la potencia eólica total que será generada por un conjunto de parques en una región, y teniendo en cuenta la posible reducción del error de predicción gracias a los efectos de suavizado frente a la predicción de parques individuales, en esta tesis doctoral se ha propuesto un modelo de predicción de la generación agregada de energía eólica en una región.

El método de predicción propuesto se basa en la idea de que situaciones meteorológicas similares en una región, darán lugar a una generación de energía eólica total similar. Así, las predicciones de potencia se calcularán en base a la comparación entre el nuevo vector de viento previsto en algunas coordenadas de la región y un registro histórico de vectores de viento previsto.

Para tal fin, se ha acudido a modelos no paramétricos de aprendizaje basados en la memoria, que usando un conjunto de datos históricos modelan de forma local la relación no lineal existente entre la potencia eólica total generada y el viento previsto en algunas coordenadas de la región. Se han tratado las aproximaciones de modelado local por suavizado de Media Ponderada y de Regresión Local Ponderada.

Estas aproximaciones modelan localmente el problema en el entorno del punto de entrada para el que se desea la predicción mediante un suavizado de los datos. Frente al uso de las funciones Kernel clásicas de suavizado, se ha comprobado que ante la complejidad de nuestro problema se obtienen mejores resultados usando un método de suavizado que trate de forma independiente la medida de distancias, la selección de datos cercanos y las funciones de ponderación de los datos seleccionados en función de su distancia y antigüedad.

Se han evaluado varias formas de medir la similitud o distancia entre vectores de predicción de viento, incluyendo medidas de distancia que tienen en cuenta factores de escala para las distintas dimensiones del problema o coordenadas de la región.

Se han propuesto y evaluado varios métodos de selección de datos cercanos, tanto de ancho de banda fijo como de ancho de banda variable, incluyendo algunas propuestas originales que tienen en cuenta la forma de la distribución de las distancias calculadas en cada instante.

También se han evaluado dos funciones de ponderación por distancia, basadas en un decaimiento exponencial con la distancia y en la inversa a la distancia. Estas funciones han sido modificadas para que incluyan una corrección del parámetro de ponderación en función de la concentración de datos cercanos en cada instante, determinada por la mediana de la distribución de distancias.

Se ha determinado que los mejores resultados, obtenidos con datos de energía eólica total el territorio peninsular español, corresponden a usar la medida de distancia absoluta estandarizada con la media, mientras que el método de selección de datos

Métodos de Predicción de la Generación Agregada de Energía Eólica

más adecuado ha resultado ser un ancho de banda variable determinado por un porcentaje fijo de datos. Por último, la función de ponderación que ha dado mejores resultados ha sido la inversa a la distancia de los datos, con una corrección del parámetro de ponderación en función de la mediana de la distribución de distancias en cada instante. Se ha comprobado también que la inclusión de una función de ponderación por antigüedad de los datos mejora las predicciones, especialmente para horizontes de predicción bajos.

Debido a que la relación no lineal existente entre las predicciones de viento y la potencia eólica total generada puede variar a lo largo del tiempo, se ha propuesto un método de estimación adaptativa de los tres parámetros del modelo de predicción, que definen la selección de datos y su ponderación relativa por distancia y antigüedad.

La estimación adaptativa de parámetros parte de unos valores iniciales ordenados en tres vectores de parámetros. Las combinaciones alternativas posibles de distintos valores de parámetros en los vectores dan lugar a un conjunto de predicciones alternativas que se organizan en un cubo. Una evaluación conjunta del error de las predicciones alternativas determina el sentido de desplazamiento de los vectores de parámetros hacia valores mayores o menores, con el objetivo de minimizar el error de predicción.

En este sentido, se ha comprobado que el método de estimación adaptativa de los parámetros del modelo de predicción proporciona mejores resultados que fijar los parámetros de forma global mediante un periodo de entrenamiento.

Para aprovechar las ventajas que presentan tanto los modelos de Media Ponderada como de Regresión Local Ponderada, en esta tesis se propone combinar las predicciones obtenidas con ambos modelos, que al ser de distinto orden se adaptarán mejor o peor al problema de predicción dependiendo de la complejidad local del mismo en cada instante.

Mediante un procedimiento de combinación en dos etapas que se puede encontrar en la literatura, se ha comprobado que se obtienen mejores resultados cuando se combinan las predicciones obtenidas con ambas aproximaciones. La idoneidad de combinar predicciones también se ha demostrado utilizando como datos de entrada al modelo las predicciones de viento de dos agencias meteorológicas distintas.

Como evaluación final, se han comparado las predicciones obtenidas con el modelo de predicción propuesto, las calculadas por la herramienta de predicción Sipleólico, y las proporcionadas por otra agencia de predicción de potencia agregada. Se ha comprobado que los errores del modelo agregado propuesto son menores que los de la otra agencia de predicción agregada, y entre un 12 y un 16 % inferiores a los obtenidos con Sipleólico. Como ventaja añadida, el tiempo de cálculo del modelo de predicción de potencia agregada resulta ser de apenas 30 segundos cuando calcula predicciones para horizontes de 1 a 48 horas, unas 40 veces inferior al empleado por la herramienta Sipleólico.

Para dar mayor utilidad a las predicciones de potencia es necesario que vayan acompañadas de una medida de incertidumbre. A tal fin, se ha propuesto un modelo de predicción probabilista basado en técnicas no paramétricas de estimación de densidades con funciones Kernel. Se propone una predicción probabilista condicionada al vector de viento previsto en la región, que utiliza las medidas de distancia y las funciones de ponderación utilizadas en el modelo de predicción de potencia. El modelo proporciona la función de distribución del error de predicción, a partir de la cual se puede obtener cualquier percentil o intervalo de predicción deseado.

Se ha comprobado finalmente que la obtención de la función de distribución condicionada a los vectores de viento es un método muy atractivo para el cálculo de predicciones probabilistas, pues proporcionando la misma fiabilidad y agudeza, se consigue aumentar la resolución de los intervalos de predicción obtenidos respecto de otros modelos probabilistas no condicionados o condicionados al nivel de potencia prevista.

10.2 Desarrollos futuros

El modelo de predicción de generación agregada de energía eólica propuesto se basa en la comparación de vectores de predicción de viento en algunas coordenadas de la región. Si bien en esta tesis se han utilizado aproximadamente 80 coordenadas repartidas por el territorio peninsular español, en ubicaciones cercanas al emplazamiento de los parques, no se ha establecido ningún método de selección óptima de coordenadas.

La selección de las coordenadas adecuadas puede ser de vital importancia para el modelo de predicción, tanto en su ubicación geográfica como en el número de coordenadas seleccionadas. Un número excesivamente elevado de coordenadas de predicción de viento puede dar lugar a problemas en el modelado por la excesiva dimensión del espacio de entrada, haciéndose necesaria una cantidad inabordable de datos de entrenamiento [129]. Además, el coste computacional de los modelos crece de forma exponencial con el número de dimensiones del problema. Este fenómeno es conocido en la literatura como “maldición de la dimensionalidad” [112],[130].

Se podría establecer un proceso de selección a partir del cual se seleccione el número mínimo imprescindible de variables de entrada, en nuestro caso coordenadas de predicción de viento en la región. El proceso de selección puede consistir en crear una relación ordenada de variables de entrada en función de los resultados obtenidos considerándolas una a una, seleccionando aquella que dé lugar al menor error de predicción (variables relevantes). Posteriormente se repite el proceso añadiendo una a una el resto de variables para seleccionar de nuevo la que disminuya en mayor medida el error. El proceso se repetiría hasta que la adición de una nueva variable no aporte mejoras. Este proceso también podría hacerse en sentido inverso, descartando una a una aquellas variables que al ser eliminadas den lugar a una mejora en las predicciones (variables redundantes) [91],[131].

Métodos de Predicción de la Generación Agregada de Energía Eólica

También se puede establecer un proceso heurístico de selección de coordenadas, basándose por ejemplo en un algoritmo genético, mediante el cual se establezcan poblaciones de individuos cuyos cromosomas indicarían en sus genes la selección o descarte de entre un conjunto de coordenadas disponibles. Con los procedimientos de selección de individuos, cruce y mutación propios de este tipo de algoritmos se podría obtener el conjunto óptimo de coordenadas en base a un periodo de entrenamiento, e incluso establecer un proceso evolutivo de selección de coordenadas que excluya o incluya nuevas coordenadas con el paso del tiempo.

En el contexto de la predicción probabilista, el algoritmo de predicción de potencia propuesto se podría utilizar para calcular intervalos de predicción, utilizando conjuntos de predicciones de viento ensembles. Estas predicciones ensemble son el fruto de ejecutar varias veces el modelo numérico de predicción atmosférica con pequeñas perturbaciones en las condiciones iniciales y de contorno, obteniéndose distintas evoluciones posibles de la atmósfera y por consiguiente del viento, que podrán ser parecidas o muy distintas entre sí dependiendo de la estabilidad atmosférica. El conjunto de predicciones de potencia resultante daría lugar a una estimación probabilista de la predicción de potencia.

Gracias a que el tiempo de cálculo del modelo de predicción agregada propuesto es muy reducido, de apenas 30 segundos, se podría abordar en un tiempo razonable el cálculo de predicciones de potencia agregada en la región utilizando como entradas la colección completa de predicciones de viento. Si por ejemplo el ensemble consiste en 50 predicciones alternativas de viento en las coordenadas, el modelo agregado tardaría tan solo 25 minutos en calcular la predicción de potencia agregada para todas ellas, y se podrían establecer entonces los intervalos de predicción de la potencia agregada basados en ensembles de predicción de viento.

Por último, teniendo en cuenta la disminución del error que ha mostrado el modelo de predicción de potencia agregada propuesto frente a la agregación de las predicciones obtenidas para los parques individuales que realiza Sipreólico, el modelo agregado podría ser empleado como complemento para mejorar la precisión de otros modelos de predicción a nivel de parque como Sipreólico.

La idea consistiría en desarrollar y evaluar un método para redistribuir entre los parques individuales el exceso o defecto de potencia total prevista por el modelo de predicción agregada respecto del modelo tradicional que calcula predicciones a nivel de parque y las suma. Esta redistribución debería tener en cuenta ciertos factores como la forma de la densidad de error de las predicciones individuales de cada uno de los parques o la potencia instalada en cada parque.

10.3 Publicaciones

- Miguel G. Lobo, Ismael Sánchez, “Aggregated Wind Power Prediction Method Based on Comparison of Weather Forecasting Vectors,” 4th PhD Seminar on Wind Energy in Europe, European Academy on Wind Energy (EAWE), Magdeburg, Alemania, 2008, 100-104.
- Miguel G. Lobo, Ismael Sánchez, “Aggregated Wind Power Prediction Methods Based on Distances Between Weather Forecasting Vectors,” *2009 International Conference on Clean Electrical Power*, Capri, Italia, 2009, 242-247.

Bibliografía

- [1] “Directiva 2001/77/CE del parlamento europeo y del consejo de 27 de septiembre de 2001 relativa a la promoción de la electricidad generada a partir de fuentes de energía renovables en el mercado interior de la electricidad,” *Official Journal of the European Commission*, vol. L 283, Oct. 2001, págs. 33-40.
- [2] N. Jenkins, R. Allan, P. Crossley, D. Kirschen, y G. Strbac, *Embedded Generation*, The Institution of Engineering and Technology, 2000.
- [3] G. Pepermans, J. Driesen, D. Haeseldonckx, R. Belmans, y W. D'haeseleer, “Distributed generation: definition, benefits and issues,” *Energy Policy*, vol. 33, Abr. 2005, págs. 787-798.
- [4] M.H. Bollen, *Understanding Power Quality Problems: Voltage Sags and Interruptions*, Wiley-IEEE Press, 1999.
- [5] R.M. Gnativ y J.V. Milanović, “Qualitative and quantitative analysis of voltage sags in networks with significant penetration of embedded generation,” *European Transactions on Electrical Power*, vol. 15, 2005, págs. 77–93.
- [6] *Global Wind 2008 Report*, Global Wind Energy Council - GWEC, 2008.
- [7] *World Wind Energy Report 2009*, World Wind Energy Association, 2010.
- [8] *Plan de Energías Renovables en España 2005-2010*, IDAE, Instituto para la Diversificación y Ahorro de la Energía, 2005.
- [9] *El Sistema Eléctrico Español. Avance del informe 2009*, Red Eléctrica de España, .
- [10] *El Sistema Eléctrico Español. Informe 2008*, Red Eléctrica de España, .
- [11] “Directiva 2009/28/CE del parlamento europeo y del consejo de 23 de abril de 2009 relativa al fomento del uso de energía procedente de fuentes renovables,” *Official Journal of the European Commission*, vol. L 140, Jun. 2009, págs. 16-62.
- [12] *Plan de Energías Renovables 2011-2020 (Informe de previsión)*, IDAE, Instituto para la Diversificación y Ahorro de la Energía, 2010.
- [13] J. Usaola, O. Ravelo, I. Sanchez, C. Velasco, J. Dominguez, M. Lobo, G. González, y F. Soto, “SIPREOLICO, A Wind Power Prediction Tool for the Spanish Peninsular Power System Operation,” *Global Windpower Conference*, Paris, France: 2002.
- [14] J. Rodriguez-Garcia, T. Dominguez, J.F. Alonso, y L. Imaz, “Large integration of wind power: the Spanish experience,” *2007 IEEE Power Engineering Society General Meeting*, Tampa, FL, USA: 2007, págs. 1-5.
- [15] D. Bunn, *Comparative Models for Electrical Load Forecasting*, John Wiley & Sons, 1985.
- [16] S. Thor y P. Weis-Taylor, “Long-term research and development needs for wind energy for the time frame 2000-2020,” *Wind Energy*, vol. 5, 2002, págs. 73-75.

Bibliografía

- [17] G. Giebel, L. Landberg, G. Kariniotakis, y R. Brownsword, “State-of-the-Art on Methods and Software Tools for Short-Term Prediction of Wind Energy Production,” *Proc. of the 2003 European Wind Energy Association Conference, EWEC'03*, Madrid, Spain: 2003.
- [18] J. Usaola, O. Ravelo, G. González, F. Soto, M.C. Dávila, y B. Díaz-Guerra, “Benefits for Wind Energy in Electricity Markets from Using Short Term Wind Power Prediction Tools; a Simulation Study,” *Wind Engineering*, vol. 28, Ene. 2004, págs. 119-127.
- [19] P.A. Østergaard, “Ancillary services and the integration of substantial quantities of wind power,” *Applied Energy*, vol. 83, May. 2006, págs. 451-463.
- [20] H. Abildgaard, D. Klaar, B. Kriszak, J. Rodriguez, y W. Winter, “European Wind Integration Study (EWIS) - Reference Study Towards a Successful Integration of Wind Power into European Electricity Grids,” *CIGRE Session*, Paris, France: 2008.
- [21] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, y others, “The NCEP/NCAR 40-year reanalysis project,” *Bulletin of the American Meteorological Society*, vol. 77, 1996, págs. 437-472.
- [22] G. Grell, J. Dudhia, y D. Stauffer, *A description of the fifth-generation Penn State/NCAR mesoscale model (MM5)*, Boulder, Colorado: National Center for Atmospheric Research, 1994.
- [23] R.A. Pielke, W.R. Cotton, R.L. Walko, C.J. Tremback, W.A. Lyons, L.D. Grasso, M.E. Nicholls, M.D. Moran, D.A. Wesley, T.J. Lee, y others, “A comprehensive meteorological modeling system—RAMS,” *Meteorology and Atmospheric Physics*, vol. 49, 1992, págs. 69-91.
- [24] “WAsP – the Wind Atlas Analysis and Application Program,” <http://www.wasp.dk/>.
- [25] M. Negnevitsky y C. Potter, “Innovative short-term wind generation prediction techniques,” *2006 IEEE Power Engineering Society General Meeting*, Montreal, Que., Canada: 2006, pág. 7 pp.
- [26] G. Kariniotakis, I. Marti, D. Casas, P. Pinson, T. Nielsen, H. Madsen, G. Giebel, J. Usaola, I. Sanchez, A. Palomares, R. Brownsword, J. Tambke, U. Focken, M. Lange, P. Louka, G. Kallos, C. Lac, G. Sideratos, y G. Descombes, “What performance can be expected by short-term wind power prediction models depending on site characteristics ?,” *2004 European Union Wind Energy Conference an Exhibition*, London: 2004.
- [27] L. Landberg, G. Giebel, H.A. Nielsen, T. Nielsen, y H. Madsen, “Short-term Prediction - An Overview,” *Wind Energy*, vol. 6, 2003, págs. 273-280.
- [28] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, y E. Feitosa, “A review on the young history of the wind power short-term prediction,” *Renewable and Sustainable Energy Reviews*, vol. 12, Ago. 2008, págs. 1725-1744.
- [29] G. Giebel, G. Kariniotakis, y R. Brownsword, “The State-of-the-Art in Short-Term Prediction of Wind Power - A Literature Overview,” *Position paper for the ANEMOS project, disponible en <http://anemos.cma.fr>*.

- [30] B.F. Hobbs, M.H. Rothkopf, R.P. O'Neill, y H. Chao, *The Next Generation of Electric Power Unit Commitment Models*, Springer, 2001.
- [31] G.E.P. Box y G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, 1976.
- [32] S.G. Makridakis, S.C. Wheelwright, y R.J. Hyndman, *Forecasting: Methods and Applications*, Wiley, 1997.
- [33] G. Kariniotakis, G. Stavrakakis, y E. Nogaret, "Wind power forecasting using advanced neural networks models," *IEEE Transactions on Energy Conversion*, vol. 11, Dic. 1996, págs. 762-767.
- [34] P. Flores, A. Tapia, y G. Tapia, "Application of a control algorithm for wind speed prediction and active power generation," *Renewable Energy*, vol. 30, Abr. 2005, págs. 523-536.
- [35] A. Sfetsos, "A novel approach for the forecasting of mean hourly wind speed time series," *Renewable Energy*, vol. 27, Oct. 2002, págs. 163-174.
- [36] T.S. Nielsen, A. Joensen, H. Madsen, L. Landberg, y G. Giebel, "A new reference for wind power forecasting," *Wind Energy*, vol. 1, 1998, págs. 29-34.
- [37] A. Joensen, G. Giebel, L. Landberg, H. Madsen, y H. Nielsen, "Model output statistics applied to wind power prediction," *Proceedings of the European Wind Energy Conference, EWEC'99*, Nice, France: 1999, págs. 1177-1180.
- [38] L. Landberg y S.J. Watson, "Short-term prediction of local wind conditions," *Boundary-Layer Meteorology*, vol. 70, Jul. 1994, págs. 171-195.
- [39] G. Giebel, L. Landberg, T.S. Nielsen, y H. Madsen, "Zephyr - The Next Generation Prediction System," *Proceedings of the 2002 Global Windpower Conference*, Paris, France: 2002.
- [40] U. Focken, M. Lange, y H.P. Waldl, "Previendo-a wind power prediction system with an innovative upscaling algorithm," *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark*, 2001, págs. 2-6.
- [41] H. Beyer, H. Heinemann, K. Mellinshoff, K. Mönnich, y H. Waldl, "Forecast of Regional Power Output of Wind Turbines," *Proceedings of the European Wind Energy Conference*, Nice, France: 1999, págs. 1070-1073.
- [42] B. Bailey, M. Brower, y J. Zack, "Short-term wind forecasting: development and application of a mesoscale model," *Proceedings of the European Wind Energy Conference, EWEC'99*, Nice, France: 1999, págs. 1062-1095.
- [43] J.W. Zack, M.C. Brower, y B.H. Bailey, "Validating of the Forewind model in wind forecasting applications," *Talk on the EUWEC Special Topic Conference Wind Power for the 21st Century*, Kassel, Germany: 2000.
- [44] I. Marti, T.S. Nielsen, H. Madsen, J. Navarro, A. Roldan, D. Cabezón, y C. Barquero, "Prediction Models in Complex Terrain," *Proceedings of the European Wind Energy Conference*, Copenhagen, Denmark: 2001, págs. 875-878.
- [45] I. Marti, D. Cabezón, J. Villanueva, M. Sanisidro, Y. Loureiro, E. Cantero, y J. Sanz, "LocalPred and RegioPred. Advanced tools for wind energy prediction in complex terrain."

Bibliografía

- [46] J. Jorgensen, C. Moehrlen, B. O. Gallaghóir, K. Sattler, y E. McKeogh, “HIRPOM: Description of an operational numerical wind power prediction model for large scale integration of on- and offshore wind power in Denmark,” *Poster on the Global Windpower Conference and Exhibition*, Paris, France: 2002.
- [47] T.S. Nielsen, H. Madsen, y J. Tofting, “Experiences with Statistical Methods for Wind Power Prediction,” *Proceedings of the European Wind Energy Conference*, Nice, France: 1999, págs. 1066-1069.
- [48] T.S. Nielsen y H. Madsen, “Statistical Methods for Predicting Wind Power,” *Proceedings of the European Wind Energy Conference*, Dublin, Ireland: 1997, págs. 755-758.
- [49] H. Madsen, *Models and methods for predicting wind power*, [Fredericia]: ELSAM, 1996.
- [50] T.S. Nielsen, H. Madsen, H.A. Nielsen, L. Landberg, y G. Giebel, “Prediction of regional wind power,” *Proc. of the 2002 Global Windpower Conference*, 2002.
- [51] B. Ernst, K. Rohrig, P. Schorn, y H. Regber, “Managing 3000 MW wind power in a transmission system operation centre,” *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark*, 2001, págs. 890-893.
- [52] B. Ernst y K. Rohrig, “Online-monitoring and prediction of wind power in German transmission system operation centres,” *Proceedings of the First IEA Joint Action Symposium on Wind Forecasting Techniques, Norrköping, Sweden*, 2002, págs. 125–145.
- [53] G. Hatzigargyriou, M. Matos, J. Lopes, G. Kariniotakis, D. Mayer, y J. Halliday, “‘MORE CARE’ advice for secure operation of isolated power systems with increased renewable energy penetration and storage,” *Proceedings of the 2001 European Wind Energy Conference, EWEC’01*, Copenhagen, Denmark: .
- [54] G. Kariniotakis y D. Mayer, “An Advanced On-Line Wind Resource Prediction System for the Optimal Management of Wind Parks,” *3rd MedPower Conference*, Athens, Greece: 2002.
- [55] G. Kariniotakis y P. Pinson, “Evaluation of the MORE-CARE wind power prediction platform. Performance of the fuzzy logic based models,” *Proceedings of the 2003 European Wind Energy Association Conference EWEC’03*, Madrid, Spain: 2003.
- [56] I. Sánchez, J. Usaola, O. Ravelo, C. Velasco, J. Domínguez, M. Lobo, G. González, y F. Soto, “SIPREOLICO—A Wind Power Prediction System Based on a Flexible Combination of Dynamic Models. Application to the Spanish Power System,” *Proceedings of the World Wind Energy Conference and Exhibition*, Berlin, Germany: 2002.
- [57] G. González, B. Diaz-Guerra, F. Soto, S. Lopez, I. Sánchez, J. Usaola, M. Alonso, y M. Lobo, “SIPREOLICO: Wind power prediction tool for the Spanish peninsular power system,” *2004 CIGRÉ 40th General Session & Exhibition*, Paris, France: 2004.
- [58] G. González, A. Lopez, E. Prieto, I. Sánchez, y M. Lobo, “SIPREOLICO: Wind Power Prediction Tool. Improvements in the Period 2005-2007,” *CIGRE*, 2008.

- [59] I. Sánchez, “Recursive Estimation of Dynamic Models Using Cook's Distance, With Application to Wind Energy Forecast,” *Technometrics*, vol. 48, 2006, págs. 61-73.
- [60] I. Sánchez, “Short-term prediction of wind energy production,” *International Journal of Forecasting*, vol. 22, 2006, págs. 43-56.
- [61] I. Sánchez, “Adaptive combination of forecasts with application to wind energy,” *International Journal of Forecasting*, vol. 24, Oct. , págs. 679-693.
- [62] “ANEMOS Project. Short-Term Wind Power Forecasting,” <http://anemos.cma.fr>.
- [63] “ANEMOS: Development of a Next Generation Wind Power Forecasting System for the Large-Scale Integration of Onshore & Offshore Wind Farms.,” *European Wind Energy Conference & Exhibition EWEC 2003*, Madrid, Spain: 2003.
- [64] H. Madsen, G. Kariniotakis, H.A. Nielsen, y P. Pinson, “A protocol for standardizing the performance evaluation of short-term wind power prediction models,” *Proceedings of the 2004 Global Windpower Conference and Exhibition*, Chicago, Illinois, USA: 2004.
- [65] T. Nielsen, H. Madsen, H. Nielsen, P. Pinson, G. Kariniotakis, y N. Siebert, “Short-term wind power forecasting using advanced statistical methods,” *Proceedings of European wind energy conference*, Athens, Greece: 2006.
- [66] G. Giebel, J. Badger, I. Marti, P. Louka, G. Kallos, y A. Palomares, “Short-term forecasting using advanced physical modeling—the results of the ANEMOS project,” *Proceedings of European wind energy conference*, Athens, Greece: 2006.
- [67] J. Tambke, L. von Bremen, R. Barthelmie, A. Palomares, T. Ranchin, y J. Juban, “Short-term forecasting of offshore wind farm production—developments of the ANEMOS project,” *Proceedings of European wind energy conference*, Athens, Greece: 2006.
- [68] S. Chan, D. Powell, M. Yoshimura, y D. Curtice, “Operations Requirements of Utilites With Wind Power Generation,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, 1983, págs. 2850-2860.
- [69] M.R. Milligan, A.H. Miller, y F. Chapman, “Estimating the economic value of wind forecasting to utilities,” 1995.
- [70] S. Watson, G. Giebel, y A. Joensen, “The Economic Value of Accurate Wind Power Forecasting to Utilities,” 1999.
- [71] B. Sorensen y P. Meibom, “Can wind power be sold in a deregulated electricity market?,” *Proceedings of the European Wind Energy Conference*, Nice, France: 1999, págs. 375-378.
- [72] P. Pinson, C. Chevallier, y G. Kariniotakis, “Optimizing benefits from wind power participation in electricity markets using advanced tools for wind power forecasting and uncertainty assessment,” *Proceedings of the 2004 EWEC conference*, London: 2004.
- [73] P. Pinson, C. Chevallier, y G.N. Kariniotakis, “Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power,” *IEEE Transactions on Power Systems*, vol. 22, 2007, págs. 1148-1156.

Bibliografía

- [74] P. Pinson y G. Kariniotakis, "On-line adaptation of confidence intervals based on weather stability for wind power forecasting," *Proceedings of the 2004 Global Wind Windpower Conference*, Chicago, Illinois, USA: 2004.
- [75] J. Usaola y M. Moreno, "Optimal bidding of wind energy in intraday markets," *2009 6th International Conference on the European Energy Market*, Leuven: 2009, págs. 1-7.
- [76] J.M. Angarita y J.G. Usaola, "Combining hydro-generation and wind energy: Biddings and operation on electricity spot markets," *Electric Power Systems Research*, vol. 77, Abr. 2007, págs. 393-400.
- [77] E. Castronuovo y J. Lopes, "On the Optimization of the Daily Operation of a Wind-Hydro Power Plant," *IEEE Transactions on Power Systems*, vol. 19, 2004, págs. 1599-1606.
- [78] E.D. Castronuovo y J.A.P. Lopes, "Optimal operation and hydro storage sizing of a wind-hydro power plant," *International Journal of Electrical Power & Energy Systems*, vol. 26, Dic. 2004, págs. 771-778.
- [79] J. Anagnostopoulos y D. Papantonis, "Simulation and size optimization of a pumped-storage power plant for the recovery of wind-farms rejected energy," *Renewable Energy*, vol. 33, Jul. 2008, págs. 1685-1694.
- [80] P. Pinson, G. Papaefthymiou, B. Klockl, y J. Verboomen, "Dynamic sizing of energy storage for hedging wind power forecast uncertainty," *Power Energy Society General Meeting, 2009. PES '09. IEEE*, 2009, págs. 1 -8.
- [81] P. Pinson y G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, 2004, págs. 119-132.
- [82] M. Lange, "On the Uncertainty of Wind Power Predictions---Analysis of the Forecast Accuracy and Statistical Distribution of Errors," *Journal of Solar Energy Engineering*, vol. 127, May. 2005, págs. 177-184.
- [83] G. Kariniotakis, E. Nogaret, A. Dutton, J. Halliday, y A. Androutsos, "Evaluation of advanced wind power and load forecasting methods for the optimal management of isolated power systems," *Proceedings of the European Wind Energy Conference*, Nice, France: 1999.
- [84] A. Luig, S. Bofinger, y H.G. Beyer, "Analysis of confidence intervals for the prediction of regional wind power output," *Proceedings of the European Wind Energy Conference*, Copenhagen, Denmark: 2001, págs. 725-728.
- [85] S. Bofinger, A. Luig, y H.G. Beyer, "Qualification of wind power forecasts," *CD-Proceedings of the Global Wind Power Conference*, Paris, France: 2002.
- [86] P. Pinson, "Estimation of the uncertainty in wind power forecasting," PhD dissertation, Ecole des Mines de Paris, 2006.
- [87] J. Bremnes, "Probabilistic wind power forecasts by means of statistical model," *Proceedings of the First Joint Action Symposium on Wind Forecasting Techniques*, Norrköping: International Energy Agency (IEA): 2002, págs. 49-58.
- [88] J.B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, 2004, págs. 47-54.
- [89] G. Giebel, J. Badger, L. Landberg, H.A. Nielsen, T.S. Nielsen, H. Madsen, K. Sattler, H. Feddersen, H. Vedel, J. T\ofting, y others, "Wind power prediction using ensembles," *Riso Report*, 2005.

- [90] H.A. Nielsen, T.S. Nielsen, H. Madsen, y K. Sattler, “Wind power ensemble forecasting,” *CD-Proceedings of the 2004 Global Windpower Conference, Chicago, US*, 2004.
- [91] N. Siebert, “Development of Methods for Regional Wind Power Forecasting,” PhD dissertation, Ecole des Mines de Paris, 2008.
- [92] P. Pinson, N. Siebert, y G. Kariniotakis, “Forecasting of regional wind generation by a dynamic fuzzy-neural networks based upscaling approach,” *Proceedings of the 2003 European Wind Energy Association Conference EWEC’03*, págs. 16–19.
- [93] L.V. Bremen, N. Saleck, y D. Heinemann, “Enhanced regional forecasting considering single wind farm distribution for upscaling,” *Journal of Physics: Conference Series*, vol. 75, 2007.
- [94] I. Martí, M. Gastón, y L. Frías, “Exploring the Limits of Wind Farm Grouping for Prediction Error Compensation,” *European Wind Energy Conference & Exhibition EWEC*, Athens, Greece: 2006.
- [95] U. Focken, M. Lange, y H.P. Waldl, “Reduction Of Wind Power Prediction Error By Spatial Smoothing Effects,” *Proceedings of the European Wind Energy Conference, EWEC’01*, 2001.
- [96] U. Focken, M. Lange, K. Mönnich, H. Waldl, H.G. Beyer, y A. Luig, “Short-term prediction of the aggregated power output of wind farms--a statistical analysis of the reduction of the prediction error by spatial smoothing effects,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 90, Mar. 2002, págs. 231-246.
- [97] C.G. Atkeson, A.W. Moore, y S. Schaal, “Locally Weighted Learning,” *Artificial Intelligence Review*, vol. 11, Feb. 1997, págs. 11-73.
- [98] D. Peña, *Regresión y diseño de experimentos*, Alianza Editorial Sa, 2007.
- [99] W. Härdle, *Applied nonparametric regression*, Cambridge University Press, 1992.
- [100] E.A. Nadaraya, “On Estimating Regression,” *Theory of Probability and its Applications*, vol. 9, 1964, págs. 141-142.
- [101] G.S. Watson, “Smooth Regression Analysis,” *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 26, 1964, págs. 359-372.
- [102] W.S. Cleveland, “Robust Locally Weighted Regression and Smoothing Scatterplots,” *Journal of the American Statistical Association*, vol. 74, Dic. 1979, págs. 829-836.
- [103] W. Cleveland y C. Loader, “Smoothing by local regression: Principles and methods,” *Statistical Theory and Computational Aspects of Smoothing*, Springer, 1996, págs. 10-49.
- [104] D. Ruppert y M.P. Wand, “Multivariate Locally Weighted Least Squares Regression,” *The Annals of Statistics*, vol. 22, 1994, págs. 1346-1370.
- [105] R.H. Myers, *Classical and Modern Regression with Applications*, Duxbury Press, 2000.
- [106] H. Muller, “Weighted Local Regression and Kernel Methods for Nonparametric Curve Fitting,” *Journal of the American Statistical Association*, vol. 82, 1987, págs. 231-238.

Bibliografía

- [107] T. Hastie y C. Loader, “Local Regression: Automatic Kernel Carpentry,” *Statistical Science*, vol. 8, 1993, págs. 120-129.
- [108] W.K. Härdle, M. Müller, S. Sperlich, y A. Werwatz, *Nonparametric and Semiparametric Models*, Springer, 2004.
- [109] H. Madsen, P. Pinson, G. Kariniotakis, H.A. Nielsen, y T.S. Nielsen, “Standardizing the performance evaluation of shortterm wind power prediction models,” *Wind Engineering*, vol. 29, 2005, págs. 475–489.
- [110] C.J. Stone, “Consistent Nonparametric Regression,” *The Annals of Statistics*, vol. 5, 1977, págs. 595-620.
- [111] V.A. Epanechnikov, “Non-Parametric Estimation of a Multivariate Probability Density,” *Theory of Probability and its Applications*, vol. 14, 1969, págs. 153-158.
- [112] D.W. Scott, *Multivariate density estimation*, Wiley-Interscience, 1992.
- [113] W.S. Cleveland, “Coplots, nonparametric regression, and conditionally parametric fits,” *Institute of Mathematical Statistics Lecture Notes - Monograph Series*, Hayward, CA: Institute of Mathematical Statistics, 1994, págs. 21-36.
- [114] C. Stanfill y D. Waltz, “Toward memory-based reasoning,” *Commun. ACM*, vol. 29, 1986, págs. 1213-1228.
- [115] D. Yankov, D. DeCoste, y E. Keogh, “Ensembles of Nearest Neighbor Forecasts,” *Machine Learning: ECML 2006*, 2006, págs. 545-556.
- [116] D. Shepard, “A two-dimensional interpolation function for irregularly-spaced data,” *Proceedings of the 1968 23rd ACM national conference*, ACM, 1968, págs. 517-524.
- [117] C.G. Atkeson, “Memory-based approaches to approximating continuous functions,” *Proceedings of the Workshop on Nonlinear Modeling and Forecasting*, Santa Fe Institute, New Mexico: 1990, págs. 503–521.
- [118] D.W. Aha y R.L. Goldstone, “Concept learning and flexible weighting,” *In Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, Bloomington, IL: 1992, págs. 534-539.
- [119] Y. Yang, “Combining Forecasting Procedures: Some Theoretical Results,” *Econometric Theory*, vol. 20, 2004, págs. 176-222.
- [120] D.W. Bunn, “Statistical efficiency in the linear combination of forecasts,” *International Journal of Forecasting*, vol. 1, 1985, págs. 151-163.
- [121] J. Usaola, “Probabilistic load flow with wind production uncertainty using cumulants and Cornish-Fisher expansion,” *International Journal of Electrical Power & Energy Systems*, vol. 31, Oct. 2009, págs. 474-481.
- [122] A.P. Dawid, “Present Position and Potential Developments: Some Personal Views: Statistical Theory: The Prequential Approach,” *Journal of the Royal Statistical Society. Series A (General)*, vol. 147, 1984, págs. 278-292.
- [123] T. Gneiting, F. Balabdaoui, y A.E. Raftery, “Probabilistic forecasts, calibration and sharpness,” *Journal of the Royal Statistical Society. Series B*, 2007, págs. 243--268.
- [124] F.X. Diebold, T.A. Gunther, y A.S. Tay, “Evaluating Density Forecasts with Applications to Financial Risk Management,” *International Economic Review*, vol. 39, 1998, págs. 863-83.

- [125] S. Akpınar y E. Kavak Akpınar, “Wind energy analysis based on maximum entropy principle (MEP)-type distribution function,” *Energy Conversion and Management*, vol. 48, Abr. 2007, págs. 1140-1149.
- [126] J. Juban, L. Fugon, y G. Kariniotakis, “Probabilistic short-term wind power forecasting based on kernel density estimators,” *Proc. European Wind Energy Conference*, 2007, págs. 7–10.
- [127] M. Carney y P. Cunningham, “Evaluating Density Forecasting Models,” *Dublin, Trinity College Dublin, Department of Computer Science*, vol. TCD-CS-2006-21, 2006.
- [128] J.B. Bremnes, “A comparison of a few statistical models for making quantile wind power forecasts,” *Wind Energy*, vol. 9, 2006, págs. 3-11.
- [129] B. Seifert y T. Gasser, “Variance Properties of Local Polynomials and Ensuing Modifications.”
- [130] R.E. Bellman, *Adaptive Control Processes: A Guided Tour*, Princeton University Press, 1961.
- [131] L.A. Fernández Jiménez, “Modelos avanzados para la predicción a corto plazo de la producción eléctrica en parques eólicos,” Tesis Doctoral, 2007.