

Predicción con incertidumbre en meteorología energética.
Aplicación en predicción eólica

Rafael Juan Nebot Medina

Máster Oficial SIANI.
Universidad de Las Palmas de G.C.

Máster Oficial SIANI. Universidad de Las Palmas de G.C.

Trabajo fin de Máster

Título: Predicción con incertidumbre en meteorología energética. Aplicación en predicción eólica.

Apellidos y nombre del alumno: Nebot Medina, Rafael Juan

Fecha : 3 de diciembre de 2012

Tutores:

Hernández Tejera, Francisco Mario

Hernández Cabrera, José Juan

Máster Oficial SIANI. Universidad de Las Palmas de G.C.

Agradecimientos

En primer lugar me gustaría dar las gracias a mis tutores, Mario Hernández Tejera y Jose Juan Hernández Cabrera por su excelente soporte e inapreciable atención siempre que ha hecho falta.

El tema de este trabajo se enfocó para intentar su aplicación, como parte de un convenio entre la ULPGC y el ITC, donde se pretende avanzar en métodos que permitan utilizar técnicas de aprendizaje máquina en predicción de energías renovables. Es de agradecer la colaboración entre ambas instituciones en este sentido, por el interés que tiene en Canarias conseguir cuanto avance -por pequeño que sea- se haga en el aprovechamiento de esta energía que pasa casi anónimamente sobre nuestras cabezas.

Me gustaría también agradecer a mis compañeros del ITC el buen entorno de trabajo a pesar de las circunstancias actuales. Especialmente a Yeray y a Ignacio como implicados directos en este trabajo. A Juan Francisco, Izzat, Eduardo, Carlos, Daniel, Rayco, Celia, Gonzalo, Salvador, que en distinta medida han puesto algún grano de arena para que este trabajo pueda salir adelante. Igualmente a los compañeros que por distintos motivos ya no trabajan en el ITC. Sobre todo a Roque, Jorge, Sergio y Agustín.

Finalmente, gracias a mi familia, especialmente a Víctor, Estrella, Eva y Diego.

Estudia el pasado si quieres pronosticar el futuro
Confucio

Máster Oficial SIANI. Universidad de Las Palmas de G.C.

Índice general

1. Introducción	3
2. Revisión de métodos	7
2.1. Métodos de predicción puntual	7
2.2. Métodos de predicción probabilista	9
2.3. Combinación de predicciones probabilistas	11
3. Método de remuestreo adaptado	13
3.1. Extensibilidad del método	17
3.2. Detalles de la implementación	18
3.3. Clasificador difuso	22
4. Aplicación experimental del método y resultados	25
4.1. Datos utilizados	25
4.2. Acondicionamiento de observaciones y predicciones puntuales	27
4.3. Preproceso	29
4.4. Análisis de los datos de entrada al método	30
4.5. Evaluación del método probabilista	32
5. Conclusiones, aportaciones y trabajo futuro	41
5.1. Conclusiones	41
5.2. Principales aportaciones	42
5.3. Trabajo futuro	44
Bibliografía	45

Capítulo 1

Introducción

La energía -como sector industrial- y los sistemas que la soportan constituyen uno de los pilares fundamentales de las economías actuales. De las distintas formas en que se puede aprovechar la energía, la eléctrica es la más versátil por la facilidad para distribuirla.

Este trabajo se interesa en las técnicas de predicción aplicadas a las tecnologías de conversión a energía eléctrica que tienen como energía primaria las denominadas energías renovables provenientes de la atmósfera (eólica y solar), y concretamente en las que predicen la incertidumbre. Conseguir predecir la producción renovable permite aumentar la cuota de renovables en las redes actuales y es uno de los soportes para las redes del futuro, las llamadas “smartgrids”. A su vez, disponer de una cuota alta de renovables (no solo eólica y solar) tiene múltiples ventajas. A nivel económico, favorece a las regiones que lo consiguen pues por su producción local, es decir, en el territorio de la propia región, se consigue independencia energética y robustez respecto a la producción basada en combustibles fósiles que depende de fluctuaciones de precios por distintos motivos o respecto a la producción nuclear con la que también habría dependencia exterior, aparte del impacto ambiental por la gestión de los residuos. Y es precisamente su escaso impacto ambiental respecto a las energías fósiles otra de sus ventajas.

En la producción de energías renovables destacan actualmente la energía eólica y solar, por su desarrollo, aunque se dispone de tecnología para aprovechar las olas (undimotriz), las mareas (mareomotriz), el calor proveniente del interior de la Tierra (geotérmica), la biomasa, etc.

Tanto la energía solar como la eólica tienen en común la variabilidad -en parte de componente aleatoria y por lo tanto impredecible- del potencial energético extraíble de la atmósfera. Por ejemplo en fotovoltaica, por el paso de una nube o en eólica por la oscilación que se produce en la zona sur de nuestras islas, cuando predomina el alisio y que produce, aguas abajo, una serie de vórtices de

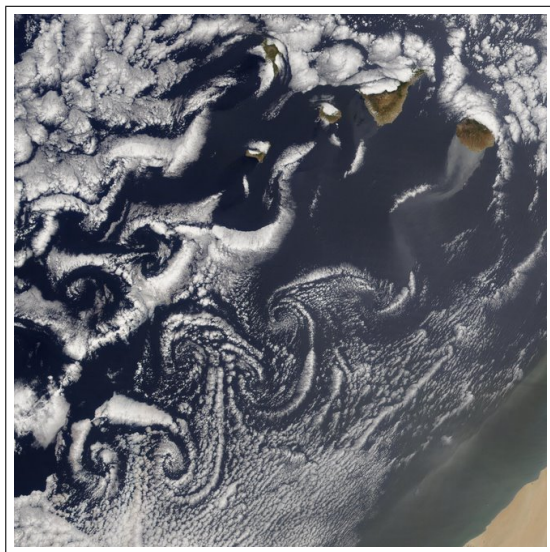


Figura 1.1: Vórtices de von Karman producidos por la influencia del alisio sobre Canarias, que ocasionan parte de la variabilidad no determinista en renovables eólica y solar

von Karman, visibles en la capa de nubes (figura 1.1).

Esta variabilidad es el principal motivo por el que hoy día las renovables siguen teniendo escasa penetración, debido a las dificultades que la incertidumbre subyacente introduce en la operación de la red.

Para reducir la incertidumbre en la producción renovable, se han desarrollado multitud de métodos para predicción de la energía, con distintos horizontes a futuro, -según la aplicación- desde minutos hasta días. Este conjunto de técnicas se estructuran en una disciplina denominada “meteorología energética”.

Existe multitud de métodos para predicción puntual de la energía. Normalmente se separa la estimación en dos etapas: predicción de las condiciones meteorológicas y modelado del sistema de conversión de energía renovable, con lo que se aprovecha la experiencia en cada una de estas áreas (meteorología por un lado y simulación de sistemas por otro). De esta cadena de modelos resulta una estimación de lo que ocurrirá dando un único valor, normalmente el valor esperado (por ejemplo, “dentro de 20 horas el parque eólico producirá energía con potencia media de 2.3 MW”).

El método analizado en este trabajo entra en la categoría de métodos de predicción denominados probabilistas, que pretenden ir un paso más allá, estimando la incertidumbre de la variable predicha. El disponer de predicciones probabilistas permite aplicar métodos estocásticos a las etapas posteriores a la predicción. Por un lado, al mercado de energía donde se deben decidir los

precios de compra y venta de cada generador y de cada consumidor. Por otro, a la operación de la red, donde se garantiza el equilibrio producción-consumo, con distintos tipos de garantía y mecanismos de salvaguarda, cada uno con un coste.

Tanto los métodos puntuales como los probabilistas se aplican a los distintos ámbitos de gestión de redes eléctricas: minimización de costes de desvío por parte de productores ¹, mercado de energía, incluyendo precios y principalmente operación (que se encarga de equilibrar en todo momento la producción y el consumo).

¹Los productores están obligados a proporcionar una estimación de su producción horaria en las próximas 48 horas. Los desvíos están penalizados con importes por MWh

Capítulo 2

Revisión de métodos

Antes de centrarnos en el método objeto de este trabajo, se hará un breve repaso por otros métodos probabilistas presentados en [8] [7].

En meteorología energética, el modelado se puede dividir en dos etapas: una de predicción de las variables meteorológicas que rodean al sistema renovable (eólico o solar) y otra de transformación de la energía por medio del sistema de conversión eólico o solar. Este trabajo se centra en la primera etapa, por lo que en la revisión solo se mencionan métodos relativos a esta primera fase. El motivo es que no se dispone de datos de producción de parques eólicos. No obstante el estudio tiene validez por dos motivos: el primer es que la mayor parte del error de una predicción meteoenergética se debe a la estimación de estado futuro de la atmósfera local, debido al carácter caótico de esta, frente al comportamiento controlado de los sistemas de conversión. El segundo es lo genérico del método, que puede ser aplicado a variables de cualquier índole.

El repaso se hará dividiendo los métodos en dos grupos: deterministas o puntuales, y probabilistas.

2.1. Métodos de predicción puntual

Existe multitud de métodos para producir predicciones puntuales de variables meteorológicas. Por un lado están los métodos basados en el análisis de series temporales: AR, MA, ARMA, ARMAX [9], redes neuronales [5], SVR [13], ... Estos métodos se pueden considerar estadísticos porque obtienen un modelo tipo “caja negra” (donde se predice el comportamiento futuro analizando el pasado, sin considerar las leyes que describen los fenómenos físicos que explicarían dicho comportamiento). Se basan en encontrar los parámetros de un modelo que minimizan algún parámetro

estadístico, tomando una muestra de valores medidos y predicciones. Una ventaja de estos métodos es que pueden ser utilizados para modelar en una sola etapa la predicción meteoenergética, utilizando medidas de la potencia producida en lugar de la variable meteorológica, en la etapa de adaptación/entrenamiento.

Un segundo tipo de métodos predice las variables meteorológicas modelando la dinámica y la física de la atmósfera. A este conjunto de métodos se les conoce como “predicción numérica del tiempo” (NWP - Numerical Weather Prediction). Según la amplitud de ámbito geográfico del que se ocupan hay de distintos tipos. Los más usados en meteorología energética son los globales y los de mesoescala. Cuando se resuelve la atmósfera global, la idea básica es resolver un problema de valor inicial: el valor futuro de las variables de estado (el estado futuro de la atmósfera) se calcula partiendo de cierto estado inicial dado -basado en observaciones- y que evoluciona siguiendo ciertas leyes expresadas como ecuaciones en derivadas parciales a las que se suman un conjunto de aproximaciones (llamadas parametrizaciones) para fenómenos no modelables a la resolución espacial/temporal en que se discretiza el dominio sobre el que se realiza la integración. Los modelos de mesoescala parten de la predicción global para resolver los fenómenos atmosféricos de una región relativamente reducida del planeta. Con los modelos globales se obtienen resultados con una resolución de por ejemplo $50 \times 50 \text{ km}^2$ o $25 \times 25 \text{ km}^2$. Los modelos de mesoescala pueden llegar a resoluciones de $500 \times 500 \text{ m}^2$. Se puede decir que los modelos de mesoescala hacen una interpolación espacial y temporal del modelo global, consiguiendo aumentar la resolución no con métodos geométricos sino respetando las leyes físicas que gobiernan la atmósfera.

Un tercer tipo de métodos de predicción se basa en tomar la predicción de un NWP más un conjunto de observaciones asociadas a predicciones pasadas y emite una nueva predicción, resultado de corregir al NWP, utilizando métodos estadísticos. A estos métodos se los denomina MOS (Model Output Statistics).

El rendimiento de los métodos estadísticos suele ser relativamente mejor que el de los modelos físicos hasta un horizonte de futuro de entre una y seis horas; a partir de este horizonte de predicción los modelos que consideran la física tienen mejor rendimiento [2] .

Todos estos métodos comparten el hecho de no tener en cuenta la incertidumbre: para determinada variable física situada en cierto punto espacio-temporal, cualquier de los modelos anteriores produce un valor puntual (escalar para variables escalares como la radiación, vectorial para el viento -velocidad y dirección-).

2.2. Métodos de predicción probabilista

Los métodos probabilistas permiten obtener predicciones según cierto formalismo relacionado con el estudio de la incertidumbre. Pueden producir desde parámetros de una distribución paramétrica, hasta distribuciones de probabilidad no paramétrica, pasando por intervalos de predicción ¹ (por ejemplo, “en 20 horas el valor de la variable x estará en el rango entre x_1 y x_2 con una probabilidad p ”).

En la bibliografía revisada, centrada en el área de la meteorología energética, las soluciones se basan en la disponibilidad de un método para obtener predicciones puntuales. Esto se explica en que existe una amplia base de métodos que obtienen predicciones puntuales (cuyos principios han sido mencionados en el apartado anterior). Y aunque existen métodos estadísticos capaces de obtener predicciones probabilistas en una sola etapa a partir de información experimental (en el área de macroeconomía y finanzas), los modelos físicos existentes hoy día no son capaces de evaluar directamente la distribución de probabilidad de las variables.

La bibliografía refiere dos tipos de método: los denominados “ensemble” (predicción por conjuntos) [15] y otros que consiguen probabilizar una predicción puntual (no se ha encontrado un nombre específico para este segundo tipo) [8]. Los métodos ensemble se basan en un mismo modelo puntual al que se aplican variantes, bien de las condiciones iniciales (la atmósfera es un sistema determinista -porque se rige por leyes físicas- pero caótico -por su dependencia del estado inicial-, lo que la convierte en no determinista a medida que nos alejamos del instante inicial de la predicción), bien de determinados parámetros del modelo. Se trata de una aplicación del método Monte Carlo para estimar la distribución de probabilidad. La aplicación de métodos ensemble requiere una capacidad de cómputo que se multiplica con cada miembro adicional del ensemble. Presentan la ventaja de que pueden determinar la predictibilidad de cada región utilizando únicamente las condiciones actuales. No obstante, la implementación de dichos modelos variando condiciones iniciales es muy compleja. De hecho actualmente los modelos ensemble variando condiciones iniciales solo los ejecutan los grandes organismos de predicción meteorológica, como el ECMWF (European Center for Medium-range Weather Forecasting) en Europa o NCEP (National Centers for Environment Prediction) en EEUU. Los modelos atmosféricos ensemble que varían las condiciones iniciales deben ser ejecutados a nivel global porque dicha variación se realiza antes de la etapa

¹Es importante distinguir “intervalos de predicción” de “intervalos de confianza”. Con el primero se hace una aserción sobre el valor futuro de la variable aleatoria mientras que el segundo la hace sobre un parámetro de la distribución de la variable aleatoria. Tienen en común que el valor real no es observable en el momento de hacer la estimación: el primero por ser una predicción y el segundo por disponer de una muestra (y no de la población).

de asimilación (tras determinar las zonas más sensibles a variaciones en el momento de ejecutar). Existe una variante menos costosa de la predicción por conjuntos (ensemble), basada en tomar predicciones sobre un instante iniciadas en distintos momentos ya pasados o utilizando modelos ya existentes provenientes de distintas fuentes. A esta variante se la denomina “Poor man’s ensemble” (PEPS - Poor man Ensemble Prediction System) [1] [10].

Existe otro conjunto de métodos (entre los que se encuentra el método de este trabajo) que suponen alguna estacionaridad en el comportamiento del método de predicción contrastada con la observación. El método desarrollado en este trabajo supone que la distribución de error al predecir distintas situaciones atmosféricas permanece constante. En [7] se evalúan varios métodos de predicción probabilista. Se introduce la posibilidad de seleccionar automáticamente las variables siguiendo criterios de información. Se sugiere la posibilidad de utilizar combinaciones de variables de manera análoga a la transformación utilizada en los métodos de máquinas kernel. Los métodos evaluados en dicha referencia se ocupan de obtener una estimación de la función de densidad de probabilidad condicional (condicionada a la información disponible en el momento de hacer la predicción probabilista, lo que puede incluir tanto predicción puntual como observaciones).

“Regresión de cuantiles usando B-Splines” (B-spline quantile regression) modela la relación entre los cuantiles y las variables explicativas usando una combinación lineal de funciones base de tipo B-Spline. El método “bosques de regresión de cuantiles” (random forests regression) se basa en la técnica de bosques aleatorios (random forests) que se utiliza para obtener regresores estables (ante cambios en el conjunto de entrenamiento) utilizando un conjunto de árboles CART, aprovechando que las observaciones ponderadas aproximan la distribución condicional. Y la técnica “estimación basada en núcleos de densidad” (Kernel density estimation) que estima la función de probabilidad condicional obteniendo una función suavizada compuesta a partir de núcleos de densidad de probabilidad localizados alrededor de las muestras.

El método de “remuestreo adaptado” se desarrolla en el próximo capítulo. Se ha seleccionado frente a los anteriores por varios motivos:

- Sencillez: como se verá la estructura del método se divide en tres pasos. Esto ha permitido trabajar en el ámbito de los métodos de predicción probabilista (que implican terminología, proceso -en entrenamiento y operación- y métodos de validación nuevos dentro del grupo de investigación) sin tener que enfrentar técnicas de optimización avanzadas o formalismos de implementación no trivial. Tras este trabajo, el esfuerzo en abordar estos métodos debe ser menor.

- Modularidad: su estructura permite el empleo de distintos métodos de aprendizaje máquina: selección automática de variables, agrupamiento automático, clasificación difusa, etc.
- Espectro de aplicabilidad: a predicción de cualquier variable, gracias al uso de distribuciones no paramétricas. Además dado que es posible introducir conocimiento experto es posible adaptarlo a ámbitos donde otros métodos no garantizar conseguir representar este conocimiento.
- Entrenamiento on-line: habitualmente en aprendizaje máquina se separa una fase de entrenamiento de una de producción. La primera suele ser bastante más costosa computacionalmente que la segunda. En este caso, el entrenamiento se realiza en cada predicción y es bastante menos costoso que un entrenamiento habitual.
- Eficiencia: el hecho de que la muestra de errores sea la propia representación de la distribución, sin requerir una regresión a una función formal, acelera los cálculos.

2.3. Combinación de predicciones probabilistas

Una vez disponible la predicción probabilista aún se puede realizar una etapa posterior. Se trata de tomar un conjunto de predicciones probabilistas obtenidas con distintos métodos o con distintas entradas y determinar la mejor manera de combinarlas utilizando rendimientos pasados.

De estudiar uno de estos métodos se encarga el Trabajo de Fin de Máster realizado por Yeyray Gutiérrez Cedrés denominado “Aprendizaje Supervisado para la combinación de modelos de predicción con incertidumbre en meteorología energética. Aplicación en predicción eólica”.

Por lo tanto, el proceso completo se compone de tres fases:

1. Predicción puntual. A partir de cierta información de partida (estado de la atmósfera, histórico en el punto donde se desea predecir) se obtiene una estimación de la variable y con antelación k , \hat{y}_{t+k}
2. Predicción probabilista a partir de predicción puntual. Se estima algún estadístico que acota la incertidumbre sobre la predicción, obteniendo una estimación de los cuantiles de la variable de interés, a partir de los cuantiles se puede obtener la distribución de probabilidad completa. Este trabajo se encarga de estudiar esta fase.

3. Combinación de predicciones probabilistas / toma de decisiones. Partiendo de varias predicciones con incertidumbre se trata de obtener una sola, considerando el rendimiento pasado de los distintos predictores utilizados como entrada.

Capítulo 3

Método de remuestreo adaptado

El método en que se basa este trabajo aparece descrito en la tesis de Pierre Pinson [10] y es revisitado en [11]. En la tesis uno de los pasos principales del método recibe el nombre de “remuestreo adaptado” (adapted resampling), y por este motivo se ha denominado así al método completo.

Como ya se ha mencionado, el objetivo del método es obtener predicciones probabilistas sobre una variable de interés (la potencia de un parque en la tesis de Pinson) a partir de predicciones deterministas o puntuales de una o más variables (variables explicativas) entre las que se suele encontrar la propia variable de interés (la velocidad del viento y la potencia del parque en la tesis de Pinson) aunque no es necesario.

En primer lugar, para entender el método a alto nivel obsérvese la figura 3.1, que refleja los pasos principales del método. En esta figura los datos se representan bien con iconos en forma de contorno de depósito, bien con gráficas. Los datos de entrada se han marcado con borde de color verdoso. Los pasos de cálculo principales, marcados con rectángulos, tienen a su izquierda la salida que producen. La salida del método está destacada con un fondo azul claro. Las flechas con fondo rojo claro indican flujo de programa mientras que las flechas de fondo blanco indican flujo de datos. Un detalle importante es que mientras que otros métodos suelen tener fases separadas de entrenamiento y operación, este método entrena y opera de manera natural en cada iteración, en un modo casi inseparable, como se puede comprobar en el diagrama. Hecha esta primera toma de contacto a continuación se describe el método con detalle formal.

El método presupone que la variable sobre la que se deben predecir los cuantiles, y_{t+k} , tendrá una distribución de probabilidad del error de predicción estimable a partir de predicciones pasadas -aquellas que sean contrastables con observaciones-. Debe disponerse para ello de un método de

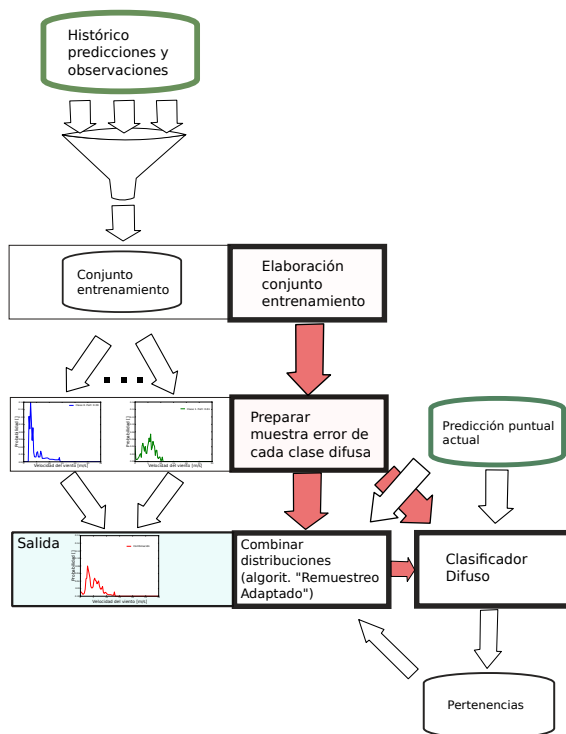


Figura 3.1: Esquema del método “remuestreo adaptado”

predicción puntual con suficiente tiempo en operación -uno o dos años-, donde añadir una nueva predicción cambiará muy poco la distribución. El instante t es aquel en que se hace la predicción, y $k = 1..$ el horizonte de predicción en unidades de tiempo. Se supone además que según las condiciones meteoenergéticas (meteorológicas en este trabajo) producidas por el sistema de predicción puntual, denotadas como $\hat{\mathbf{p}}_{t+k}$ (donde $\hat{y}_{t+k} \in \hat{\mathbf{p}}_{t+k}$ en general formada por varias variables aparte de la variable de interés), se tendrán distribuciones de error diferentes (por ejemplo, si se predice un comportamiento no lineal, la distribución de error tendrá mayor dispersión que si se predice comportamiento lineal). Una muestra de errores $\{e_j\}_{t+k}^i$ se representará por el conjunto E_{t+k}^i donde i representa cierta clase difusa, que está asociada a una condición meteoenergética tipo.

Diseñada esta partición, una predicción puntual $\hat{\mathbf{p}}_{t+k}$ podrá pertenecer a una o varias de estas clases difusas y se debe diseñar la partición de manera que una predicción sea miembro de al menos una clase difusa. El cálculo de pertenencia difusa se realiza con un clasificador difuso que recibe como entrada la predicción puntual y emite una tupla con el grado de pertenencia de la predicción a cada clase $\{m_i\} \leftarrow \mathbf{C}(\hat{\mathbf{p}}_{t+k})$. El clasificador difuso $\mathbf{C}()$ se implementa con una función de pertenencia difusa combinada (varias variables), calculada usando funciones de pertenencia trapezoidales que sirven como entrada a los predicados difusos que definen cada una de las clases.

En la etapa de preparación de muestras de error se toma cada muestra del conjunto de entrenamiento (una muestra está formada por la predicción puntual y el error observado) y se pasa por el clasificador difuso. La muestra se asocia a las clase difusas donde el grado de pertenencia de la muestra es mayor que 0.9. Es decir, si el grado de pertenencia de una predicción es mayor que 0.9, el error “observación - predicción variable interés” se asociará a la clase difusa. La distribución de error característica de la clase estará formada por el conjunto de los errores (de alguna manera es una definición “por extensión”).

Es importante destacar que el método no asume ninguna forma paramétrica para la distribución del error, simplemente se toma como aproximación a la distribución real la muestra de errores elaborada según se ha descrito en el párrafo anterior, sin suponer una forma concreta. Esto habilita al método para ser usado en predicción probabilista de variables de distinta naturaleza.

En la etapa de operación se toma la predicción y se introduce en el clasificador. En esta ocasión los grados de pertenencia serán transformados en pesos que sumando darán 1 (dividiendo cada pertenencia por la suma de pertenencias) para combinar las distribuciones de error de cada clase difusa. Para realizar la combinación y obtener los cuantiles de la variable de interés se utiliza la técnica denominada “remuestreo adaptado”. El esquema de esta técnica se detalla en el algoritmo 1. En la línea 12 se normalizan los pesos (para que su suma de 1) a partir de las pertenencias difusas. En la línea 13 se inicializa la variable de salida, puesto que hay 99 percentiles (de 1 a 99) esta variable tendrá una dimensión, con 99 valores. En la línea 14 empieza el bucle que implementa el remuestreo utilizando “bootstrapping”, con “B” repeticiones. En la línea 15 se elabora una muestra de la distribución mezcla, con tamaño “N”, a partir de elementos error extraídos de las muestras asociadas a las clases difusas. El número de elementos a extraer de cada muestra será $w_{t+k}^i \cdot N$. La función “rnd” extraerá esta cantidad de elementos de E_{t+k}^i tomando al azar -con distribución uniforme- y con reemplazo los valores de la muestra sintética. A continuación (línea 16) se toman los elementos únicos de S y se ordenan de menor a mayor. Tomar los elementos únicos tiene sentido porque computacionalmente es posible que se repitan errores. En la línea 17 se obtiene la función de distribución acumulativa, suponiendo que la probabilidad de cada elemento de la muestra S es $1/N$ y en la línea 18 se calculan los cuantiles. Termina el bucle la línea 19, acumulando las B repeticiones del remuestreo. En la línea 21 se divide la cantidad obtenida por el número de repeticiones del algoritmo de bootstrap, quedando en la variable la estimación del cuantil (percentil en este caso) para el horizonte “k”.

El método se puede resumir en tres pasos según el algoritmo 2. En la línea 1 se lee la predicción puntual a probabilizar y el conjunto de entrenamiento que permitirá hacerlo, formado por predic-

Algoritmo 1 Remuestreo adaptado

- 1: **ENTRADA:**
 - 2: E_{t+k}^i : Conjuntos con las muestras de error por clase “i”
 - 3: $\{m_{t+k}^i\}$: Conjunto con el grado de pertenencia de la predicción \mathbf{p}_{t+k} a cada clase “i”
 - 4: $y_{t+k} \in \mathbf{p}_{t+k}$: Predicción de la variable de interés
 - 5: N : N° de elementos muestra de la distribución mezcla (obtenida sintéticamente)
 - 6: **SALIDA:**
 - 7: \hat{y}_{t+k}^α : Estimación de cuantiles
 - 8: **PARAMETROS:**
 - 9: N : Tamaño de la muestra sintética (300 en el trabajo)
 - 10: B : Número de repeticiones del bucle de “bootstrapping” (40 en el trabajo)
 - 11: **INICIO:**
 - 12: $w_{t+k}^i = \frac{m_{t+k}^i}{\sum_j m_{t+k}^j}$
 - 13: $\hat{y}_{t+k}^\alpha = \mathbf{0}, \alpha = 1\% \dots 99\%$
 - 14: **para** $b = 1 \dots B$ **hacer**
 - 15: $S \leftarrow \bigcup_{i=1}^{card(\mathbf{w})} rnd(E_{t+k}^i, w_{t+k}^i \cdot N)$
 - 16: $U \leftarrow$ Elementos únicos de S , ordenados de menor a mayor
 - 17: $F(u_j) \leftarrow \frac{card(S \leq u_j)}{card(S)}, j = 1 \dots card(U), u_j \in U$
 - 18: $q^\alpha \leftarrow F^{-1}(\alpha), \alpha = 1\% \dots 99\%$
 - 19: $\hat{y}_{t+k}^\alpha \leftarrow \hat{y}_{t+k}^\alpha + y_{t+k} + q^\alpha, \alpha = 1\% \dots 99\%$
 - 20: **fin para**
 - 21: $\hat{y}_{t+k}^\alpha = \frac{\hat{y}_{t+k}^\alpha}{B}$
-

ciones pasadas y las observaciones correspondientes. En la línea 2 se obtienen las muestras de error, aplicando el clasificador difuso. En la línea 3 se aplica el clasificador difuso a la predicción puntual para obtener la pertenencia a cada clase difusa. Finalmente se aplica el remuestreo adaptado (del algoritmo 1) que recibe como entrada las muestras de error, la predicción de la variable de interés y las pertenencias difusas.

Algoritmo 2 Pasos para obtener predicción con incertidumbre según método “remuestreo adaptado”

- 1: Leer predicción puntual \mathbf{p}_{t+k} y conjunto de entrenamiento
 - 2: Obtener muestras de error E_{t+k}^i
 - 3: Aplicar clasificador difuso a predicción puntual \mathbf{p}_{t+k} obteniendo $\{m_{t+k}^i\}$
 - 4: Aplicar remuestreo adaptado (ver algoritmo 1)
-

3.1. Extensibilidad del método

Observando el algoritmo 2 se puede apreciar que es posible aplicar distintas técnicas de aprendizaje de máquina, basadas en una de las ventajas del método que es que permite utilizar de manera indistinta conocimiento experto humano o extraído de manera automática -tanto en la publicación de Pinson como en este trabajo se ha utilizado la aproximación de conocimiento experto humano-.

Esas técnicas son por ejemplo:

- Selección automática de variables. Aplicable para obtener el conjunto de variables de decisión, las que permiten situar una situación meteorológica en una clase.
- Elaboración del conjunto de entrenamiento. La elaboración del conjunto de entrenamiento implementada toma casos pasados de predicciones que cuenten con observaciones, repitiendo el proceso para cada una de las horas del horizonte de predicción. El número de casos se limita para evitar sobreentrenamiento. Un valor entre 200 y 300 puede ser razonable. La extensibilidad de esta parte surge de que es posible filtrar los casos, por ejemplo según la estación del año, o según la situación meteorológica predicha o ambas simultáneamente. Filtrar según la estación tiene sentido porque las situaciones atmosféricas de una estación suelen ser similares, con lo que la distribución de errores podría también serlo.
- Clasificador difuso. Es donde se materializa el conocimiento experto. En el presente trabajo se realiza una implementación adaptada a la información disponible, basada en la estructura

propuesta por P. Pinson. Sin embargo, el concepto de clasificación difusa aplicado al ámbito del aprendizaje automático, abre la puerta a un conjunto de métodos, incluyendo una etapa previa de identificación de clases. Un grupo de muestras (“cluster”) es más coherente o afín cuanto menor sea el espacio intercuartil según el atributo “error” (“observación - predicción”), mejorando su evaluación si se consigue un número grande de muestras y separación clara de otros grupos.

Es conveniente observar que se pueden ver las etapas de obtención del conjunto de entrenamiento y clasificación difusa como dos fases de un mismo proceso de preselección. La primera fase agrupa en conjuntos clásicos -en dos conjuntos: está en el conjunto de entrenamiento o no lo está- mientras que la segunda lo hace en conjuntos difusos -clasificando las muestras del conjunto de entrenamiento en clases más cercanas al dominio del problema-.

Actualmente es necesario cierto esfuerzo en adaptar el método a un problema particular. La combinación de métodos de aprendizaje automático y el uso de un algoritmo de optimización global (como pueden ser los algoritmos genéticos) que los sintonice, podría permitir la reducción o eliminación de parámetros al introducir problemas nuevos. Lo que se sugiere es que quizá, mediante la aplicación de métodos de aprendizaje automático, combinados con un método de optimización global se podría llegar a convertir el método en “sin-parámetros” (parameter-less).

3.2. Detalles de la implementación

Aunque como primer prototipo se realizó una implementación en Matlab, la versión utilizada se ha escrito para el lenguaje Python. El motivo principal es que soporta de modo más natural la orientación a objetos y suple bastante bien la potencia de Matlab, con sus familias de paquetes “numpy”, “scipy” y “matplotlib”.

En la implementación se siguió el diagrama de clases de la figura 3.2. Se puede apreciar una clase controladora (en el centro de la figura) con un método PRINCIPAL que implementa el algoritmo 2 siguiendo el diagrama de secuencia 3.3. Hay dos interfaces: IteradorPredicciones y ClasificadorDifuso. La primera se encarga de acceder a las predicciones puntuales y a las observaciones. Define dos funciones: una de preparación y otra de iteración por cada una de las predicciones. Esta interfaz es válida tanto para un entorno de pruebas -como el necesario para realizar un trabajo basado en datos históricos como este-, como para funcionar en operación. En este caso, cuando se inicia el programa se prepara (“prepararIteraciones”) y luego se llama a la función “obtenerSiguiente” que bloquea el hilo hasta que haya una predicción producida por el sistema de predicción puntual. Esta

clase se encarga de mantener una conexión con el sistema de predicciones puntuales. La versión implementada tendrá una caché de manera que en caso de aparecer una nueva predicción no sea necesario cargar todo el conjunto, solo las nuevas predicciones, eliminando las predicciones más antiguas. La ventaja es que esta gestión se realiza de manera transparente al resto del programa. La interfaz `ClasificadorDifuso` tiene un método que recibe una predicción puntual y retorna un conjunto con los valores de pertenencia de la predicción a cada una de las clases difusas. Es en este método donde se empaqueta el conocimiento experto en forma de reglas difusas, en el caso de este trabajo y en el de Pinson, y sería este método el encargado de implementar la parte de operación de un clasificador automático (sería necesaria la implementación de un nuevo método encargado de entrenar el clasificador en una etapa previa). El segundo método de `ClasificadorDifuso` devuelve simplemente el número de clases difusas que reconoce el clasificador, a efectos de organizar las estructuras de datos. En la clase controladora existe una función auxiliar que se encarga de llamar múltiples veces al clasificador, una vez por cada muestra del conjunto de entrenamiento, con el fin de obtener las muestras de error asociadas a cada clase que serán utilizadas en el remuestreo adaptado.

El método de mezcla de distribuciones basada en remuestreo adaptado se encuentra en la clase `RemuestreoAdaptado`, que se encarga no solo de mezclar sino también de obtener los percentiles (del 1 al 99), invocando el método “probabilizar”, que se encarga de transformar una predicción puntual -recibida como entrada- en predicción probabilista, a partir de otras dos entradas: la muestra de error (para cada clase) y la pertenencia difusa de la predicción puntual a cada clase difusa.

En la figura 3.4 se observa la presencia de una implementación de la interfaz “ClasificadorDifuso” específica para este trabajo, llamada “ClasificadorTrabajoMaster”. La implementación pretende aprovechar los datos disponibles (ver el apartado 4.1). Entre estos datos no se dispuso de datos de energía eólica. El método se ensaya prediciendo la velocidad del viento, teniendo como datos observaciones de la velocidad del viento y predicciones puntuales -también de viento- realizadas para la zona. Se decidió desarrollar un clasificador que utilizara como entrada no solo la velocidad del viento prevista sino una estimación de la predictibilidad. La predictibilidad de una situación atmosférica hace que aumente o disminuya la varianza de la distribución de error de las predicciones de dicha situación. Mayor predictibilidad implica distribuciones más estrechas y al contrario, una menor predictibilidad acarrea dificultades al predecir lo que se traduce en una distribución de error más ancha. Tras examinar las predicciones de viento y radiación comparándolas con observaciones, se observó que el modelo MM5 suele acertar en la aparición de inestabilidades

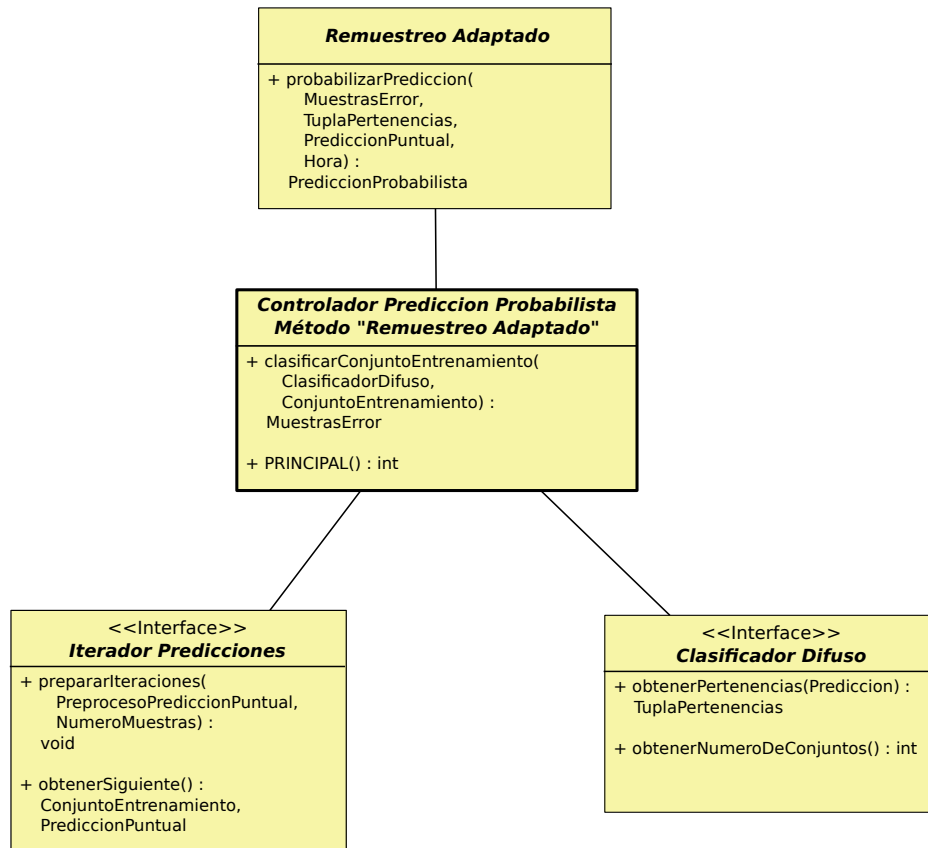


Figura 3.2: Diagrama de clases de la implementación del método “remuestreo adaptado”

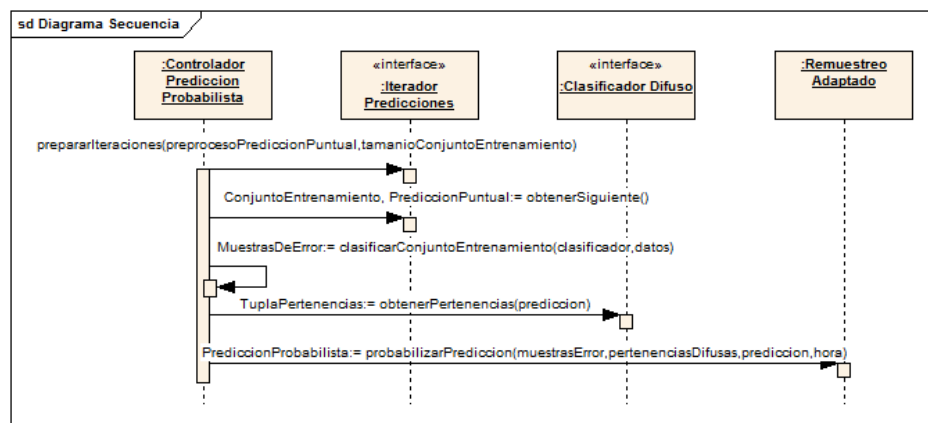


Figura 3.3: Diagrama de secuencia del método “PRINCIPAL”. Se aprecia la interacción con los tres tipos de objeto involucrados

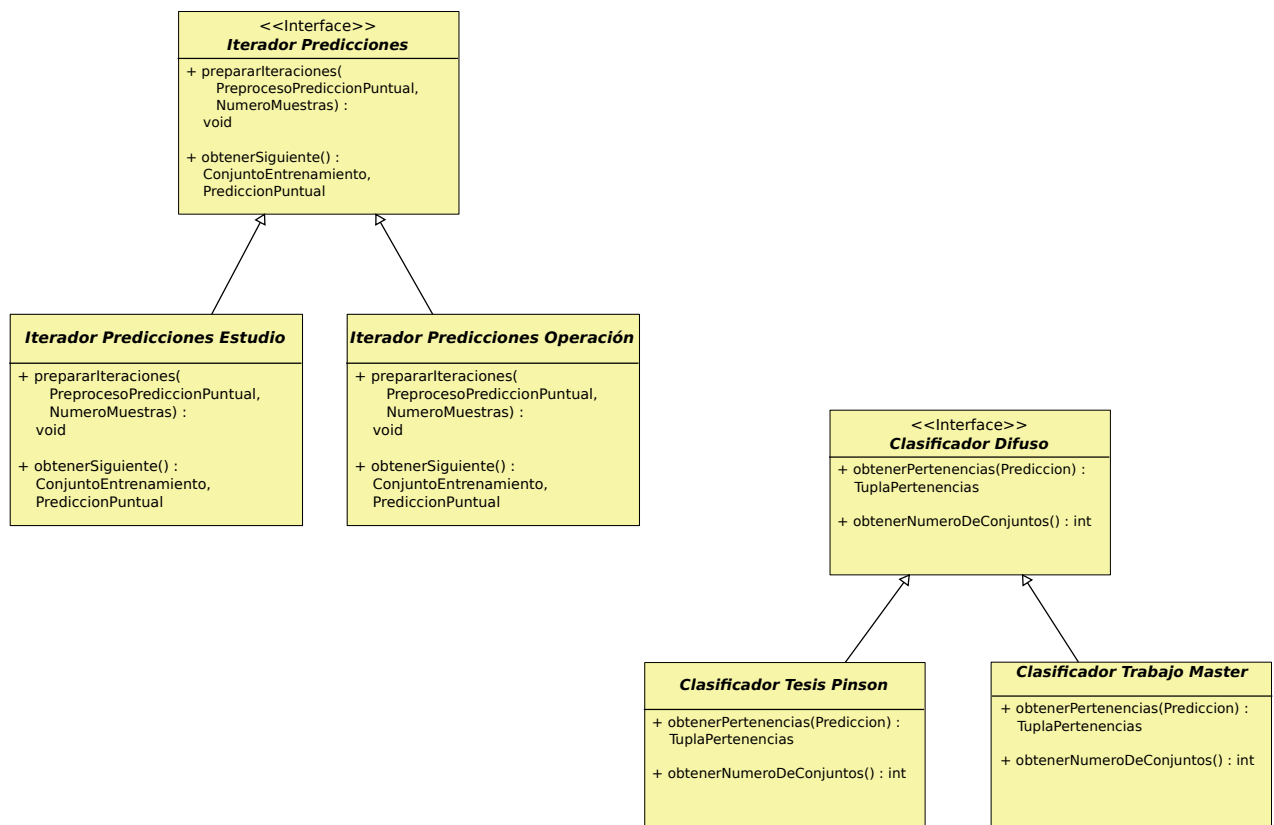


Figura 3.4: Diagrama de clases mostrando las implementaciones de las interfaces

atmosféricas pero no exactamente en la hora -“desfase”-. La predictibilidad disminuye cuando hay inestabilidad en la atmósfera. Y cuando hay inestabilidad, la nubosidad sobre la zona aumenta, lo que se refleja en una disminución de la radiación. A partir de estos hechos se diseñó un índice de predictibilidad calculado según la fórmula:

$$P_d = \frac{\sum_i G_0^i}{\sum_i G_{cd}^i} \quad (3.1)$$

Los i valores acumulados son los 24 valores del día d . Por lo tanto, P_d es un único valor para el día d . Los valores G_0^i son la radiación predicha por el sistema de predicción puntual (MM5). G_{cd}^i son los valores de un modelo de radiación de cielo despejado [14], en el que se ha estimado la radiación incidente sobre un punto a partir de la geometría sol-tierra, considerando la localización del punto sobre la superficie terrestre y considerando una turbidez habitual en el archipiélago canario. El valor resultante para P_d oscila entre 0.4 y 1.3. Se puede interpretar como el porcentaje de radiación que se producirá respecto al que habría con cielo despejado. Si el valor es uno indica cielo despejado todo el día. P_d puede ser mayor que uno, lo que indica que la turbidez ha sido menor de la habitual, algo que sucede en días posteriores a lluvias, donde se está recuperando la estabilidad. Si es menor que cierta cantidad, indica presencia de nubes y por lo tanto inestabilidad. Esta relación directa se ha utilizado para asimilar este factor con la predictibilidad y se ha realizado el clasificador según se especifica en el siguiente apartado.

3.3. Clasificador difuso

En la propuesta de Pinson se construye un clasificador basado en la previsión de velocidad del viento y de producción eólica. Sin embargo, este resultado no es aprovechable porque no se dispone de datos de potencia eólica.

Por este motivo y para experimentar con más de una variable, además de la velocidad del viento se introdujo como característica para la clasificación la predictibilidad P_d (ecuación 3.1).

Aparte de la predicción de velocidad se disponía de dos variables más: temperatura ambiente y radiación. En el momento de realizar este trabajo no se conoce una relación útil entre la temperatura y la velocidad del viento. En cuanto a la radiación, se ha utilizado indirectamente -en el cálculo de la predictibilidad-.

El clasificador difuso parte de la definición de los conjuntos difusos de las dos variables predictoras empleadas. En ambas se han utilizado funciones de pertenencia trapezoidales. Para la

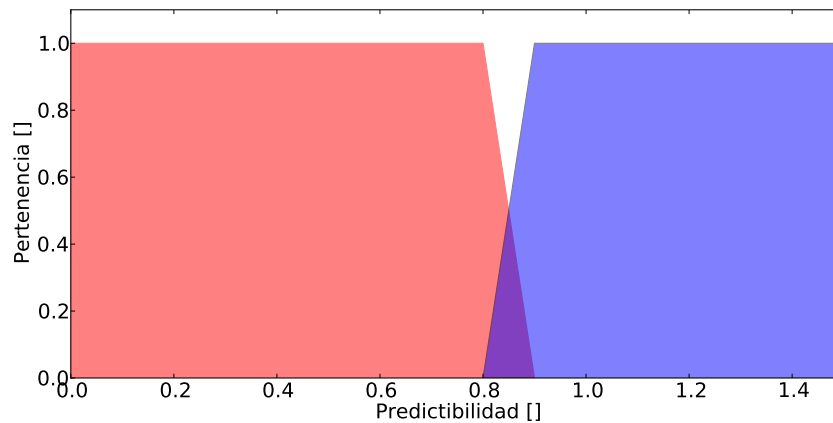


Figura 3.5: Conjuntos difusos a partir de la variable “predictibilidad”

predictibilidad, se ha considerado que cuando el valor es menor que 0.8 el día es poco predecible y predecible cuando se supera 0.9. Entre 0.8 y 0.9 hay una transición. Los conjuntos han sido etiquetados “baja predictibilidad” y “alta predictibilidad”. En la figura 3.5 se observa el resultado. El corte se ha basado en el histograma de valores de predicción obtenidos 4.7

En cuanto a la velocidad del viento, se ha dividido en tres conjuntos difusos (ver figura 3.6): baja velocidad, velocidad de transición y velocidad meseta, en referencia a las tres zonas principales de una curva de potencia. La primera zona es para velocidades donde la producción es 0 o casi 0. Las velocidades de transición son velocidades donde una máquina eólica típica produce según el cubo de la velocidad, con lo cual es una zona muy delicada porque el error también va al cubo. Por último, se define una zona “meseta” que corresponde a una zona donde la máquina funciona a potencia nominal y es capaz de regular sus subsistemas para producir potencia de manera constante. Un error dentro de esta zona no es tan importante. Una buena introducción sobre producción eólica es [12].

A partir de aquí se definen seis reglas difusas que forman seis conjuntos difusos en los que se clasificarán las muestras y las predicciones del conjunto de entrenamiento, que están formadas por tuplas “(Predictibilidad, Velocidad del viento)” y por el error “Velocidad observada - Velocidad predicha”:

- Clase 0. Baja velocidad del viento y día predecible
- Clase 1. Velocidad de transición y día predecible
- Clase 2. Velocidad “meseta” y día predecible

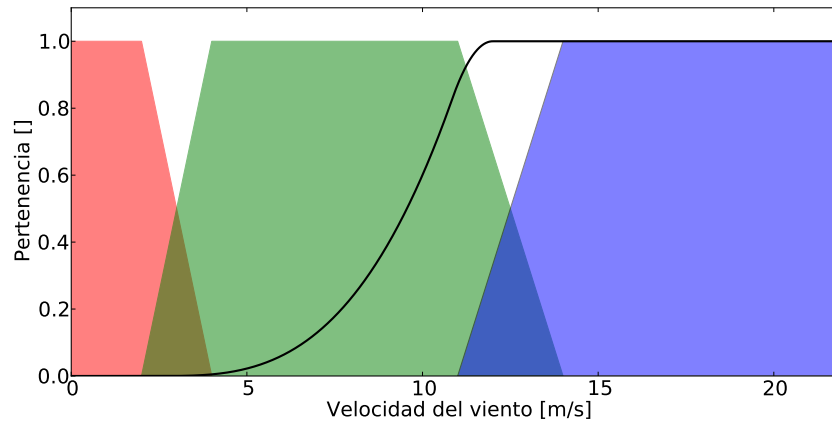


Figura 3.6: Conjuntos difusos a partir de la variable “velocidad del viento”. Se superpone una curva de potencia estándar con velocidad de arranque 2.5 m/s y de meseta 12 m/s

- Clase 3. Baja velocidad del viento y día poco predecible
- Clase 4. Velocidad de transición y día poco predecible
- Clase 5. Velocidad “meseta” y día poco predecible

La pertenencia de un par (Predictibilidad, Velocidad) a cada una de estas reglas es el producto de la pertenencia a los conjuntos difusos base.

Capítulo 4

Aplicación experimental del método y resultados

El método introducido en los capítulos anteriores se ha aplicado a la predicción del viento en un punto de la geografía canaria del que se disponen observaciones de calidad y a la vez predicciones de la misma variable en una serie larga, obtenidas utilizando el sistema NWP “MM5” (muy conocido y utilizado por ser de código abierto).

Además, se tienen también predicciones de la radiación incidente, que serán utilizadas para calcular un estimador de la predictibilidad.

Velocidad del viento y predictibilidad serán las entradas del clasificador desarrollado para el trabajo.

4.1. Datos utilizados

Las observaciones utilizadas han sido tomados por anemómetros ubicados en las instalaciones del ITC (Instituto Tecnológico de Canarias) en Pozo Izquierdo (Santa Lucía), en el punto de Gran Canaria que se señala en la captura de Google Earth mostrada en la figura 4.1.

En la figura 4.2 (captura de Google Street View), se puede observar la torre donde se encuentran instalados los sensores. Aunque la torre cuenta con anemómetros a distintas alturas, se han seleccionado los dos sensores ubicados a 10 m. de altura (dos: uno orientado al sureste otro al noroeste), porque es a esta altura donde se obtiene la predicción de viento de MM5 (MM5 emite otra variable 3D con la estimación de la velocidad del viento a distintas alturas de presión, que requiere un tratamiento para su conversión a metros más una interpolación en caso de que no se

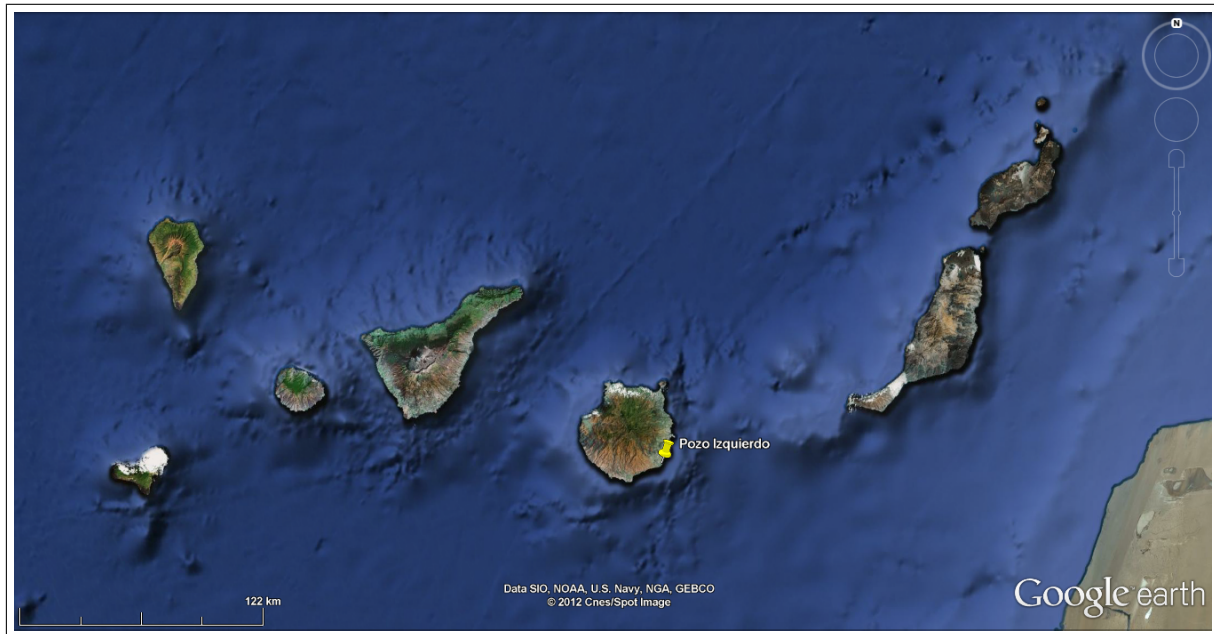


Figura 4.1: Ubicación del punto donde se realiza la medición. (c) Google 2012

tenga el resultado a la altura deseada).

Las predicciones de velocidad del viento son la salida del modelo MM5, y han sido cedidas también por el ITC. Se dispone de registro de la velocidad a 10 metros sobre el suelo que coincide con las observaciones elegidas. El modelo MM5 resuelve el valor de velocidad del viento (y otras muchas variables) en puntos de una malla cuasi-rectangular. De esta malla, se ha tomado el punto más cercano a la ubicación de la torre como predicción de lo que sucederá en este punto. Esta aproximación se considera razonable porque el punto se encuentra en una zona llana de la isla, donde además se ha hecho un estudio que concluye que mediciones en puntos situados a 2 kilómetros tienen una correlación muy alta. Las predicciones son horarias, con un horizonte de hasta 120 horas de adelanto (5 días). Se tiene una predicción por día, que empieza siempre a las 00:00 del día. Los datos se organizan en forma matricial, con tantas filas como predicciones y 120 columnas (una para cada hora de adelanto), para facilitar su proceso. Las celdas sin predicción se marcan con un “Not a Number” (NaN).

También se consideran las predicciones de radiación, igualmente horarias y con la misma forma matricial que las predicciones de velocidad del viento. Se utilizarán para estimar la predictibilidad según se ha descrito en 3.3.



Figura 4.2: Torre anemométrica en instalaciones del Instituto Tecnológico de Canarias en Pozo Izquierdo (Gran Canaria). (c) Google 2012

4.2. Acondicionamiento de observaciones y predicciones puntuales

El objetivo es trabajar con observaciones y predicciones de frecuencia horaria. Con las predicciones no hay que hacer nada porque están ya a esa frecuencia. Sin embargo, las observaciones se tienen cada minuto o cada tres minutos, dependiendo del subperiodo considerado. Además las observaciones se toman desde dos sensores: uno orientado hacia el sureste y el otro hacia el noroeste, preparados para medir el viento predominante en la zona, el alisio.

Para obtener una sola serie de observaciones de velocidad del viento horaria a 10 metros, se promedian las dos series con la resolución temporal de partida y posteriormente el resultado se diezma, promediando los valores de cada hora. Si en una hora falta alguno de los datos, se descarta la hora completa (esto es posible hacerlo por la cantidad de observaciones disponible; en caso de que hubiera menos información habría que hacer una media horaria extrapolando el valor a partir de los datos disponibles).

Para facilitar el proceso posterior, los datos observados se ponen en forma matricial con la misma forma que las predicciones: una fila por predicción (con tantas filas como días se ha hecho la predicción) y 120 columnas, una para cada horizonte horario de predicción. Los elementos (i, j) de las matrices de predicción y observación se hacen coincidir en el tiempo. De esta manera se simplifica la implementación de los cálculos. Por ejemplo para obtener los errores, mediante el uso de lenguajes como Matlab o Python es tan sencillo como escribir la sentencia “ $E = O - P$ ”

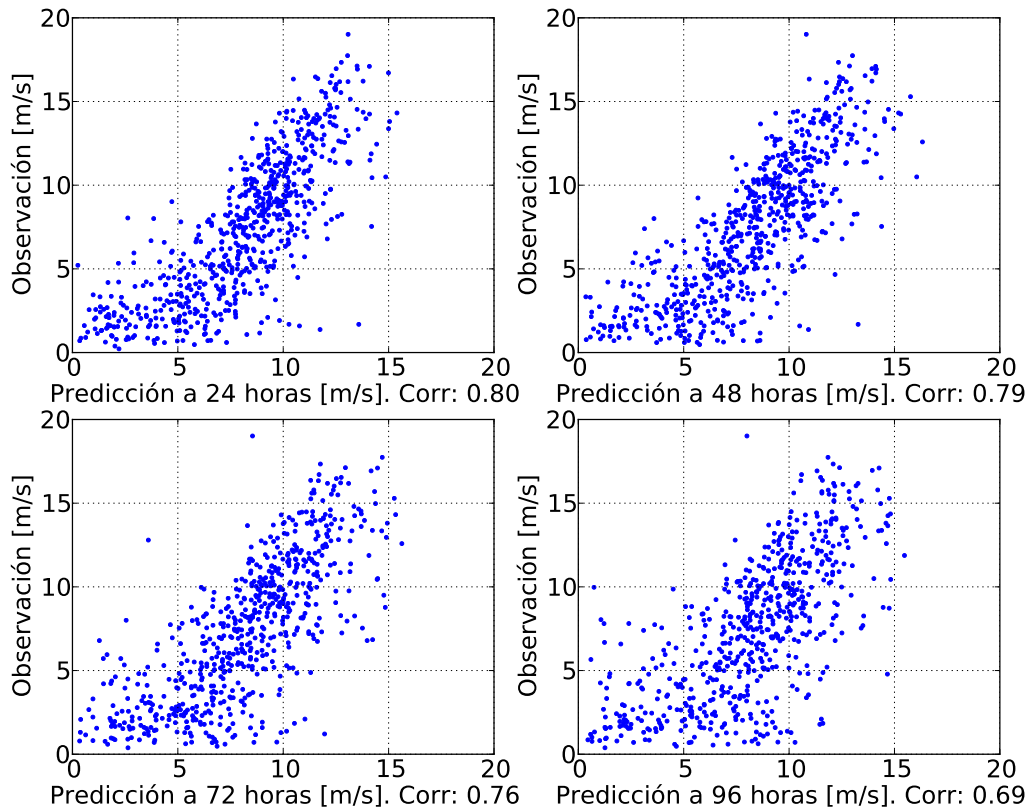


Figura 4.3: Nube de puntos “predicciones vs observaciones”. Se muestran distintos horizontes de predicción.

donde O es la matriz de observaciones, P la matriz de predicciones y E la matriz resultante con los errores.

El periodo temporal abarcado por los datos es de más de dos años. Sin embargo, hay que considerar los huecos sin información. Teniendo en cuenta que un par predicción - observación no es utilizable si falta cualquiera de los miembros del par, el número de datos se reduce. Aún así, los datos aprovechables son 78357, un 65 % de los potenciales en el periodo.

En la figura 4.3 se puede apreciar la correlación entre observaciones y predicciones, para distintos horizontes de predicción. Se observa que la nube se dispersa a medida que nos alejamos del punto de la predicción.

	+1 a +24	+25 a +48	+49 a +72	+73 a +96	+97 a +120
d_{i-3}					
d_{i-2}					
d_{i-1}					
d_i					

Figura 4.4: Configuración de predicciones para hacer una combinación “ensemble”

4.3. Preproceso

En la figura 4.3, donde se representan observaciones frente a predicciones en nubes de puntos de diferentes horizontes de predicción, se aprecia el comportamiento del sistema de predicción original basado en MM5. Antes de pasar a la etapa de predicción probabilista es posible aplicar dos mejoras que de manera sencilla consiguen un resultado mejor.

La primera es corregir el sesgo (“bias”). Para ello se toman los errores del conjunto de entrenamiento y se calcula el promedio, realizando la operación por separado para cada horizonte de predicción k . La corrección se consigue sumando a la predicción original este promedio.

La segunda corrección consiste en combinar -con una suma ponderada y lineal- varias predicciones puntuales, realizadas en días distintos. Esta operación se puede considerar un conjunto (“ensemble”) porque se están tomando varias predicciones que han sido realizadas con condiciones iniciales diferentes, consiguiendo un “poor man’s ensemble”. Al hacer esta predicción por conjuntos hay que renunciar a horizonte de predicción, ya que el miembro más antiguo del conjunto deja de tener información. Se pasa de tener una predicción a 120 horas a tener la predicción hasta 48 horas. En la figura 4.4 se observa la configuración elegida para el ensemble. Se obtendrá la predicción de +25 a +48 tomando un promedio ponderado de las predicciones señaladas en gris, que predicen los mismos instantes en días distintos. Como se observa, la predicción resultante podría abarcar 48 horas. Pero las 24 primeras no serán útiles, porque como ya se ha mencionado, es necesario disponer de la predicción antes de las 10 de la mañana para hacer el despacho diario. Como contrapartida, el intervalo de +25 a +48, especialmente útil para realizar el despacho diario, se ve mejorado.

Se pueden obtener otras dos correcciones combinando la corrección de sesgo y el ensemble en distinto orden. Puesto que ambos operadores son lineales el orden de aplicación de las correcciones no influye, con lo que la predicción resultante es exactamente igual en ambos casos y mejora las dos aplicadas por separado.

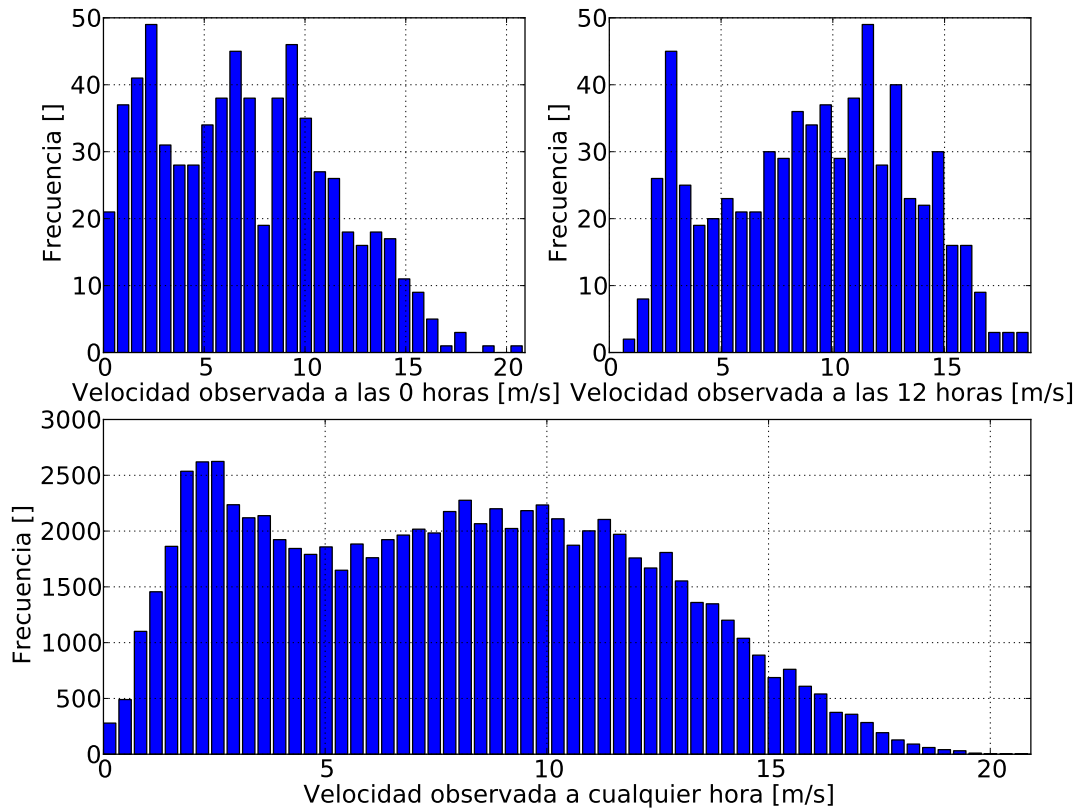


Figura 4.5: Distribución de la velocidad del viento. Arriba izda., velocidad a medianoche; arriba dcha., velocidad a mediodía. Abajo, velocidad tomando todos los datos

4.4. Análisis de los datos de entrada al método

Antes de pasar a experimentar el método con los datos, tanto observaciones como predicciones puntuales (crudas o preprocesadas), se representan algunas de sus características estadísticas.

Las observaciones de velocidad del viento se distribuyen según se puede observar en la figura 4.5. Mientras que en las gráficas superiores, donde se muestran los histogramas de los datos a las 00 y a las 12, no se aprecia una distribución clara, en la gráfica inferior, con un histograma de todas las observaciones (sin importar la hora), se puede apreciar el carácter bimodal de la velocidad del viento en Pozo Izquierdo. Además, en las gráficas superiores se puede apreciar que a mediodía hay más velocidad del viento que a medianoche y que a mediodía hay cierta bimodalidad.

Con el fin de hacer una comparación cualitativa con las predicciones, en la figura 4.6 se puede apreciar la distribución de velocidades sintetizada por el sistema de predicción MM5, tomando las

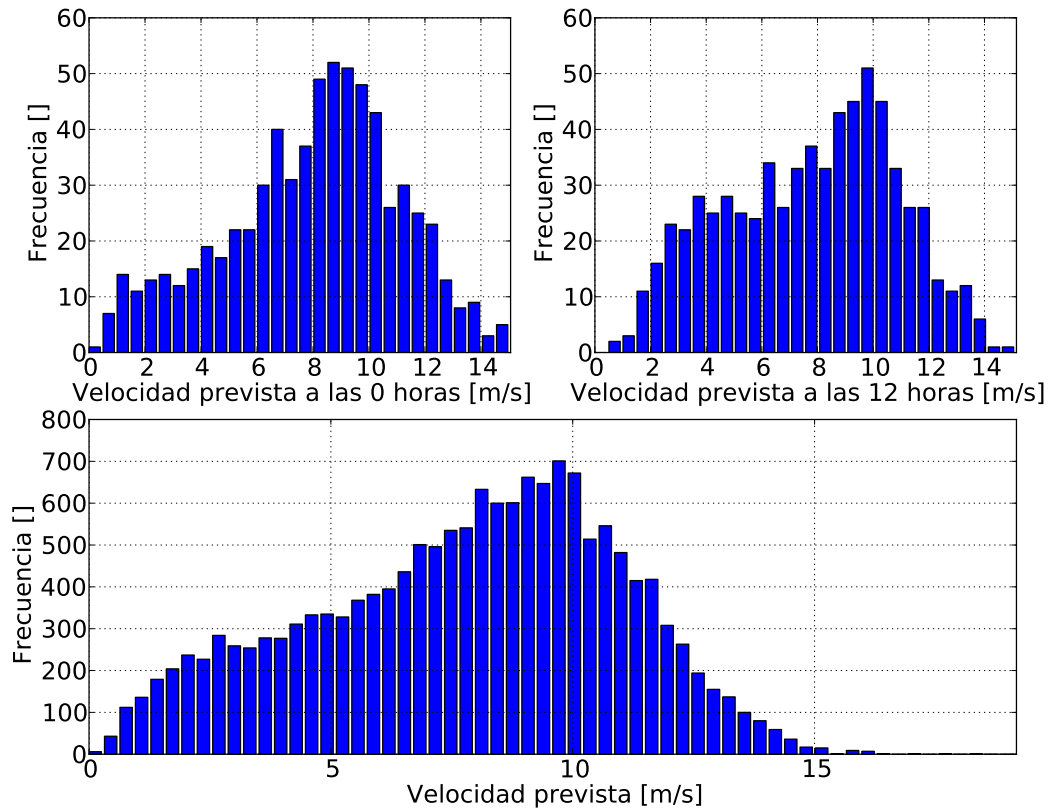


Figura 4.6: Distribución de la predicción de velocidad del viento. Arriba izda., velocidad a medianoche; arriba dcha., velocidad a mediodía. Abajo, velocidad tomando todos los datos

predicciones entre +25 y +48 horas. La predicción es capaz de coincidir con la parte principal del comportamiento, y parece reflejar vagamente la bimodalidad.

También se muestra la distribución de la variable “predictibilidad” (figura 4.7). Recuérdese que en este caso se tiene un único valor para cada día, repetido las 24 horas de ese día. En la distribución se aprecia una zona de crecimiento leve hasta que en 0.9 la frecuencia sube bruscamente, bajando de nuevo rápidamente al alcanzar 1. Esta información se ha utilizado para diseñar el clasificador.

El RMSE horario de las distintas predicciones puntuales se muestra en la figura 4.8, para los horizontes horarios entre +25 y +48 horas. Se observa que el método con menor RMSE es el que toma las predicciones originales y les aplica corrección de sesgo y luego un promedio ponderado de predicciones puntuales.

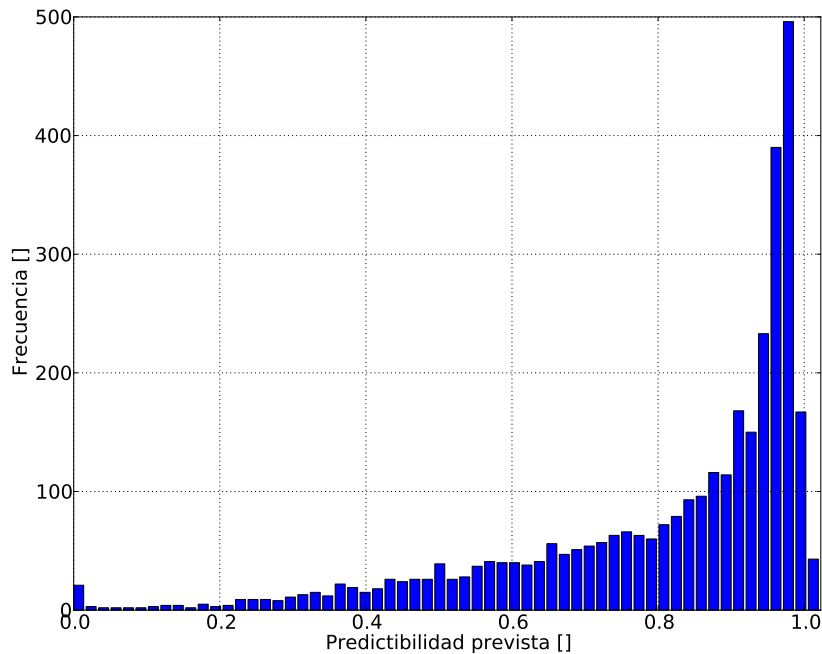


Figura 4.7: Distribución de la predicción de predictibilidad

4.5. Evaluación del método probabilista

El cálculo de predicciones probabilistas fue aplicado a los distintos conjuntos de predicciones puntuales (la original con o sin preproceso). El resultado es una matriz bidimensional con la misma estructura que las predicciones puntuales, pero donde cada celda tiene 99 valores, que son los percentiles 1 % a 99 % de la variable de interés, la velocidad del viento.

Las distribuciones de error por clase (resultantes de aplicar el clasificador difuso al conjunto de entrenamiento) tienen un aspecto como el de la figura 4.9. En esta misma figura se pueden apreciar las distribuciones mezcla (resultado de aplicar el “remuestreo adaptado”). Se aprecia que aunque la mezcla se realiza correctamente las distribuciones de partida son muy parecidas, lo que apunta al diseño del clasificador que debería ser capaz de obtener distribuciones base al menos con distinta varianza.

En las figuras 4.10, 4.11 se representa el resultado de la predicción probabilista. En tonos de gris se marcan los intervalos de predicción en los que se moverá la variable velocidad del viento. El tono de gris más claro marca el intervalo 90 %, mientras que el tono más oscuro define el intervalo 10 %. Entre estos intervalos y con tonalidades de gris cada vez más claro, están los intervalos 20 %,

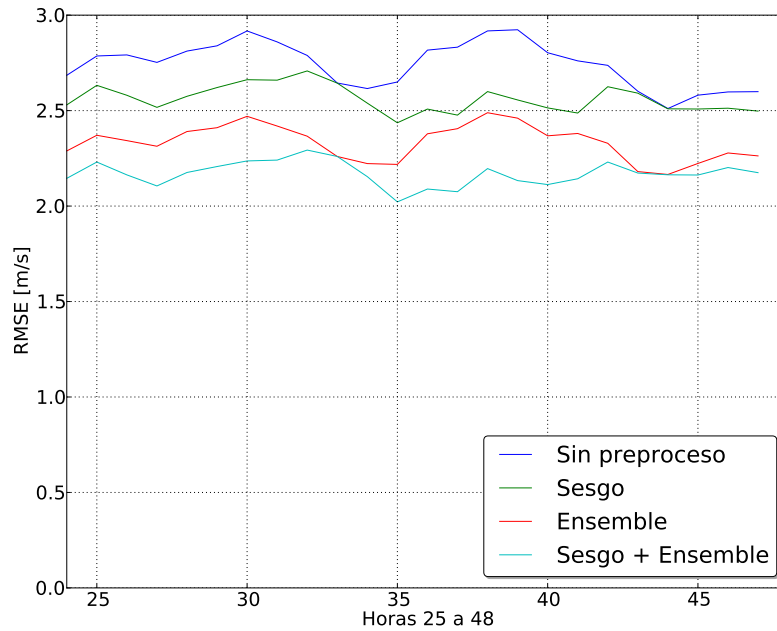


Figura 4.8: RMSE horaria (+25 a +48 horas) de las diversas predicciones puntuales (la original más tres correcciones -sesgo, ensemble, sesgo+ensemble-)

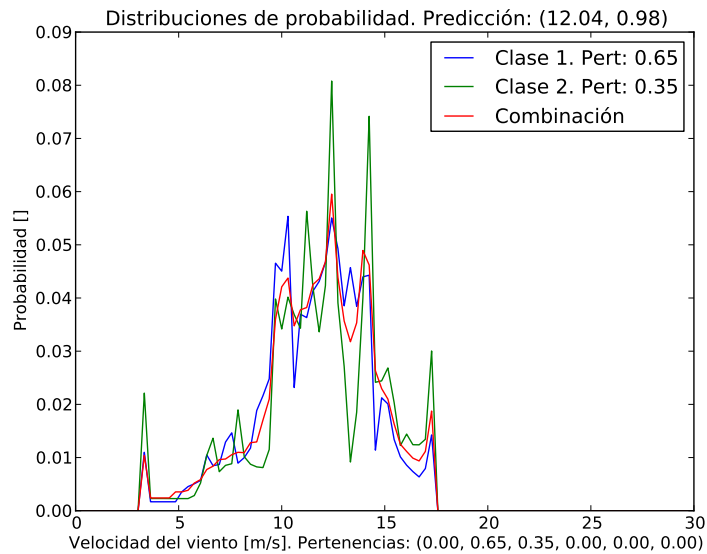


Figura 4.9: Distribuciones de probabilidad de las dos clases difusas activadas al tener una predicción puntual para cierta hora de 12.04 m/s (viento) y 0.98 (predictibilidad). Se muestra también la distribución combinada, resultante tras aplicar remuestreo adaptado (en rojo). Clase 1 es “predecible y velocidad de transición”, Clase 2 es “predecible y velocidad meseta”

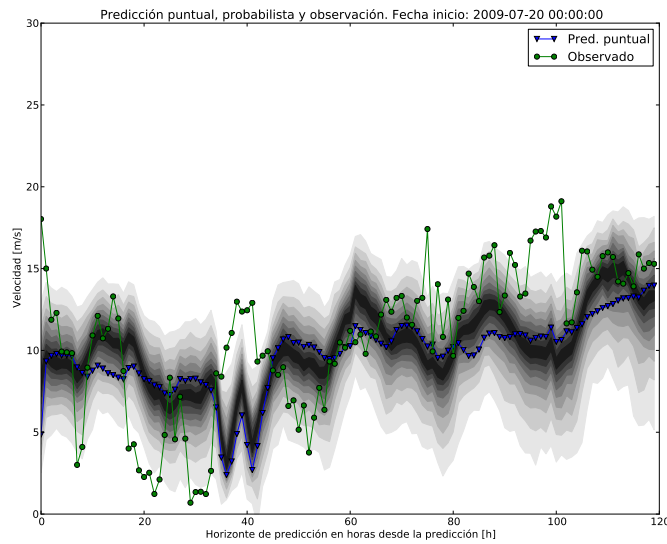


Figura 4.10: Predicción probabilista (desde intervalo 90% -gris claro- hasta intervalo 10% -gris oscuro-), predicción puntual (azul), observación (verde)

30%, ...

En [15], [6], [11], [10], [3] se encuentran estadísticos que permiten validar predicciones probabilistas. Un diagrama de fiabilidad permite ver cuán calibrado está cada uno de los percentiles. Se calcula contabilizando las veces en que se cumple el percentil, es decir, que $F(v^\alpha) \approx \alpha$ donde v^α es el percentil de nivel α y $F()$ es la distribución de probabilidad acumulada. A este valor se le resta el nivel α antes de representarlo. Si el offset es 0, la calibración es perfecta. El diagrama de fiabilidad del sistema de predicción utilizando remuestreo adaptado se observa en la figura 4.12. Se aprecia que los percentiles tienen en general un descalibramiento negativo.

A continuación, se presentan resultados con estadísticos que tienen en cuenta varias cualidades aparte de la fiabilidad. Estas son nitidez -o “sharpness”- (directamente relacionada con el ancho de la distribución resultante) y resolución (capacidad de aprovechar la información disponible a priori para producir resultados distintos).

En primer lugar se aplicará un factor de rendimiento utilizado en la publicación [11]. Está pensado para ser utilizado como “skill score”, término que se puede traducir como factor de habilidad pero que siempre es relativo a un modelo de referencia. La referencia suele ser un modelo trivial, que en este contexto puede ser bien la climatología bien la persistencia. Se ha utilizado la climatología: previamente se ha calculado la distribución de viento por hora en un conjunto de datos y

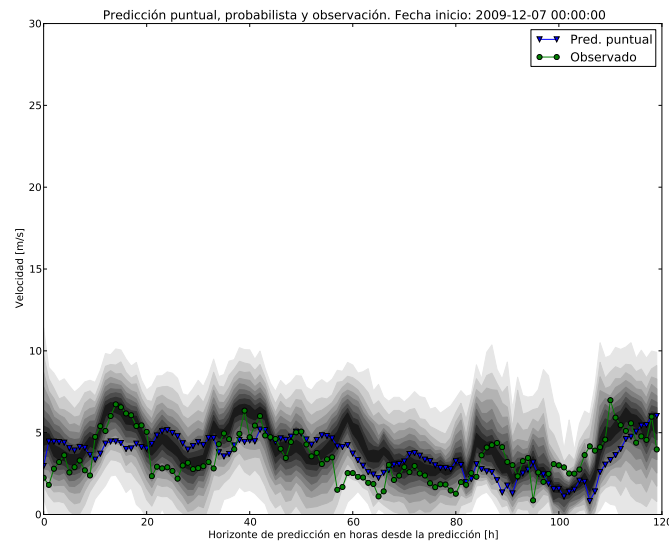


Figura 4.11: Predicción probabilista (desde intervalo 90 % -gris claro- hasta intervalo 10 % -gris oscuro-), predicción puntual (azul), observación (verde)

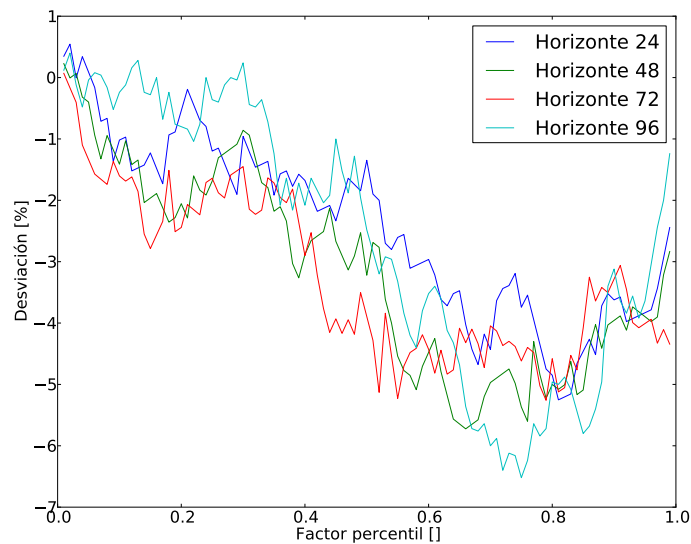


Figura 4.12: Diagrama de fiabilidad, a distintos horizontes horarios

esa ha sido la predicción para cada hora del conjunto de prueba. El factor se calcula de la siguiente manera:

$$Sc(\hat{f}_{t+k}, y_{t+k}) = \sum_{i=1}^m (\xi^{\alpha_i} - \alpha_i)(y_{t+k} - \hat{q}_{t+k}^{\alpha_i}) \quad (4.1)$$

donde ξ^{α_i} es 1 si $\hat{q}_{t+k}^{\alpha_i} > y_{t+k}$ y 0 en caso contrario, m es el número de cuantiles que se proporcionan y $\hat{q}_{t+k}^{\alpha_i}$ el cuantil de nivel α_i . \hat{f}_{t+k} es la predicción probabilista de la que se extraen los cuantiles. Se obtiene un único valor Sc_k promediando todos los Sc de la hora k :

$$Sc_k = \frac{1}{N} \sum_{t=1}^N Sc(\hat{f}_{t+k}, y_{t+k}) \quad (4.2)$$

El “skill score” para la hora k es:

$$SSc_k = 100 \frac{Sc_k^0 - Sc_k}{Sc_k^0} \quad (4.3)$$

donde Sc_k^0 es el factor obtenido para la climatología.

El resultado comparativo se muestra en la figura 4.13. El modelo propuesto mejora claramente a la climatología. Para poder valorar la parte de mérito que corresponde al método y la parte que corresponde al modelo de predicción puntual, se representa también el skill score de la predicción puntual (se ha convertido esta a una predicción probabilista donde todos los percentiles están en el valor predicho).

Como segunda medida se utilizará “CRPS” (Continuous Ranked Probability Score) (consultar [4]). La idea de esta medida es sumar (integrar) los cuadrados de las diferencias entre las distribuciones acumuladas de predicción y observación, como medida de distancia. Puesto que la observación es puntual, su distribución acumulada es una función escalón (“Heaviside”). Cuanto más cercano a cero se encuentre el valor, será mejor. Para una predicción probabilista y su observación concurrente, el CRPS se calcula:

$$crps = \int_{-\infty}^{\infty} [F(v) - H(v - v_{obs})]^2 dv \quad (4.4)$$

siendo $F()$ la distribución acumulada -predicha- y v_{obs} la observación. $H(x)$ es la función de Heaviside, que es 0 si $x < 0$ y 1 en caso contrario (función escalón). Se obtiene uno de estos valores para cada hora de cada día.

En la gráfica 4.14 se muestra el valor de promedio de CRPS, por horizonte horario. Se muestran



Figura 4.13: Factor de habilidad. Valores superiores a 0 mejoran al modelo de referencia basado en la climatología.

tanto el resultante de aplicar el método como el que resulta de aplicar un modelo de predicción basado en la climatología. También se muestra la predicción puntual (CRPS en este caso se reduce al MSE -Mean Squared Error-, se obtiene al sustituir $F(v)$ por $H(v - v_{pred})$). Por un lado se aprecia que el método mejora tanto la climatología como la predicción puntual (siendo esta última peor con diferencia, incluso respecto a la climatología) y por otro que el modelo tiende a empeorar ligeramente con el horizonte temporal, aunque no lo hace de manera monótona, apreciándose ciclos diarios.

Finalmente, se incluye una medición del éxito del método a la hora de acotar el error. Lo que se hace es contabilizar el número de ocasiones en que la observación está dentro del rango intercuartil (entre los percentiles 25 y 75). Se ha hecho primero un promediado por horizonte horario (figura 4.15 y luego un histograma con el número de éxitos por día (figura 4.16). En la evolución horaria no se aprecia una regularidad imputable claramente a algún fenómeno. En cuanto al histograma si se aprecia una cierta forma de distribución normal centrada en valores del entorno de 60. Esto indica que habitualmente se consigue acotar 60 o más de las predicciones.

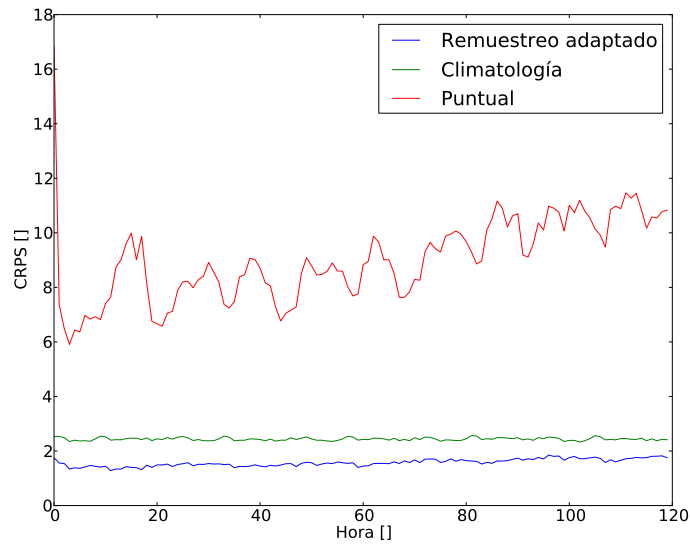


Figura 4.14: CRPS horario de la predicción

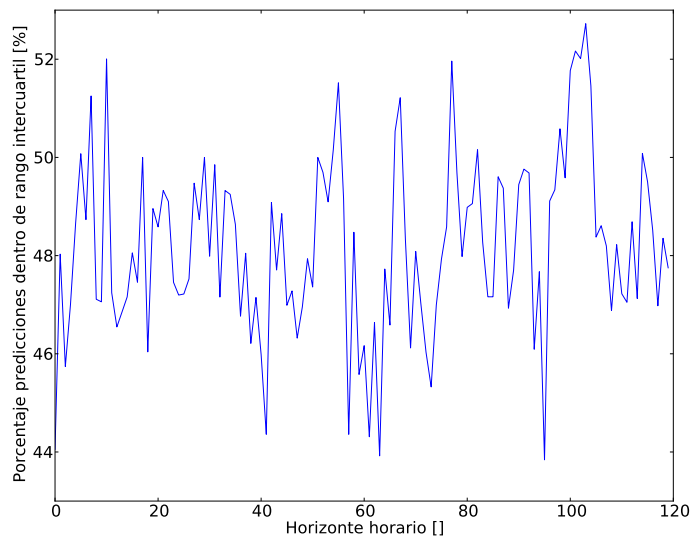


Figura 4.15: Porcentaje de ocasiones en que la predicción está dentro del rango intercuartil

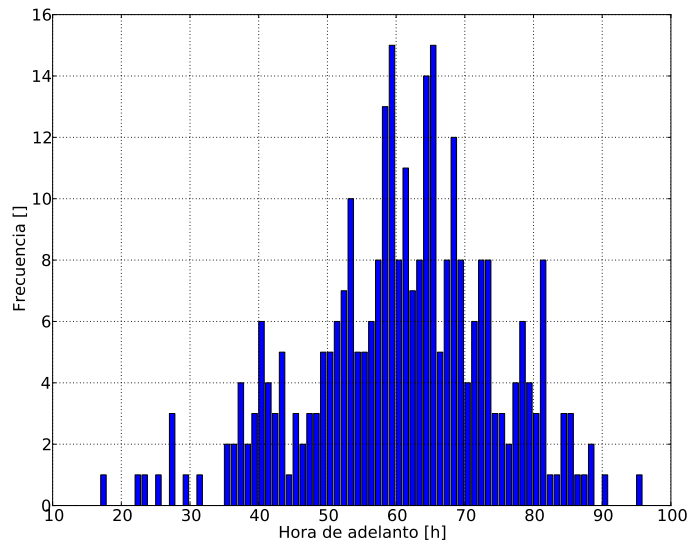


Figura 4.16: Histograma de número de ocasiones en que cada predicción de un día está dentro del rango intercuartil

Capítulo 5

Conclusiones, aportaciones y trabajo futuro

5.1. Conclusiones

Se ha hecho una incursión en los métodos probabilistas de predicción para meteorología energética como herramienta necesaria para conseguir aumentar la penetración renovable. La disponibilidad de estos métodos posibilita la aplicación de métodos estocásticos de operación, que completan la cadena de explotación de información para una gestión de la red óptima. Y disminuir en lo posible el uso de combustibles fósiles gracias también a los mecanismos que empiezan a aparecer con las redes inteligentes, que permiten almacenar energía y gestionar la demanda, y con la futura aparición de procesos industriales (como desalación) y de transporte adaptados a la variabilidad.

Se ha estudiado con detalle el método denominado “remuestreo adaptado”, previa revisión de distintos métodos de igual naturaleza sobre los que mejora a priori en rendimiento (según [11]), y a los que aventaja en determinados aspectos (mencionados en 2.2).

Durante la experimentación se ha visto como enlazar métodos de predicción puntual con el método probabilista, se han probado distintos métodos de proceso previos a la probabilización y se ha desarrollado un clasificador difuso adaptado a la información disponible, con reglas según la experiencia.

Tras los ensayos se observó que la clave para conseguir mejores rendimientos está en identificar situaciones con una distribución lo más estrecha posible y conseguir las empaquetar en clases. En este sentido, al elaborar las muestras de error, inicialmente se incluían todas las muestras con pertenencia difusa mayor que cero. Pero se observó que incluyendo solo aquellas con pertenencia

mayor que 0.9 se conseguía afinar la distribución. Aún así con el clasificador ensayado no se detectó una diferenciación apreciable en los rangos intercuartiles de las distribuciones base de cada clase.

Como ventajas del método en si (sin considerar el rendimiento obtenido), se han añadido a las identificadas a priori la claridad en cada uno de los pasos, el hecho de que el entrenamiento se hace on-line, es decir, como parte del mismo proceso de predicción probabilista, y que no se ha requerido de módulos de cálculo de gran complejidad, siendo el tiempo para la obtención de las predicciones muy constante.

Se comprueba el potencial que tiene el método al poderse aplicar multitud de técnicas del ámbito del machine learning tanto en la etapa de selección de variables explicativas del clasificador difuso como en la propia elaboración de este. También la etapa de elaboración del conjunto de entrenamiento implica una primera labor de clustering. Por ejemplo, estadísticamente el verano es más predecible que el invierno. Con un histórico suficiente, se puede hacer que si se está prediciendo algún día de verano se entrene usando solo predicciones pasadas realizadas para días de verano.

En uno de los ensayos se experimentó el uso de la mediana como un método de preproceso, igual que se realizó la corrección de sesgo y el de ensemble. Sin embargo, se comprobó que la mejora introducida tras aplicar alguno de estos métodos de preproceso no es significativa. El método permite mejorar la predicción puntual sin más, pero si la predicción puntual ha sido preprocesada la mejora es insignificante, incluso empeorada ligeramente algunas horas, como se puede apreciar en la figura 5.1.

5.2. Principales aportaciones

La principal dificultad, y de alguna manera se considera aportación, ha sido experimentar en un área de aplicación (predicción probabilista para meteorología energética) donde no había experiencia previa local, partiendo eso si de la experiencia en dos áreas afines al método abordado: aprendizaje máquina y predicción puntual para meteorología energética. Trabajar con el formalismo probabilista acarrió una dificultad inicial que se espera traiga resultados beneficiosos en futuros ensayos.

Se ha estructurado la implementación desde un punto de vista de ingeniería del software, descomponiendo los pasos en distintas clases e interfaces. Al hacerlo se ha visto que la estructura permite introducir y experimentar distintos clasificadores, y distintos métodos para selección del conjunto de entrenamiento. Por este motivo, el desarrollo del clasificador puede verse apoyado por

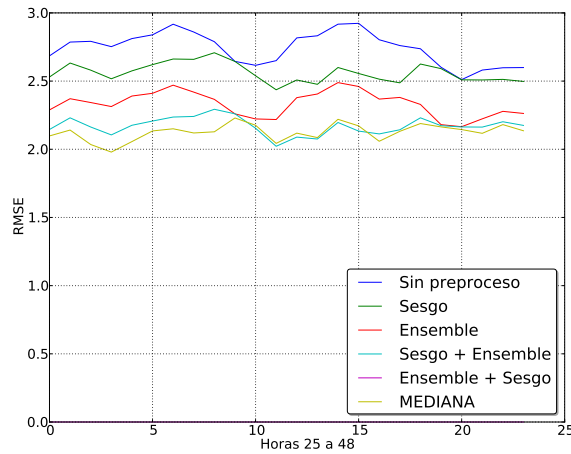


Figura 5.1: RMSE de los distintos preprocesos y mediana -percentil 50%- tras haber utilizado como entrada la predicción con el sesgo corregido y combinada usando un ensemble

técnicas de selección de variables, de clustering, de clasificación, etc. La estructuración permite también disponer fácilmente de un modo “pruebas”, como el empleado en este trabajo -y que facilitaría la experimentación-, y de un modo “operación”, que se aplicaría en un hipotético caso de uso real.

Se han implementado medidas para validación de predicciones probabilistas aplicables a métodos de este tipo, que tienen en cuenta distintos aspectos: fiabilidad, nitidez y resolución (diagrama de fiabilidad, CRPS y el skill score propuesto en [11], además de otras medidas como el número de veces que se consigue que la observación esté dentro del rango intercuartil).

Se han ensayado distintas variantes de corrección de las predicciones puntuales de entrada, como un posible paso previo que permite mejorar el resultado a pesar de no ser probabilista.

Se ha diseñado un clasificador difuso adaptado a los datos disponibles, utilizando conocimiento de las características climáticas de la zona, de energía eólica, del funcionamiento del modelo de predicción puntual utilizado y de la aplicación de un modelo de radiación con cielo despejado.

Se realizó una implementación utilizando software abierto, funcionando de modo eficiente.

La implementación ha quedado preparada para el ensayo de métodos de combinación de predictores probabilistas y para experimentar con toma de decisiones estocástica.

5.3. Trabajo futuro

Las etapas iniciales del método permiten la aplicación de técnicas de aprendizaje máquina.

En primer lugar, la selección de variables automática se puede aplicar para definir las entradas para el clasificador difuso. Los modelos físicos producen cantidad de variables que son susceptibles de ser aprovechadas en este proceso.

La selección del conjunto de entrenamiento, previa a la clasificación difusa, se encargaría de seleccionar muestras de situaciones atmosféricas similares. Desde variaciones estacionales hasta situaciones atmosféricas distintas. Se puede tratar como un primer clasificador.

Una vez se dispone del conjunto de entrenamiento, se trata de dividirlo según los rangos de las variables predichas. Esta etapa que ya pertenecería a la clasificación difusa, obtendría una clasificación diferente según el tipo de situación de alto nivel (etapa anterior) en que se encuentre la atmósfera. Para conseguir esto se podrían aplicar técnicas de agrupamiento automático, definiendo criterios que premien la afinidad y un rango intercuartil lo más pequeño posible, garantizando un mínimo de muestras.

Para operación, donde puede ser necesario hacer predicciones de varios sistemas, se puede aprovechar la correlación existente entre puntos meteorológicamente cercanos (con características similares).

Como ya se ha comentado es posible aplicar el método a otras variables y aprovecharlo para predecir radiación y predicciones de potencia tanto eólica como fotovoltaica, las principales renovables en Canarias.

Las distribuciones de probabilidad empíricas resultantes tienen una forma con picos. Partiendo de la hipótesis de que las distribuciones reales aún siendo no paramétricas puedan ser de formas suaves, se podría ensayar pasar la distribución por un filtro pasa bajos y analizar el efecto que esto tendría sobre el rendimiento.

Con el método elaborado se puede construir un sistema que encaje varias predicciones meteorológicas o meteoenergéticas y las transforme en probabilistas, y que sirva a su vez de entrada a la tercera etapa de predicción, el tomador de decisiones (sobre qué predicción tomar -o combinación de estas-) que devuelve también una distribución de probabilidad, pero basándose en el rendimiento pasado de distintos previsores probabilistas.

Actualmente MM5 ya no se está utilizando, se empieza a utilizar WRF y se pretende ensayar con predicciones del ECMWF. Se debería ensayar el método con estas predicciones, porque tienen mejores rendimientos.

Aprovechando técnicas de aprendizaje máquina y de optimización global, se podría plantear el desarrollo de un método sin parámetros de manera casi automática. Para cada sitio se puede automatizar la etapa de selección de variables, la selección de conjuntos de entrenamiento y la clasificación difusa, que sería optimizada y contrastada con un conjunto de prueba.

Finalmente, otra vía es experimentar métodos alternativos para predicción probabilista como los métodos referidos en [7] y [8] lo que permitirá comparar rendimientos.

Bibliografía

- [1] A. Arribas, K. B. Robertson, and K. R. Mylne. Test of a poor man’s ensemble prediction system for short-range probability forecasting. *Monthly Weather Review*, 133, July 2005.
- [2] G. Giebel, L. Landberg, G. Kariniotakis, and R. Brownsword. State-of-the-art on methods and software tools for short-term prediction of wind energy production. In *European Wind Energy Conference & Exhibition EWEC 2003*, 2003.
- [3] Robin Girard, Pierre Pinson, Jeremie Juban, and Georges Kariniotakis. Towards the definition of a standardized evaluation protocol of probabilistic wind power forecasts. Technical report, Anemos Plus - Advanced Tools for the Management of Electricity Grids with Large-Scale Wind Generation - 6PM European Commission - Priority 6.1: Sustainable Energy Systems, January 2009.
- [4] Tilmann Gneiting, Adrian E. Raftery, Anton H. Westveld III, and Tom Goldman. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, 133:1098–1118, May 2005.
- [5] Simon Haykin. *Neural Networks and Learning Machines*. Pearson, third edition, 2009.
- [6] Ian T. Jolliffe and David B. Stephenson, editors. *Forecast Verification. A Practitioner’s Guide in Atmospheric Science*. Wiley, 2003.
- [7] Jérémie Juban, Lionel Fugon, and George Kariniotakis. Probabilistic short-term wind power forecasting based on kernel density estimators. In *Probabilistic wind power forecasting - European Wind Energy Conference*, May 2007.
- [8] Jérémie Juban, Lionel Fugon, and George Kariniotakis. Uncertainty estimation of wind power forecasts. In *European Wind Energy Conference & Exhibition EWEC 2008*, April 2008.

- [9] Henrik Madsen. *Time Series Analysis*. Chapman & Hall/CRC, 2008.
- [10] Pierre Pinson. *Estimation of the Uncertainty in Wind Power Forecasting*. PhD thesis, Ecole des Mines de Paris, March 2006.
- [11] Pierre Pinson and George Kariniotakis. Conditional prediction intervals of wind power generation. *IEEE Transactions on Power Systems*, 25(4):1845–1856, November 2010.
- [12] J.L. Rodríguez Amenedo, J.C. Burgos Díaz, and S. Arnalte Gómez. *Sistemas eólicos de producción de energía eléctrica*. Rueda, S.L., 2003.
- [13] Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels - Support Vector Machines - Regularization, Optimization and Beyond*. The MIT Press, 2002.
- [14] Marcel Suri and Jaroslav Hofierka. A new gis-based solar radiation model and its application to photovoltaic assessments. *Transactions in GIS*, 8(2):175–190, February 2004.
- [15] Daniel S. Wilks. *Statistical Methods in the Atmospheric Sciences*. Academic Press, second edition, 2006.