Universidad Complutense de Madrid

Facultad de Ciencias Físicas

Departamento de Astrofísica y Ciencias de la Atmósfera

Centennial Simulation Of Wind Power Output Through Soft Computing Algorithms

Simulación Secular de la Potencia Eólica Generada Mediante el Empleo de Algoritmos Soft-Computing

Tesis Doctoral

Nicolás Kirchner Bossi

Madrid, 2014

Universidad Complutense de Madrid

Departamento de Atrofísica y Ciencias de la Atmósfera

Tesis Doctoral

Nicolás Kirchner Bossi

Centennial Simulation Of Wind Power Output Through Soft Computing Algorithms

Simulación Secular de la Potencia Eólica Generada Mediante el Empleo de Algoritmos Soft-Computing

Dirigida por

Luis Prieto Godino

Ricardo García Herrera

Ricardo García Herrera, Profesor Catedrático del Departamento de Astrofísica y Ciencias de la Atmósfera de la Universidad Complutense de Madrid y Luis Prieto Godino, Doctor en Física por la Universidad Complutense de Madrid

CERTIFICAN:

Que la presente memoria 'Centennial Simulation Of Wind Power Output Through Soft Computing Algorithms' ha sido realizada bajo nuestra dirección en el Departamento de Astrofísica y Ciencias de la Atmósfera de la Facultad de Ciencias Físicas de la Universidad Complutense de Madrid por D. Nicolas Kirchner Bossi y que constituye su tesis para optar al grado de Doctor en Ciencias Físicas.

Y para que conste, firman la presente certificación en Madrid, a 31 de Enero de 2014.

Dr. Luis Prieto Godino Prof. Dr. Ricardo García Herrera

Agradecimientos

Me gustaría agradecer en estas líneas a las muchas personas que me han ayudado y animado durante estos años de doctorado.

En primer lugar, quiero dar las gracias a Emiliano, quien me dio la oportunidad de empezar este viaje y quien es el director honorífico de esta tesis. Quiero dar las gracias a mis directores Luis y Ricardo, por haberme brindado la oportunidad y los medios para llevar a buen puerto esta tesis. Complementariamente en el tiempo y en el método, de ellos he recibido permanentemente sus inestimables enseñanzas y apoyo durante todo este tiempo, sin importar el lugar o la hora cuando lo he necesitado. Gracias por su paciencia y comprensión infinitas, y su ayuda decisiva en cada paso.

Quiero expresar mi gratitud de manera especial a Ricardo Trigo, por haberme acogido durante mi estancia en Lisboa, durante la cual encontré su permanente disposición para atender mis dudas, así como por haberme ayudado en momentos importantes.

Quiero dar las gracias especialmente a Sancho, José Antonio y Leo, por haberme enseñado tanto en el campo de la computación, evitando la tentación de desistir en el intento, y haberme involucrado en valiosísimas discusiones científicas.

A David y Natalia, quienes siempre han estado ahí para escuchar mis inquietdues, ayudandome en muchos detalles que sin duda están reflejados en esta tesis.

A mis compañeros, con los que he compartido todo este tiempo en el laboratorio, primero arriba, luego abajo y luego otra vez arriba. Ya véis, esto de la ciencia tiene sus altibajos, como la vida misma. Espero que en estos años que vienen todos podamos encontrar de un modo u otro nuestra oportunidad para aplicar todo lo que hemos aprendido en nuestro grupo, STREAM. Especialmente, quiero dar las gracias a Miguel, quien siempre lo ha dado todo para que el laboratorio funcione, técnica y humanamente. A Gabi, Maddalen, Mar, Froila y Fernando, por su continua predisposición, sus certeras respuestas a cada una de mis dudas o planteamientos, y sobre todo por su simpatía y complicidad.

A Fidel, por sus consejos en serenas veladas en la facultad, y a Marisa y a toda esa gran familia que es Palma, por su amistad y haberme acogido en adopción, casi como un palmípedo más. Gracias por haberme regalado tan buenos ratos. Además de todo, gracias especiales a Etor, por estar siempre dispuesto a renegar de su tiempo y echar un cable incondicional.

En cuanto a la sección de Astrofísica, a Nicolás por su ayuda desinteresada con problemas estadísticos, y a Raffaella, Lucía, Nestor y Alex, por su compañía en los descansos, aprovechados para, seguro, hablar de otras cosas.

Pero esta tesis, como cualquier otro desafío en mi vida, no habría sido posible sin el afecto y apoyo de mi familia y amigos, unos tan cerca, otros desde el Sur del mundo. Gracias a mis padres, Beatriz y Fernando, por habermelo dado todo y haberme traído hasta aquí. A mi hermano Patricio, siempre ejemplo de pragmatismo y optimismo en mi vida. A Germán, por ayudarme en todo tipo de dudas científicas y darme consejo en los momentos más difíciles de esta tesis. A mis tíos y primos del otro lado del charco, por ser mis únicos oídos a las cuatro de la mañana, especialmente Juan Pablo, con quien tantos ratos he compartido en la distancia. Por último, a Roni, quien me ha estado llamando cada mes desde tan lejos para darme un gran abrazo y darme consejo y aliento en mi camino. A Nawfel, Raúl, Saúl, Iván, Marc y Toni, por estar simepre ahí cuando he necesitado encontrar respuesta a mis preocupaciones, algo de complicidad o refortalecer mi espíritu. Y finalmente a Elia, que con su alegría y su fortaleza tan gigantes por existir y amar esta vida, me ha llenado de energía y esperanza, demostrandome especialmente que, como decía Rousseau, el ser humano no puede ser malo por naturaleza (un caso contrario es suficiente para negar el teorema). Esta tesis es vuestra.

A mis padres

'*Of all the forces of nature, I should think the wind contains the largest amount of motive power (...) Take any given space of the earth's surface, for instance, Illinois, and all the power exerted by all the men, beasts, running water and steam over and upon it shall not equal the 100th part of what is exerted by the blowing of the wind over and upon the same place. And yet it has not, so far in the world's history, become properly valued as motive power. It is applied extensively and advantageously to sail vessels in navigation. Add to this a few windmills and pumps and you have about all. As yet the wind is an untamed, unharnessed force, and quite possibly one of the greatest discoveries hereafter to be made will be the taming and harnessing of it.*'

Abraham Lincoln,
Lecture "*Discoveries and Inventions*", 1860.

Contents

List of Acronyms and Symbols **XV**

Summary **XXI**

Resumen **XXV**

1 Introduction **1**

1.1 Wind Power and Wind Conditions 2

1.2 Scales of the Atmospheric Dynamics 3

1.2.1 Synoptic Circulation over the Iberian Peninsula 4

1.2.2 Wind Forcings at Local Scale 8

1.3 Wind Variability Scales and its Application to the Wind Power Industry . 10

1.4 The Long-Term Wind Variability 13

1.4.1 Observations 14

1.4.2 Numerical Models 15

1.4.3 Downscaling Techniques 16

1.5 The Statistical Downscaling of Wind 18

1.5.1 The design of an appropriate predictor for the local wind speed . . 18

1.5.1.1 The classification of the synoptic circulation as a wind predictor 19

1.5.1.2 Circulation patterns Clustering and Soft-Computing . . . 20

1.5.2 Sources of Uncertainty in a Statistical Downscaling 22

1.6 The Wind Power Long-Term Variability 23

1.6.1 Centennial Wind Power Inter-annual Series: Application Outlook . 25
1.7 Objectives 25

2 Clustering and Statistical Downscaling of Wind by Employing Different Soft-Computing Tools 27

2.1 Introduction 27
2.2 Geostrophic-Based Wind Type Classification for the Statistical Downscaling of Wind 28
2.2.1 SLP Clustering Optimization Problem and Wind Types Classification 29
2.2.2 Optimization of the Clustering Process through different Soft-Computing Algorithms 32
2.2.2.1 Geostrophic Clustering Optimization through an Evolutionary Algorithm, FE 33
2.2.2.2 Geostrophic Clustering Optimization Through a Greedy Algorithm (FG) 36
2.2.3 Statistical Downscaling through Wind Type Classifications as a Predictor 36
2.3 Statistical Downscaling through an Evolutionary Algorithm of Pressure Differences (PD) 38
2.3.1 Cost Function 39
2.3.2 Evolutionary Architecture of PD 39

3 Performance of the Downscaling methodology 45

3.1 Introduction 45
3.2 Input Data 46
3.2.1 Wind Observations 46
3.2.2 Reanalysis Data 48
3.3 Cross-Validation of the Developed Methods 49
3.3.1 Daily and Monthly Estimation Performance of the Introduced Methodology 51
3.3.2 Examples on Daily Wind Reconstruction 54
3.3.3 Homogeneity within the Intra-Class SLP Daily Fields 58

3.3.4 Comparison with Previous Work . . . 59
3.4 Obtained Wind Types . . . 60
3.4.1 FG and FE Wind Types: Geostrophic vs local Wind . . . 61
3.4.2 Synoptic Dynamics and Monthly Variability associated to the obtained Wind Types . . . 65

4 Long-Term Wind Variability **71**
4.1 Introduction . . . 71
4.2 Wind Speed Long-Term Variability . . . 72
4.2.1 Obtention of the Wind Speed Series . . . 72
4.2.2 Climatological Wind Speed Variability . . . 75
4.2.3 Multidecadal Wind Speed Variability . . . 76
4.3 Wind Speed PDF variability . . . 78
4.4 Wind Direction Variability . . . 80
4.5 Wind Type frequency variability associated to the reconstructed wind speed 82

5 Long-term Wind Power Output Variability **91**
5.1 Introduction . . . 91
5.2 Remarks on the Procedure and the Data Employed . . . 92
5.3 Wind Power Output and Wind Speed . . . 93
5.4 Wind Power Centennial Series . . . 96
5.5 Long-Term Variability of the Simulated Output Power . . . 98
5.6 Impact of the Teleconnection Patterns on the Simulated Output Power . . 99
5.7 Incidence of the technical wind speed limits of a wind turbine on the Simulated Output Power . . . 104
5.8 Impact of the k and c Weibull parameters on the Simulated Output Power 107

6 Conclusions **113**

References **117**

List of Acronyms and Symbols

$\%Var_f$	percentage of explained variance of the annual variability of the Wind Types frequency
20CR	20th Century Reanalysis
$\alpha(z_0)$	exponential coefficient of the power law
$\alpha(F)$	direction of the geostrophic index
A	Anticyclonic Wind Type
$\mathcal{A}$	set of angular borders
a_i	angular border
AEG	Area del Estrecho de Gibraltar
AMO	Atlantic Multidecadal Oscillation
ANN	Artificial Neural Network
AT	Atlantic European Region
BMA	Bayesian Average Model
c	scale parameter of the Weibull function
C	Cyclonic Wind Type
c_κ	generic class
C_p	Coefficient of Power
CA	Cluster Analysis
CCA	Canonical Correlations Analysis
CDF	Cumulative Distribution Function
CF	Capacity Factor
CFD	Computational Fluid Dynamics
CIP	Central Iberian Plateau
CPC	Composed Principal Component
$\mathbf{d}_{c_\kappa}$	mean wind speed observations at a certain class c_κ
$\mathbf{d}_i$	wind speed observations at the day i
DJF	December, January and February

Dp	Differences of Pressure
E	East
E(v)	kinetic Energy
e_i	border before mutation
e'_i	border after mutation
EA	Eastern Atlantic pattern
EC	Evolutionary Computation
ENSO	El Niño Southern Oscillation
EOF	Empirical Ortogonal Function
Φ	cost function
φ	coordinate longitude
F	geostrophic index of the flow intensity
F_{c_κ}	average geostrophic index of the flow intensity within a class c_κ
F_i	geostrophic index of the flow intensity at a day i
f	Coriolis Parameter
F_W	Weibull Probability Density Function
FE	Geostrophic - Evolutionary Algorithm
FG	Geostrophic - Greedy Algorithm
FL	Fuzzy Logic
G	Geostrophic wind
$G_{u\lambda}$	zonal component of the Geostrophic wind at a certain latitude λ
$G_{v\lambda}$	meridional component of the Geostrophic wind at a certain latitude λ
GCM	Global Circulation Model
GR	Greedy Algorithm
GSA	Gibraltar Strait Area
HA	Hybrid-Anticyclonic Weather Type
HC	Hybrid-Cyclonic Weather Type
i_s	a generic day of the period τ_s
IFEP	Improved Fast Evolutionary Programming
IP	Iberian Peninsula
JJA	June, July and August
k	shape parameter of the Weibull function
L	randomly generated population
l	generic geographic location
λ	latitude
M1-M3	Wind Type radial magnitudes

MAE	Mean Absolute Error
$MAE_{\|V\|}$	Mean Absolute Error of the wind speed
MAE_V	Mean Absolute Error of the wind vector
MAM	May, April and March
MED	Mediterranean Sea Basin
MIC	Meseta Iberica Central
MOS	Model Output Statistics
NAO	North Atlantic Oscillation
NC	No-classification procedure
NCAR	National Center for Atmospheric Research
NCEP	National Center for Environmental Prediction
NE	North East
NW	North West
Ω	angular speed of the Earth rotation
P	pure directional Weather Type
p	statistic of significance
P_i	probability at the Cumulative Distribution Function of the Weibull distribution
$\mathcal{P}_i$	set of considered grid points
P_m	Probability of mutation
P_o	wind turbine annual power output
P_t	daily synoptic-scale pressure field
P1,P2	annual wind power model for T1,T3
PBL	Planetary Boundary Layer
PC	Principal Component
PCA	Principal Component Analysis
PCI	frequency of wind speeds higher than the $Cut\text{-}In$
PCO	frequency of wind speeds higher than the $Cut\text{-}Out$
PD	Pressure Differences - Evolutionary Algorithm
PDF	Probability Density Function
POT	Point Over Threshold
r	Pearson Correlation Coefficient
R	Earth radius
$\mathcal{R}$	set of radial borders
r^2	explained variance
r_i	radial border
RC	random classification

RCM	Regional Circulation Model
ρ	air density
RPCA	Rotated Principal Component Analysis
S	generic area
S_D	four-dimensional space of pressure differences
SC	Soft-Computing
SCAND	Scandinavia pattern
SD	Statistical Downscaling
SE	South East
SF	southern component of the flow intensity geostrophic index
σ_M	average standard deviation of the pressure fields within the classes obtained through the FG or FE methods
σ_{RC}	average standard deviation of the pressure fields within the classes obtained through a random classification
σ_{RR}	relative ratio between σ_M and σ_{RC}
${\sigma_i}^2$	sample variance
$\bar{\sigma}_y^2$	population variance
SLP	Sea Level Pressure
sNAO	summer NAO
SON	September, October and November
SST	Sea Surface Temperature
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SW	South West
SZ	southern component of the absolute vorticity geostrophic index
T	period
T_{obs}	observational period
T_{rec}	reconstruction period
T1-T6	meteorological towers
τ_r	training period
τ_s	test period
$\mathrm{T_p}$	Towers representative of the Iberian Plateau
u	zonal component of the wind speed
V	wind speed
v	meridional component of the wind speed
v_a	wind speed at height z_a

V_{c_κ}	average wind speed representative for a certain class c_κ
V_{i_s}	wind speed at a certain day of the test period
v_i	interval wind speed at the Weibull Function
W	West
WdT	Wind Type
WF	western component of the flow intensity geostrophic index
W-NCAR	10 m wind from the NCAR reanalysis
WT	Weather Type
WZ	western component of the absolute vorticity geostrophic index
Z	absolute vorticity geostrophic index
z_0	roughness length
z_a	reference height at the power model

Summary

The wind power output variability comprises different time scales. While there exists a wide knowledge on either the short-term variability or the climatological conditions of wind and its corresponding wind power, the study of the decadal and multidecadal variations is still in its early stages. This can be attributable to different reasons, as the relatively recent importance of wind power industry, the relevance of shorter time scales or the difficulties in obtaining representative long time series of wind. However, interannual to multidecadal variability timescales can substantially modify the performance of a wind power plant during its lifetime.

In this context, this work develops a methodology to obtain wind and wind power centennial series with daily resolution at a given point, thus contributing to shed some light on the interannual wind and wind power variability. To do this, large-scale reanalysis pressure data were employed as predictor of wind observations (predictand) from different wind farms over the Iberian Peninsula (IP). By employing these data series, different tools for the statistical downscaling of wind are designed, by employing distinct Soft-Computing techniques as Greedy and Evolutionary Algorithms. Apart from downscaling daily wind conditions, they allow clustering local wind regimes, which is useful to better understand the circulation mechanisms governing surface wind variability.

When cross-validating the method, the three proposed approaches perform better in terms of daily wind speed uncertainty reduction when compared to other approaches, such as the Circulation Weather Types (WT, Jenkinson and Collison (1977)), or the wind signal at the NCEP/NCAR (NCAR, Kalnay et al. (1996)) reanalysis. Correlation between estimates and observations are 10 and 30% higher than WT and NCAR respectively, while monthly results are even more competitive. Although comparison with other works is not straightforward, this methodology shows better performance than studies performed over sea.

With the new methods, 139-yr reconstructions of daily wind conditions were obtained

over different locations on the IP. Results on its variability analysis show a long-term negative trend of wind speed over the Central Iberian Plateau (CIP) in the second half of the 20th century, and a positive trend at the Gibraltar Strait Area (GSA). One of the most outstanding results on wind speed variability showed for the first time different statistically significant ($p<0.05$) periodicities around the 25-yr band. Decadal and annual wind speed distributions and wind roses were computed, showing statistically significant variability in both wind speed or wind direction. Finally, an *a posteriori* Principal Components analysis on the annual clusters frequency allowed to understand which meteorological regimes most contribute to the annual wind speed variability.

A wind power simulation from a market wind turbine (Vestas V-82 1.65 Mw, 50 Hz) was carried out from the reconstructed wind series. This time the 56 ensemble members of the 20thCR (Compo et al., 2011) reanalysis was employed as wind predictor. The computed annual output power shows a very different long-term variability depending on the location. In this way, a significant long-term power output decrease (increase) was observed in CIP (GSA). These trends became particularly significant during the second part of the 20th century which, if sustained in time, must be considered in a wind farm layout or in further repowerings. On a seasonal basis, although DJF is typically the season with the highest contribution to power output (and so occurs at CIP), GSA does not show a significant seasonal cycle.

The availability of 139 years of output power allowed the study of the long-term relationships existing between the variability of the annual power output and different factors related with it. In first place, it was observed that large-scale circulation patterns exert a remarkable impact on the wind power interannual variability. The power output variance explained by them rises up to 70% and 45% in SON for CIP and GSA respectively. Although the coupling is also important in JJA or DJF at CIP, this relation has been proven as non-stationary. There, wind power showed statistically significant negative correlations with NAO at all seasons, although no relationships were detected at GSA. The opposite behavior of SCAND (positive at CIP, negative at GSA) is consistent with its spatial structure over Iberia, where an anticyclonic (cyclonic) system in its negative (positive) mode is located over Spain, explaining thus its connection to the prevailing easterly winds at Southern Spain.

Secondly, the reconstruction of the wind speed frequencies beyond the lower (*Cut-In*) and upper (*Cut-Out*) wind turbine technical limits showed a high consistency with the power output results. Results on the relationship between these parameters and the

output power show an overall higher influence of Cut-In than Cut-Out, for any considered location.

Finally, the possibility to obtain a wind speed PDF for 139 years allowed the assessment of the empirical relationship between the interannual variability of the annual wind power output and that of the shape (k) and scale (c) Weibull parameters, once the PDFs where adjusted to that function. The long-term variability analysis of these parameters provided trends and periodicities closely related to the wind power output, specially in the case of c. A multilinear regression between annual power and Weibull parameters was computed for each tower. The evidenced linearity of the obtained expression reflected the high dependence of power with these parameters (explained variances of 98%). This approach improved the estimation of the simulated output power as compared with that computed through the PDF fit to a Weibull distribution. This resulted in a model specially appropriate when only k and c data is available, without the need to await for a good fit of the PDFs to a Weibull distribution.

The introduced methodology has allowed a realistic approach to the long-term wind clustering and its subsequent statistical downscaling with daily resolution, by exclusively employing public domain SLP reanalysis data and with a very low computational cost. It has been shown that it can be implemented for the reconstruction of centennial wind series at the wind farm locations, where long-term wind measures are rarely available. This contribution on the wind and wind power variability represents a new outlook on their variational analysis, specially appropriated within the study of the multidecadal variability range of the wind power performance.

Resumen

Introducción

La variabilidad de la producción de energía eólica comprende diferentes escalas de tiempo. Si bien existe un amplio conocimiento ya sea sobre la variabilidad a corto plazo o las condiciones climatológicas de viento y su energía eólica correspondiente, el estudio de la variabilidad desde la escala interanual a la multidecenal se encuentra todavía en sus primeras etapas. Esto puede atribuirse a diversas razones, como la relativamente reciente importancia de la industria de la energía eólica, la relevancia de escalas de tiempo más cortas o las dificultades en la obtención de series temporales de viento suficeintemente representativas. Sin embargo, escalas de variabilidad desde la interanual a la multidecenal pueden modificar sustancialmente el rendimiento de un parque eólico durante su vida útil.

Objetivos

En este contexto, el objetivo general de esta tesis es analizar la variabilidad a largo plazo de la energía eólica en determinados puntos de especial relevancia energética, mediante la reconstrucción de sus condiciones de viento a lo largo de más de cien años.

En el proceso de consecución de este propósito, se ha pretendido asimismo la obtención de un método objetivo para la clasificación de la circulación a escala sinóptica sobre el punto de interés y el *downscaling* estadístico de las condiciones locales de viento a escala diaria. Ello permitirá reconstruir las condiciones diarias del viento a lo largo de un período secular, lo que posibilitará el análisis de su variabilidad a largo plazo o escala multidecadal. Finalmente será posible, para el mismo período, la simulación de la potencia de salida anual generada por un modelo de mercado de aerogenerador, permitiendo a su vez el análisis pormenorizado de su variabilidad interanual y multidecadal.

Métodología

Para el diseño de un clasificador objetivo de la circulación sinóptica y su consiguiente empleo como predictor en un downscaling estadístico se emplearon datos de presión de reanálisis a gran escala como predictor, y observaciones del viento de diferentes parques eólicos sobre la Península Ibérica (IP) como predictando. De esta forma se llevaron a cabo tres aproximaciones diferentes mediante el empleo de distintas técnicas de Soft Computing, como algoritmos voraces (*greedy*) y evolutivos.

Dichos métodos permiten la agrupación del vientos local en una serie de clases o *tipos de viento*, resultando así útiles para comprender mejor los mecanismos dinámicos que rigen la variabilidad del viento superficial. La obtención de una clasificación diaria acorde a estas clases permite la esitmación de las condiciones de viento en aquellos días o períodos sobre los que sólo se disponga de información sobre la circulación a escala sinóptica.

Aportaciones Fundamentales

La validación cruzada del método revela que las tres desarrolladas reducen la incertidumbre de la velocidad del viento diaria en comparación con otros métodos, como la circulación de los tipos de tiempo (WT, Jenkinson and Collison (1977)), o la señal de viento en el reanálisis NCEP / NCAR (NCAR, Kalnay et al. (1996)). La correlación entre las estimaciones y observaciones son 10 y 30% mayores que WT y NCAR respectivamente, mientras que los resultados mensuales son aún más competitivos. Aunque la comparación con resultados de otros trabajos no es sencilla, esta metodología muestra, por ejemplo, un mejor rendimiento que algunos estudios realizados sobre océano.

Los nuevos métodos se emplearon para la reconstrucción de 139 años de las condiciones de viento diario en diferentes lugares de la IP. Los resultados en el análisis de la variabilidad a largo plazo muestran sobre la Meseta Ibérica Central (MIC) una tendencia de la velocidad del viento negativa en la segunda mitad del siglo 20, así como una tendencia positiva en la Area del Estrecho de Gibraltar (AEG). Uno de los resultados más sobresalientes revela periodicidades estadísticamente significativas ($p<0.05$) alrededor de la banda de variabilidad de los 25 años, un resultado que no se había observado previamente y que ha sido constatado recientemente mediante series independientes a las aquí empleadas (Barriopedro et al., 2013). Se calcularon distribuciones de velocidad y rosas de viento anuales y decadales, mostrando diferencias estadísticamente significativas tanto entre diferentes décadas como en años dentro de una misma década. Por último, un Análisis por Componentes Principales de la frecuencia anual de las clases obtenidas permitió

entender qué tipos de viento contribuyen en mayor medida a la variabilidad anual de la velocidad del viento.

A partir de las series reconstruidas de viento se llevó a cabo la simulación de la energía eólica producida por un modelo de aerogenerador específico (Vestas V-82 1,65 Mw, 50 Hz) para los lugares considerados. Para ello se utilizaron los 56 miembros del reanálisis del siglo XX (Compo et al., 2011) como predictores. La potencia de salida anual muestra una variabilidad a largo plazo muy diferente dependiendo de la ubicación. De esta manera, se observó una tendencia negativa (positiva) en MIC (AEG). Estas tendencias resultaron particularmente significativas durante la segunda parte del siglo XX, lo cual, de mantenerse en el tiempo podrían implicar, especialmente en AEG, la necesidad de considerarlas en el diseño del parque eólico o en próximos repowerings. Atendiendo a la variabilidad interestacional, aunque DJF suele ser la estación de mayor contribución a la producción de energía (y así sucede en MIC), AEG no muestra un ciclo estacional significativo.

La disponibilidad de 139 años de potencia de salida permitió el estudio sobre la relación a largo plazo existente entre la variabilidad de la potencia anual generada y diferentes factores relacionados con ella. En primer lugar, se observó que los patrones de circulación a gran escala ejercen un impacto notable en la variabilidad interanual de la energía eólica. Así, durante SON estos patrones explican un 70% y un 45% de la varianza de la potencia en MIC y AEG respectivamente. Aunque en MIC el acoplamiento también es importante en JJA o DJF, esta relación se ha manifestado no estacionaria. Allí, la energía eólica mostró correlaciones negativas estadísticamente significativas con la NAO en todas las estaciones, aunque no se detectaron relaciones en AEG. El comportamiento opuesto del patrón SCAND (positivo en MIC, negativo en AEG) es coherente con su estructura espacial sobre Iberia, donde tiene lugar un sistema anticiclónico (ciclónico) en su modo negativo (positivo), lo que explica su conexión con los vientos predominantes del este en el sur de España.

En segundo lugar, la reconstrucción de la frecuencia de velocidad por debajo (*Cut-In*) y por encima (*Cut-Out*) de los límites técnicos de funcionamiento del aerogenerador mostró un alto nivel de coherencia con los resultados de la salida de potencia. Los resultados muestran una mayor influencia de estos parámetros sobre MIC, estando *Cut-In* más fuertemente relacionado con la señal de potencia de energía que el *Cut-Out* en cualquier emplazamiento.

Por último, la posibilidad de obtener una distribución de velocidad (PDF) para 139 años permitió la evaluación de la relación empírica entre la variabilidad interanual de

la producción de energía eólica y la de los parámetros de forma (k) y escala (c) de la distribución de Weibull. El análisis de la variabilidad a largo plazo de estos parámetros proporcionó tendencias y periodicidades estrechamente relacionados con las de producción eólica, especialmente en el caso de c. A partir de la regresión multilineal entre la potencia anual y dichos parámteros se evidencia la alta dependencia de la primera con los últimos (con varianzas explicadas del 98% no solapadas). Este modelo mejora la estimación de la potencia de salida simulada en comparación con otras formas de calcular la potencia, como el ajuste de la PDF a una distribución de Weibull. Este hecho da como resultado un modelo especialmente apropiado para situaciones en las que sólo se dispone de datos de k y c, sin necesidad de esperar a que exista un buen ajuste de la PDF anual a una distribución de Weibull.

Conclusiones

La metodología introducida permite la clasificación de las condiciones diarias de la circulación sinóptica, pudiendo esta ser empleada en el *downscaling* del viento a resolución diaria, para ello empleando datos de presión de dominio público mediante técnicas Soft-Computing. Los métodos obtenidos se muestran capaces de ofrecer un gran rendimiento en la estimación de las condiciones de viento local a resolución diaria, cuando se lo compara con el obtenido por dos aproximaciones de referencia como el viento del reanálisis NCEP/NCAR v.2 o el metodo de los tipos de tiempo (WT).

La implementación de estos métodos en la reconstrucción de las condiciones de viento diarias a lo largo de 139 años así como la simulación de la potencia eólica generada por parte de un aerogenerador real revelaron por primera vez la existencia de periodicidades estadísticamente significativas en torno a la banda de variabilidad de 25 años para ambas regiones. Dichas series mostraron asimismo la existencia de un comportamiento contrario entre ambas zonas consideradas, con una tendencia positiva (negativa) en AEG (MIC) durante la segunda mitad del siglo XX. Se observó un fuerte acoplamiento de la potencia anual a los principales patrones teleconectivos de la dinámica a gran escala, especialmente en MIC mediante la NAO y el patrón escandinavo. Finalmente, se ha verificado la validez de un modelo empírico para la estimación de la potencia anual contando unicamente con los parámetros de la distribución de Weibull, sin necesidad de esperar a que dicha curva se ajuste correctamente a la distribución de velocidades real.

Chapter 1

Introduction

The kinetic energy of an air mass has been used by mankind since ancient times to satisfy different mechanical needs. It has been so since the first vestige of sailing, dated in 5000 BC in Mesopotamia. That technology derived into more sophisticated elements as rotating sails over an horizontal axis for water pumping (India, 4th century BC) or grain mills (Persia, 2nd century BC). Wind energy technological development remained latent until 1890, when in Denmark the first wind turbine was designed to generate electricity (Mathew, 2006). However, from the 1940's this technology has experienced a rapid and continuous evolution (Ackermann et al., 2005). Thus, from the mid 1990's the generation of electricity from wind power began to represent a real alternative as a large-scale energy source. As an example, in Spain wind energy supplied a monthly maximum of 26% of electricity share (April), overcoming 61% of the production at hourly scale (REE, 2012). In 2013, more than 21% of the electricity consumption was provided by wind power, becoming for the first time the largest year-round electricity source in the country (REE, 2013). These production levels place wind power as a real solution for the energetic problem, increasingly fenced by climate change and the exhaustion of fossil fuels.

The planning of a wind power project must take into account the different variability scales of the potential wind power output that real wind turbines would experiment at the selected site. As it will be shown along this introduction, while there exists a wide bibliography on the short-term variability (hours) or the climatological conditions, the study of the long-term wind power variability (at decadal and multidecadal scales) and

their governing mechanisms are still in an early phase of development. This is mostly attributable to the difficulties in obtaining wind power measurements in a given location. Thus, most of the decision making process is based on data sets covering a few years. This work develops a methodology to estimate wind power variability at multidecadal scales, thus contributing to overcome this type of problems.

1.1 Wind Power and Wind Conditions

Usually, when wind farms are planned no wind turbines have been previously installed in the scheduled location. This means that probably records of measured wind turbine output power are not available or, at best, they comprise only a few years of data. Thus, in order to analyze its long-term behavior a power-related variable must be considered. Variables like air density are related with the temporal rate of kinetic energy E extractable from a moving air mass of density ρ through a certain area S. However, its wind speed v is the primary meteorological variable involved in it (power, P), as it is shown through the definition of kinetic energy:

$$P = dE/dt = \frac{1}{2}\rho S v^3 \tag{1.1}$$

It must be pointed out that the real efficiency of modern wind power turbines make that in practice the value of usable wind power through a wind power device is better related to the square of wind speed (Anderson and Bose, 1983). In any case, since analyzing multidecadal variability of wind power observations results unfeasible in most cases, considering long-term wind variability at the location of wind power interest results an appropriated approach to infer wind power conditions. In this way, different procedures that simulate wind power by considering wind speed conditions have been implemented (Şahin and Aksakal, 1998; Celik, 2004; Sailor et al., 2000). Among them, Veigas and Iglesias (2012) or Earl et al. (2013) perform a long-term wind power variability approach by considering wind observations from weather stations.

Although long range time series of wind observations can span more than a hundred years back at certain meteorological stations, there exists a high probability that wind observations are not available exactly at the point where a wind turbine is to be placed. Since the high spatial variability of wind can make that wind conditions are extremely different within a distance of a few hundred of meters (Pérez-Landa et al., 2007; Mott and

Lehning, 2010; Etienne et al., 2010), the need to account for precise *in situ* information becomes critical. By assuming that surface wind is the outcome of the interaction between local and large-scale mechanisms, one solution to the lack of observations can be attained by approaching local wind conditions from data representative of higher scales. The different tools to perform these approach are based on their capacity to detect, retain and harness these interactions.

1.2 Scales of the Atmospheric Dynamics

The atmospheric dynamics allows differentiating among different temporal and spatial variability scales (Fiedler and Panofsky (1970); Orlanski (1975), Fig. 1.1). However, the interactions among those scales, specially under baroclinic conditions (Tribbia and Baumhefner, 2004), can show a high level of complexity (McFarlane, 1987; Cebeci and Bradshaw, 1977), resulting in non-linear and non-stationary relationships among scales. An accurate comprehension of these interactions can allow a better characterization of the dynamics occurring at a certain scale where its variability is only partially known (Cai and Mak, 1990; Katsoulis, 1996).

In the case of wind, its local variability at a certain point with a wind power relevance can be properly described by studying how the atmospheric circulation at synoptic scale interacts with different local mechanisms (Whiteman and Doran, 1993; Garratt, 1994; Jungo et al., 2002). Nevertheless, before relating both scales an exhaustive knowledge of the wind variability within each scale is required. The next two sections describe the main particularities of the large-scale (synoptic) wind circulation over the region considered for this thesis, the Iberian Peninsula, as well as the main processes involved in the microscale (local) dynamics.

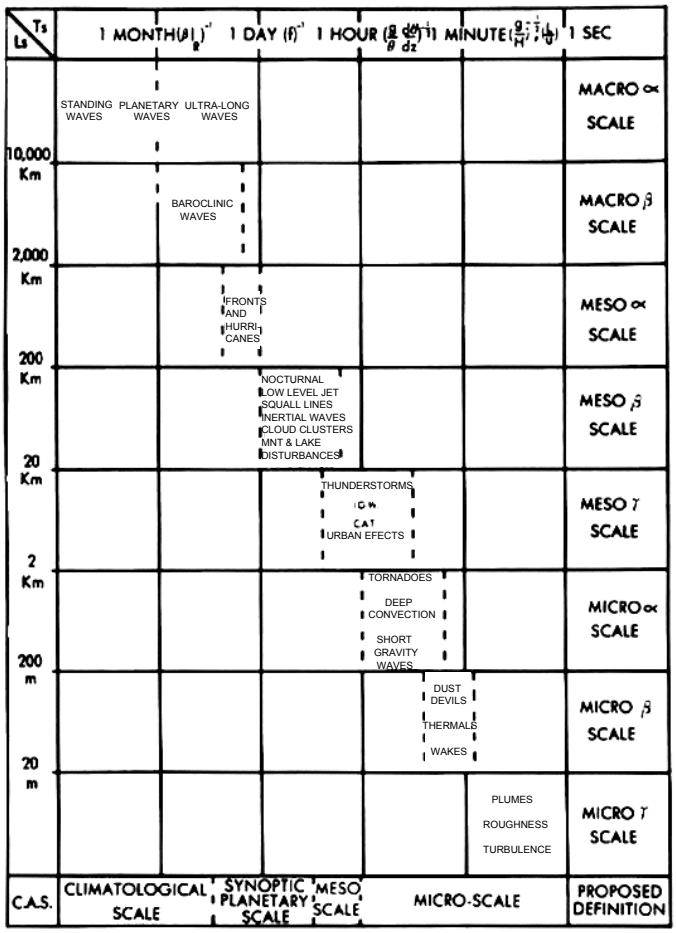

Figure 1.1: Spatiotemporal Scales of the atmospheric circulation (adapted from Orlanski (1975)).

1.2.1 Synoptic Circulation over the Iberian Peninsula

The Iberian Peninsula (IP) is the area of study of this work. The highest persistence mode in the IP dynamics is characterized by the preponderant influence of the Azores anticyclone, generally producing a weak flow when placed over the IP. The disturbances of this steady-state of variability are mainly ruled by the confluence of the climatological conditions from the Atlantic Ocean (AT) and the Mediterranean basin (MED). Other remote regions can also have an impact over the IP at specific moments, evidenced through large-scale teleconnections at both regional (Scandinavia or Russia (Barnston and Livezey, 1987)) and planetary (ENSO, (Rocha, 1999)) scales.

The synoptic dynamics over AT is characterized by hosting a clearly defined *storm track* over mid-latitudes. This semi-permanent flow is conditioned by a moderate south-west to north-east component caused by the Gulf Stream (Joyce et al., 2009), as well as other phenomena as baroclinic processes (Hoskins et al., 1985), linear instability (Zhang and Held, 1999), downstream development (Chang, 1993), barotropic modulation (Lee, 1995), and diabatic heating (Chang et al., 2002). This path of instability systems exerts an intermittent influence over the IP (Seierstad et al., 2007), with great persistence during winter, in turn almost nonexistent in summer (Trigo et al., 2000; Rogers, 1997), and its incidence entails moderate to high winds and a clearer relationship between synoptic conditions and local wind.

The North Atlantic Oscillation (NAO, 2003) represents the best descriptor of the mid- and low-frequency variability over the Atlantic. It influences the storm track and results a useful pattern to describe the intra and interannual variability of the AT influence on the PI. This descriptor measures fluctuations in pressure anomalies on the dipole formed by the depression of Iceland and the Azores anticyclone. In its positive phase it reflects a northeastward-oriented storm track, with high pressure conditions in the IP and Mediterranean Europe. In its negative phase, however, the southward excursion of the Icelandic depression causes the storm track deviation in that direction, reversing the rainfall conditions between northern and southern Europe and leading to negative SLP anomalies over the IP. Sometimes, the predominant role of the NAO is replaced by the Eastern Atlantic (EA) pattern, with a southern dipole spanned towards the tropics, regulating the IP precipitation in spring (Rodriguez-Puebla et al., 1998). Jerez et al. (2013) performed a study of the impact of NAO on wind speed over the IP (winter), observing a negative phase associated to winds 10-15% higher than at the positive phase, as can be seen in figure 1.2.

As for the MED, its high amount of latent heat retention and markedly irregular coastal topography (Lionello et al., 2006a) make this region actively involved in the circulation regime over the PI. In this environment, the storm track during the extended winter penetrates between the Alps and Pyrenees as an intense flow on the north-eastern IP and the Balearic Islands, reinforcing the semi-stationary genoese depression (Trigo et al., 2000). In addition, during this season a teleconnective pattern centered over Scandinavia (SCAND) seems to be a good descriptor of the cyclonic circulation over the Western Mediterranean basin (Trigo et al., 2008). During spring and summer the genoese low pressure system induces cyclogenesis precesses, reinforced by the thermal diurnal forcing at the western are of MED (Trigo et al., 2002). In summer, however, the great amount of

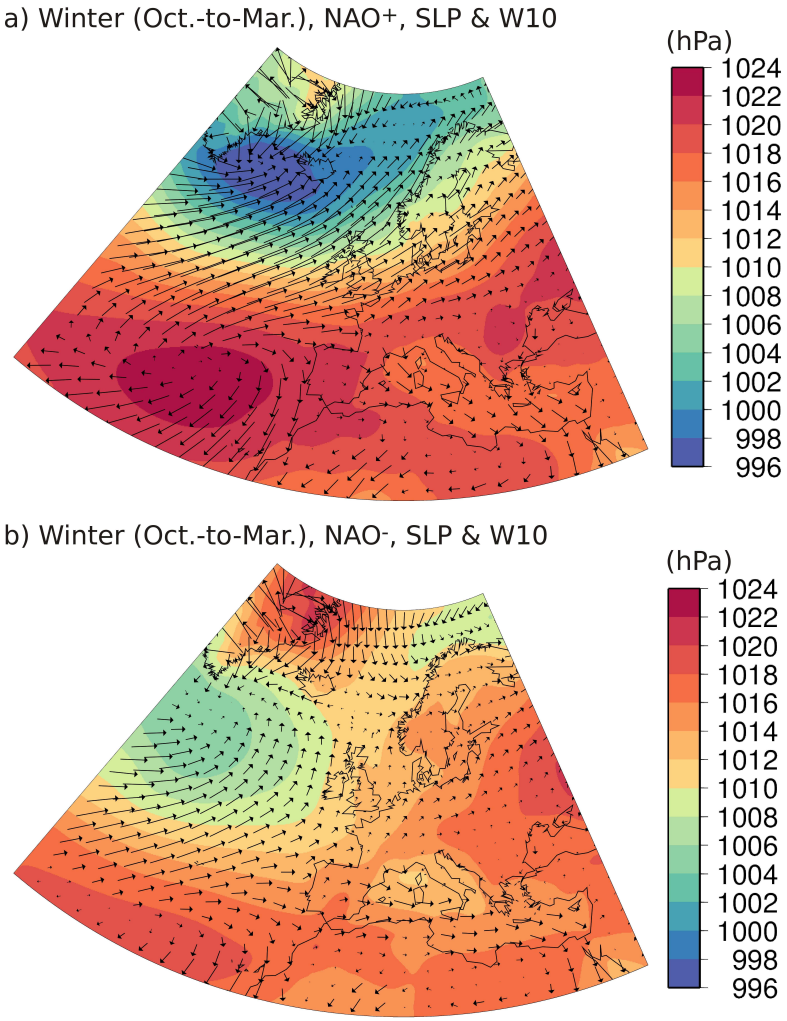

Figure 1.2: Extended winter (October-March) climatology of sea level pressure (SLP by shaded colors, hPa) and 10m wind (arrows) fields during (a) positive and (b) negative phases of NAO. Averaged period: 1959-2007. Data for SLP and W10 were obtained from the ERA-40 reanalysis (Uppala et al., 2005) for the period 1959-2002 and from the ECMWF analysis for the period 2003-07. Data for the NAO index were obtained from the NOAA/CPC (adapted from Jerez et al. (2013)).

acquired power in the form of sensible and latent heat promotes the arise of mesoscalar low pressure systems. These thermal depressions, ruled by the daily thermal gradient, induce in turn convective circulation patterns over northern Africa and the eastern half of the IP (Lionello et al., 2006b).

Under the influence of these factors of so different nature, the synoptic wind conditions on the IP are characterized by a specially high variability and seasonality. The region receives the highest energy input between NW and SW directions (Gastón et al., 2008), associated with the storm track penetrating direction during the extended winter (Oct-May), when it presents a significantly negative correlation with the NAO (Jerez et al., 2013), becoming almost non-existent in summer. A second synoptic-scale wind pattern on Iberia with N and NW flows can be associated with positive pressure anomalies over the Atlantic (García-Bustamante et al., 2012), enhancing the flow over the extreme eastern IP (Mistral and Tramontana) and the Gibraltar Strait (Bormans et al., 1986), there occurring with a NE and E component (Jerez and Trigo, 2013). Although it can take place year-round, this is the only example of high flow pattern over the IP in the summer months, when it appears associated with a sNAO (*summer NAO*, Bladé et al. (2012)) positive index. Figure 1.3 shows the NAO impact on the Iberian wind (March), evidencing the consistence between the Gibraltar Area positive impact and a positive pressure anomaly over the Atlantic. A variation of this pattern occurs when the pressure gradient over the Ebro valley is maximized and the cold air advection from the Atlantic generates a passage of strong winds (de Pedraza and Nacional, 1964), mainly in autumn and winter (*cierzo*).

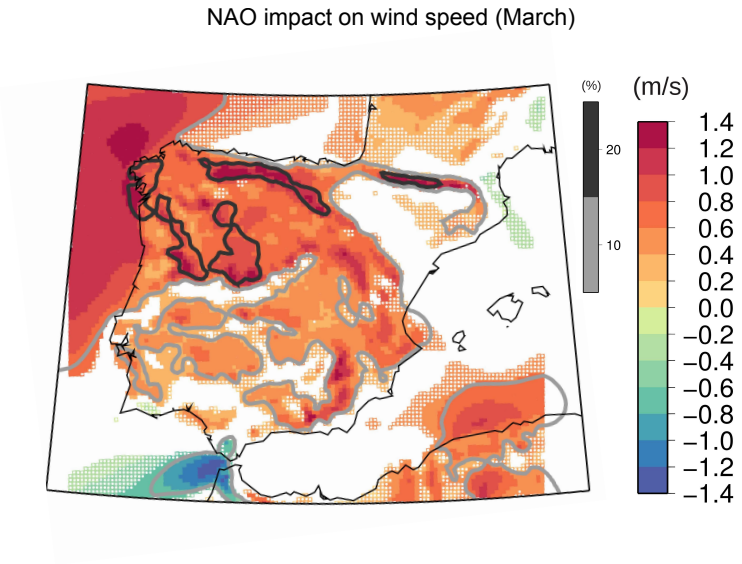

Figure 1.3: NAO impact on wind speed (March): differences in mean wind speed between negative and positive phases of NAO, statistically significant at the 10% level and supported by significant temporal correlations between wind speed and NAO series. Gray-shaded contours depict the differences expressed as a percentage. The superimposed white dots indicate a signal-to-noise ratio that is below 0.75. Notice the differential signal around the Gibraltar Strait area as a contrast to the rest of the IP (adapted from Jerez et al. (2013)).

The great diversity within the synoptic-scale circulation over the IP, together with a high mesoscale variability and its complex terrain make this region particularly interesting when analyzing the variability of surface wind.

1.2.2 Wind Forcings at Local Scale

When the spatiotemporal scale of the atmospheric circulation is reduced, various elements with a local nature progressively take a stronger role and significantly modulate the atmospheric dynamics shown at larger scales (Cebeci and Bradshaw, 1977). At an altitude lower than 150 m (average maximum height of the turbines) the flow is permanently within the Planetary Boundary Layer (PBL), a layer characterized by hosting the dynamic processes originated at the surface (Schlichting et al., 1968). Fundamentally, two principles of local action can influence the circulation. On one hand lie those processes with a thermodynamic nature, like evapotranspiration caused by differences in radiation and humidity. Since they are highly coupled with the solar radiation incidence, they usu-

ally show a remarkable 24-hour cycle (Holtslag and De Bruin, 1988), usually accompanied by higher frequency cycles until the one second band (Fiedler and Panofsky, 1970). On the other hand, other mechanisms caused by the surface orographic conditions, entail a permanent (sea, land) or quasi-permanent (ice sheet, deciduous forests) forcing on wind (Bellecci et al., 2001).

The turbulent flux originated by diabatic heat can provide a vertical component to the flow by forming micro convection cells or eddies, specially under unstable atmospheric conditions in the PBL. The size of these elements will depend, among other factors, of the radiation over surface and the amount of water vapor which can be extracted, either by vegetal or marine evapotranspiration (Dunst, 1982). These elements, have a great influence on the circulation characteristics transmitted from higher levels (free or stratified atmosphere), promoting its deceleration due to the energy dissipation caused by differences in humidity and temperature. In addition, evapotranspiration and condensation processes can generate pressure differences capable of modulating the circulation characteristics (Makarieva et al., 2010). Since these processes depend on the soil composition and its use, these will constitute a factor of influence on the local dynamics, specially if there exists a high water content, depending thus on the crop or land use (Mahrt and Ek, 1993).

Regarding the mechanical forcings, the orography is one of the main reasons for the spatial variability of the surface wind, which is higher than other meteorological variables such as temperature, humidity or pressure. Thus, local topography will be able to modulate the surface circulation intensity and direction, especially in areas with significantly irregular conditions as mountain ranges, canyons, cliffs or straits (Corby, 1954; Wood, 2000). In turn, the roughness of the terrain constitutes a local feature with a substantial influence on the flow intensity. This factor, which can be considered as the microscale topography, plays a key role in the conservation or dissipation of momentum within the near surface fluid. Since wind speed must be zero at the ground level according to mass continuity, roughness plays a major role in the increase that speed experiences with height. The most extended approximations of this increase are the logarithmic (Larsén et al., 2012) and the power (McIntyre et al., 2011) approaches. The first one is based in the theoretical description of the wind shear and results more appropriate at neutral atmospherical conditions, as the latter is an empirical law which results more suitable during stable or instable conditions (Kubik et al., 2011).

Thus, there is a big amount of mechanisms modulating local wind conditions. However,

most of them are not detectable by a synoptic scale atmospheric model, since they have a scale smaller than the spatial resolution of these models. Hence, if the surface wind is to be estimated, an accurate characterization of the local variability will be required in order to establish robust relationships among the different scales. The most employed methods for solving these local processes at high resolution consist on employing the information from the coarser scale either through specific resolution modules -pbl (Fan and Sailor, 2005), microphysics (Hong and Lim, 2006), floor model (Verseghy, 1991), etc- or either on dynamic (Baik et al., 2003; Skamarock et al., 2005) or statistical (Zorita and Von Storch, 1999; Jakob Themeßl et al., 2011) downscaling techniques. In the next section these and other methods are detailed in the context of wind power industry applications.

1.3 Wind Variability Scales and its Application to the Wind Power Industry

The wind speed at a given point is the result of the interaction of its variability at all time scales. However, certain frequencies of its spectrum present energy density peaks, which can be associated with different aspects in wind power research (Figure 1.4).

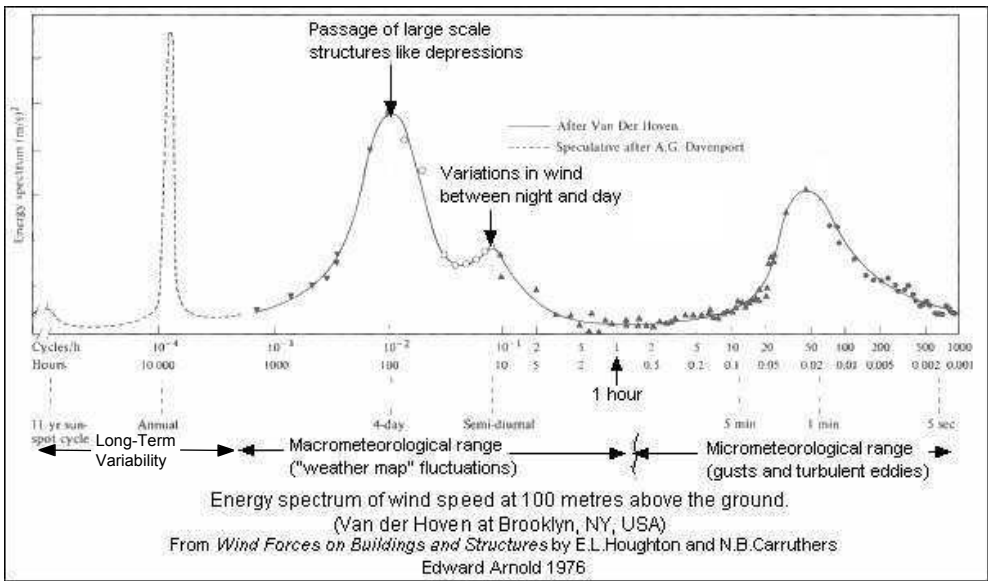

Figure 1.4: Energetic spectral density model of wind speed (modified from Houghton and Carruthers (1976a)).

In the short-term wind speed variability different spectral density energy peaks can be observed, at the one-minute and one-day frequency bands (Van der Hoven, 1957). In turn, in the mid-term variability a maximum is found on the four-day frequency. Finally, the frequency associated with the annual cycle shows a high prominence (Houghton and Carruthers, 1976b). However, the references to surface wind cycles with frequencies lower than the annual range are scarce (Barriopedro et al., 2013). Throughout this temporal variability spectrum, the wind industry application scope is large and diverse. Four main timescales can be distinguished:

- *Intra-daily Variability (very short-term)*

  The band with periods lower than 24 hours corresponds mostly to the turbulent effects of surface dynamics take place. However, the electrical grid system can not admit a power fluctuation beyond a certain threshold due to possible instabilities in the electricity supply system. Thus, the knowledge of the variability and wind forecast in this range is vital from a purely operational perspective. Usually, the power producer must submit to the electric distribution company an hourly-resolution estimate of the generated power, with the corresponding penalty for the producer if this estimate is wrong. In order to optimize the short-term wind prediction different statistical methods have been implemented during the last decade, such as Artificial Neural Networks (ANN, Cadenas and Rivera (2009); Salcedo-Sanz et al. (2009)), Support Vector Machines (Salcedo-Sanz et al., 2011) or Kalman Filters (Giebel et al., 2011; Crochet, 2004).

- *Inter-daily Variability (sort-term)*

  The four-day cycle of the wind variability is associated with the average duration of cyclonic synoptic-scale weather processes in mid-latitudes. The economic value of this range forecasts lies on the ability to anticipate meteorological events which can pose a danger for the structural safety of the facility, allowing preventing measures such as stopping turbine operations. Other situations that can entail a stopping process can occur when winds weaken until falling below the wind turbine operation minimum speed ($Cut\text{-}In$). By attaining an accurate forecasting, operational and maintenance shutdowns can be planned in those situations, thus minimizing the power loss. This level of variability is captured by the numerical models, which are usually enhanced with different statistical resources (Galanis et al., 2006; Kariniotakis et al., 2006; Louka et al., 2008).

- *Intra-annual Variability (mid-term)*

  The sub-annual changes in the surface flow has been analyzed through the monthly and seasonal teleconnective impact of main large-scale patterns as ENSO (Harper et al., 2007) or NAO (Brayshaw et al., 2011). Since wind power electricity volume represents a great ratio on the electricity quotation, the analysis of these variability scale is particularly relevant for commercial operators of the energy sector, which have to bid in advance for contracts in the electricity market auctions, with a quarterly frequency in most countries. Different works have been performed at these variability range, as those on the comparison of seasonal wind variability between hemispheres (Sandwell and Agreen, 1984) or those analyzing its implications on wind production (Weisser and Foxon, 2003).

- *Inter-annual and longer Variability (long-term)*

  The climatic characteristics of wind has been considered since the beginning of the wind industry as the key element in deciding the location of a wind farm. In this regard, wind atlases have been computed through the implementation of *in situ* observations or reanalysis data into models like WAsP (Wind Atlas Analysis and Application Program, Mortensen et al. (1993)) or Regional Circulation Models (RCM Yim et al. (2007); Lo et al. (2008); Byrkjedal and Berge (2008)). There, the main objective is to estimate the spatial distribution of the average values of wind conditions (see Fig. 1.5 as an example). However, the low frequency wind variability has been scarcely considered. Although several contributions provide estimations on wind conditions at the end of the 20th century under different forcings from climate change scenarios (Sailor et al., 2000; Najac et al., 2009; Pryor et al., 2005a), the multidecadal wind variability in the past has been addressed in few works (Ekström, 2002; Pryor and Barthelmie, 2003; García-Bustamante et al., 2012; Earl et al., 2013). The following section discusses in more detail this type of variability.

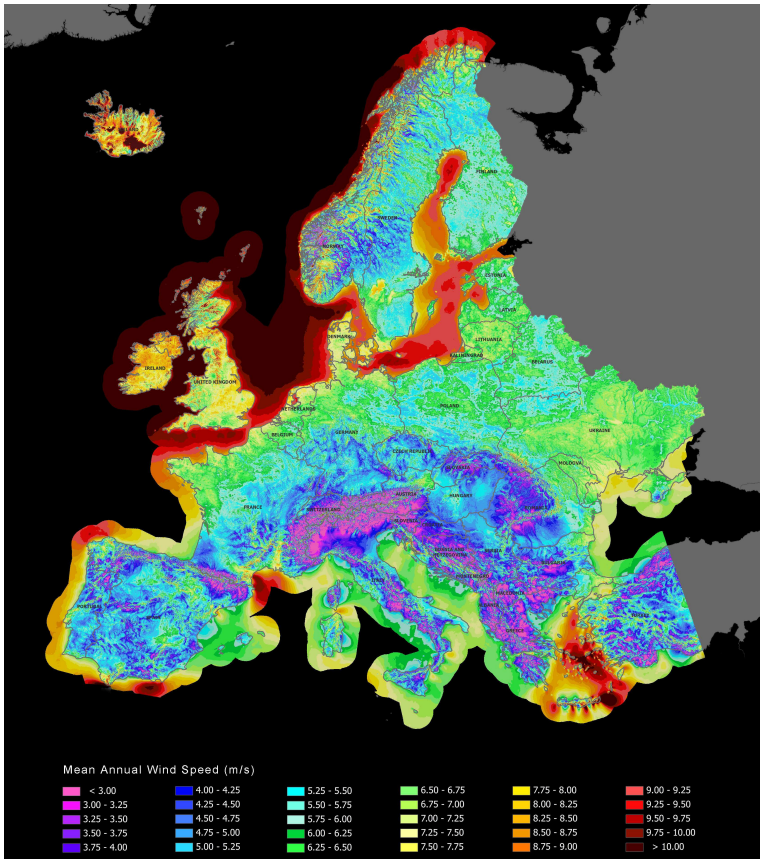

Figure 1.5: Wind resource map for Europe, with a resolution of 200m for a 80m height (Adapted from AWS Truepower (2012).

1.4 The Long-Term Wind Variability

The wind industry has very recently paid attention to the wind variability affecting the long-term perspective of the existing or planned facilities. This need rises from the long planned time range of a wind project, where the wind turbines lifetime can be guaranteed up to twenty-five years. In addition, the project can be continued beyond that time through the wind *repowering*, consisting on the progressive renewal of wind turbines for new and more advanced and efficient models. This time scale, covering from inter-annual to centennial changes, will be referred hereafter to as *multidecadal*.

The available information from wind variability at multidecadal scales depends on both the length and the temporal resolution of the data series. Its resolution will determine the maximum analyzable frequency, while its length will determine the minimum variability

frequency that can be considered. In this sense, the study of the multidecadal changes of wind requires time series that at least double the time range length of the timescale to be analyzed, with a resolution in turn high enough to study inter-annual changes. Next, different ways to obtain a detailed wind time series from a given point of wind power relevance are described.

1.4.1 Observations

The most immediate possibility to analyze the long-term behavior of wind is by considering observations (Baker et al., 1978; Martner and Marwitz, 1982; Wendland, 1982; Archer and Jacobson, 2003; Veigas and Iglesias, 2012). However, these contributions are usually limited to assess the climatological wind capacity of the site without analyzing its variability. Earl et al. (2013) represents an exception since individual cases of inter-annual variability in the UK are analyzed throughout 30 years from wind observations at forty weather stations.

As it has been mentioned before, the high spatial and temporal wind variability adds difficulties to this approach. The relatively recent introduction of quality wind measurements compared with other variables such as temperature or pressure, and to the low probability that a location with a high wind power interest holds a meteorological measuring station beforehand makes that usually wind observations can not be considered as representative from a multidecadal variability perspective. In addition, even considering a flat terrain, and if observations existed, they must undergo a quality control (Eischeid et al., 1995; DeGaetano, 1997), with the possibility of not gathering the expected conditions. This requires *in situ* observation campaigns with a duration between one and five years. These campaigns, although able to provide a first approximation of the wind regime at the site, are insufficient for a multidecadal analysis.

Useful tools are the micro-scale models like WAsP (Landberg et al., 2003), which only need nearby observations to estimate the surrounding wind conditions. They allow mapping wind regimes in a simple way from the extrapolation of observations under roughness and topography forcings. Despite being wind data-dependent methods, they provide a good performance when a high spatial density of observations is available (Hošek et al., 2004; Landberg et al., 2003). Although this kind of models solves the dynamical governing equations in a simplified way, they result a competitive alternative to analyze the spatial variability of the average wind regime in a given area, reflected through a so-called *Wind Atlas* (Troen and Lundtang Petersen, 1989; Mortensen et al., 2006, 2012). An alternative

to these models has emerged during last years through the CFD (Computational Fluid Dynamics, Van Wachem et al. (2001)) models thanks to the increasing computing power. Contrary to WAsP, these tools are able to solve the governing flow equations in a very limited space domain (Launder et al., 1975; Periera et al., 2010).

1.4.2 Numerical Models

Thanks to advances in computational capacity, the limitations due to the scarcity of observations can be partially mitigated by developing models based on the numerical integration of the equations of motion for a given time period and space resolution. Simple models of synoptic scale weather forecasting began its development in the 50s (Aspray, 1990). Parallel to them, the first climate models, with very low resolution (35000 km) but globally run (Global Climate Models, GCM), started its implementation (Phillips, 1956). These approaches at different scales evolved for decades, up to offer higher performance solutions with different purposes, either meteorological (NAM, 2006, GFS, 1989, around 40 km resolution) or climatic (CESM,2013, HadCM3, 2005, 100-200 km res.). Although numerical methods were primarily used in the development of predictive models, in recent decades they have been implemented for *a posteriori* analysis, through different *reanalysis* projects. They recalculate the dynamical conditions of the atmosphere by considering as input all kind of observed meteorological data -stations, balloons, radiosondes, etc- and being forced to them through data-assimilation procedures. This allows to study the past weather conditions through the increase and standarization of the spatial and temporal resolution of the considered inputs. Thus, several projects that offer gridded data every six hours have been developed (Kalnay et al., 1996; Uppala et al., 2005), including those covering centennial periods (20CR2, 2011 ; HadSLP2, 2006). Finally, Pryor and Barthelmie (2003) provide a detailed analysis for Scandinavia by assessing the long-term wind variability and its influence in wind energy from the NCEP/NCAR (Kalnay et al., 1996) reanalysis wind outputs.

In general, such models are able to provide an extensive spatial coverage, but its resolution is insufficient for analyzing microscale wind conditions, being necessary to consider alternatives which reduce the spatial scale (i.e. *downscaling*).

1.4.3 Downscaling Techniques

In order to increase the spatial resolution of the variables obtained from GCMs at a particular region, different dynamical downscaling techniques have been developed since the nineties (von Storch et al., 1993; Wilby and Wigley, 1997). Regarding wind, Frey-Buness et al. (1995) and others (Bergström, 1996; Sandström, 1997) run a mesoscalar wind model from defining a steady-state of the atmosphere in a number of clusters. However, the most extended downscaling technique consists on improving the resolution of GCMs by approaching the flow equations through the finite differences method (Richardson, 1911) at a certain number of nested domains, where the boundary conditions are obtained from the immediately larger domain previously resolved. The implementation of these techniques in addition to the increased computing capacity and the implementation of resolutive modules of specific problems (microphysics, PBL, soil model, etc) conform the Regional Circulation Models (RCMs), which allowed to increase the spatial resolution until less than 10 Km. A wide amount of works have been performed with RCMs to analyze a given forcing or the variability of a given feature at a specific area or height. Numerous contributions perform a dynamical downscaling focussing on wind, either at a localized area (Jiménez et al., 2012; Jerez et al., 2013; Jerez and Trigo, 2013) or a wider region (Walter et al., 2006; Kanamitsu and Kanamaru, 2007).

In general, the dynamical downscaling models perform well when describing the mesoscale circulation, but their spatial resolution of several kilometers show some troubles in explaining the mechanisms ruling the surface wind in a specific location (Grotch and MacCracken, 1991; Gates et al., 1996; Huth, 1999). This is evidenced in the troubles when describing features as abrupt roughness changes (Hahmann et al., 2012), or when results are employed as input for microscale approaches (Badger et al., 2010).

The *Model Output Statistics*, (MOS 1972) has been a widely used statistical technique for the treatment of the atmospheric data for the computation of mid-term predictions (Sailor et al., 2000). It was used for the preparation of the post-processing of numerical prediction models, and combines the outputs of the dynamical model with statistical estimates (Klein and Glahn, 1974). The MOS uses numerical forecasts based on the atmospheric physics to predict the large-scale climate patterns and then uses equations of statistical regression to resolve details of the surface at a smaller scale, obtaining a performance usually much higher than that of any option executed separately. Following a similar procedure, the so called empirical or *statistical downscaling* methods (SD) have been developed in recent decades. These methodologies apply the statistical relationships

observed between flow characteristics at different variability scales. These resources assume these interactions as valid for other periods where no information on the smaller scale is available. This makes possible to increase the resolution at a specific site and time range, for which only a limited period of *in situ* observations are available.

Since they do not need to resolve the governing equations of flow, SD proceed at a very low computational cost. Thus, they result particularly appropriate when solving complex relationships between scales. In this way, such techniques are appropriate for problems in which dynamical computing does not show a high performance, either because the local scale encloses a special complexity (Salameh et al., 2009; van der Kamp et al., 2012), the variable to be studied does not share an explicit relationship with the larger-scale dynamics (as happens with phenological series (Matulla et al., 2003)), the reaction time needed is particularly short (as in the short-term estimations described in Section 1.3) or the period to be simulated is particularly large (García-Bustamante et al., 2012).

To perform a statistical downscaling two sets of time series belonging to different sources are needed. On the one hand a series of the variable to be estimated at local scale must be considered (predictand). On the other, a data set of one or more series of larger-scale variables with some predicting value (predictors). These two sets need to match for a period large enough so that the existing relationships between them can be identified. This process is known as the calibration or *training* of the predictor. After this, and before the retained relationships can be extended to periods where only predictor data is available, the method will have to be cross-validated, within a *test* period not previously employed at the train.

The way to relate the two pairs of data within the training has adopted a wide amount of forms. For instance, statistical downscalings have been developed from canonical correlation analysis (CCA, Von Storch and Zwiers (2001)) to estimate local values of temperature (Barnett and Preisendorfer, 1987) and precipitation (Busuioc et al., 2008), singular value decomposition (SVD, Bretherton et al. (1992)) or Artificial Neural Networks (ANN, Crane and Hewitson (1998); Trigo and Palutikof (1999); Cavazos (2000) while others have estimated temperature and precipitation by linear statistical (Hanssen-Bauer et al., 2005) or stochastic (Hundecha and Bárdossy, 2008) procedures. In turn, methods on circulation patterns clustering have been applied to analyze temperature (Hewitson and Crane, 1992; Hewitson, 1994; Schubert and Henderson-Sellers, 1997), precipitation (Noguer, 1994; Corte-Real et al., 1995; Saunders and Byrne, 1996) and sea level variability (Heyen et al., 1996).

The implementation of a statistical downscaling of wind is particularly appropriate to analyse the long-term at a given site, a problem where the computational cost is specially penalized under a dynamical approach. As it will be seen in the next section, there exist a wide number of papers which present different methods on statistically downscaling wind (either correlation- or regression-based). However the approaches that focus on applying it to analyze its multidecadal variability is rather sparse (Palutikof et al., 1992; Ekström, 2002; García-Bustamante et al., 2012).

1.5 The Statistical Downscaling of Wind

1.5.1 The design of an appropriate predictor for the local wind speed

The development of an appropriate predictor from the large-scale circulation for describing the local wind conditions at a specific point results in critical relevance in order to obtain an accurate model for the surface wind.

A major trouble in the training process of a statistical downscaling, specially when considering non-linear processes, is the risk that the large-scale predictor experiences overfitting (Hawkins, 2004; Babyak, 2004). Should this happen, it would mean that the predictor has assimilated the characteristics of the local wind in a too high level of detail (Salcedo-Sanz et al., 2005). Hence, in absence of a predictand it will describe very accurately those situations which resemble the processes experienced during the training, but it will not be able to reproduce properly the overall wind conditions. On the other hand, a too generalistic algorithm (i.e. linear regression) can result in a low performance when estimating the surface wind since wind conditions usually can not be considered linear. In this way, an efficient predictor has to be able to compromise accuracy with generality. The consideration of a wind predictor that classifies the temporal series in different types or groups according to the synoptic characteristics contributes to prevent the problem of overfitting (Monahan, 2012), due to the constriction of the predictor information into a few modes of variability. A clustering is thus useful when considering a training period long enough, so that the established similarities among the elements within a group can be maintained in the local wind, regardless of wether the relationship between the different groups can be non-linear. In the next section the different methods on atmospheric circulation clustering are described.

1.5.1.1 The classification of the synoptic circulation as a wind predictor

A wide amount of methods have been employed for the statistical downscaling of wind, all of them by considering different information and parameterizations on the large-scale atmospheric circulation as predictor. Most of them include a clustering process at a certain step. Different methods on clustering synoptic circulation have been developed from the 1950s to classify the regional dynamical conditions into a number of prevailing patterns, usually with a daily resolution. Initially this methodology was carried out in a subjective way (Hess and Brezowsky, 1952; Lamb, 1972), where the meteorologist decided which type or class deserved every synoptical situation, basing its decision on the position of the pressure systems, the flow direction or the vorticity. The first objective methods arose by means of correlation techniques (Lund, 1963; Kirchhofer, 1974). Later other objective methods appeared, highlighting a geostrophic-based approach first proposed by Jenkinson and Collison (1977). They developed an automatized classification based on the intensity and vorticity of a geostrophic approach, computed through certain pressure differences around the center of a fixed point over the British Isles. This technique, which became popularized as the *Jenkinson-Collison* types or Circulation Weather Types (WT), was later applied to other locations like Scandinavia (Linderson, 2001) or the IP (Trigo and DaCamara, 2000; Spellman, 2000). Autovector-based classifications were developed (*Prinicpal Components Analysis*, PCA, Davis and Kalkstein (1990); Esteban et al. (2006)) including studies over the IP (Romero et al., 1999; Paredes et al., 2006; Queralt et al., 2009; Jiménez et al., 2009). Other classifications are based on log-linear models (Prieto et al., 2004) or Euclidean distances over the considered space, either through k-means (García-Valero et al., 2012) or k-medoids (Hopke and Kaufman, 1990) from a CCA (Beranová and Huth, 2008).

Regarding the classification of the atmospheric conditions exclusively attending to the wind conditions, Green et al. (1992) performed a PCA of the vector wind series in the analysis of transport of pollutants, while Burlando (2009) developed a wind field classification of the Mediterranean Area by performing a Cluster Analysis (CA) through a k-means classificator (Hartigan and Wong, 1979) and the *Minimum Variance Method* (Ward Jr, 1963) on the Euclidean distances of the wind speed space.

For the statistical downscaling of wind, most of the literature has considered different techniques on the classification of the low-resolution circulation as predictor. Thus, Palutikof et al. (1992) considered the types of the Lamb classification (Lamb, 1972) which best correlated with the geostrophic wind in order to perform correction factors, Mengelkamp

(1999) performed a CA (Anderberg, 1973) on the geostrophic wind components and the vertical temperature gradient. In turn, Goubanova et al. (2011); van der Kamp et al. (2012); Davy et al. (2010); Monahan (2012); García-Bustamante et al. (2012) performed a PCA of the large-scale selected descriptors (usually geopotential heights or pressure fields), while Bogardi and Matyasovzky (1996); Salameh et al. (2009) and Najac et al. (2009) applied a k-means classificator to the obtained PCs.

Additionally to preventing overfitting, the clustering of the different states of the atmospheric conditions is also useful as a stand-alone tool that provides a clear description on the variability of the different wind conditions. This is most clear in Salameh et al. (2009) and Najac et al. (2009), who describe the variability of the main circulation patterns associated with the local wind conditions in several points in France.

There are a few works which directly apply a transfer function during the training, either through a series of multiple linear regressions and correlations of the wind components (Ekström, 2002; Achberger et al., 2002), cumulative distribution functions of wind speed (Michelangeli et al., 1995), or an implicit transfer function of weights within an Artificial Neural Network (Sailor et al., 2000). This last method belongs to the so called *Soft Computing* field, represented by a series of tools which allow considering an optimization (clustering) or regression (wind speed estimation) problem without the need of facing all its possible solutions, allowing to obtain an approximate result of a *hard* computational problem.

1.5.1.2 Circulation patterns Clustering and Soft-Computing

The realization of a PCA on the field where the wind predictor is to be extracted is a recurrent technique in the literature. This is mainly due to its ability to extract a high amount of information in few modes or patterns, while providing coefficients for a multiple linear regression in which its components will be linearly independent. These features are highly useful to mitigate redundant information in terms of explained variance by the predictor. However, the realization of a PCA can imply certain drawbacks. First, it requires the subjective consideration of the domain where the modes of variability will be extracted. Considering a small number of PCs to prevent overfitting implies the dismissing of a significant amount of information. This can imply that only 50% (van der Kamp et al., 2012; Monahan, 2012) to 70% (Goubanova et al., 2011) of the predictor variability is considered. Finally, the condition of independence between the predictor PCs does not always ensure that only one type of local wind variability is retained on a certain

PC. These factors can lead to the situation where the retained PCs do not represent the optimal predictors of local wind within the information considered.

Furthermore, the techniques on circulation patterns clustering sometimes entail some problems associated to the high amount of information considered. This is the case when daily pressure field gridded data are considered. This is possible that a clustering problem, based on the similarity within the group and the differentiation between groups, contains implicit more efficient solutions than those found by the classical clustering techniques, but they remain neglected due to the large number of clustering possibilities. On the other hand, in the case of wind, a statistical model on a monthly basis entails limitations (Pryor et al., 2005c; Monahan, 2012; Curry et al., 2012), being preferable to consider higher resolutions as the daily scale. By doing so the amount of points of the considered time series becomes significantly higher, thereby facilitating the obtention of longer training periods, which ensures more robust results.

Soft-Computing techniques (SC) are able to address this type of problems. Such methodologies, initiated at the beginning of the 90s, offer an efficient solution to a regressive or an optimization problem with a small computational cost without dismissing information. This is possible because by definition, they do not need to cover all the possible solutions. Instead, at each iteration they follow a certain path by making use of different properties that emulate behaviors found in the nature, offering thus an efficient outcome. These methods include the problem solving by Fuzzy Logic (FL, Mamdani and Assilian (1975)), Artificial Neural Networks (ANN, Hopfield (1988)), Support Vector Machines (SVM, Osuna et al. (1997), Greedy Algorithms (GR, DeVore and Temlyakov (1996); Bednorz (2008) or the Evolutionary Computation (EC, Back et al. (1997); Fogel (2006).

In meteorology soft-computing techniques have been employed to produce pressure patterns through methods as SANDRA (Philipp et al., 2007; Küttel et al., 2011). In turn, *Random Forests* (Breiman, 2001) are employed for the statistical downscaling of wind (Faucher et al., 1999; Davy et al., 2010), where a selection of the best predictors is extracted from a list of different variables related to the circulation at low resolution. Mohandes et al. (2004) employ SVMs within wind speed data-only to perform a daily wind prediction over 4 years of test, while Bouzgou and Benoudjit (2011) do so by considering temperature and temporal and spatial coordinates. Finally, Sailor et al. (2000) perform an ANN for the daily reconstruction of an average year from 30 years of data, and then compare it with climate change projections. Evolutionary algorithms have been applied

in multiple areas with positive results. These include problems on cloud microphysics and aerosols (Krekov and Sukhanov, 2012), the prediction of atmospheric corrosion (Fang et al., 2008), the resolution of the variational analysis of ocean dynamics (Barth, 1992) or the development of temperature profiles (Buczak and Barrett, 1998). In the study of wind these algorithms have been used to estimate wind in the lower atmosphere through Doppler pulse radar (Chen et al., 2001) or for wind power short-term predictions (Jursa and Rohrig, 2008). Finally, the greedy algorithms have been used to reduce information in atmospheric data assimilation (Ochotta et al., 2005; Oehlert, 1996). However, evolutionary algorithms or greedy algorithms have not been applied previously in the classification of a wind predictor or its statistical downscaling.

1.5.2 Sources of Uncertainty in a Statistical Downscaling

Since statistical downscalings are based on a series of empirical relationships, the quantification of their estimation uncertainty results particularly important. In these kind of approaches the uncertainty is caused mainly by two different sources. The first gathers all kinds of uncertainty arising from the predictor. In the case of a statistical downscaling of wind from reanalysis data, all sources of uncertainty related to data assimilation must be considered. This includes the uncertainties due to instrumental error in the observations used as input, and the error in the integration of the equations of movement through the method of finite differences (Arakawa, 1966). In turn, the uncertainty due to the initial and boundary conditions will propagate and expand the errors contained in the previous sources (Lorenz, 1965). The second source of uncertainty lies on the statistical exercise itself. In the case of wind, implementing clustering methods has the advantage of reducing the dimension of the problem, but leads to the uncertainty associated with the dispersion that exists in any clustering exercise. The method can not guarantee that the statistical relationships are kept stationary in time, with the implicit risk to increase the uncertainty in other periods away from the training range. Finally, the employment of a non-deterministic method as an evolutionary algorithm requires its launching a repeated number of times (usually 30, but will depend on the problem) in order to ensure the robustness of the method. This will thereby generate one additional source of (systematic) uncertainty due to the dispersion produced by the different results obtained at every launch.

Most of the sources of uncertainty involved in a statistical downscaling can be quantified within the error values obtained through a cross-validation process. Because of its

low computational cost, performing a cross-validation or *testing* results particularly simple in a statistical downscaling, offering the possibility to provide a value for the error or uncertainty of the wind estimates.

The only uncertainty which can not be attained through a cross-validation process lies on the possible non-stationarity of the relationships between the large-scale circulation and the local scale in a given period. This is a feature that can not be guaranteed for other temporal ranges away from training. Although some efforts have been recently undertaken to analyze this issue by combining statistical approaches and numerical resources (Etemadi et al., 2013), this still seems to represent an implicit uncertainty component inherent to the model.

1.6 The Wind Power Long-Term Variability

Once a certain wind speed series is obtained, different aspects need to be considered when planning a long-term simulation of wind power generated by a real wind turbine in a specific location. The empirical discrepancies between wind speed and wind power output from a real turbine might justify a wind power statistical downscaling by directly employing wind power data as predictand instead of wind data. Although that approach could contain a valid rationale, its implementation presents a series of problems, as pointed out in 1.1. Indeed, it is very likely that wind power data are not available beforehand. However, in the case that such records existed, these will usually embrace a short period, since modern wind turbine installations can not be much more than a decade old. If still this kind of data was considered, a number of factors should be taken into account, especially when monthly and annual power regimes are considered. These items include from wake effects in relation to other wind turbines nearby (Manwell et al., 2002) to shutdowns by maintenance or damage, until periods where wind speed goes beyond the operating range (minimum or maximum, Jerez et al. (2013)). Although several works consider directly the actual measurements of power output, most of them focus on the assessment of the wind power variability in the short term (Lee and Baldick, 2012; Frehlich, 2013) or on the analysis of different operational features of wind farms (Fertig et al., 2012) or turbines (Ribrant and Bertling, 2007).

Multiple works have simulated wind power through different power models with wind speed data as input (Şahin and Aksakal, 1998; Celik, 2004; Sailor et al., 2000; Earl et al., 2013), including those which performed nearly-centennial analyses of the wind power po-

tential (Veigas and Iglesias, 2012) from wind speed observations at several meteorological stations. All of them employed the usual procedure to estimate wind power, consisting on considering the theoretical wind power-wind speed cubic model, after adjusting the wind speed probability distribution. Others, as Palutikof et al. (1987) evaluated wind power over 52 meteorological stations through the Cliff curves (Cliff, 1977) obtained from wind observations. These curves are simple power models that assume a wind speed Rayleigh distribution and consider the average speed at the hub height as well as the turbine shutdown speed.

Different distributions have been used to adjust the wind speed Probability Density Function (PDF). Several distributions have been considered in the literature, as the Rayleigh (Hennessey Jr, 1978), von Mises (Qin et al., 2010) or Bayesian Average Models (BMA, 2010). However, the maximum likelihood estimation performed on a Weibull (Weibull et al., 1951; Rehman et al., 1994) distribution is the most extended approach (Rehman et al., 1994; Stevens and Smulders, 1979; Seguro and Lambert, 2000). Thus, Pryor et al. (2005c) harnesses the generality of a Weibull fit in its statistical downscaling to generate estimates of decadal wind power density from the cubic model. In turn, Harper et al. (2007) evaluate the long-term impact of ENSO on central North America, through a power model based on a fourth-order polynomial fit of wind speed data. In turn, Brayshaw et al. (2011) assessed the impact of the NAO on the UK wind power generated from wind observations through the cubic model and a number of requirements related to the minimum and maximum speeds supported by a turbine. The same reasoning is followed by Ruiz-Arias et al. (2012); Jerez et al. (2013); Jerez and Trigo (2013). Finally, García-Bustamante et al. (2013) developed a model of statistical downscaling of monthly wind power from instrumental measurements of power, constituting thus an exception to this procedure.

However, different caveats have been reported either on the cubic model(Wilson et al., 1976; Vries, 1979; Anderson and Bose, 1983) or the Weibull distribution (Tuller and Brett, 1984; Jamil et al., 1995; García-Bustamante et al., 2008). These circumstances motivate to seek alternative ways to obtain a series of wind power from wind data which shows robust validation results.

1.6.1 Centennial Wind Power Inter-annual Series: Application Outlook

The availability of a long simulation of the generated wind power opens the door to analyze different coarsely known aspects, such as the impact of the large-scale circulation patterns on wind power, or the parametrization of empirical aspects at the multi-decadal scale not covered by the analytical expressions.

In the last few years this type of variability has been associated to large-scale circulation regimes (Klink, 2007; Harper et al., 2007; Spears and Jones, 2010; Brayshaw et al., 2011; García-Bustamante et al., 2012). Since the typical life span of most wind farms reaches up to three decades, the multidecadal variability of wind can reach significant impact in wind power production (Brayshaw et al. (2011) but specially Jerez et al. (2013); Jerez and Trigo (2013)). Obtaining a multidecadal or centennial simulation of wind power allows analyzing the impact that large-scale spatial patterns of climate variability exert on the generated power from a low frequency perspective, or whether this influence is stationary over time. In this context, the analysis of this multidecadal scale impact can be addressed by considering the main spatial patterns at different states of variability, as shown in Brayshaw et al. (2011).

The empirical relationship between the annual generated power and other elements such as the shape and scale parameters of the Weibull distribution are also of interest. Despite sharing an implicit relationship with wind power, the degree of long-term incidence of these elements has not been deeply evaluated, basically due to the lack of available wind power long series providing robust enough results. As a result to this, exploring new expressions capable of describing the generated wind power output from known parameters in a simple way becomes feasible if long series are available.

1.7 Objectives

The overall objective of this work is to analyze the long-term variability of wind power at given points of special relevance in terms of energetic resource, by the reconstruction of its wind conditions along more than a hundred years.

Specifically, this work will:

1. Design an objective method for the synoptic circulation clustering at daily scales

through the implementation of an algorithm capable to identify wind patterns in the absence of wind observations.

2. Develop a methodology capable to perform daily wind speed estimates at a specific point of wind power relevance with a competitive performance in terms of uncertainty reduction.

3. Build a reconstruction of daily wind regimes, long enough to carry out the analysis of its interdecadal and multidecadal variability.

4. Implement, from the reconstructed wind speed series, simulations of the generated wind power by a real wind turbine, analyzing its interdecadal and multidecadal variability.

5. Analyse the long-term impact of the main large-scale variability modes (*teleconnections*, TC), as well as other wind power-related variables and parameters exert on wind power output.

In Chapter 2 a new methodology is presented, consisting on the development of three different wind-type clustering methods from the daily synoptic situation, through the application of different soft-computing techniques to daily pressure fields (objective 1). In the same chapter a strategy is proposed for the use of the obtained classifications as predictors for a statistical downscaling of daily wind with different observational series as predictand (objective 2). Chapter 3 accounts for two distinct parts. In the first part of the chapter,the cross-validation of the methodology developed is performed, for different points of wind power relevance on the IP. In the second part of the chapter a descriptive analysis of the different obtained wind classes or types is performed.

In Chapter 4 (objective 3) a daily wind reconstruction for more than a hundred years is carried out for the locations considered. This will allow the analysis of the multidecadal variability of the wind speed and direction. From this basis, in Chapter 5 a simulation of the daily generated wind power is performed. Such simulations, obtained by considering the technical characteristics of a real wind turbine, allow analyzing the long-term variability of wind power (objective 4). The second half of the chapter computes and analyzes the degree of impact that various elements have on such variability (objective 5).

Finally, Chapter 6 contains the main conclusions that can be drawn from this thesis, including some considerations on potential future work.

Chapter 2

Clustering and Statistical Downscaling of Wind by Employing Different Soft-Computing Tools

2.1 Introduction

In this chapter three different methods for the statistical downscaling of wind from the synoptic scale circulation to the local wind conditions are introduced. As described in the introductory chapter, most works consider the classification of the wind predictor, mostly by retaining only few EOFs within a PCA, although it implies the loss of certain information. Here, this problem is solved by classifying the large-scale circulation through different soft-computing (SC) algorithms, which allow considering the whole variability of the atmospheric circulation with a low computational cost. Within the SC tools, *Genetic* or Evolutionary Algorithms (EA, Fogel (2006)) are a sub-category of Evolutionary Computation that employ different principles of biological genetics to rule out or accept a set of possible solutions. These algorithms provide a solution for an optimization problem characterized by keeping a non-deterministic basis (such as ANN), leaving a small part of the iterative process to randomness. This implies that an exact solution can not always be ensured, providing a certain solution at every launch. However, this offers a competitive output with a minimal computational cost. Thus, given an optimization problem, an EA implements a set of evolutionary operators (Bäck and Schwefel, 1993), such as mutation,

selection or crossover, to keep or discard the different *individuals* (solutions) at each *generation* (iteration). In this way they are able to provide a valid solution with extremely short computing time. With such characteristics, these tools seem *a priori* suitable for minimizing differences in a large data set, such as a pressure field in the context of its classification as a predictor of wind. Complementarily to EAs, Greedy Algorithms (GR) provide a deterministic solution. Within this kind of tools, the optimal choice in each local step of the process is chosen, with the assumption that an overall optimal solution will be thus obtained at the end of the process. Although different works perform a statistical downscaling of wind by employing SC tools (Faucher et al., 1999; Sailor et al., 2000; Davy et al., 2010), EA or GR have not been previously applied.

Although the three techniques presented here consider the same synoptic circulation data for the obtention of a wind predictor -a gridded daily SLP field-, this information is processed differently depending on the method. Thus, the first two techniques (Sect. 2.2) are based on the classification of the synoptic circulation through a modified version of the geostrophic approach defined by Jenkinson and Collison (1977) for their Circulation Weather Types (WT). In this way, the predictor is obtained independently from the predictand (wind observations). One of the main goals of these predictand-independent designed methods is the ability to develop stand-alone classifications of the atmospheric circulation into different classes according to their flow characteristics, hereafter called Wind Types (WdT). On the other hand, the last method (Sect. 2.3) classifies the atmospheric circulation according to the difference of pressure between a set of pairs of points that best cluster the wind speed observations. This means that the predictor clustering process is coupled to the predictand (wind observations), which has been the employed option in several previous papers (Salameh et al., 2009; Davy et al., 2010; García-Bustamante et al., 2012). Thus, with an algorithm that classifies the synoptic scale through continuously pivoting on the local wind observations without experiencing overfitting the uncertainty reduction should be minimized.

2.2 Geostrophic-Based Wind Type Classification for the Statistical Downscaling of Wind

This section describes the process followed to construct a statistical downscaling of wind through the development of a daily classification of the synoptic conditions of the surface flow as a predictor. First, a geostrophic-based automatized parametrization of the syn-

optic circulation is described. Then, the clustering optimization of the obtained indices is presented by means of two different soft-computing tools. Finally, the training process for the statistical downscaling of wind with the obtained wind type classifications as predictors is described.

2.2.1 SLP Clustering Optimization Problem and Wind Types Classification

The classification of the synoptic conditions is based on the clustering optimization of two geostrophic indices, F and Z. They can be computed directly from any gridded SLP field. The first one is directly proportional to the geostrophic wind speed, while the second refers to the absolute vorticity. F and Z are related to the spatial SLP field through the geostrophic wind G. Thus, F can be determined through its relationship with the zonal ($G_{u\lambda}$) and the meridian ($G_{v\lambda}$) components of G in the following terms:

$$G_{u\lambda} = -\frac{1}{\rho f}\frac{\partial p}{\partial y} \approx -\frac{1}{\rho f_\lambda R \Delta\lambda} WF \tag{2.1}$$

$$G_{v\lambda} = \frac{1}{\rho f}\frac{\partial p}{\partial x} \approx \frac{1}{\rho f_\lambda R\, cos(\lambda) \Delta\varphi} SF \tag{2.2}$$

where WF and SF stand for the zonal (westerly) and meridian (southerly) components of F, λ and φ are the latitude and the longitude of the measured grid point and $\Delta\lambda$ and $\Delta\varphi$, their differences. f stands for the Coriolis parameter, given by the expression $f_\lambda = 2\Omega sin(\lambda)$ (Ω represents the angular speed of the Earth), while R is the Earth radius and ρ is the density of the air. This shows how F can be determined by differences of pressure only.

In turn, the relationship between Z and G comes from the partial derivatives of the G components within the x and y directions in the Cartesian plane:

$$\frac{\partial G_{u\lambda}}{\partial y} \approx \frac{1}{R}\frac{\Delta G_u}{\Delta\lambda} = -\frac{1}{R^2 \rho f_\lambda (\Delta\varphi)^2} WZ \tag{2.3}$$

$$\frac{\partial G_{v\lambda}}{\partial x} \approx \frac{1}{2Rcos(\lambda_\lambda)}\frac{\Delta G_v}{\Delta\varphi} = \frac{1}{R^2 \rho f_\lambda (\Delta\varphi)} SZ \tag{2.4}$$

where WZ and and SZ represent the Z components produced respectively by the West-East and the the South-North pressure differences.

F can be considered as a proper approximation (Palutikof et al., 1992) for the observed real wind, V. The vorticity given by Z provides some complementary information on the wind conditions when the flow intensity given by F is weaker than a given threshold.

In the employed parametrization (similar to the developed by Jenkinson and Collison (1977)) F and Z are obtained for a given grid point J from the interpolation of certain pressure values adjacent to it when a gridded SLP field is provided. The actual arrangement of the employed grid points on a certain location J is displayed in Figure 2.1.

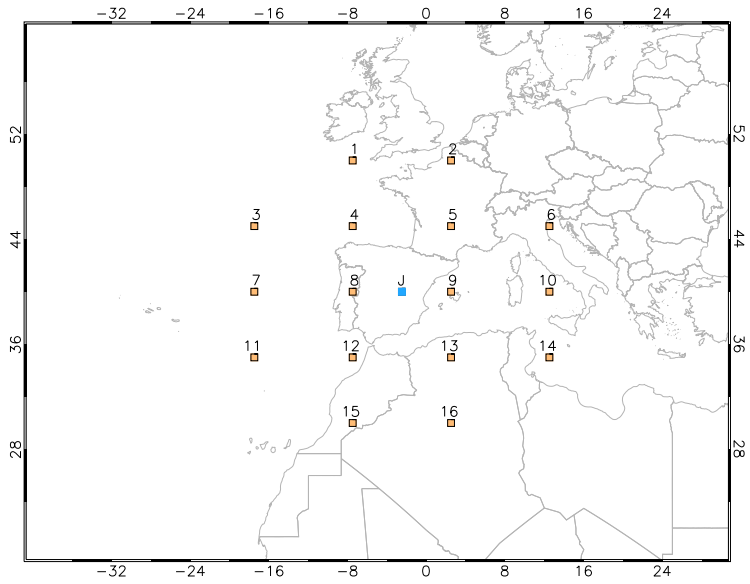

Figure 2.1: Example of the considered SLP grid structure employed in the definition of the F, Z and α indexes at the introduced algorithms for a generic point J.

The following expressions show how F and Z, calculated for J at a latitude $\lambda = 40^{o}$, can be computed exclusively using the 16 grid point pressure values shown in Figure 2.1:

$$WF = \frac{1}{2}(p_{12} + p_{13}) - \frac{1}{2}(p_4 + p_5) \tag{2.5}$$

$$SF = \frac{1}{cos(40^o)} \left[\frac{1}{4}(p_5 + 2p_9 + p_{13}) - \frac{1}{4}(p_4 + 2p_8 + p_{12}) \right] \tag{2.6}$$

so

$$F = \sqrt{(WF)^2 + (SF)^2} \tag{2.7}$$

In turn, for Z these relationships are:

$$WZ = \frac{sin(40^o)}{sin(35^o)} \left[\frac{1}{2}(p_{15} + p_{16}) - \frac{1}{2}(p_8 + p_9) \right] - \frac{sin(40^o)}{sin(35^o)} \left[\frac{1}{2}(p_8 + p_9) - \frac{1}{2}(p_1 + p_2) \right] \tag{2.8}$$

$$SZ = \frac{1}{2cos^2(\lambda_{40^o})} \left[\frac{1}{4}(p_6 + 2p_{10} + p_{14}) - \frac{1}{4}(p_5 + 2p_9 + p_{13}) - \right.$$
$$\left. - \frac{1}{4}(p_4 + 2p_8 + p_{12}) + \frac{1}{4}(p_3 + 2p_7 + p_{11}) \right] \tag{2.9}$$

Hence, as Z is a scalar value,

$$Z = WZ + SZ \tag{2.10}$$

Finally, a value for the angle of the F vector direction is obtained as follows:

$$\alpha = atan \frac{SF}{WF} \tag{2.11}$$

Within the WT classification method, which has been used as a benchmark for our results, a set of rules concerning F, Z and α are established in order to derive 26 Circulation Types: 8 of them are purely directional related to the 8 wind rose main directions, 2 are rotational classes defining pure cyclonic and anticyclonic patterns respectively, and 16 are considered hybrid classes, produced by the mixture between pure directional and each one of the rotational classes (Trigo and DaCamara, 2000; Spellman, 2000; Linderson, 2001).

Unlike WT, the classification criteria applied here are based in the dispersion minimization of the vectorial distances of the considered F values. Let F_i, $r = 1 \ldots, i$, be

a series of a daily (i) F vector derived for a given location J, along a given period of time τ_r. Under these conditions, the values F_i are clustered into a series of classes c_κ whose dispersion is, in average, minimized. This concept can be represented through the minimization of a *cost function* Φ, expressed in the following expression:

$$\Phi_J(F) = \frac{1}{\tau_r} \sum_{\kappa=1}^{n} \sum_{i \in c_\kappa} | F_{c_\kappa} - F_i | \tag{2.12}$$

where i stands for a generic day of the period considered, and F_{c_κ} stands for the average of F values within a class c_κ. This clustering problem is based in the same geometric concept of the Euclidean distances employed by Weber and Kaufmann (1995), Kaufmann and Whiteman (1999), Gerstengarbe et al. (1999) and others to develop their wind field spatial similarity methods within their clustering processes. However, in this work the optimization for the reduction of F dispersion is independent from observed wind data V, allowing to characterize wind conditions in those points or time series where no real observations are available.

A relevant issue on the classificator design lies on the number of clusters in which the F space will be divided. The classification scheme follows the criteria employed by the WT, Bergström (1996) or Sandström (1997), which divided the flow space into 8 directions. In order to compare our methodology with the WT approach, here the geostrophic flow characterization will result into 26 classes. This is achieved by assigning 3 flow intensities (i. e. 24 classes), as well as two additional classes to the central (those close to zero) flow values, as it will be detailed in next sections. Other widely extended SLP-based classifications of the synoptic circulation perform 29 (Hess and Brezowsky, 1969; Gerstengarbe et al., 1993; James, 2007) and 27 (Lamb, 1972) classes, making their characteristics comparable.

2.2.2 Optimization of the Clustering Process through different Soft-Computing Algorithms

The minimization of Φ is faced through two soft-computing algorithms, which are performed in a given period of time τ_r (r for "training") over a certain grid pint J which lies nearest to the meteorological tower. One of the approaches applies an Evolutionary Computing method (Yao et al., 1999), which allows a high computing performance when a large range of possibilities is considered. The other one consists on a greedy (Bednorz,

2008) algorithm. Since it is designed to obtain a deterministic solution for the problem, its results can be easily compared to those from WT. Through these algorithms, daily classifications according to the geostrophic wind conditions set by F can be obtained, employing in turn Z values to distinguish among weak flow conditions.

2.2.2.1 Geostrophic Clustering Optimization through an Evolutionary Algorithm, FE

The iterative procedure on clustering optimization within FE is performed by means of an Evolutionary Algorithm (EA). Given an optimization problem, this evolutionary algorithm starts from an initial set (population) of random solutions (individuals). These solutions are computed through a set of evolutionary operators (Eiben and Smith, 2010), which evolve and finally retain or dismiss them. This process is applied repeatedly, in a set of loops (generations). Since individuals are selected according to the quality of the solution they represent, a fitness operator is applied to each individual of the population. Hence, the individuals with the best values of fitness are more likely of being selected for replication and survival. The selected individuals are reproduced by means of crossover and mutation operators. While crossover exchanges some genetic material between two or more individuals, mutation changes parts of individuals with a small probability, preventing the algorithm to keep restrained in local minima. By applying this iterative procedure, the EA explores the whole space of possible solutions, without the need to compute all of them.

The eight angular borders $\mathcal{A} = [a_1, \cdots a_8,]$, $a_i \in [0, 360^o]$ must be established in order to define eight angular sectors. At the same time, every sector is split into 4 *slots* by 3 radial magnitudes ($r_{a,1} < r_{a,2} < r_{a,3}$), so three sectorial classes are defined by the 3 highest slots, while a near-to-zero area is defined by the lowest slot of every sector. This area for low F values determines in turn two classes depending on wether Z is positive (cyclonic calm) or negative (anticyclonic calm). This way a matrix $\mathcal{R}$ of size $[3, 8]$ is defined, with an independent condition that splits the weak flow in two additional classes. Eventually, all $\mathcal{A}$ and $\mathcal{R}$ elements are adjusted in each iteration so that a consistent solution is obtained. Figure 2.2 shows a result of this stratification on the $F = [WF, SF]$ space. There, each dot represents the daily F vector, in terms of the origin of the flux.

Unlike Bergström (1996); Sandström (1997) or WT methods, with a fixed classification criteria, here the advantages of adaptative tools as cluster analysis (Mengelkamp, 1999), k-means (Bogardi and Matyasovzky, 1996; Salameh et al., 2009; Najac et al., 2009) or

random forests (Faucher et al., 1999; Davy et al., 2010) are maintained. In this way, the cluster means and the frequencies of each cluster are determined by the distribution of the original data. This is achieved through the implementation of an EA.

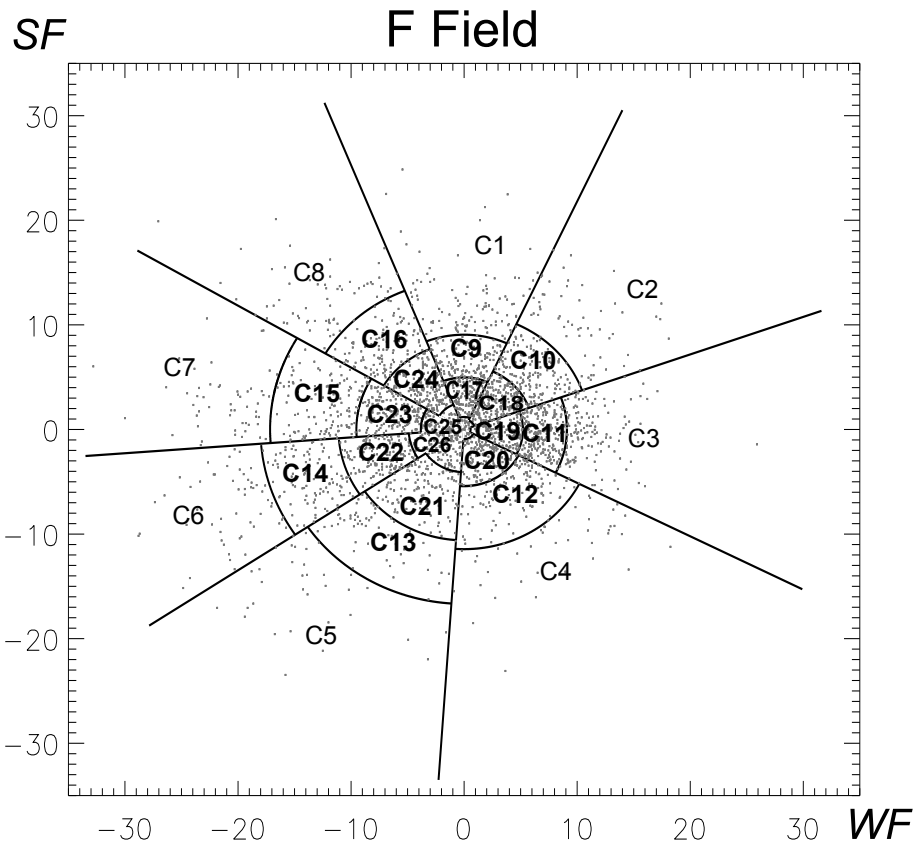

Figure 2.2: Example of the class-stratification structure performed over the F field by the FE algorithm, where angular and radial borders are defined.

Regarding the EA structure, at each generation a population of a given number N of individuals is developed, every one of them characterized by a certain $\mathcal{A}$ and $\mathcal{R}$ matrix. Throughout the generations, the efficiency on the minimization of the cost function (Eq. 2.12) is assessed. The final solution is then fixed when a stopping condition is reached. This happens when the result does not change more than a certain value compared to the previous generation, or the limit of generations has been reached. At the first generation, a completely random generation of the $\mathcal{A} \times \mathcal{R}$ matrix ($a_i \in [0, 360^{\text{o}}]$ and $r_i \in (0, \text{inf})$) is considered for every individual. From that, each one of the following generations can be described through a series of steps or operators:

1. **Repair**: In order to keep the increasing order at both angles and the radial borders in the encoding, elements are sorted from the smallest to the largest value.

2. **Fitness**: Each individual is associated with a value of fitness (i. e. performance) obtained from the cost function (Equation (2.12)).

3. **Selection**: The algorithm selects those individuals whose F average class-dispersion (or bias) is lower than the population's average dispersion. The rest of individuals will not survive for the next generation, and they will be replaced by new elements created through crossover and mutation of selected individuals.

4. **Crossover**: Two individuals from the selected population are taken at random to generate a new one. The configuration of the elements of the new individual, for both the vector $\mathcal{A}$ and matrix $\mathcal{R}$, is performed within a multi-point crossover procedure. In this crossover, each a_i an r_i from a parent has a probability of 0.5 of being transmitted to the new individual.

5. **Mutation**: The new individuals have a small probability of being mutated. For FE this probability is 0.05. The mutation consists of modifying 50% of the elements from the vector $\mathcal{A}$ and matrix $\mathcal{R}$ as:

$$\mathbf{e'}_i = \mathbf{e}_i + 0.1 \cdot \mathbf{N_1}(\mathbf{0}, \mathbf{1}) \tag{2.13}$$

   where e_i stands for a given element of $\mathcal{A}$ or $\mathcal{R}$ before mutation, e'_i is the element after mutation and $\mathbf{N_1}(\mathbf{0}, \mathbf{1})$ represents a normalized Gaussian distribution of random numbers with $\mu = 0$ and $\sigma = 1$. Note that we keep the constraints $a_i \in [0, 360^{\text{o}}]$ and $r_i \in (0, \text{inf})$.

For this work, this process has been performed over 3000 generations for a population of $N = 1000$ individuals.

As FE employs random sequences at some of its operators, it works as a non-deterministic application, and it obtains a different solution each time that it is launched. FE is launched 30 times for every point J. At each one of them, only the best outcomes on dispersion minimization are considered in the results.

2.2.2.2 Geostrophic Clustering Optimization Through a Greedy Algorithm (FG)

A second SC algorithm has been conceived to provide a deterministic solution of the cost function minimization. To achieve this, the F Field Optimization has been performed through the design of a Greedy (Bednorz, 2008) algorithm (called hereafter FG), an heuristic tool (Polya, 2008; Pearl, 1984) that assumes a local solution as a valid result at each step, with the hope that the final solution becomes a global optimum. FG introduces two differences in the conditions of the algorithm structure, compared to FE:

First, the angular borders which determine the classes at the F space have been kept fixed, so that the defined sectors coincide with the 8 cardinal wind rose directions (similarly to WT). With this, the physical interpretation of the synoptic dynamics behind the obtained classes is facilitated.

Second, FG tracks the dispersion minimization for every sector individually. By having fixed the angular borders, this procedure does not imply any loss of performance. This time the algorithm fixes the three classes at a given sector by exploring all the possibilities in the combination of $r_{a,1}$,$r_{a,2}$ and $r_{a,3}$ within that sector. First, the search is performed with a tracking accuracy of 1 m/s. Once the local optimum has been found, the resolution is increased to 0.1 m/s, and a new optimum local solution is then pursued, thus reflecting the greedy behavior of the algorithm.

Through these features, with a similar Z condition to describe the weak flow, the resulting algorithm structure coincides with that of FE, defined by a $(\mathcal{A}, \mathcal{R})$ matrix with dimensions 3×8.

2.2.3 Statistical Downscaling through Wind Type Classifications as a Predictor

A predictor considered for a statistical downscaling needs to be calibrated or *trained* with local observations. When this predictor consists on a series of clusters, the relationships with the surface wind are usually established by means of a series of transference functions within each one of the obtained clusters. In this way, several works (Palutikof et al., 1992; Najac et al., 2009; Goubanova et al., 2011; Monahan, 2012; van der Kamp et al., 2012) perform multivariate linear regressions for each one of the considered PCs as large-scale predictors, while Faucher et al. (1999) and Davy et al. (2010) employ random forests to build different possible predictors from those PCs. In turn, Salameh et al. (2009) devel-

oped a series of generalized additive models to design non-linear spline-based regression models, as others directly perform transference functions based in autoregressive models (Bogardi and Matyasovzky, 1996). In other occasions, correlation-based techniques are employed (García-Bustamante et al., 2012), where Canonical Cluster Analysis is performed between the obtained synoptic clusters and a set of wind series obtained from several masts. Other works directly apply a transference function to each one of the considered (non-clustered) predictors. This can be done again through multivariate linear regressions, either of pressure gradients and relative vorticity (Pryor et al., 2005c,b; Curry et al., 2012) or approximations of the geostrophic wind (Ekström, 2002; Achberger et al., 2002), either through cumulative distribution functions of wind speed (Michelangeli et al., 1995) or implicit transfer functions within ANN (Sailor et al., 2000).

In this work the classifications obtained by FE and FG are directly employed as wind predictors in the statistical downscaling of wind. Unlike a linear regression, here the transfer function between the designed predictors and the surface wind observations consists on assigning a specific wind value V_{c_κ} to each one of the obtained pressure-derived clusters c_κ. This prevents the downscaling process from overfitting episodes. The specific wind vector V_{c_κ} which will characterize each class in the plane of speeds $[u, v]$ is defined by the cartesian barycenter of the vector pointers representing the mean daily wind values corresponding to the set of days $i_r \in c_\kappa$ within a given period τ_r.

Once the training process is concluded, each one of the 26 wind types has been characterized according to the real observed wind in terms of speed and direction. Thus, that information can be employed for estimating local daily wind conditions in those periods where only SLP daily data are available. In Figure 2.3 a fluxogram of the process is depicted. Details are also provided in Kirchner-Bossi et al. (2013).

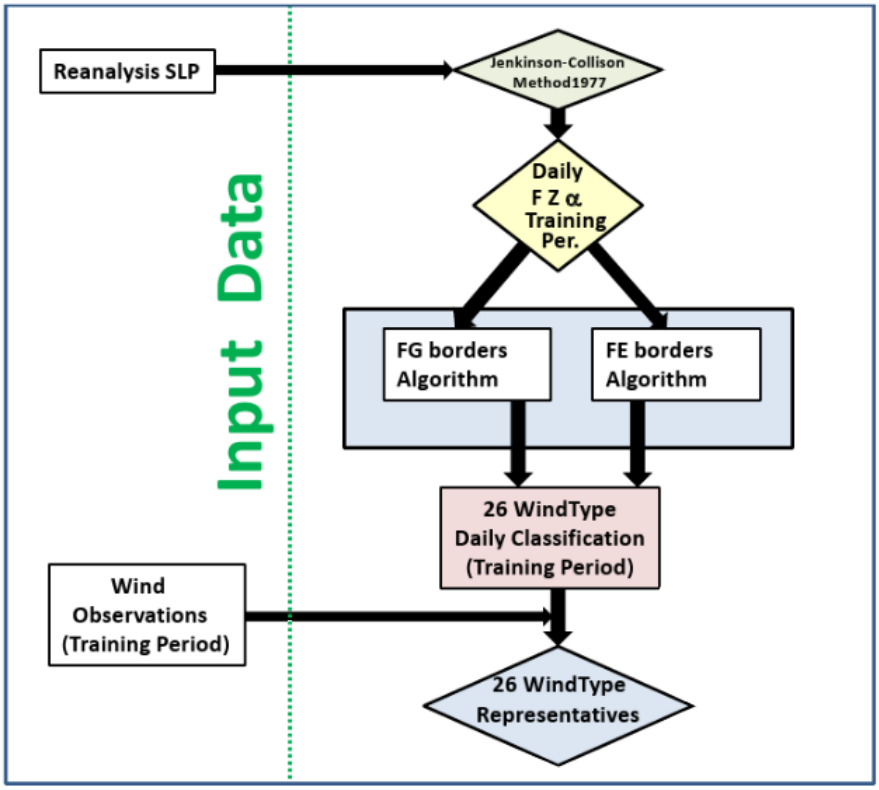

Figure 2.3: Fluxogram of the classification and training process (for either FE or FG).

2.3 Statistical Downscaling through an Evolutionary Algorithm of Pressure Differences (PD)

Before applying the designed FE and FG methods, a third technique for the statistical downscaling of wind will be introduced. Pressure-related gridded fields at synoptic scale contain a large amount of information which can be very useful for the establishment of a predictor for the daily surface wind (Monahan, 2012). However, the number of possibilities within such a high amount of data is a handicap when trying to find an appropriate one. In an effort to consider the whole number of gridded points within an extensive enough domain centered over a given location with a small computational cost, a third clustering algorithm has been designed. Unlike the first two models, which classify the atmospheric circulation on the basis of its own SLP field information (the F and Z indices), the Algorithm of Pressure Differences (PD, introduced in Carro-Calvo et al. (2011)) clusters

the synoptic dynamics according to the dispersion minimization of an external dataset, the wind observations. This means that in this case the obtention of a wind predictor from the SLP gridded field can not be considered independent from wind observations, and a wind type classification can not be obtained without them. In the counterpart, these conditions maximize the accuracy in the estimation of the daily wind.

2.3.1 Cost Function

The cost function of the clustering optimization has the same basis as in the geostrophic approaches, but this time its assessment is directly performed over the observed wind, instead of F or Z. In addition, now two different cost functions are considered, depending on the measure to be estimated, the vector or the speed of wind. Let $\mathbf{d}_i,\ t = 1 \ldots, T$, be a series of daily wind vectors measured in a given point, for a given period of time T. In turn, let $P_t,\ t = 1 \ldots, T$, be the daily synoptic-scale pressure fields corresponding to those $\mathbf{d}_i$ values. The pressure pattern extraction consists in forming a set of N clusters of daily pressure field (space of pressure P_t) in such a way that the dispersion of its associated wind values $\mathbf{d}_i$ in each cluster is, in average, minimized. With the intention of optimizing the accuracy of the wind speed estimations, here (unlike FE and FG) the wind speed minimization associated to a given location l is obtained through a cost function based on the unidimensional distances at the wind module space, described as:

$$\Phi_l(\mathbf{d}) = \frac{1}{T} \sum_{\kappa=1}^{N} \sum_{i \in c_\kappa} ||\mathbf{d}_{c_\kappa}| - |\mathbf{d}_i|| \tag{2.14}$$

where N represents the amount of clusters considered, $\mathbf{d}_i$ represents the wind at the set of days belonging to a given class c_κ, and $\mathbf{d}_{c_\kappa}$ stands for the mean value of them. This expression coincides with the definition of the *Mean Absolute Error* (MAE), which measures the average absolute bias between estimations and observations. This point is sometimes important, since wind farms are designed to be optimum for an average wind coming from a given direction.

2.3.2 Evolutionary Architecture of PD

Given the minimization problem at Eq. 2.14, it will be faced independently through a series of concatenated generations of a certain population within an evolutionary algorithm

until an optimum solution is found.

Before describing the evolutionary process, the problem encoding must be defined. In this way, two parameterizations on the pressure patterns classification will be evolved in parallel. One refers to the grid points of the pressure field to be considered, while the other one faces the definition of the N clusters. For the first one, we will consider a matrix of 13×14 surface pressure values (182 values), measured in a grid surrounding the case of study, as shown in 2.4. From the 182 grid points ($\mathcal{P}_i$), 4 pairs of points are selected (8 points $\mathcal{P}_i$), and the difference of pressure Dp is measured within each pair. This means that the 182-dimension initial problem is reduced into a four-dimensional space S_D.

Regarding the second parametrization, under these four Dp coordinates the S_D space is split into N clusters, each one of them parameterized by a specific four-coordinated centroid, initially distributed at random. Then, the centroids will characterize the dimensions of the clusters according to the principles of a Voronoi diagram (Voronoi, 1908). This way, an individual can be fully described through the expression:

$$[\mathcal{P}_1, \ldots, \mathcal{P}_8 \mid Dp_{11}, Dp_{12}, Dp_{13}, Dp_{14}, \ldots, Dp_{N1}, Dp_{N2}, Dp_{N3}, Dp_{N4}] \tag{2.15}$$

where the left part concerns the considered 8 grid points and the right part represents the 4 coordinates for the N defined clusters, so an overall amount of $4N$ values is obtained.

Figure 2.4 shows a sample individual at the S_P space, represented by four segments defined by its 8 $\mathcal{P}_i$ grid points. The extremes of each one of these segments will define the four pressure differences Dp. After the individual has been fully characterized by its 8 $\mathcal{P}$ grid points and its resulting N clusters, the S_D space is filled by an amount T of daily points, each one of them characterized by the four daily Dp values. Once this has been done and the set of days has been distributed throughout the N established clusters, the corresponding daily wind observations, similarly clustered, are introduced in the cost function, and the average wind dispersion per cluster is measured.

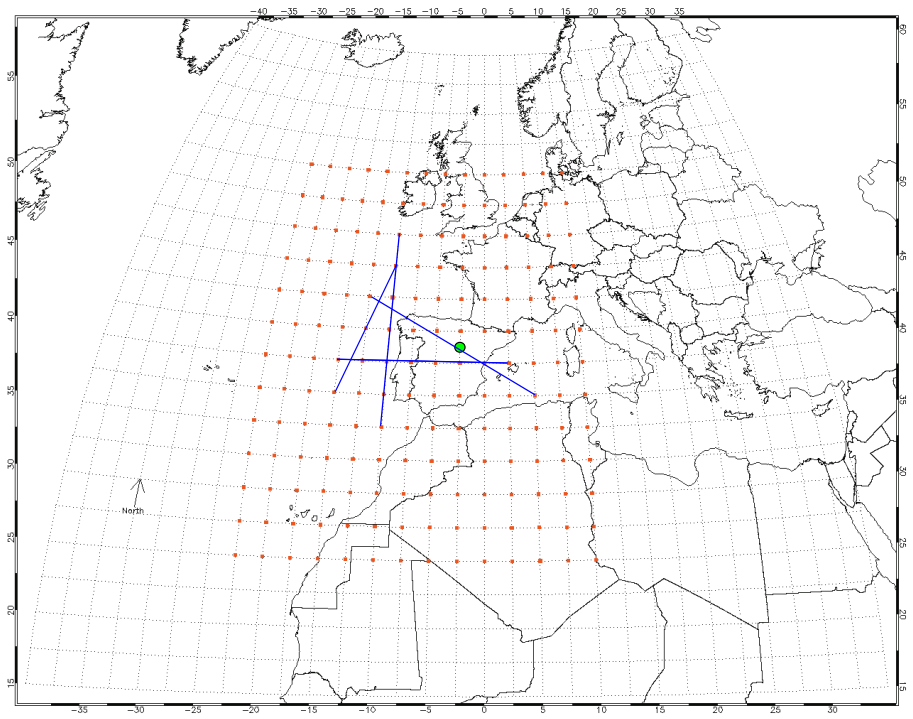

Figure 2.4: Grid employed as input for the PD algorithm (182 orange dots), around an hypothetical location of wind speed observations (green circle). The four blue segments represent a sample individual. The pressure values at the extremes of each one of the segments ($\mathcal{P}_i$s) are subtracted to produce the four differences of pressure Dp.

Regarding the evolutionary process, at the first generation the algorithm defines a population of L randomly generated individuals ($\mathcal{P}_i \in \mathbb{N}(1, 182)$, $Dp_i \in \mathbb{R}(-5, 5)$). From then until a stopping condition is found (like FE), each of the generations can be described through different operators, characterized through the following steps:

- *Crossover operator*

  After forming L random couples with two individuals within the population, a crossover operator is implemented for every couple over each one of the two parts of the individual separately, one for the selected grid points $\mathcal{P}_i$ and one for the Dp centroid coordinates. At each coordinate, each parameter of the individuals has a probability P=0.5 of being swapped with that from the other member of the couple. The crossover operator for the Dp centroids coordinates, however, is applied in two different modes, each of them with a probability P=0.5. One considers each one of

the Dp coordinates separately, so new centroids are generated by combining existing ones. Instead, the second mode considers the complete centroid (its four Dp coordinates) at a time, so the centroids are kept unaltered and are fully interchanged between the two individuals.

- *Mutation operator*

  After the new $2L$ individuals have been obtained, a mutation operator is applied within every coordinate of each individual. Mutation operator is applied with a very low probability ($P_m = 0.01$). The mutation procedure is divided into two different versions, depending on whether it is applied on the $\mathcal{P}_i$ grid points or to the Dp coordinates of the individual. Since $\mathcal{P}_i \in \mathbb{N}$, their mutation is carried out by means of an integer randomized values in the interval $[1, 182]$. In turn, the mutation of the Dp coordinates ($Dp \in \mathbb{R}$) is carried out by adding a value of uniform noise in the interval [-5,5].

- *Selection operator*

  Finally, a tournament selection is applied to the joint population formed from merging the initial and offspring populations. The result of the selection operator will be a single population, of size L, which will be the parents of the next generation of individuals. Once the complete joint population of parents and offspring is formed, the standard tournament selection, as described in Yao et al. (1999), has two main steps:

  1. Pairwise comparisons are conducted over the union of parents and offspring: p opponents are chosen uniformly at random from all the parents and offspring. For each comparison, one individual receives a "win" if its fitness is better than the opponent's, which in this work will mean that the value of the cost function is lower (optimizad clusterization).
  2. The L individuals out of the union of parents and offspring that have most "wins" are selected to be parents of the next generation.

  Through this procedure, the remaining L individuals act as the parents of the next generations, and the crossover and mutation operators are applied again in a loop fashion, until the maximum number of generations are reached or an optimal result is obtained.

In Figure 2.5 the fluxogram of the evolutionary algorithm is shown. Different experiments were performed in order to explore the performance obtained (in terms of wind estimation accuracy) depending on the number of clusters considered, and the highest yield was obtained for a value around N=25, although this amount is not absolute, and could sometimes vary due to the non-deterministic nature of the method. For comparison purposes with the other methods, the number of classes has been fixed in N=26.

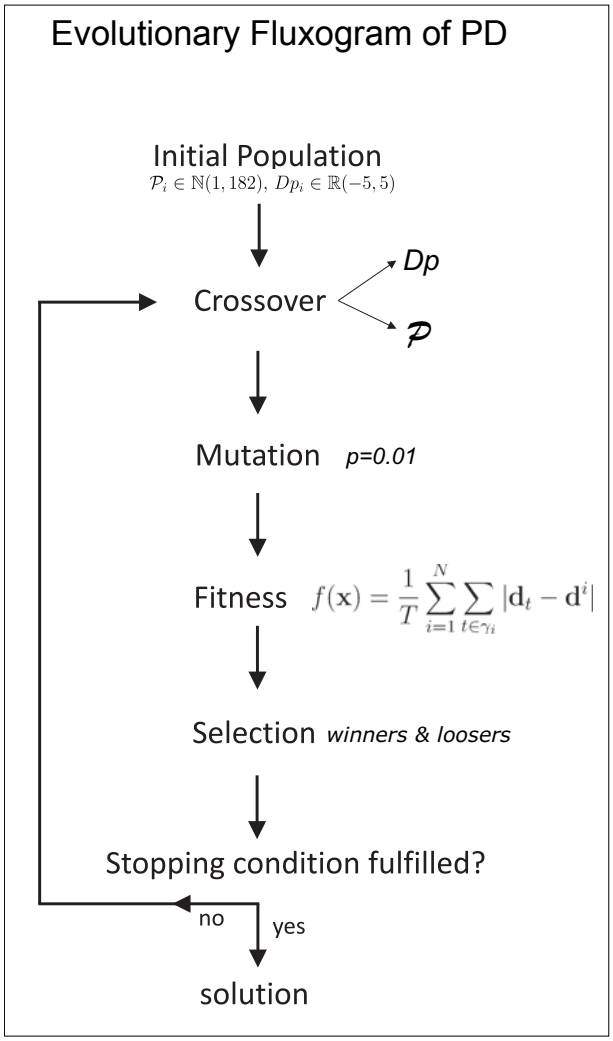

Figure 2.5: Steps at the evolutionary process within the PD algorithm.

Unlike FG and FE, this time the classes are not arranged according to intensity or direction. This is mainly because the non-deterministic nature of the algorithm will lead the evolution of the solution through a different path each launch. In addition,

2.3. Statistical Downscaling through an Evolutionary Algorithm of Pressure Differences (PD)

PD has a cost function in a separate space than the space of wind speed vectors $\mathbf{d}_t$, and the transference function is based in a non-evident relationship between wind and Dp. These issues will yet mask the relationship between classes. However, the fact the clustering method permanently pivotes on the observed wind speed should maximize the performance of the considered algorithm in order to optimize the wind estimations accuracy.

Chapter 3

Performance of the Downscaling methodology

3.1 Introduction

With any of the three introduced methods, the parametrization of the obtained circulation classes by means of a wind representative allows to estimate the local wind conditions in periods when only data on a large-scale predictor are available. However, it needs to be cross-validated, by assessing their performance in terms of different measures of goodness at the considered location. Once the methods are validated, the main features of the resulting Wind Types can be described, so a better understanding on the obtained classifications can be provided. This chapter describes the observed performance of the statistical downscaling methods developed in the previous chapter.

First of all, the data employed to construct a large-scale predictor (pressure fields) and a local wind predictand (wind observations) at daily scale are described (Sect. 3.2). Then (Sect. 3.3), different measures of goodness are computed, either at daily and monthly scales, and the obtained values are compared to other well known approaches. The third part of the chapter (Sect. 3.4) focusses on describing the main features of the Wind Types obtained through the FG and the FE methods.

3.2 Input Data

As mentioned in Chapter 1, the Iberian Peninsula circulation shows a high spatial and temporal variability, due to its complex terrain and the mixed influence of the Mediterranean and the Atlantic. Here six different sites spread over the region are considered. As in any statistical downscaling, two independent sets of data have been considered. Wind observations from the considered locations have been employed as wind predictands, while a series of SLP reanalysis fields have been considered to define different wind predictors, depending on the method.

3.2.1 Wind Observations

Wind speed and direction data with a 10-minute frequency were considered from meteorological masts located at six different wind power projects belonging to *Iberdrola Renovables Energa*. The six locations selected where those which provided the longest time ranges. Figure 3.1 shows their distribution throughout the Iberian Peninsula. The anemometers at the masts lie at a height between 26 and 41.5 m (see details on Table 3.1). T1 and T2 are located in central-east Spain, relatively close to each other. In turn, T3 is located just at the Gibraltar Strait Area, while T4 lies west, next to the Portuguese border, and T5 and T6 stand respectively north and east of the IP. This high geographical spread of the considered locations will ensure the assessment of the wind estimation methodology over sites with different orographic features and governed by distinct synoptic patterns, as will be shown later.

Wind observations range between 1995 and 2009 (T4), although most of them cover the period 1999-2009. Since the downscaling methodology is based on a daily resolution, the assimilation process of the wind observations consisted on computing the daily mean wind, from averaging the 10-min measures belonging to the same day throughout more than 10 years.

Data quality was controlled through the detection of missing data and the persistence of zero values. In the first case, a few periods without data were detected along the observations. None of the 6 towers showed more than 13% of the series with missing values (Table 3.1), this amount ranging from 2% (T1) to 12.9% (T5). In a second exercise, zero wind values representing abrupt transitions from high winds were dismissed.

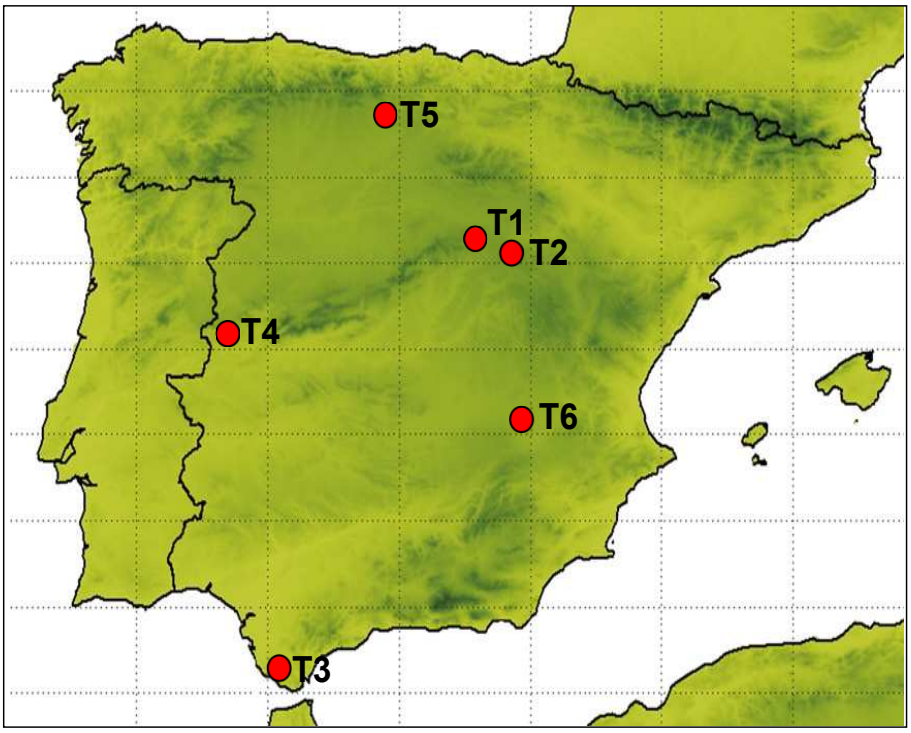

Figure 3.1: Geographical location of the six considered wind measurement towers in the PI.

This was done when continuous zero values were surrounded by previous and subsequent wind speeds higher than 5 m/s. Next, only those days which preserved more than half of their 144 daily values were included in the analysis, and their measures averaged. After these considerations, 6 series of daily mean wind with an average length of nine years and two months were obtained (they range between 7 (T3) and 12 years (T4)).

Table 3.1: Main features on the considered observational wind datasets.

Tower	Anemometer height (m)	Period (m/yr)	Missing data ratio (%)	Valid time range (years)
T1	41.5	09/1999 - 08/2009	5.6	9.5
T2	41.0	07/1999 - 10/2008	1.8	9.2
T3	40.0	08/2001 - 01/2009	4.9	7.2
T4	26.0	12/1995 - 10/2008	8.1	11.9
T5	41.0	10/1999 - 01/2009	12.9	8.2
T6	41.5	08/1999 - 08/2009	10.1	9.0

3.2.2 Reanalysis Data

SLP gridded data were retrieved from two different Reanalysis Projects in order to build the synoptic scale predictors, the National Center for Environmental Prediction/National Center for Atmospheric Research Reanalysis Project (NCEP/NCAR, 1996) and the the second version of the "Twentieth Century Reanalysis" Project (Compo et al. (2011), hereafter $20CR$). In order to ensure a common data usage for the six locations, a specific domain was considered over the IP. For the geostrophic approaches a domain conciding with the map in figure 2.1 waas considered, while those points represented in 2.4 where considered in PD.

NCEP/NCAR consists on a data-assimilation process set over the GFS NCEP model, spanning from 1948 through four daily values, at 0000, 0600, 1200 and 1800 UTC. The considered data include all sort of available observations, including those from planes, stations, radiosonde and satellite observations and others. After an exhaustive quality control, the selected data are implemented through the 4-D Var data-assimilation method into the T62L64 GFS model. This model version considers a grid with a 2.5×2.5 degree (a grid length circa 210 km) resolution, solving 28 sigma levels. It provides more than 80 variables, including geopotential height, temperature, relative humidity, U and V wind components, etc., diagnostic terms as radiative heating or convective heating, and accumulative variables like precipitation rate. Although its characteristics might be comparable to those of other reanalysis as the ERA-Interim (Dee et al., 2011), the NCEP/NCAR data set spans longer (1948 instead of 1979). Since averaging SLP fields from different hours within a day can induce some artificially smoothed situations, mainly over weather systems, the 1200 UTC SLP field has been retained as representative for the day. Daily SLP fields have been retained for the period 1948-2009. Within this range, data along 1995-2009 have been employed for the statistical downscaling process, and the rest is employed for reconstruction purposes. Many works have employed this reanalysis within the study of surface wind (Michelangeli et al., 1995; Pryor et al., 2005a; McVicar et al., 2008), including those building a surface wind predictor of wind from it (Ekström, 2002; García-Bustamante et al., 2012; Goubanova et al., 2011).

In order to span the wind estimations beyond the NCEP/NACR period (1948-2009), $20CR$ has been employed to complement the wind predictor time range. With a resolution of 2×2 degrees, the $20CR$ Reanalysis was computed by applying an Ensemble Kalman Filter to the background 'first guess' estimations from the GFS numerical weather prediction model. By slightly modifying its initial conditions, an ensemble of 56 forecast

outputs are obtained for every 6-hour forecast, and the Kalman Filter selects each time the ensemble member which best suits with observations. It has been run from 1871 to present, representing the only available 3-D reanalysis dataset reaching the 19th century with such resolution. By generating an ensemble of likely upper-air fields that are dynamically consistent with concurrent and previous pressure observation fields, it provides a useful estimation of the uncertainty in the analysis fields at each analysis time. It assimilates only surface pressure reports and observed monthly sea-surface temperature and sea-ice distributions as boundary conditions, which does not represent a drawback for an SLP-based wind predictor as the proposed here. This dataset results particularly robust for the early decades at middle latitudes, with plenty of SLP data at the time (i.e. Europe and North America), being less accurate outside these regions Compo et al. (2011); Wang et al. (2012). Although the publication of this reanalysis data set is relatively recent (2011), it has been used in several works on wind (Bronnimann et al., 2012; Candlish et al., 2012; Dafka et al., 2013).

3.3 Cross-Validation of the Developed Methods

The requirement for cross-validating a model arises from the need to provide a measure of its accuracy for eventual estimates performed with new predictor data. This accuracy is usually quantified in terms of uncertainty (or error). The cross-validation must be performed by considering predictor and predictand data sets which have not been previously employed during the training. First, the new predictor data are employed to obtain a certain estimate according to the already defined transference functions. Then, this estimate is compared with observations in order to measure its estimation goodness.

A cross-validation was performed over a period called τ_s (*test*), whose data were not considered at the training (τ_r) process. Within τ_s, daily SLP data are distributed into the 26 classes c_κ previously defined through the classificators (i.e. the transference functions), according to their F_s and Z_s (for FE and FG) or the daily values of pressure differences Dp_i (for PD) during the training (τ_r). Once the new data have been classified, they are characterized with their corresponding class-representatives V_{c_κ}. The validation can be determined by measuring the differences between V_{c_κ} and the corresponding observations at τ_s, V_{i_s}. The uncertainty is defined for each considered tower through the *Mean Absolute Error* (MAE), which can be applied in the form of the vectorial objective function (MAE_V) or the module objective function ($MAE_{|V|}$), which can be calculated through

the expressions:

$$MAE_{|V|}(T) = \frac{1}{\tau_s} \sum_{c_\kappa=1}^{26} \sum_{i_s \in c_\kappa} ||V_{c_\kappa}| - |V_{i_s}|| \tag{3.1}$$

$$MAE_{V}(T) = \frac{1}{\tau_s} \sum_{c_\kappa=1}^{26} \sum_{i_s \in c_\kappa} |V_{c_\kappa} - V_{i_s}| \tag{3.2}$$

where i_s stands for a generic day of the period τ_s. These expressions for the wind speed uncertainties represent a measure of goodness for our methods (FG, FE and PD) and reflect the degree of accuracy reached when daily wind reconstructions are performed. This means that it is possible to estimate the wind conditions with a measure of its uncertainty for each class c_κ at a certain period where the SLP field data are available. With respect to PD, its nature makes that the arrangement of the classes is not related between one launch and another. Thus, it is not possible to compare the objective function values obtained by the module and the vectorial approach class by class. However, it is possible to analyse the average values obtained in the whole test set.

Due to the non-linear empirical relationships that can be established through a soft-computing algorithm, these tools specially entail the risk on artificially generating favorable results during the training, i.e. the appearance of overfitting. In this way, apart of properly measure the goodness of the method, a cross-validation is crucial to measure the extent to which a certain regression or classification method can be still improved. Thus, if its estimation error is significantly higher than those performed at the training, then overfitting is occurring. This indicates that it is still possible to improve the performance of the model at the test period. This improvement, and the subsequent removal of overfitting, can be made by providing more generality to the model. This is usually done by reducing, for example, the number of variables or clusters acting as predictors. When limiting the model in order to increase its generality, the accuracy at τ_s will be lowered. However, with the reduction of overfitting the accuracy at τ_s will be increased. This way, this process of increasing the model generality can be carried out until a compromise between generality and overfitting reduction is reached. This will make that the obtained uncertainty at τ_s approaches its minimum attainable value. Next, results on the cross-validation are described, either within a daily and a monthly resolution.

3.3.1 Daily and Monthly Estimation Performance of the Introduced Methodology

The wind speed accuracy of the developed methods can be analyzed by comparing at the test period the V_{c_κ} obtained for every day with the observational wind series. The values for the uncertainty have been computed in the form of the Mean Absolute Error for both wind speed ($MAE_{|V|}$, 3.1) and wind vector (MAE_V, 3.2) variables. In addition, Pearson Correlation Coefficients between estimated and observed wind speed have been calculated for the comparison between real and reconstructed wind series. The average for the six towers has been computed by weighting the result at each tower with its corresponding series length, then computing the mean.

In Table 3.2 results on the MAEs obtained for every method are shown for every tower. Results have been compared with two independent wind speed estimations, used as reference approaches. Firstly, wind speed estimations were computed by using the WT (Circulation Weather Types, Jenkinson and Collison (1977)) classificator. Secondly, gridded wind values directly retrieved from NCEP/NCAR (Kalnay et al., 1996) reanalysis (hereafter called W-NCAR) were considered. In order to associate W-NCAR to the observations (as it has been done within the rest of the methods), a linear regression between the reanalysis and the observed wind was performed at the τ_r period, and then computed over W-NCAR at τ_s. Among the reanalysis grid points surrounding a tower, that with the highest Pearson correlation value between W-NCAR and observations was computed. Finally, the observed climatological mean wind speed has been considered for control purposes. Thus, the average of the entire wind speed series within τ_r has been computed. This constant value is then validated in τ_r. Since it can be conceived as a 1-class approach (thus considered as a *no-classification*, NC) the MAE obtained with this method can be considered as an upper bound of the implemented methods' accuracy.

Results show that, among the six considered approaches, FG, FE and PD present the smallest module and vectorial errors for all the considered locations. The developed methods show an average $MAE_{|V|}$ reduction between 19 (FG) and 21% (FE and PD) with respect to WT, and between 13 (FG) and 16% (FE and PD) compared with W-NCAR. The highest performances in average for the six towers in terms of $MAE_{|V|}$ and MAE_V reduction are provided by PD. FE presents very similar results for $MAE_{|V|}$ ((an average error just 0.7% higher), providing even a lower uncertainty than PD in T1 (-6%), T2 (-13%) and T6 (-5%), where PD is also outperformed by FG.

Table 3.2: Comparison of test results of obtained error by FG, FE and PD methods for the six daily wind speed data sets considered. Weighted averages within each entire classification (26 classes) for both $MAE_{|V|}$ and MAE_V are shown. For comparison purposes, values for WT and W-NCAR are also depicted. In addition, results on the no-classification (1 class, NC) are also detailed.

statistic	method	T1	T2	T3	T4	T5	T6	mean
$MAE_{\|V\|}$ (m/s)	NC	2.17	1.79	2.93	2.13	1.87	2.75	2.27
	W-NCAR	1.59	1.34	2.43	2.20	1.41	1.30	1.71
	WT	1.65	1.49	1.99	2.55	1.76	1.51	1.82
	FG	1.32	1.22	1.83	2.00	1.36	1.12	1.48
	FE	1.29	1.18	1.72	1.99	1.37	1.12	1.44
	PD	1.38	1.36	1.60	1.71	1.34	1.18	1.43
MAE_V (m/s)	NC	6.39	6.37	7.96	6.58	6.55	6.37	6.70
	W-NCAR	4.35	3.41	3.85	4.17	4.21	3.67	3.94
	WT	4.65	4.05	3.97	5.09	4.58	4.26	4.48
	FG	4.08	3.50	3.65	4.46	4.10	3.85	3.97
	FE	3.99	3.33	3.40	4.37	3.90	3.41	3.77
	PD	3.74	3.26	2.82	3.58	3.76	3.16	3.39

Regarding MAE_V, PD values are lower than FE for all towers (between -3% and -18%), and FE values are always slightly lower than FG (between -2% and -11%). In turn, FG, FE and PD show average error values 11, 16 and 24% lower than WT. With respect to W-NCAR, FE and PD show 4 and 14% lower uncertainties. FG shows an average error 1% higher than W-NCAR in 2 of the 6 towers.

The highest average value of the Pearson coefficient (Table 3.3) is again obtained by PD (0.76), showing FE and FG a similar average value of 0.73. These values are all clearly higher than W-NCAR (0.67) and WT (0.53). Finally, the tower which presents a higher improvement with PD is T3.

Table 3.3: Test results for the Pearson Correlations between observed wind speed and values obtained by FG, FE and PD methods at a daily resolution. Results for W-NCAR and WT are also depicted. All results showed statistically significant values at a confidence level of 1%.

method	T1	T2	T3	T4	T5	T6	mean
W-NCAR	0.67	0.70	0.55	0.63	0.77	0.70	0.67
WT	0.61	0.57	0.65	0.39	0.50	0.52	0.53
FG	0.77	0.73	0.71	0.68	0.72	0.77	0.73
FE	0.78	0.74	0.74	0.68	0.71	0.77	0.73
PD	0.80	0.68	0.78	0.81	0.74	0.79	0.76

A monthly timescale validation has been carried out by averaging wind speed values according to the calendar months (tables 3.4 and 3.5). As it was expected, the uncertainty decreased in most cases with respect to the daily resolution, due to the lower variability when decreasing the time resolution. The ratio between $MAE_{|V|}$ monthly values at the different methods is similar to those at daily scale at all towers excepting T5, where PD is overcome by the other methods except WT. The same happens for r, where PD again shows an anomalous performance at T5. These decreases with respect to the daily scale (specially within r results) can be attributable to the possible existence of certain non-stationarity at monthly scale. This would provoke that monthly averaged values at those cases are coupled to an artificially increased monthly variability, explaining thus their performance loss.

The best performances are respectively provided by FG (0,54 m/s) and PD (0.82). $MAE_{|V|}$ overall results show that the new methods outperform W-NCAR between 24 (FE and PD) and 28% (FG), and WT between 30 (FE and PD) and 33% (FG), while the introduced methods correlate substantially higher than WT and W-NCAR. The maximum performance is reached at T6, with $r = 0.92$ (PD) and $MAE_{|V|}$=0.34 m/s (FG), with very similar results for FE. For all this, there is no clear preference for FG, FE and PD, since their performance are similar and depend on the considered statistic, tower and averaging period. Similar relationships are obtained for MAE_V, again W-NCAR performing particularly well at T5, with an accuracy only lower than PD.

Table 3.4: Monthly $MAE_{|V|}$ and MAE_V through FG, FE and PD methods at the test period for the six wind speed data sets considered. For comparing purposes, values for WT and W-NCAR are also depicted. Mean values are weighted according to the length of the considered time series at each tower. All obtained r values are statistically significant ($p<0.01$).

statistic	method	T1	T2	T3	T4	T5	T6	mean
$MAE_{\|V\|}$ (m/s)	W-NCAR	0.66	0.54	1.16	0.81	0.87	0.54	0.75
	WT	0.72	0.63	0.73	0.91	1.19	0.72	0.81
	FG	0.44	0.32	0.79	0.62	0.89	0.34	0.54
	FE	0.45	0.40	0.70	0.62	0.95	0.40	0.57
	PD	0.51	0.45	0.58	0.52	1.02	0.43	0.57
MAE_V (m/s)	W-NCAR	1.96	1.58	1.47	1.87	2.01	1.40	1.71
	WT	1.87	1.49	1.65	1.74	2.05	2.04	1.81
	FG	1.44	1.09	1.44	1.63	1.52	1.58	1.45
	FE	1.51	1.14	1.44	1.71	1.59	1.55	1.49
	PD	1.50	1.25	0.93	1.21	1.53	1.37	1.30

Finally, r values obtained with the introduced methods show less spatial variability than those derived from the reference approaches (table 3.5, which suggests that the new methods performance is less sensitive to local factors. Summarizing, in overall terms all three methods perform better than the reference approaches at both daily and monthly scales.

Table 3.5: Same as 3.3 but for the monthly scale.

method	T1	T2	T3	T4	T5	T6	mean
W-NCAR	0.67	0.80	0.41	0.61	0.79	0.75	0.67
WT	0.74	0.71	0.72	0.20	0.57	0.32	0.55
FG	0.84	0.89	0.71	0.73	0.75	0.91	0.81
FE	0.84	0.84	0.74	0.75	0.75	0.90	0.80
PD	0.84	0.85	0.83	0.91	0.52	0.92	0.82

3.3.2 Examples on Daily Wind Reconstruction

In this section some examples on the behavior of the developed methodology are shown. Additionally, wind roses are also computed for the test period, so that the angular configuration accuracy can be qualitatively assessed. It must be emphasized that this last

exercise represents a useful tool when assessing the wind resource at the wind power industry, and will be employed later in the wind reconstruction.

Figure 3.2 shows two estimated wind speed series at daily and monthly scales at T6 with all the considered approaches compared with observations. Figure 3.2 (a) shows 100 days within τ_s. It confirms the higher reconstruction performances using FG and FE, with an overall better track than WT with respect to the observed wind. Figure 3.2b shows the monthly wind speed test reconstruction applied throughout the whole considered instrumental time range, where FG and PD performed best.

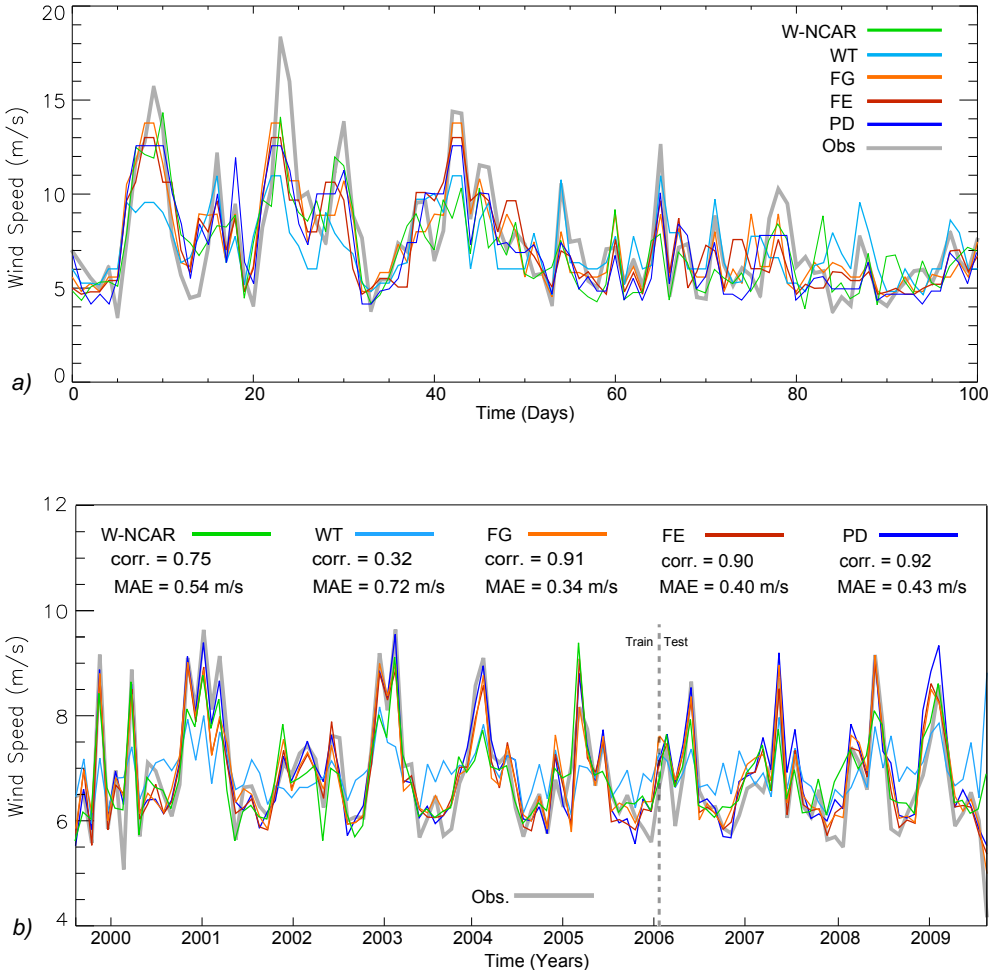

Figure 3.2: 100 days sample (a) of the daily wind speed reconstruction at the τ_s period performed by WT, FG and FE methods for T6 compared to the observed wind speed signal (from February 9 to May 20, 2006) and monthly (b) observed and estimated wind speed for the period 1999-2009 by the same methods and tower.

Wind roses have been computed through FG and FE for the six considered locations, for τ_s. The data cross-validation procedure to build the wind rose consisted on a similar procedure as for wind speed. First of all the wind rose for each obtained class (frequency and speed) is calculated in the training τ_s period. Then, the weighted average of these roses is computed depending on the frequency of each class within the considered time range. The obtained values by PD are displayed in Figure 3.3, where wind rose reconstructions are compared with the real data in τ_s period. Figure 3.4 shows the same approach for FG and FE. Results show a consistent similarity between the observed and reconstructed series both in wind intensity (color) and wind frequency (shape) for each sector of the wind rose for the three methods.

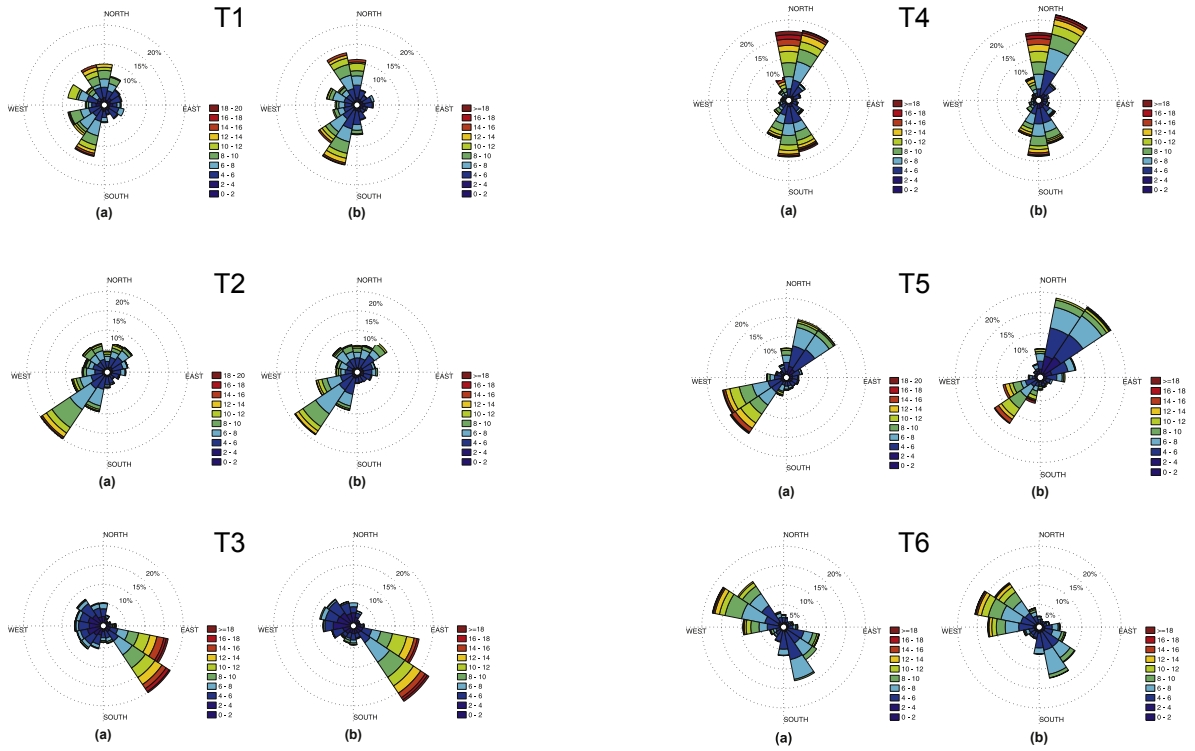

Figure 3.3: Comparison of the real and reconstructed wind speed using the proposed PD, for Towers T1-T6, respectively.

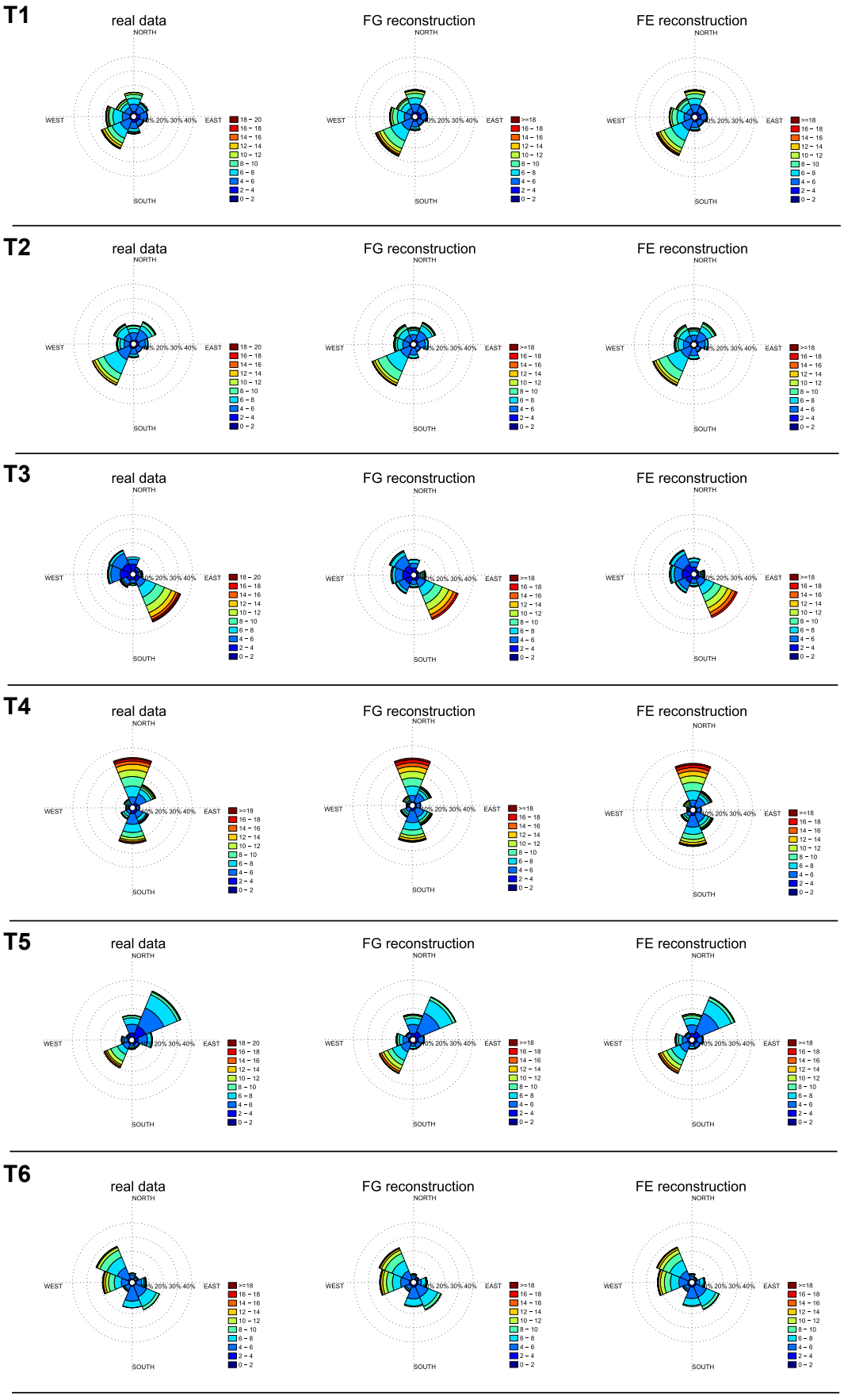

Figure 3.4: Wind roses computed from observations (left), FG (centre) and FE (right) for the 6 towers at the test period.

3.3.3 Homogeneity within the Intra-Class SLP Daily Fields

Each obtained class or Wind Type within the developed methods can be associated to an SLP synoptic situation that represents its circulation features. In this work each one of these pressure maps has been computed by averaging the SLP field from those days belonging to a given class. The class-representativeness of these SLP synoptic composites associated to each one of the computed classes has been assessed as an additional measure of goodness. To do this, the SLP dispersion (in terms of Standard Deviation σ) within all the daily pressure fields forming the composite has been measured, for every grid point of the considered domain (that employed for FG and FE, see figure 2.1), for every given class. Finally, the average dispersion per grid point and class has been computed, and weighted depending on the number of elements per class, for each one of the six considered locations. Results have been compared to the average SLP field dispersion provided by the WT method.

The average grid point SLP σ_M's have been compared to those (σ_{RC}) obtained by a random classification (RC). In this way, σ_{RC} has been measured over a set of 1000 artificially generated classifications whose elements within each class have been randomly chosen. To perform a proper comparison, the amount of elements per class has been set to coincide as in those classifications obtained through WT, FG, FE and PD. The relative ratio between the average SLP dispersion (σ_M) of every method with respect to the average σ of RC (σ_{RC}) has been computed, represented as:

$$\sigma_{RR} = \frac{\sigma_{RC} - \sigma_M}{\sigma_{RC}} \tag{3.3}$$

Figure 3.4 shows σ_{RR} for each tower and method. Results show that the overall dispersion of FE and FG per grid point is lower than that at WT. FG shows the best performance in all locations, while FE is better than WT in all sites except T3, where a slight difference (0.3%) occurs. A Chi-square Homogeneity test has been performed to assess the statistical significance of the SLP dispersion reduction. Thus, the amount of grid points (%) with a significant ($p < 0.05$) σ_{RR} was 48, 50 and 56 for WT, FG and FE respectively. These results show a robustness in FG and FE similar or even higher than WT to explain the synoptic circulation, specially when a large spatial scale is considered. Finally, σ_{RR} shown by PD is clearly higher than that of the other methods, although still overcoming WT at several towers. This is consistent with the nature of this method, not designed to cluster the synoptic circulation, but to estimate the local wind speed

with the minimum uncertainty. These results suggest that the obtained classes are more unambiguously and better related with V than WT, which implies, apart from a better characterization of wind, the possibility of developing important climatic applications, being their analysis beyond the scope of this thesis.

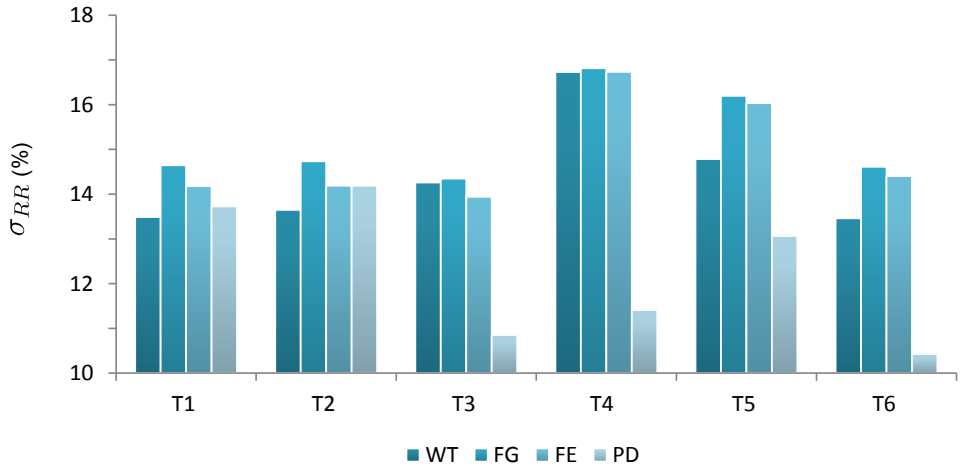

Figure 3.4: Obtained values for σ_{RR}. Histograms show results as spatial averages for each classification versus the mean SLP dispersion considering a random classification with same number of elements per class.

3.3.4 Comparison with Previous Work

The obtained results have been compared with a wide amount of statistical downscalings of wind listed in the introduction. However, a strict comparison could not be done with the design made in this work, because in some cases a numerical measure of goodness is not provided (Palutikof et al., 1992; Michelangeli et al., 1995; Bogardi and Matyasovzky, 1996), results are shown within an intra-daily (6h) timescale (Faucher et al., 1999; Davy et al., 2010), a cross-validation is not performed (or it is not specified wether it has been done) (Salameh et al., 2009; Mengelkamp, 1999; Pryor et al., 2005c,b), or the orographic conditions are clearly different. In this sense, our performance is observed to be in the same order or higher than works with oceanic (van der Kamp et al., 2012; Curry et al., 2012; Monahan, 2012) or manifestly simple (Bergström, 1996; Sandström, 1997; Ekström, 2002) orography.

Results from Najac et al. (2009) can be more easily compared. There, a maximum (minimum) RMSE of 2.8 m/s (1.1 m/s) and a maximum (minimum) r=0.75 (0.43) are re-

spectively obtained for daily wind speed series from 78 meteorological stations in France. Daily values obtained by our methodology (considering FG, FE and PD) show similar RMSE ranges (maximum/minimum of 2.6/1.4 m/s) and higher performances for r (max/min 0.81/0.68). Finally, García-Bustamante et al. (2012) show an average monthly wind speed linear correlation lower (0.7) than ours (0.80-0.82), although with a different orography.

Regarding those methods employing soft-computing tools, Sailor et al. (2000) obtained for three inland locations a climatological year constructed from 30 years of data (27 for training), with a daily linear correlation with observations of 0.85. The same exercise for every one of our considered methods and towers showed values between 0.82, with less than 7 training years in average. Results also appear to be in the order of Bouzgou and Benoudjit (2011), Mohandes et al. (2004), and Torres et al. (2005). Finally, it has been observed that PD outperforms for all locations (Carro-Calvo et al., 2012) other evolutionary configurations (IFEP in Yao et al. (1999)).

Since the orographic and the circulation regimes at the different studies can differ substantially, these comparisons must be considered carefully. For instance, it must be taken into account that the selected towers for this thesis are located at places with a wind power relevance, with high mean wind speeds (i.e. higher uncertainties), instead of the average wind conditions at a regular meteorological station. For the same reason, the considered sites are placed generally at the top of hills, entailing notable complex orographic conditions, not comparable with others with much simpler conditions or just over sea. Finally, most studies consider locations with a high prevalence of the mid-latitude westerlies or the Atlantic storm track (Rogers, 1997). Surface winds at these and other regions with a clear large-scale circulation can result more easily parameterized than those with a more weakly defined geostrophic flow.

3.4 Obtained Wind Types

The high quality of the validation performance allows analyzing the resulting Wind Types, represented by their corresponding class-composites at the SLP field, with some depth (annual frequency, wind features, etc). This is a useful exercise to understand the meteorology and climatology associated to the wind speed variability. It must be noticed, however, that the PD method was not designed to cluster the SLP field, but to directly estimate local wind. This makes that PD does not show an homogeneous magnitude and direc-

tion difference between classes. Moreover, this method produces a class-arrangement that varies from one launch to another, with no possible connection between the class identification obtained at different launches. However, in general terms PD groups the pressure patterns in classes which are easily recognizable as typical meteorological situations, what is, intuitively, a good point of reference.

3.4.1 FG and FE Wind Types: Geostrophic vs local Wind

As explained in chapter 2, F represents a valid approach for the real wind, V. In this section this relationship between the geostrophic index F and the observed wind V is analyzed for FE (the FG relationship between F and V showed a similar behavior). The daily values for F and V vectors can be represented by the vector pointers describing the origin of the flow. Among the six studied locations, only T1 and T3 will be described here. The reason for this is that T3 showed a very different wind regime compared to the rest of the towers, with a highly prevalent easterly flow. Since the other towers showed a similar relationship between F and V, T1 has been selected as representative for them. Figures 3.5 and 3.6 show F and V values obtained for T1 and T3 respectively.

In the figures, each box depicts Fs and their corresponding Vs values for each resulting class. As it can be observed, at both T1 and T3 V is characterized by a weaker intensity (a 24% smaller in average) and a counterclockwise tending direction (31º in average) than F due to drag effects from geostrophic to the local approach. However, V is closer to F at T1 than at T3 (specially within E, SE and S sectors), which corresponds with a minor disturbance of the synoptic circulation by lower scale forcings. This can be observed in the difference between grey (F) and black (V) dots in figure 3.5. In turn, T3 shows very clear differences between synoptic (F) and local (V) flow, the latter showing a particularly high prevalence of classes representing flow from the east. This evidences the influence of the particular funnel shape of the Gibraltar Strait at smaller-than-synoptic (mesoscalar or local) scales (Dorman et al., 1995; Bormans et al., 1986) which, unlike most of the IP, originates the prevalence of strong *levanters* (easterlies). In turn, T1 can be considered as representative for the overall Central Iberian Plateau flow regime, where westerlies are more prevalent (Jerez et al., 2013).

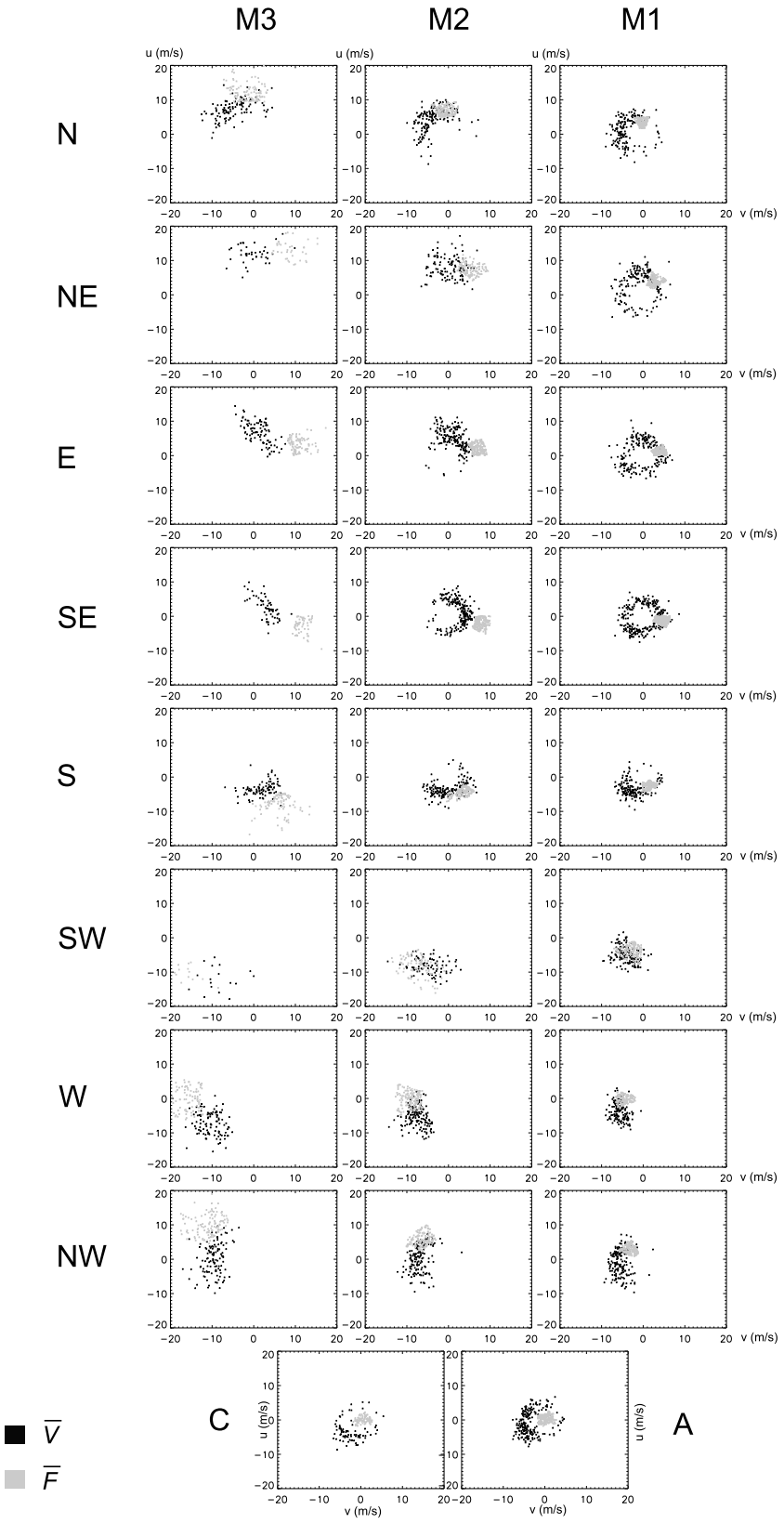

Figure 3.5: F and V values for each class obtained by the FE algorithm for tower T1.

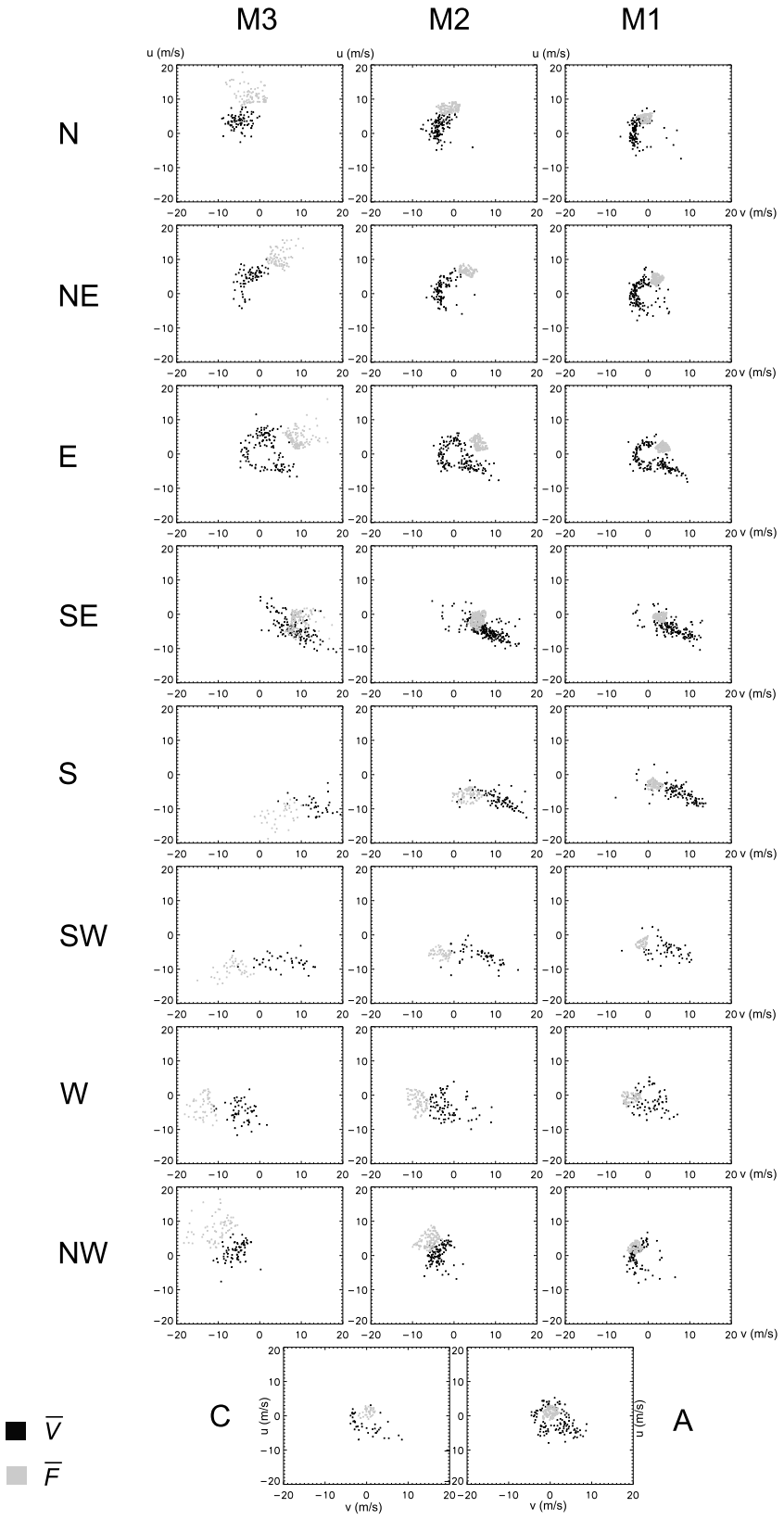

Figure 3.6: F and V values for each class obtained by the FE algorithm for tower T3.

Although differences between F and V can be larger in T3 than in T1 due to these distortions of the synoptic flow, results evidence, independently of the location, an overall capability in both methods to retain similar wind features within elements belonging to the same class. The clustering capability within V can was measured by calculating a dispersion measure as the *radius* of a class, i.e. the average Euclidean distance to the class representative V_{c_κ} in the speed space. This magnitude can be derived through Equation 2.12 (by employing this time V instead of F). The average vectorial radius per class is clearly smaller for FG and FE methods, compared to WT, for the six considered towers. In average WT showed for all towers a vectorial *radius* of dispersion 12 and a 17% bigger than that observed in FG and FE respectively. FE showed the best performance in T3, with an improvement of a 17% compared to WT, while FG improved 9%. This can be explained through the fact that FE is the only method which considers the angular borders as a variable to be adjusted, allowing the predominant directions to be fixed more accurately. When speaking of wind speed module, performances are similar. The obtained average distances for all towers was 1.70 m/s for WT, with values of 1.40 and 1.38 m/s for FG and FE respectively, a clustering capability more than 23% higher compared to WT.

3.4.2 Synoptic Dynamics and Monthly Variability associated to the obtained Wind Types

The main characteristics of the obtained classes through FG and FE are now analyzed. These include the synoptic circulation SLP class-composites over Iberia as well as their monthly frequencies and wind speeds.

Results are shown for T4, since it holds the largest data set compared to the other towers (over 50% more data). Results on the pressure composite, monthly frequency and monthly wind speed for every class are shown for FG and FE in figures 3.8 and 3.9 respectively, and compared to those obtained through WT (Fig. 3.7). Dark regions in maps denote statistically significant ($p<0.05$) values for σ_{RR} (see Sect. 3.3.3), i.e. the SLP pressure dispersion is significantly small against a class randomly computed.

The class arrangement described in sect. 2.2 for FG and FE is represented by SLP class-composites consisting on a sectorial (N,NE,E,SE,S,SW,W and NW) and radial (M1, M2, M3) stratification, complemented by a cyclonic (C) and an anticyclonic (A) calm Wind Types. In general terms FE and FG circulation patterns show, compared to WT, more clearly differentiated SLP fields among the three magnitudes of a given sector (most clear at S, W or NW sectors). This can be explained attending to the differences in the criteria of the methods to arrange the classes in a certain sector. In WT patterns within a given sector are established through hybridization criteria between the pure (P) directional type and the two cyclonic types (C and A) to form HC and HA, without a special need to differentiate the flow intensity between HC and HA. In contrast, FG and FE define a specific criteria for each one of the three patterns within a sector (M1, M2 and M3). Within this module stratification it can be observed that the circulation patterns associated to two classes from the same sector can keep bigger differences between them, compared to WT. These differences of classification criteria also illustrate a more balanced distribution of days frequency per class for the new methods, implying a better characterization of the area: while WT concentrates 50% of the days in only 5 classes, FG and FE need 8 and 10 classes to reach that ratio. This happens because WT anticyclones (A) hold nearly a quarter of the considered period, while FG and FE anticyclones, which refer only to calm circulations, include respectively 15 and 5% of the days.

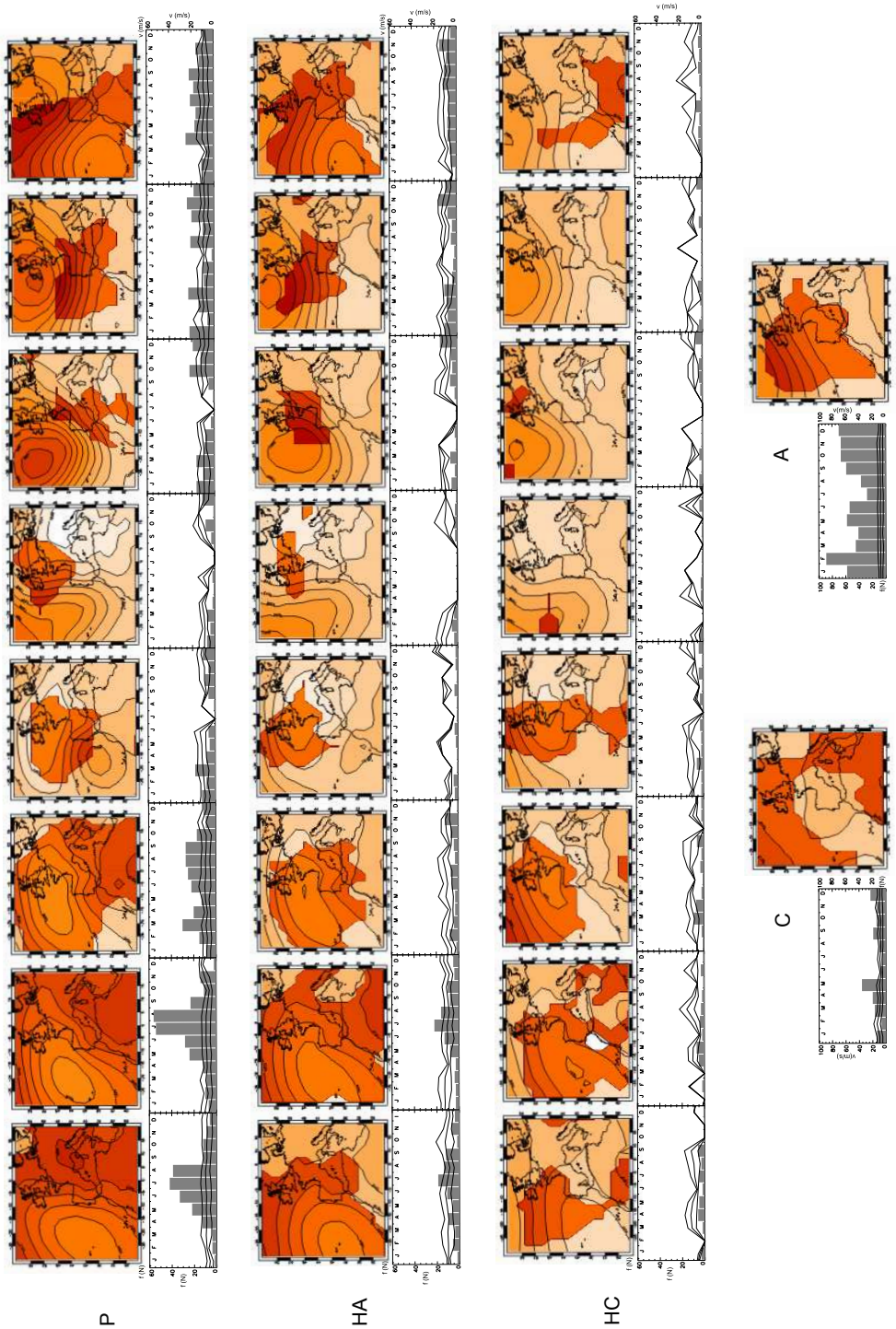

Figure 3.7: Synoptic (map) and intra-annual (histogram) characterization for the twenty-six WT patterns obtained for the period 1999-2009 at T4. Plots are horizontally distributed according to the pure directional (P), hybrid directional-cyclonic (HC) and hybrid directional-anticyclonic (HA) WT types, as the vertical alignment stands for the 8 sector wind rose. The two last plots stand for the pure cyclonic (C) and anticyclonic (A) types respectively. Darker areas at the synoptic composites represent regions with a similar SLP field ($p < 0.05$) among the elements conforming the pattern. Histograms show the class monthly frequency, where lines stand for the class-monthly wind speed (thick) and its σ deviation (thin ones).

A strong seasonality is observed at different classes (histogram bars under the composites in figures 3.7, 3.8 and 3.9). It must be noticed that in many cases there is a strong annual cycle, with a single maximum (normally in winter or in summer), or two maxima, normally in spring and autumn. By focusing on those classes with a high wind speed (M3 flows), some circulation pattern typology can be observed. Thus, a marked winter cycle for strong flows at sectors NW, W, SW, S and SE can be noticed. By considering their related high wind speed and their spatial distribution, the first three classes are consistent with the features of strong storms in winter (Trigo, 2006; Goyette, 2011), in agreement with their associated high wind speed (Allen et al., 2010). On the other hand, SE represents a clear example of cold anticyclone at Central Europe (Lopez-Bustins et al., 2008), while S results in a hybrid that feeds from both situations.

The maximum frequency at FG and FE calm flow patterns (C and A) is observed in summer for all locations. Wind Type C is characterized by a weak, cyclonic horizontal flow with a seasonal cycle perking in August and September. These features coincide with different prevailing meteorological situations over the Iberian Peninsula in late summer, related with convective precipitation precipitation regimes (Mosmann et al., 2004) associated to the Iberian *thermal low* depression (Hoinka and Castro, 2003). On the other hand, the anticyclonic calm (A) Wind Type presents the highest occurrence frequency in summer, reflecting the prevailing situation where the influence of the Azores anticyclone extends further into the Iberian Peninsula (weak flow). This circulation type is different from the WT pure anticyclonic class, which is most prominent in winter. This difference evidences that, although both WT or FG/FE A types describe an anticyclonic situation, FG and FE retain those days with an explicit calm flow (with wind speeds 20% lower than A at WT).

Other characterizations can also be found for M1 and M2 magnitudes. In this regard, FE and FG W-M2 class (and to some extent SW-M2) presents a seasonal distribution with maxima in spring and autumn, related to westerly moderate perturbations in the wettest months (April and November). In turn, in N, NE and E sectors a very strong summer-centered seasonality is observed for magnitudes M2 and M1, which provide medium-low winds. As it can be seen in the corresponding figures (FG and FE), this set of classes is strongly influenced by the Azores high, characterized by a weak flow in summer over Iberia (Davis et al., 1997). Finally, when the center of this anticyclone remains stationary west of the IP over Azores, N, NE and E flows over Iberia become stronger (M3), and the seasonal cycle is no longer seen.

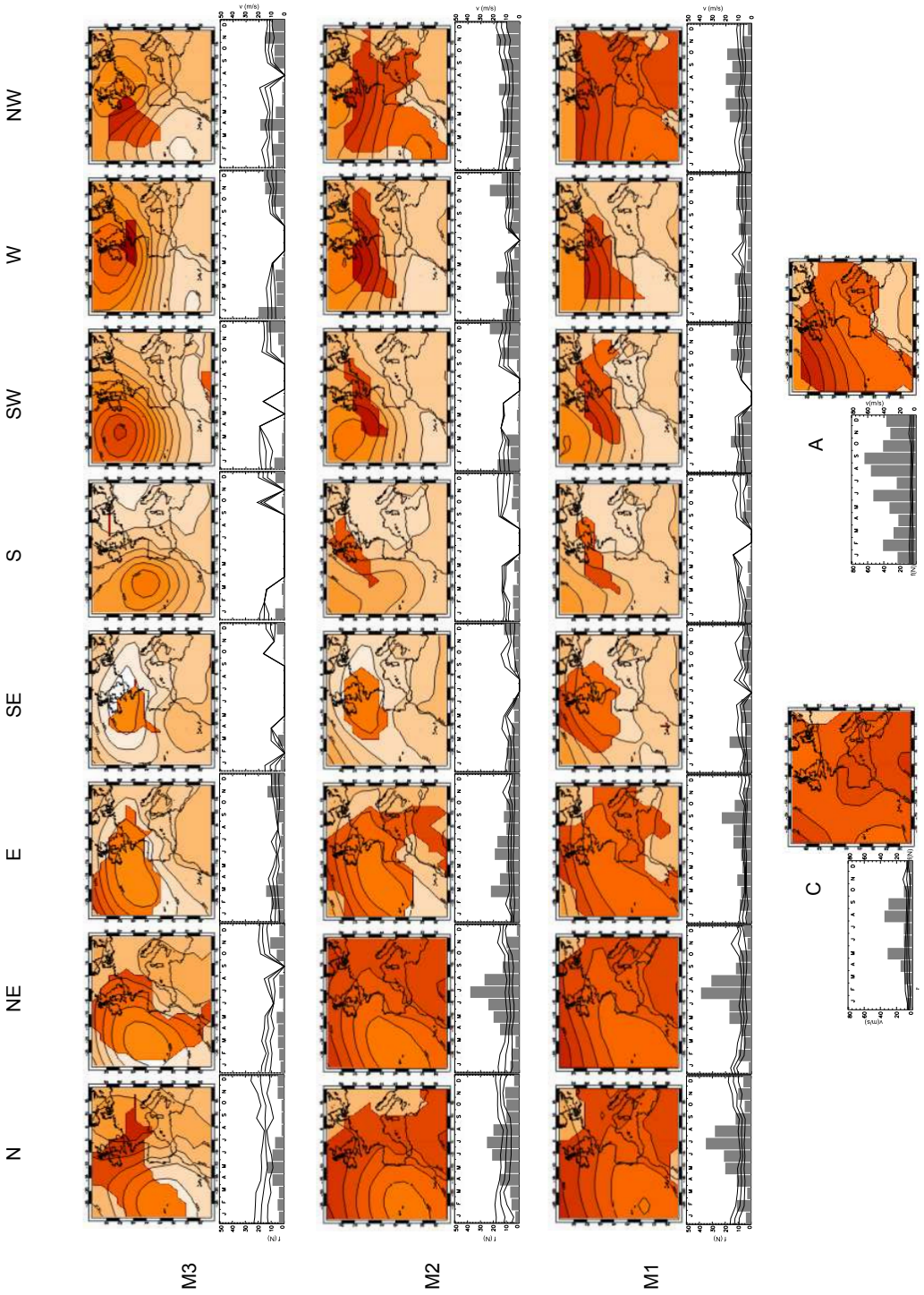

Figure 3.8: Same as Fig. 3.7 for FG. This time the horizontal distribution corresponds to the three sector intensities of F, M3 (biggest), M2 and M1 (smallest) set in the FG algorithm criteria.

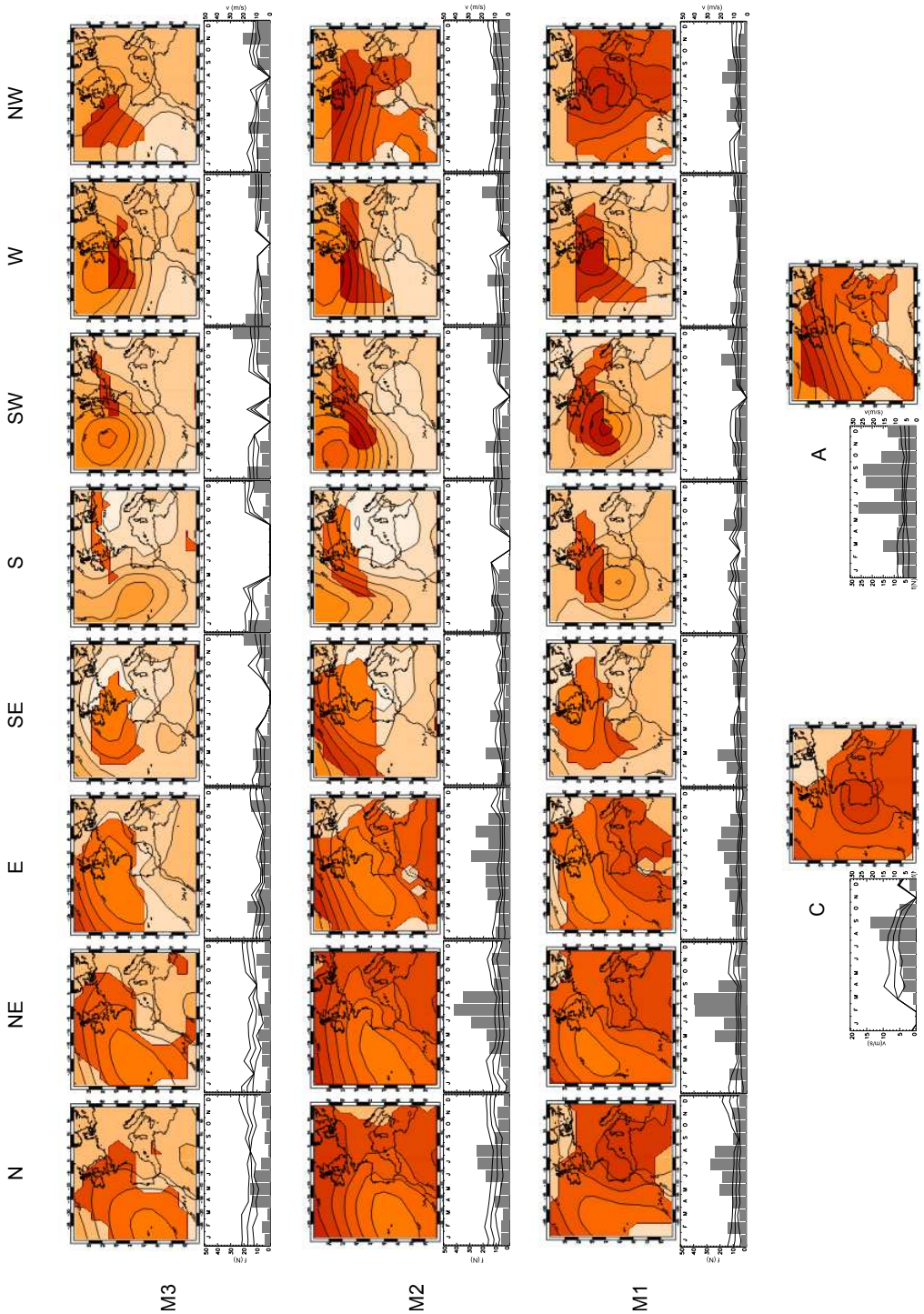

Figure 3.9: Same as Fig. 3.8 for FE. Note that since FE sectors are not predefined, the 8 flow orientation labels might not completely coincide with an 8 sector wind rose.

Chapter 4

Long-Term Wind Variability

4.1 Introduction

In this chapter the wind conditions at the six towers are analyzed from a long-term (i.e. multidecadal) perspective, similarly as in Kirchner-Bossi et al. (2013) is done for T1. To do that, the daily wind is reconstructed for the period 1871-2009 by applying the introduced statistical downscaling methodology to a 139 year daily SLP data set ranging between 1871 and 2009. The NCEP/NCAR is employed to obtain a daily wind classification for the period 1948-2009, while the 20CR ensemble mean is employed to span the daily pressure field from 1947 until 1871. This construction from two different data inputs has been also performed by Jones et al. (2012), arguing that NCEP/NCAR data offers a slightly higher reliability for the period when it is available. Apart from analyzing the long-term variability of the derived wind speed time series (Sect. 4.2), a variability analysis is performed in terms of the annual wind speed Probability Density Functions (PDF, Sect. 4.3) and the annual wind roses (Sect. 4.4). In the last section of the chapter (Sect. 4.5) the reconstructed annual frequency of the obtained Wind Types is considered to analyze their incidence onto the reconstructed mean annual wind speed variability along the 139 years computed.

FG is used in this chapter because its performance is similar to the other methods, but it has a deterministic, fixed N, NE, E, SE, S, SW, W and NW flow directions and magnitudes, so it appears as the most robust option for a Wind Type incidence analysis.

4.2 Wind Speed Long-Term Variability

4.2.1 Obtention of the Wind Speed Series

The daily F, Z and α values derived from the SLP daily field set 1871-2009 (20CR+NCEP/NCAR) were used as predictors in the FG classification conditions that had been fixed at the training period, to compute a daily 26-Wind Type classification for each one of the 6 considered locations. Then, every day was identified with its class-representative V_{c_κ} and an uncertainty value, which corresponds to the class-MAE values computed at the test period. The computed daily wind speed reconstructions can be observed at figures 4.1 and 4.2). Annual reconstructed and observed values are shown jointly with an 11-year running mean. The wind speed range between percentiles 2.5 and 97.5 is also shown for the 11-year and the 1-year running means. To calculate these magnitudes we assumed that the daily uncertainty computed in the test process follows a normal distribution, i.e. the MAE equals the percentile 50 of that distribution, so the standard deviation can be expressed as σ=MAE/0.68. In turn, the population variance (1 or 11 years, $\bar{\sigma}_y^2$) can be derived from its relationship with the sample variance (1 day, ${\sigma_i}^2$), which is described as follows:

$$\bar{\sigma}_y = \frac{\sqrt{\sum_{i=1} \sigma_i^2}}{n-1} \tag{4.1}$$

Finally, the 2.5 and 97.5 percentiles are estimated by taking into account its relationship with $\bar{\sigma}$ in a gaussian distribution, which corresponds to $\pm$ 2.25$\bar{\sigma}$. The higher uncertainty observed at T4 might be associated to the existence of a certain degree of non-stationarity in the relationships between scales throughout the whole training period. In addition, this tower shows an abrupt wind speed peak at the 20CR-NCEP/NCAR transition period (centered in 1948), (Fig. 4.2, T4). This kind of variability around 1940-1960 is much smoother in T1, T2, T5 and T6, without any apparent inconsistency, so the origin of such prominent change at T4 remains unclear.

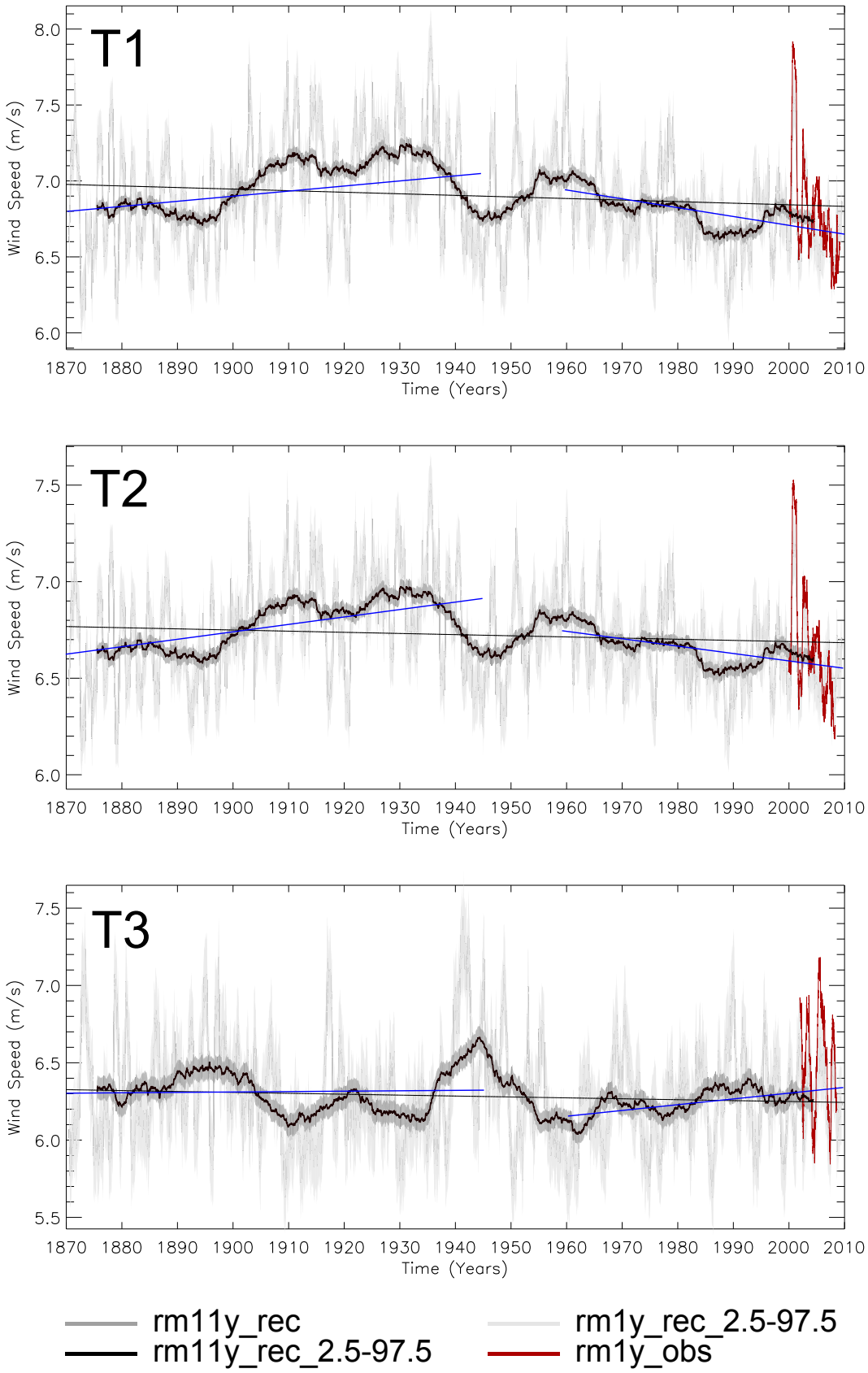

Figure 4.1: 1871-2009 wind speed reconstruction for T1, T2 and T3, represented in terms of a 11-yr running mean (black). Light and dark gray represent the 1-yr and 11-yr running mean of the uncertainty band comprised between the 2.5% and the 97.5% percentiles. Grey line represents the linear regression for the whole period, as blue ones respond to the 1871-1945 and 1960-2009 linear trends. The observations 1-yr moving average (red) are also depicted.

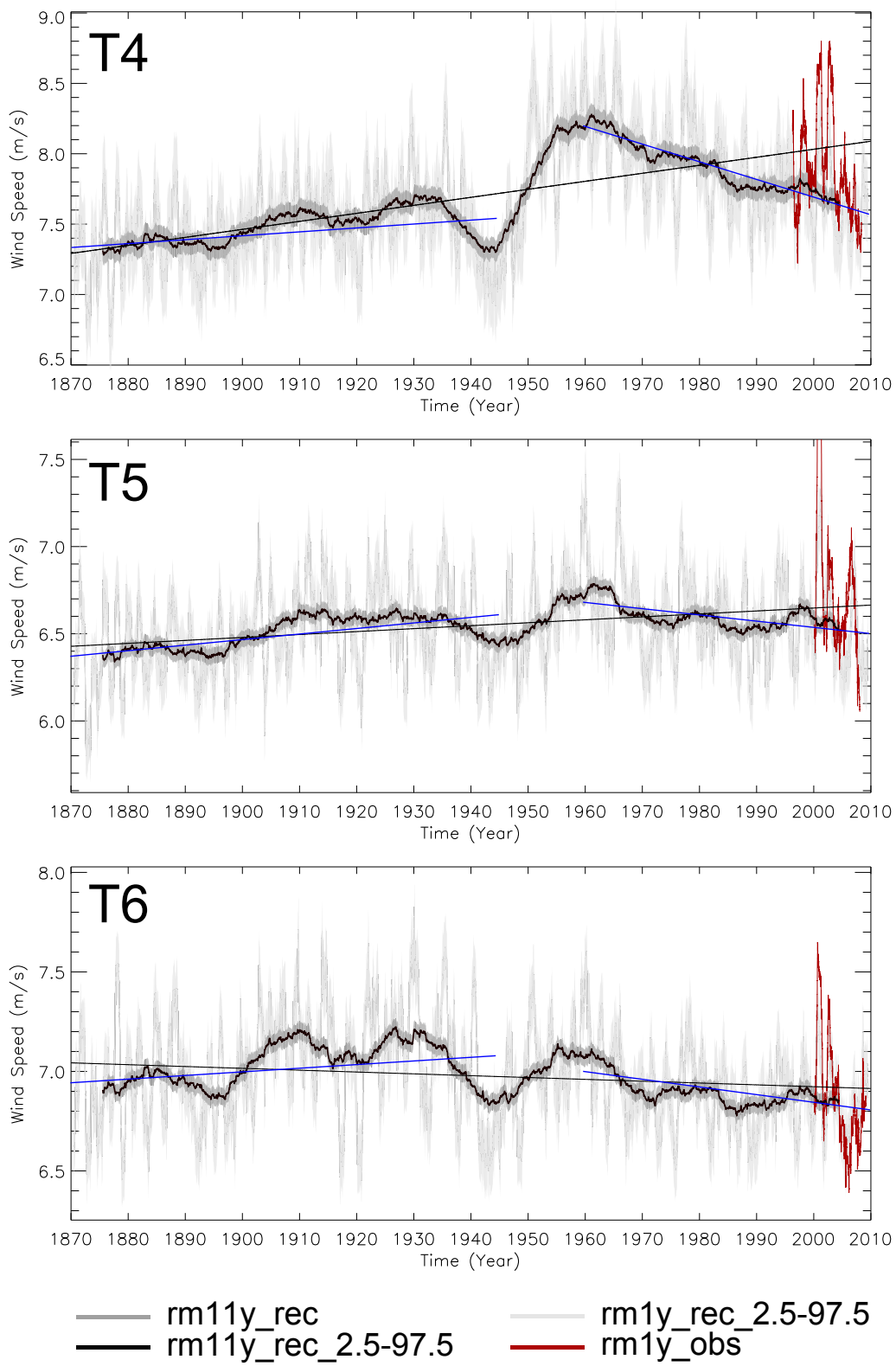

Figure 4.2: Same as 4.1 for towers T4, T5 and T6.

4.2.2 Climatological Wind Speed Variability

All towers except T3 show a statistically significant ($p<0.05$) positive correlation among them (Table 4.1). In turn, T3 shows a negative significant correlation with respect to the rest. These results confirm the fact that, as suggested in the previous chapter, all towers except T3 share some similar flow features. This behavior allows to consider all towers except T3 as representative for the Central Iberian plateau (T_p) wind regimes, yet retaining their distant location among each other (except T1 and T2, which are only few kilometers away). Among them, T1, T2, T5 and T6 are most closely correlated. In turn, T4 shows, although still significant, an overall lower correlation with the rest of the towers. As shown before, T4 shows an anomalous behavior around 1950 and it has the highest annual variability in terms of standard deviation and annual wind speed range. Results on these and others annual variability statistics towers are summarized in Table 4.2.

The negative (significant) correlation that T3 holds with the other towers evidences that a non negligible part of the variability of the circulation surrounding the Gibraltar Strait is being ruled by the same mechanism governing flow at the Central Iberian plateau, which in turn exhibits an opposite incidence at both regions.

The wind variability at T3 shows, similarly to T4, a range between the maximum (7.26m/s in 1941) and the minimum (5.66 m/s in 1986) annual wind speed 48-18% higher than the rest of the towers. Regarding the mean annual wind speed, T3 shows the lowest wind speed value. However, its high wind speed range makes that, at an annual scale T3 shows a maximum annual speed value higher than some other locations (T1 and T5).

Table 4.1: Inter-annual Pearson correlation coefficients between the wind speed reconstruction in terms of annual means at the six considered locations for the period 1871-2009. All values are statistically significant ($p<0.05$).

tower	T1	T2	T3	T4	T5	T6
T1	1.00	0.99	-0.42	0.43	0.74	0.87
T2	-	1.00	-0.39	0.45	0.74	0.88
T3	-	-	1.00	-0.35	-0.40	-0.35
T4	-	-	-	1.00	0.67	0.44
T5	-	-	-	-	1.00	0.59
T6	-	-	-	-	-	1.00

Table 4.2: Statistics of the annual mean of the wind speed reconstructed series at the six considered locations.

statistic	T1	T2	T3	T4	T5	T6
mean (m/s)	6.90	6.72	6.29	7.66	6.54	6.97
SD (m/s)	0.30	0.23	0.30	0.36	0.22	0.23
max. (m/s)	7.72	7.36	7.26	8.60	7.24	7.66
min. (m/s)	6.36	6.28	5.66	6.97	6.02	6.51
range (m/s)	1.36	1.08	1.60	1.63	1.22	1.15

4.2.3 Multidecadal Wind Speed Variability

In order to detect possible low frequency wind variations, a spectral analysis was performed on the annual wind speed series through a Fourier transform. Results revealed a statistically significant ($p < 0.025$) variability cycle within the 23 year frequency band for all towers except for T5, while a periodicity of 46 and 69 years is also detected for T3 and T4 respectively. Figure 4.3 shows the power of the annual spectra at the six towers.

A linear trend analysis on the whole period showed slight but statistically significant negative linear trend of around 0.1 m/s every 100 years for T1, T2, T5 and T6, while T4 showed an increase of 0.45 ms^{-1}/100yr, although the particular issues described above related to this tower must be taken into account. However, during the 1871-1945 and 1960-2009 periods, linear changes are higher. In this way, through the first interval all towers except T3 showed a significant (positive) trend (between 0.20 and 0.45 ms^{-1}/100yr). During 1960-2009, however, this behavior becomes opposite, and all towers but T3 show significant negative trends, between -0.37 (T6) and -1.27 (T4) ms^{-1}/100yr. In turn, T3 presented a significant positive trend of 0.36 ms^{-1}/100yr.

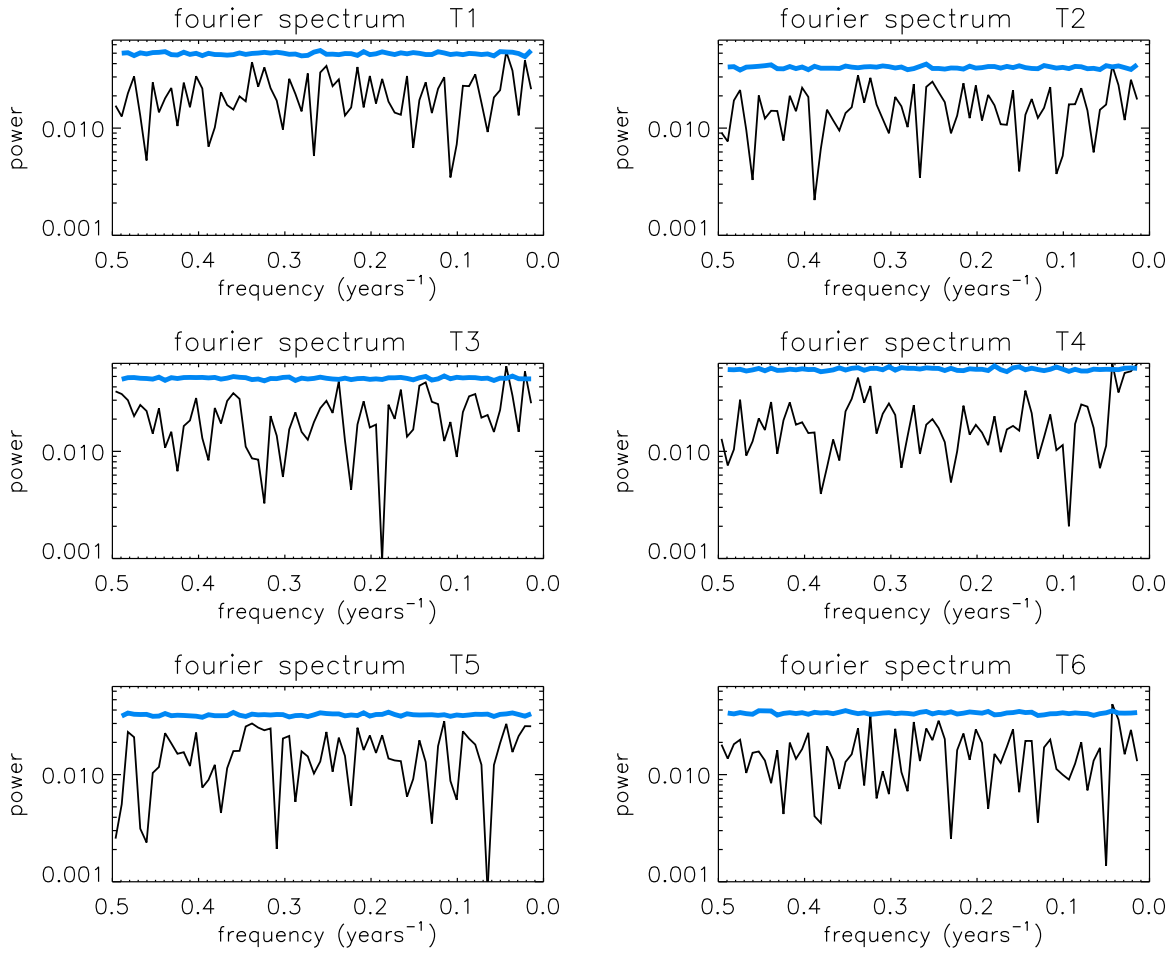

Figure 4.3: Power density for the variability spectrum of the six considered towers (black). Statistically significant ($p<0.025$) values are those which exceed the blue label.

All these results allow to interpret the wind farm performance reached in the last years into a multidecadal context. In this way, at T_p towers, the last period of observations appears located within a minimum phase of the multidecadal variability, which could induce an underestimation of the long-term wind speed resource at the Iberian plateau region. On the other hand, T3 appears to be in a period where wind speed is tending to increase.

4.3 Wind Speed PDF variability

In order to analyze the differences among different years in a decade and different decades along the whole period (1871-2009), the daily wind speed frequency distribution has been computed for different periods at decadal and annual scales with a 1 m/s resolution.

To provide an even higher robustness to the reconstructed PDFs, the obtained Wind Type (WdT) classification is employed here by considering, for a given class, the complete wind speed distribution computed at the observations period (T_{obs}), instead of just its V_{c_κ} class-representative. The reconstruction process for every annual wind speed PDF can be summarized as follows:

1. For every day of T_{obs} we computed the daily wind speed PDF from the 10-min. observations.
2. The corresponding daily WdTs of T_{obs} were identified.
3. The daily PDFs of those days with the same WdT were merged into a WdT-PDF. This was performed for each one of the 26 classes, so 26 WdT-PDFs were obtained.
4. Every reconstructed day was characterized with a certain WdT-PDF, according to the WdT classification.
5. For each year, a PDF was computed from its daily PDFs.

Since all T_p locations showed a comparable variability, in this section T1 has been selected as representative for the analysis of the long-term wind speed variability for the central Iberia plateau, while T3 represents the existing wind regime at the Gibraltar area.

Figure 4.4 shows the T1 wind speed frequency distribution at those decades with the highest (1926-1935) and the lowest (1979-1988) average wind speeds. Additionally, the decade with highest interannual variability (1936-1945) is also shown. For these three decades, the years with the highest and the lowest wind averages have been selected to analyze possible significant changes among them. A Chi-squared test of homogeneity showed that 1926-1935 and 1979-1988 PDF distributions are significantly different ($p<0.05$). This inhomogeneity is mostly attributable to the 3-6 m/s interval, where the 1979-1988 period shows a frequency 10% higher than the 1926-1935 decade (156 days), and to the 9-14m/s interval, where the last shows a frequency 20% higher than the first (127 days). For the period 1936-1945, although the PDF differences between low (1945)

and high annual winds (1936) were not statistically significant, some remarkable changes can be seen. 1945 shows a 30% higher frequency than 1936 in the first quartile (3-4 m/s). In turn, 1936 shows frequencies 35% higher than 1945 in the third quartile (8-9 m/s).

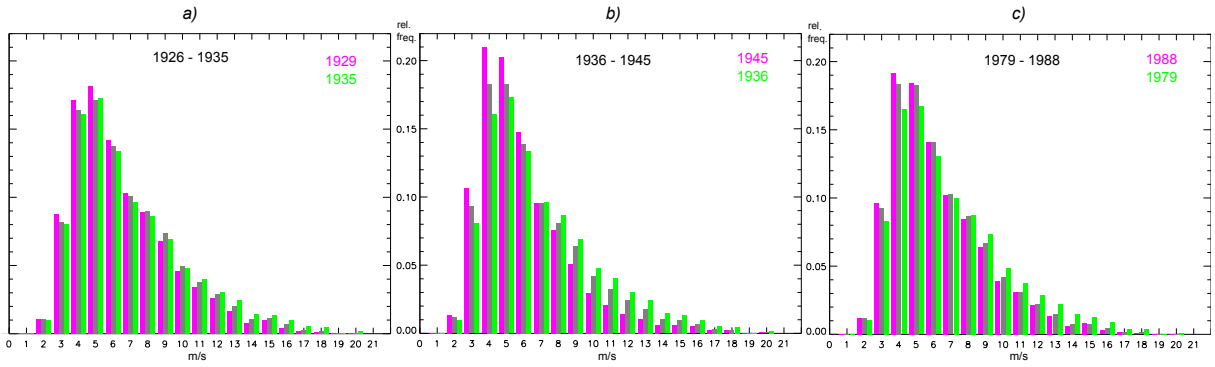

Figure 4.4: Reconstructed wind speed frequency distribution at T1 for the decadal (grey bars) periods 1926-1935 (a), 1936-1945 (b) and 1979-1988 (c). Green (magenta) bars represent the frequency distribution of the years with the highest (lowest) wind speed average of the decade, 1935, 1936 and 1978 (1929, 1945 and 1988).

The PDFs variability in T3 is higher (Figure 4.5). The windiest decade was 1940-1949, while the calmest occurred during 1957-1966, their wind speed PDFs being significantly ($p<0.05$) different through a Chi-test of homogeneity. This inhomogeneity is most evidenced in the 4-5 m/s interval, where the 1957-1966 period presents a frequency 29% higher than the 1941-1950 period (106 days), and in the 8-15 m/s interval, where this decade shows a frequency 40% higher (290 days). In addition, at the decade with the highest interannual variability (1941-1950), the windiest (1941) and the calmest (1950) years show statistically significant differences. The frequency at the first quartile (3 to 4 m/s) is 31% higher for 1950, while at the third quartile (9 to 10 m/s) 1941 shows frequencies 112% higher.

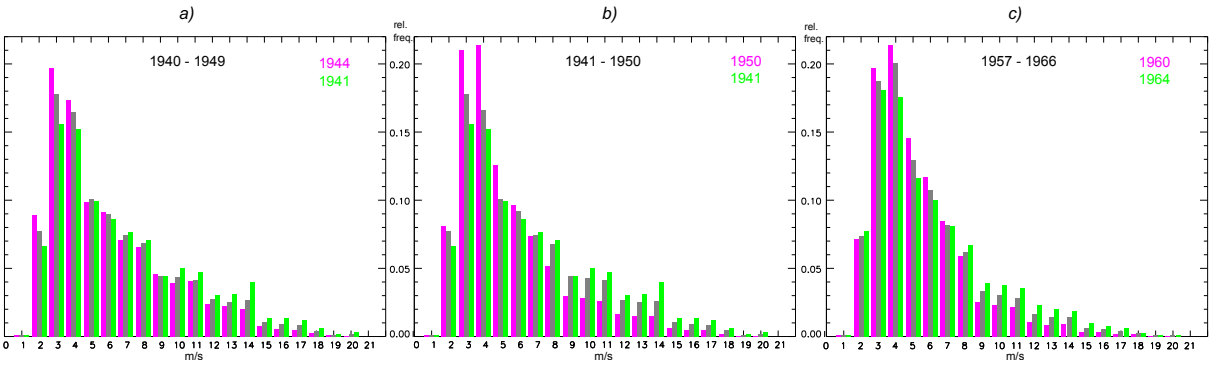

Figure 4.5: As Figure 4.4 but within T3, for decades1940-1949 (a), 1941-1950 (b) and 1957-1966 (c).

4.4 Wind Direction Variability

Similarly to PDF analysis, the inter-annual and decadal wind rose variability has been analyzed for T1 and T3. The same inputs and method than those employed for the wind speed PDFs (steps 1-4) where employed to perform a running decadal 16-sector wind rose reconstruction throughout the period 1871-2009. Thus, every wind rose was obtained by first computing 26 roses, one for each class, according to their wind features during the observations period (T_{obs}). Thus, every one of them was computed according to its class-frequency and -wind speed distribution at every sector along T_{obs}. Then, every reconstructed day was characterized with a wind rose according to its WdT. Finally, for a certain year the reconstructed wind rose was obtained by considering all the daily wind roses of that year, similarly to the class-PDFs previously obtained.

Within T1, SW is the most prevailing sector throughout the whole period. It has been observed that when the annual average wind speed increases above the average, the SW sector frequency is restrengthened, while N and E directional quadrants become weakened. Thus, a statistically significant ($p<0.05$) Pearson correlation of 0.50 is obtained between SW annual wind average frequency and speed. These results can be observed also in decadal averages. In Figure 4.6 the wind roses are depicted for the the considered decades. There, the frequency of the strongest sector (SW) is 30% higher during 1936 than in 1945. The maximum annual frequency differences within the SW sector reached 50% between 1962 (min.) and 1968 (max.).

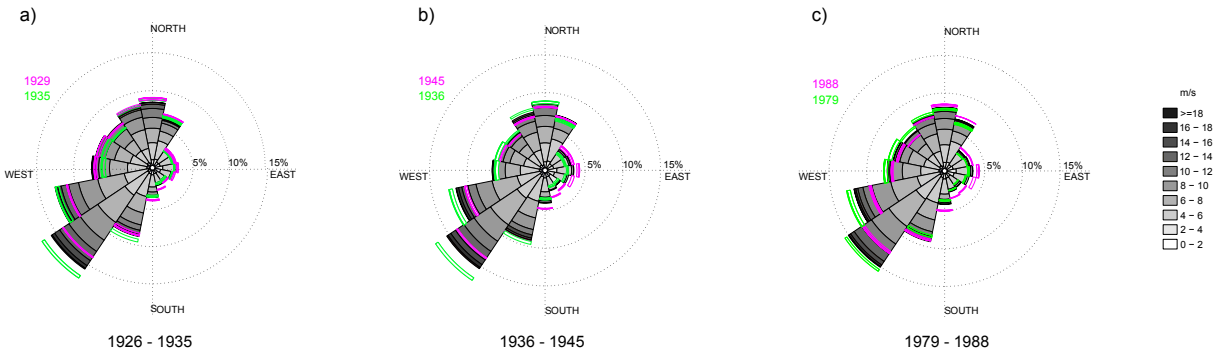

Figure 4.6: Wind rose reconstructions at T1 for the decades with the highest(a) the lowest (c) and the biggest differences (b) wind speed average. Green (magenta) profiles stand for the relative frequency per sector for the year with the highest (lowest) average wind speed of the decade.

In T3, the great frequency during all the period (0.52) of E-SE and SE directions (SE) reflects its particular wind regime. Figure 4.7 shows wind roses for the selected decades. Along the whole period (1871-2009), strong frequency changes of SE can be observed, associated to variations of the annual wind speed. In this way, wind frequency from SE increases during the windiest years, the rest of the directions doing so when the annual wind speed is lower (Fig. 4.7 b). This is corresponded with a significant ($p<0.05$) difference of speed averages between the 10 years with highest and lowest SE frequency values. This difference is equivalent to the 60% of the maximum observed decadal speed range. Nevertheless, this relationship loses stationarity in the last years of the series. Thus, the correlation coefficient between the annual wind speed and SE frequency is 0.51 during the period 1871-1960, while it is lower than 0.30 during 1961-2009. Thus, although T3 wind speed appears to be rising in the last decades, this increase can not be attributable to an increase of easterly winds. This is also evidenced with results from the observational period (Fig. 3.4 c), where wind speed shows SE frequencies lower than that at 1940-1949 (Fig. 4.7a), although similar annual wind speed values are obtained.

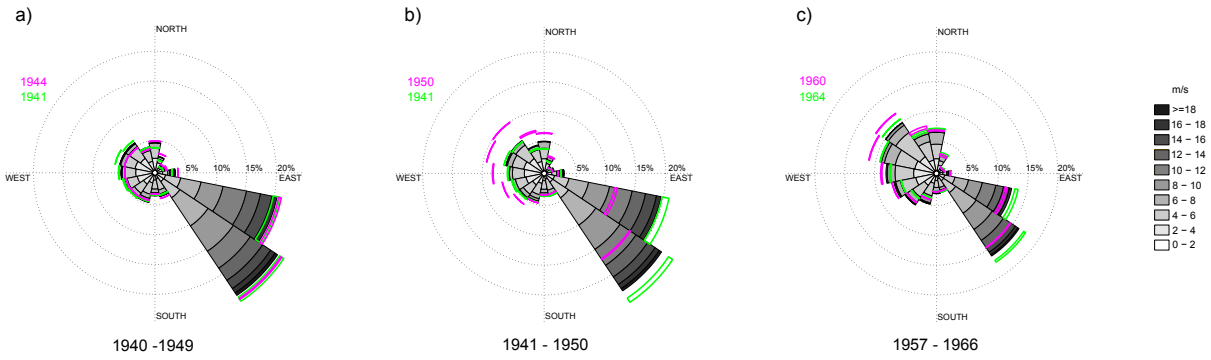

Figure 4.7: As Fig. 4.6 but for T3.

4.5 Wind Type frequency variability associated to the reconstructed wind speed

This section analyses the way the different daily Wind Types obtained through the FG method contribute to the reconstructed annual wind speed outcome, V_a. To do this, a Principal Component Analysis (PCA) was performed on the annual frequency of occurrence (f) of the 26 obtained Wind Types throughout the period 1871-2009. This way a series of Principal Components with a length of 139 years were obtained, representing the annual evolution of the most important variability modes (EOFs) of the Wind Types' frequencies. The PCs obtained and their corresponding EOFs were arranged according to their percentage of explained variance of f, so called $\%Var_f$.

In order to identify which PCs retain more information on the annual wind variability, each one of them was compared with the reconstructed annual wind speed series, and the rate of variance shared among them r^2(PC,V_a) was computed. Among the 26 obtained PCs, only those which showed a statistically significant ($p<0.05$) r^2(PC,V_a) were considered. In this context, 1 PC was considered in T1, T2 and T6, two in T4 and T5 and 3 in T3, as seen in table 4.3. It also shows that in most cases (T1, T2, T4, T6) just a single PC was enough to explain more than 60% of the annual wind speed variability. At each one of the towers with more than one significant PC, the considered PC secores were computed to obtain a composed PC (CPC) from the initial ones. Figures 4.8a)-4.13a)

illustrate the six obtained PCs or PC compositions (CPC).

Table 4.3: PCs retained for each tower. Var_f is the rate (%) of annual Wind Type frequency variability obtained at the annual Wind Type frequencies' PCA, and the shared variance between PCs and wind speed, r^2(PC,V_a). PCs are labeled (1,2,3) according to the amount of %Var$_f$ observed at the PCA.

Tower	PC	Var_f	r^2(PC,V_a)
T1	PC2	14	74
T2	PC2	15	70
T3	PC1	31	11
	PC2	12	31
	PC3	9	26
T4	PC1	41	69
	PC2	9	17
T5	PC1	26	42
	PC2	12	25
T6	PC2	15	60

Results for the six tower EOFs are shown in Figures 4.8b)-4.13b). At every EOF, each Wind Type shows the loading of its contribution (positive or negative) to the annual wind speed variability. The PC composition for T3, T4 and T5 allowed to compute more complete associated EOFs (i.e. containing more information) at those locations. Each EOF shows the incidence that each Wind Type exerts on the annual wind speed variability.

To a high degree, all locations within T_p show a similar Wind Type contribution pattern, with westerlies (easterlies) showing a positive (negative) contribution to the annual average wind. T4 and T5 reflect a higher positive contribution within strong (M3) westerlies and a bigger negative input from weaker (M1) easterlies, as in T1, T2, T6 the contribution is equally distributed through all magnitudes (M1-M3). Regarding calms, all T_p towers showed negative contributions to annual wind, with a higher negative incidence of cyclonic calms at T1, T2 and T6, and a higher contribution of anti-cyclonic calm at T4 and T5.

The pattern (EOF) at T3 is clearly different. Here, SE Wind Types provide the highest inputs, while NEs show the biggest negative coupling. According to calms, in contrast with the other towers a positive contribution of anti-cyclonic calms is observed. This is due to the fact that this class entails, although moderate, an easterly flow, which is

prevalent in this tower.

A clear display of the contribution of the daily wind regimes onto the average annual wind speed can be illustrated by considering the wind observations at the instrumental period and its daily classification (Figs. 4.8 c-4.13 c). There, each daily wind vector is represented with a dot at the (u,v) space, and is characterized through a color scale according to its Wind Type loading at the EOF. Results show a clear W-E bipolar contribution to the average annual wind at all towers, with again an opposite behavior at T3 compared with the rest of the towers. Figures clearly depict surface wind as the result between the interaction of large-scale circulation with surface forcings. In those cases where the prevailing direction shows a W-E dipole, the contribution at every pole is clear, with a positive input to the annual variability (red dots) at westerlies in T_ps and within easterlies at T3. However, some times the prevailing direction dipole does not show a clear positive-negative pattern, appearing divided if the local strengthening is N-S and the prevailing synoptic direction is W-E (T4).

The fact that so few PCs explain high percentages of the wind speed variance shows that the classification derived from the SLP data is an appropriate wind predictor, and the SLP information employed on reconstructing wind (considered PCs) appears clearly isolated from the rest of the SLP variability (rest of PCs).

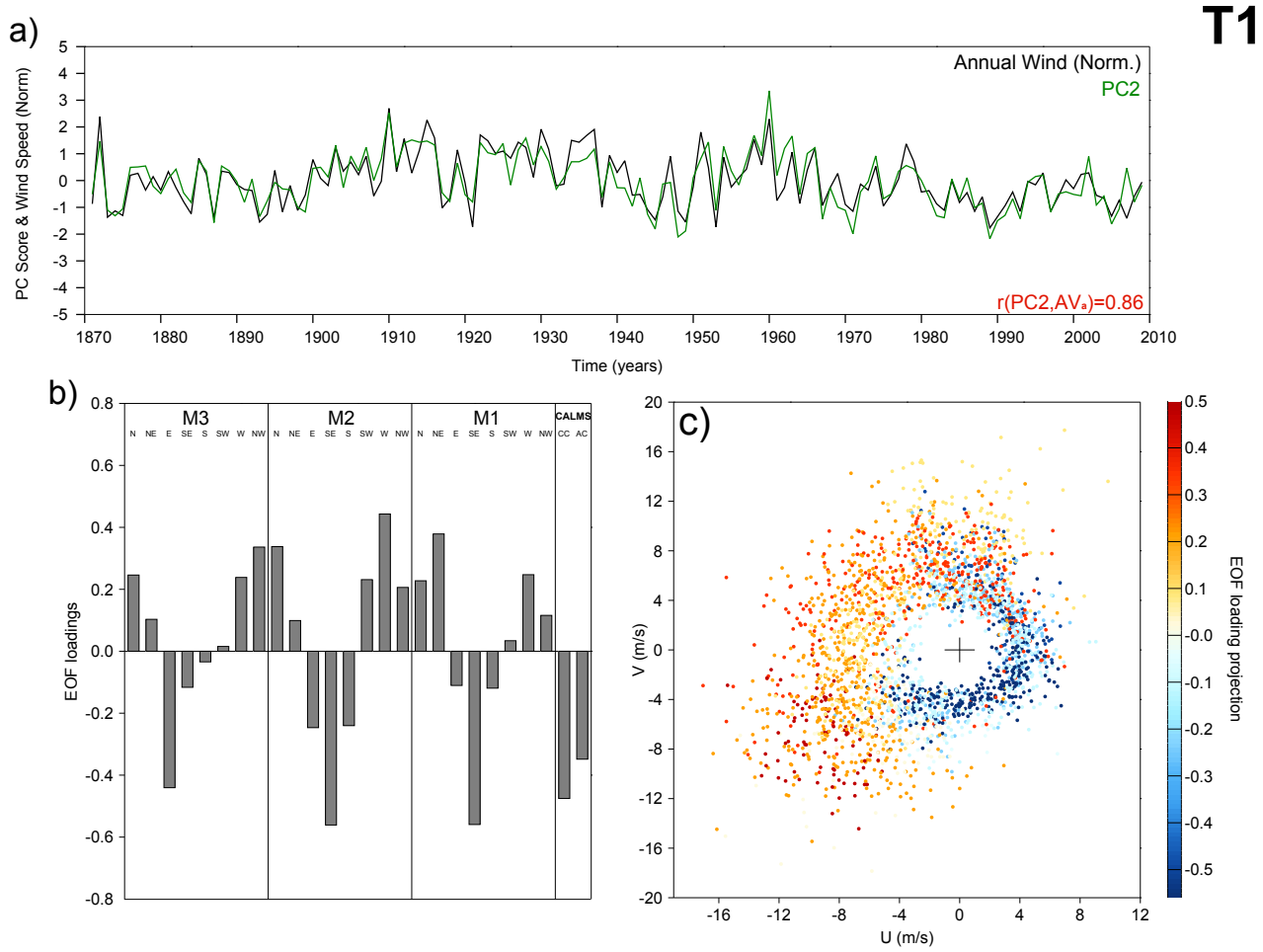

Figure 4.8: Scores of the considered PC of the Wind Types annual frequency variability for all the period at T1, compared to the average annual reconstructed wind (normalized) (a), EOF (eigenvectors) corresponding to the considered PC (b) and projection of this over the wind field at the observational period, according to the derived Wind Types at that time range (c).

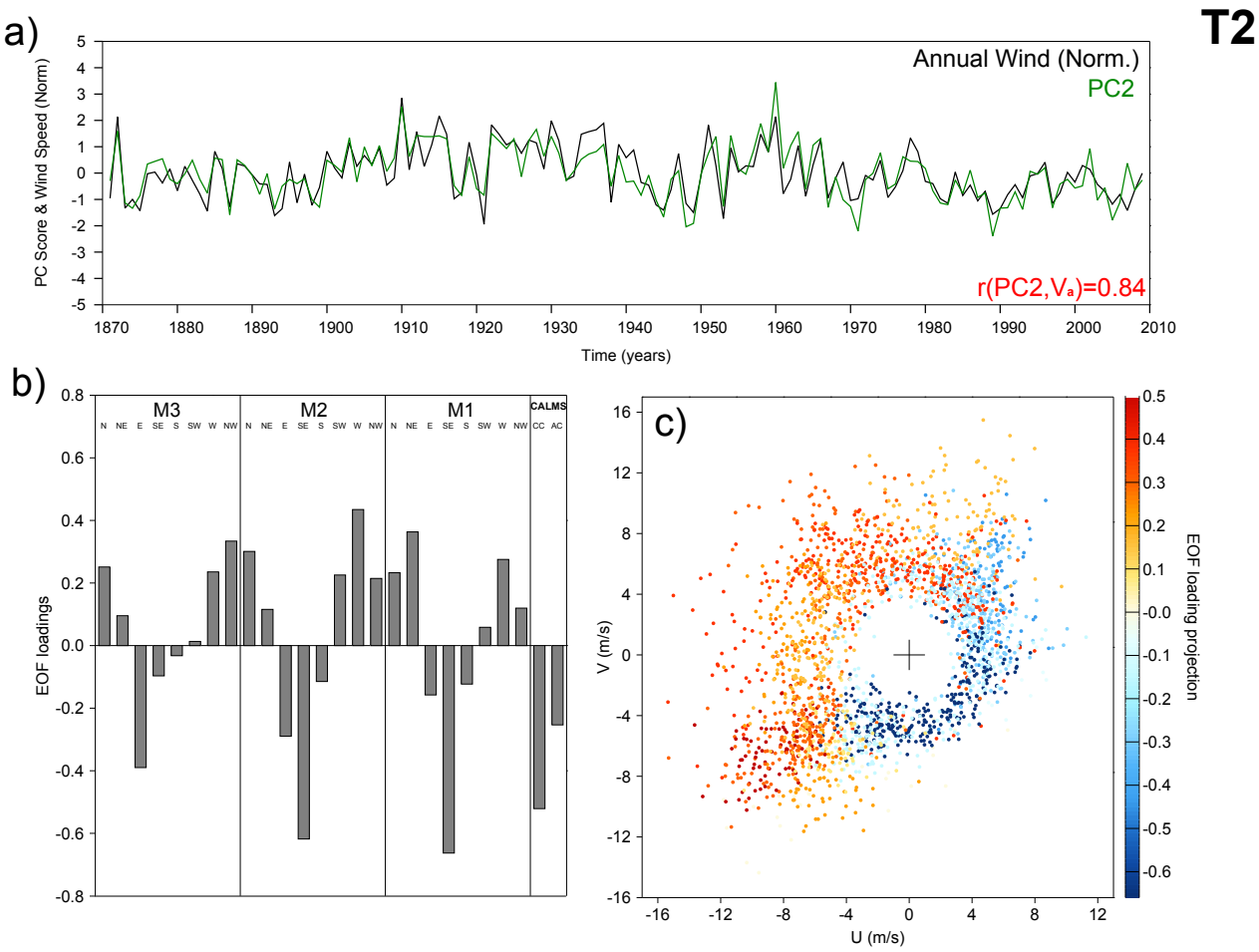

Figure 4.9: Same as 4.8 but for T2.

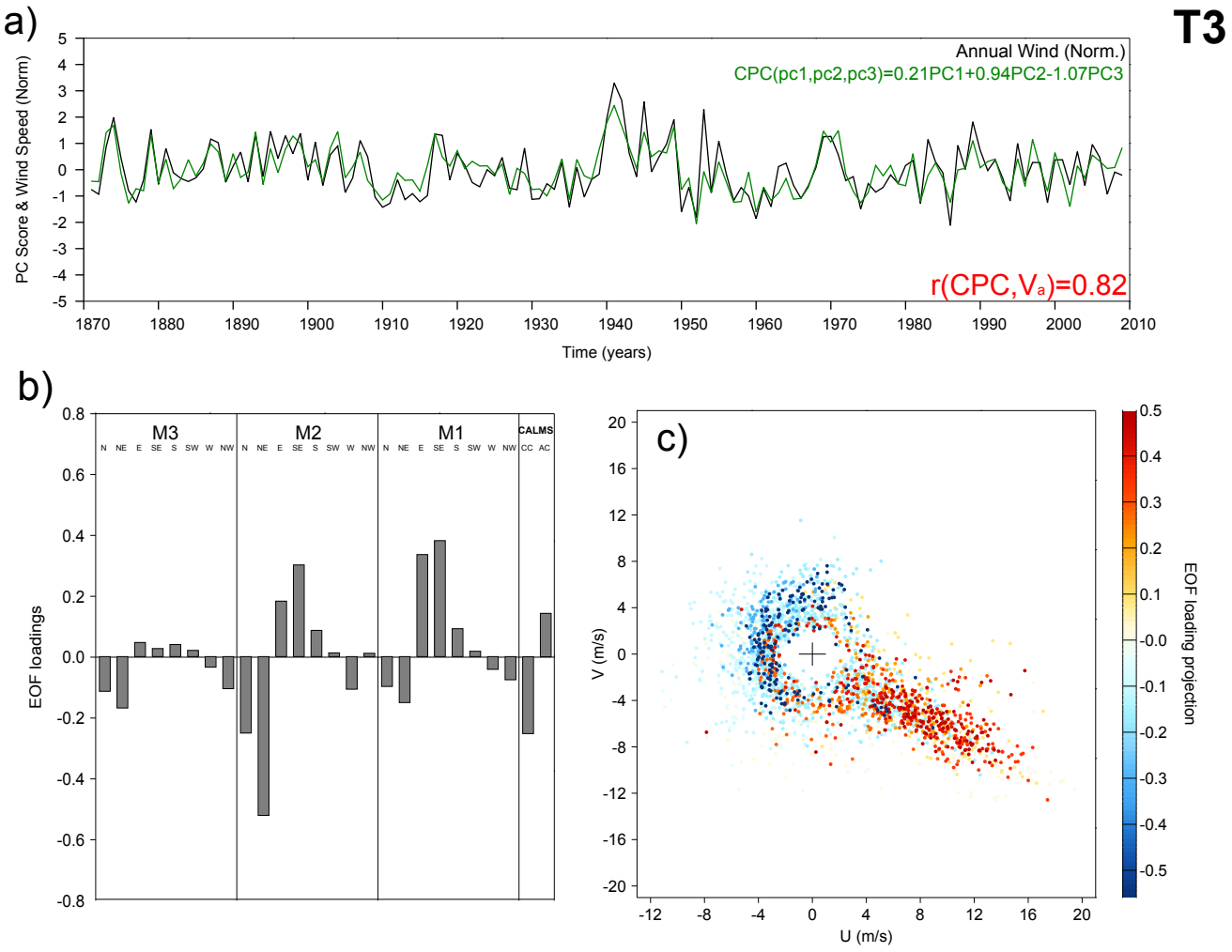

Figure 4.10: Same as 4.8 for T3, but this time the resulting t-mode component (CPC) is obtained by composing the 1st, 2nd and 3rd PCs (by means of a multiple linear regression).

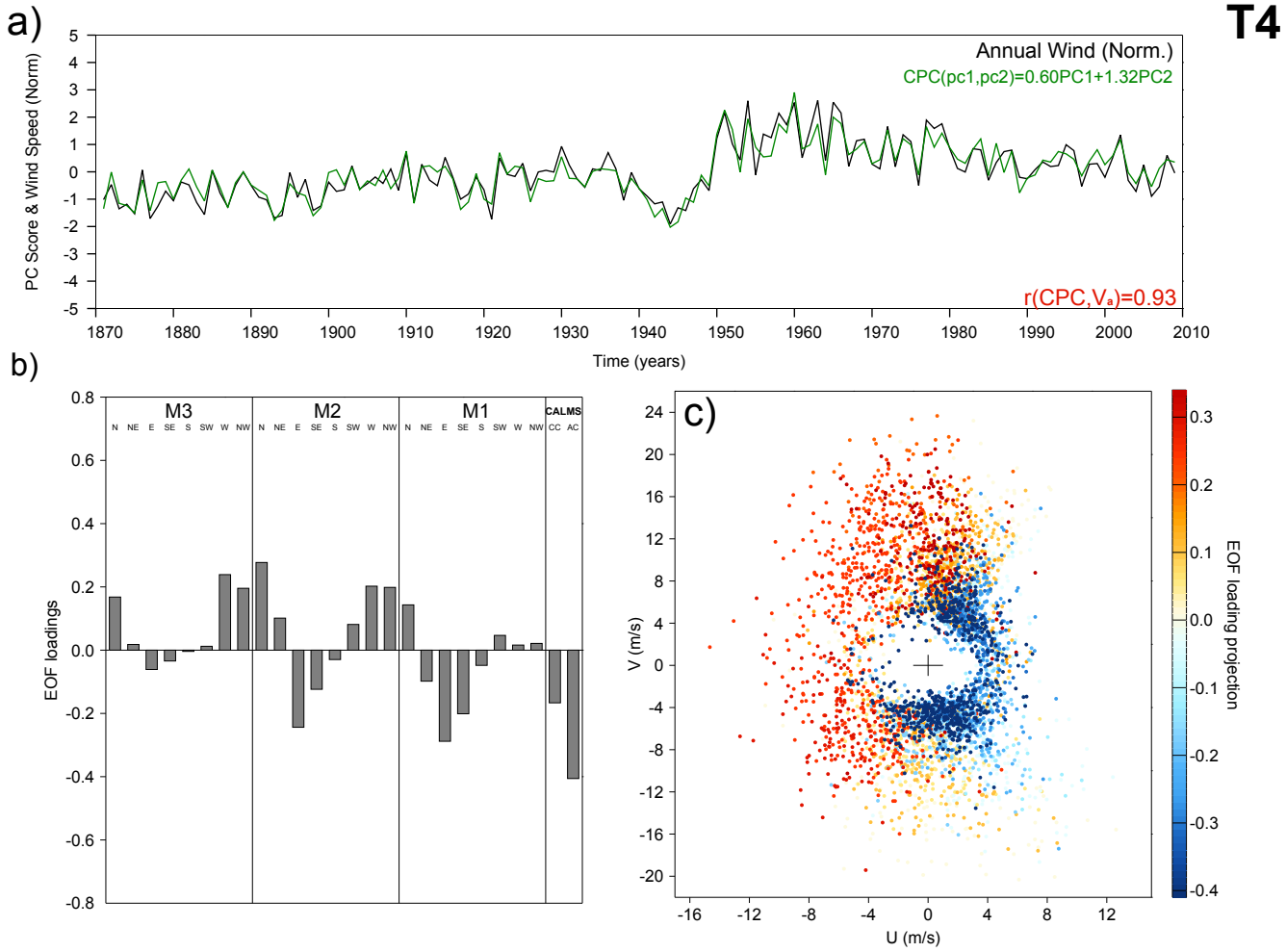

Figure 4.11: Same as 4.10 but for T4.

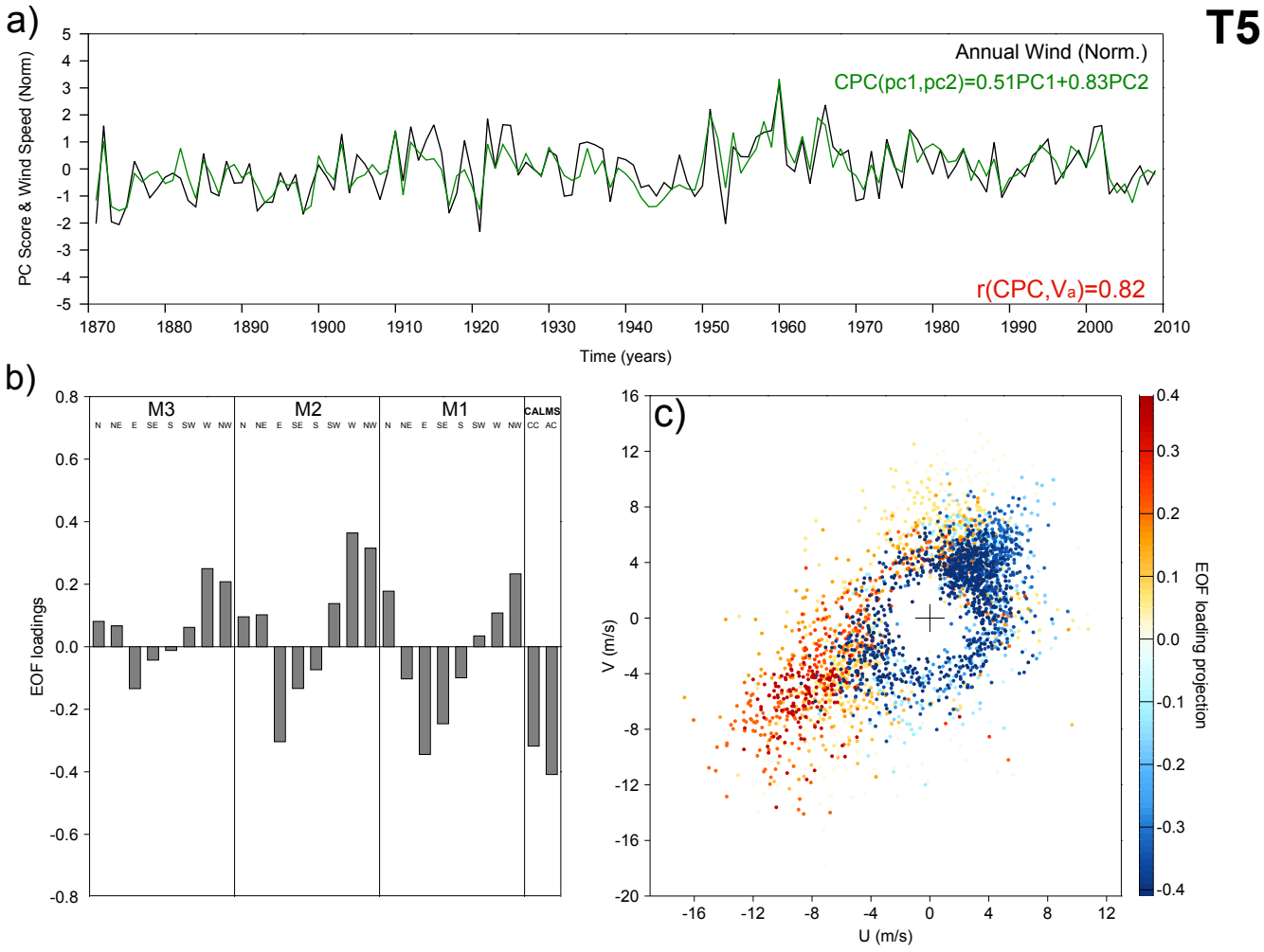

Figure 4.12: Same as 4.10 but for T5. This time the resulting CPC is obtained from two PCs.

4.5. Wind Type frequency variability associated to the reconstructed wind speed

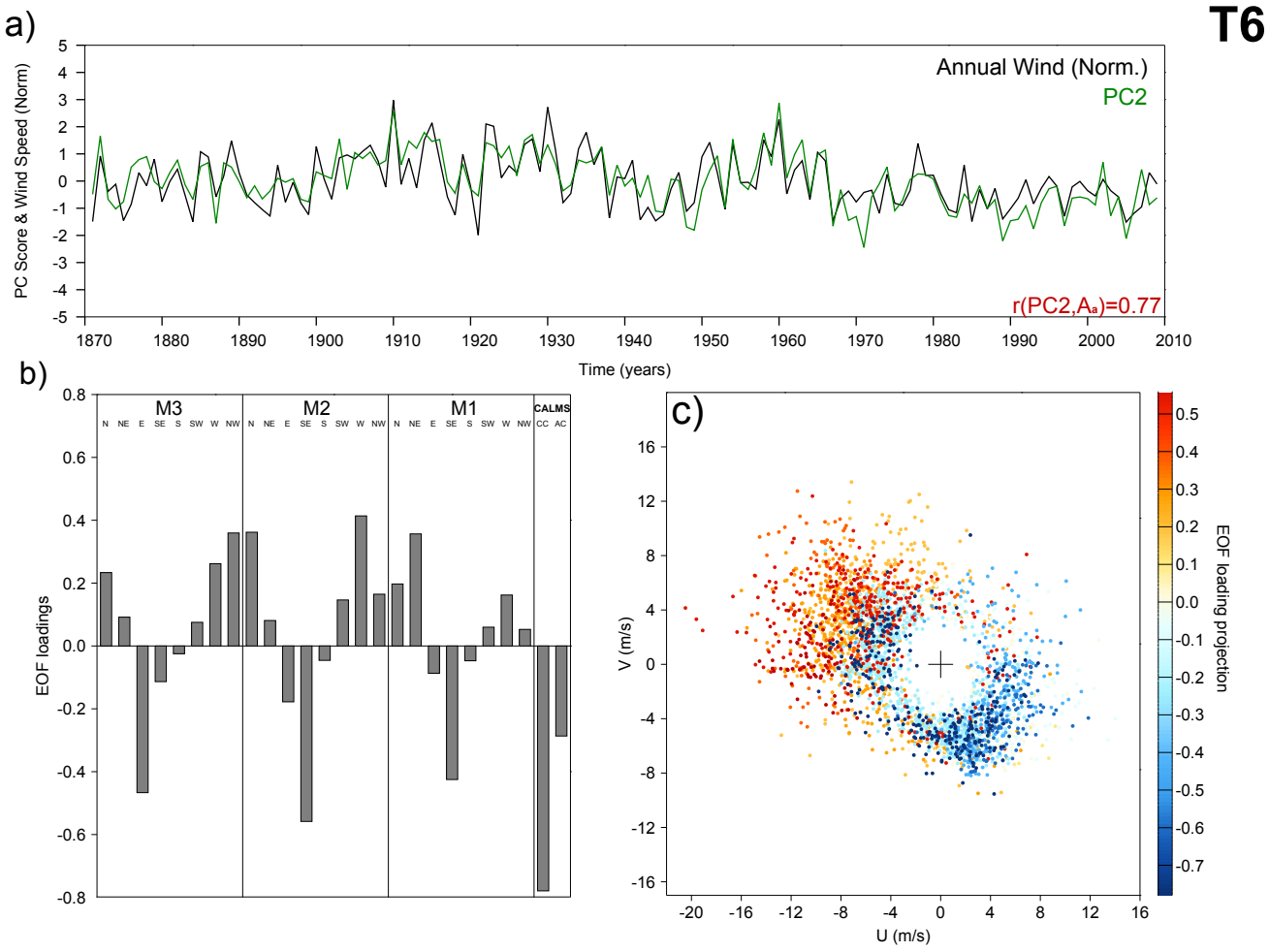

Figure 4.13: Same as 4.8 but for T6.

Chapter 5

Long-term Wind Power Output Variability

5.1 Introduction

In the previous chapter different trends and significant multidecadal changes in wind speed have been analyzed for the considered locations of wind power interest. Here, this analysis is extended to the wind power output. To do this, a simulation on the annual wind power output (P_o) generated by a market wind turbine is performed for the period 1871-2009 at the center (T1) and the south (T3) of the IP (see also Kirchner-Bossi et al. (2014)).

The fact of accounting for a long-term wind power simulation allows to assess the impact that different factors exert on its variability in a multidecadal context. Several works have recently evaluated the relationship between the wind power and atmospheric large-scale teleconnection patterns for different locations, all of them revealing a high incidence of the main regional (Harper et al., 2007; Spears and Jones, 2010; Brayshaw et al., 2011) and global (Klink, 2007; Goubanova et al., 2011) teleconnection modes on wind power. In the IP, García-Bustamante et al. (2013) performed a statistical downscaling from wind power measurements to analyze the long-term relationship between power and the circulation over the Atlantic in north-eastern Iberia. In turn, Jerez et al. (2013) and Jerez and Trigo (2013) assessed the influence of the NAO and the other main large-scale

indices on wind power for the overall region through a dynamical downscaling, showing also a high impact on wind power at monthly scale. Here, the long-term impact of these teleconnections on wind power and its temporal stationarity are assessed for specific points.

Two other factors ruling the long-term variability of wind power can be faced through a centennial wind power series, which to the best of our knowledge have not been previously considered in the literature. In first place, the incidence that the operational wind speed thresholds of a wind turbine exert in the annual wind power output can be analyzed. These thresholds define respectively the minimum (Cut-In) and maximum (Cut-Out) velocities of the wind speed range at which a wind turbine can operate. Although these parameters are obviously adjusted to optimize the generated power, a certain pair of values are set according to the different existing wind regimes. Since wind conditions are far to be stationary in time (as seen in the previous chapter), the extent to which these technical limitations can impact on wind power output can significantly vary through the different decades. Within a long-term wind speed and wind power series all these considerations can now be studied.

The second matter addresses the influence of the wind speed PDF morphology on wind power. A Weibull fit of the estimated wind speed PDFs allows to define it in terms of two single values, the PDF parameters of shape k, and scale c (Celik, 2004). According to this, the fact of accounting for a long wind power and k and c annual series allows to explore if the annual wind power variability can become fully described empirically in terms of k and c exclusively. Should this happen, the possibility to establish a simple expression for the annual wind power output as a linear function of them can be considered, attending the results from García-Bustamante et al. (2013).

5.2 Remarks on the Procedure and the Data Employed

In this chapter, we employ wind observations from T1 and T3. However, in order to estimate the wind power output, it was necessary to consider the characteristics of a real wind turbine. We chose the Vestas V-82 1.65 Mw (50 Hz) (Vestas, 2005), which has been employed in several other works (McIntyre et al., 2011; Tenguria et al., 2011; Schubel and Crossley, 2012) and it can be considered as a reference in wind power research applications. It has a hub height of 78 m, while the 10-minute wind speed observations were performed

at 41.5 (T1) and 40 m (T3). The *Wind Profile Power Law* model extrapolates wind speeds at a certain height where no data are available (Counihan, 1975). It has been applied to adjust the wind speed observations to the turbine hub height:

$$v(z) = v_a \left(\frac{z}{z_a}\right)^{\alpha}, \alpha = 0.096(log_{10}(z_0)) + 0.016(log_{10}(z_0))^2 + 0.24 \quad (5.1)$$

where $v(z)$ is the wind speed at height z, and v_a that at height z_a, while α denotes the contribution of the site roughness length (z_0) to the speed vertical gradient of the atmospheric boundary layer. In our case, T1 and T3 locations present an α value of 0.20.

The employed reanalysis SLP data are slightly different in this chapter. It has been recently reported that the fact of considering the SLP ensemble mean of the 56 simulations could entail a certain loss of information (Donat et al., 2011), due to smoothing effects specially within storms (Wang et al., 2012), being preferable to consider the 56 ensemble members separately and only combining their outcomes at the end of any variable estimation or model computation. In order to overcome this problem, all the $20CR$-derived variables computed in this chapter have been obtained by considering each one of the 56 ensemble components independently. Only after having computed the final derived variables (wind power), their mean and deviation are calculated from the 56 outputs, as it will be shown in the following sections. In this way, here wind speed reconstructions are obtained for each one of the 56 ensemble members by performing the FE classification, in the same wind speed was reconstructed in the previous chapter. We have used this method because here no detailed analysis on the Wind Typing is required, so a non-deterministic algorithm can be suitable. As seen at the validation (Sect. 3.3.1), differences between FE and PD daily wind estimation performances were very similar.

5.3 Wind Power Output and Wind Speed

As mentioned in the introduction, the output power generated by a fluid of density ρ through a wind turbine of surface S is related to its speed v according to the expression:

$$P = \frac{1}{2} C_p \rho S v^3 \quad (5.2)$$

where C_p represents the *Power Coefficient* of the wind turbine, with a theoretical maximum value of 16/27 (known as the Betz limit). However, this expression is not

realistic for a wind turbine, mainly due to friction and sustentation features of the turbine blades (Wilson et al., 1976; Vries, 1979), and power results adjust better to the square of speed (Anderson and Bose, 1983). This makes that C_p becomes a non-linear function of wind speed (Monroy and Alvarez-Icaza, 2006; Kjellin et al., 2011) and wind power (Lanzafame and Messina, 2010). Beyond the consideration of this theoretical expression, a realistic estimation of the wind power from wind speed can be made if the *Power Curve* (Hau, 2000) of a wind turbine is considered. This curve is developed by the manufacturer along a theoretical-empirical performance test of the product, consisting on the implementation of an aeroelastic model to a theoretical curve, which later must be tested experimentally. It describes the experimental relationship observed between a series of 10-minute wind speed averages and the mean P_o actually measured. Output power can be directly estimated from wind speed by means of the matrix product between its Probability Density Function (PDF) and the Power Curve of the considered wind turbine (Carta et al., 2009).

As it was mentioned in Chapter 1, the fit of the wind speed PDF over a Weibull distribution is a widely extended procedure to parameterize the wind speed conditions at a given location. It allows to derive a smoothed curve described through only two parameters (shape k and scale c). Particularly, a statistical fit is appropriated if wind data are recorded at intervals longer than one hour, because wind speed can present large short-term fluctuations (Yeter et al., 2012). Unfortunately, the goodness-of-fit of this adjustment is not always as accurate as desired (García-Bustamante et al., 2008), specially when the orthogonal components of the horizontal wind differ of a normal distribution (Tuller and Brett, 1984), mainly in low wind speeds (Jamil et al., 1995). A solution to the problems presented by a Weibull fit consists on generating a wind speed PDF by considering the experimental data from wind. Here, since the reconstructed daily wind speed has been computed from a high time resolution (10-minute data) homogeneous data, the annual PDFs have been directly computed from the obtained annual wind data.

The reconstructed wind speed PDF series that allowed to compute the experiments in this chapter were obtained according to the steps 1-4 at section 4.3. They were applied to each one of the 56 ensemble components of the $20CR$ SLP dataset. Hence, a set of

56 wind speed PDFs was obtained for every year/season. For each year, a PDF was computed by merging its daily PDFs. The 56 reconstructed wind series resulted into 56 annual PDFs, and 56 values of P_o could be computed for every year. So, the average P_o and its corresponding uncertainty was computed for every year, producing a series of 139 annual P_o with a certain dispersion.

5.4 Wind Power Centennial Series

Figure 5.1 shows the evolution of the annual *(a-b)* and seasonal *(c-j)* simulated P_o for the period 1871-2009. The observed P_o dispersion is represented in the figures through the 10, 25, 50, 75, 90 percentiles and the maximum and minimum values of the simulated P_o ensembles. The reduction in the P_o dispersion with time is clearly evidenced. The ensemble outputs are increasingly similar according to the accuracy improvements of the $20CR$ inputs with time. In fact, this spread is partially related with the quality of the 20CR reconstruction for a given year, largely dependent on the number of stations providing pressure data in the area (Compo et al., 2011). Therefore, it is expected that errors increase as we go back into the early decades of the 20 CR dataset.

The wind speed estimation accuracy for the method employed here was successfully evaluated at Kirchner-Bossi et al. (2013), so no additional validation was required. We have compared the simulated P_o obtained through the wind reconstruction with the annual values obtained directly from the wind observations, along the T_{obs} period. The obtained mean P_o values at T1 were 666 Kw (obs.) and 661 Kw (rec.). At T3 these values were respectively 582 Kw (obs.) and 585 Kw (rec.). These results imply annual MAE of 26.59 Kw (T1) and 31.64 Kw (T3) with respect to the observations. In turn, the *Capacity Factor* (CF, the ratio between the real mean P_o and the nominal rated power of the device, 1650 Kw in this case) according to the observations was 0.36 (T1) and 0.32 (T3). These values were obtained by considering an additional 10% power loss due to maintenance stops, electrical grid losses and wake effects (Manwell et al., 2002). Simulations through reconstructed wind provided identical CF values.

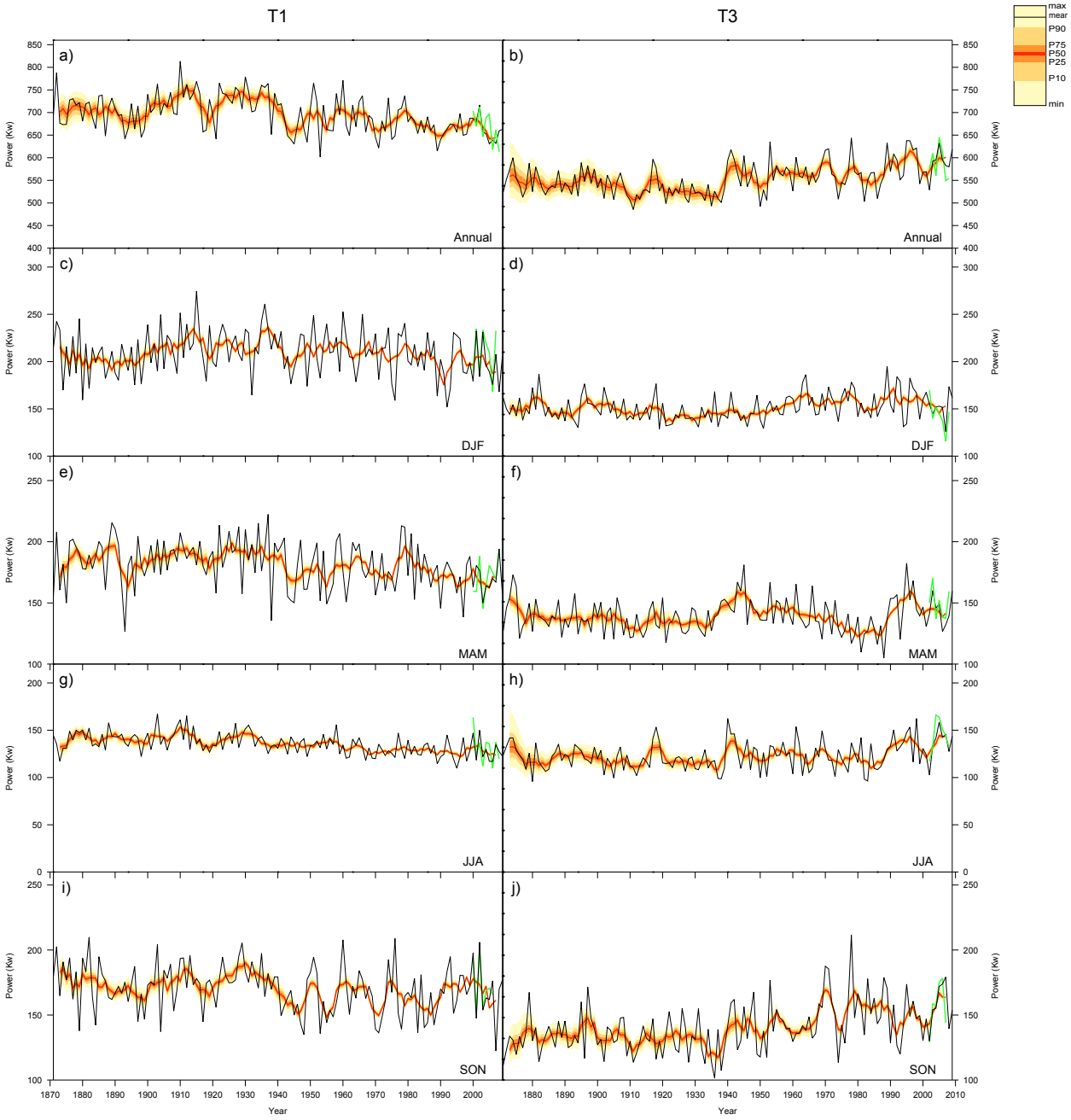

Figure 5.1: Annual estimated P_o (1871-2009) for the Vestas V-82 turbine at T1 and T3 sites through the FE method. A-B describe the overall year-round signal, c-j stands for the seasonal series. Since each one of the 56 ensembles produced a certain value of P_o, some dispersion was obtained for each year. The black line represents the annual mean, as dispersion is represented in terms of the percentiles shown in the top right scale, through a smoothing given by a 5-yr running mean. Dispersion is observed to grow as time goes back, as expected attending the dispersion of the 20CR ensemble with time (Compo et al., 2011). The green line denotes the P_o estimated strictly from observations.

5.5 Long-Term Variability of the Simulated Output Power

Both towers exhibit high interannual variability superimposed on long-term patterns. In T1 P_o reached its peak in the 1910s (813 Kw) and has experimented a continuous decline since then. In fact, it shows an annual (statistically significant) negative trend of -41 Kw (-5.6%) every 100 years (hereafter all trend changes refer to this period), with a minimum value in 1953 (602 Kw). Seasonally, the biggest decrease was observed at JJA, with a significant trend of -13 Kw/100 years, with all seasons showing negative trends (also significant in MAM). In T3 P_o shows the opposite behavior, with the minimum values at the beginning of the 20th century (485 Kw in 1911) and highest values after 1940, peaking in the last 30 years (maximum in 1978, 644Kw). This increase is mostly attributable to SON, which shows a positive trend of 37 Kw (+16.0%), although all seasons presented positive trends (all significant except MAM). Particularly, the last 10 years show even a higher growth at MAM and JJA, with a JJA increase of 12,9 Kw (+7.6%) with respect to the previous 130 years. In turn, the increase at SON and DJF is most evident during the second half of the 20th century.

The wind power seasonal cycle typical at mid-latitudes was strongly evidenced in T1, with considerable higher P_o values in DJF (30% of the overall P_o) than in JJA (19.5%). However, this seasonality is less pronounced at T3 (27 and 22.2% for DJF and JJA resp.). This anomalously high summer contribution is consistent (CL20 at Esteban et al. (2006)) with the anticyclonic circulation regime in summer over Iberia (Fernández-Montes et al., 2012; García-Herrera et al., 2005), which induces relatively high wind conditions (prevailingly easterly) over this location compared to other points of Iberia.

Multidecadal periodicities modulate the observed trends. A spectral Fourier analysis showed the existence of several statistically significant interdecadal and multidecadal cycles at T1 (27 yr) and T3 (46 and 13 yr). At T1 the annual signal seems to be spread among all seasons, as no long-term significant cycles were found for any specific season. In turn, T3 periodicities evidenced the main contribution of JJA (46 and 13y) and SON (13y) to these multidecadal fluctuations. In the case of JJA, the 46y cycle could be directly contributing to the last decade increase.

5.6 Impact of the Teleconnection Patterns on the Simulated Output Power

To investigate the role of the teleconnection patters on the P_o variability we have considered the leading modes over the North Atlantic/European Region, the North Atlantic Oscillation (NAO), the Eastern Atlantic (EA) pattern and the Scandinavia (SCAND) pattern. All three indices where obtained from NOAA (Bell et al.). There, a Rotated Principal Component Analysis (RPCA) is performed to the 500mb field from the NCEP/NCAR (Kalnay et al., 1996) reanalysis for the period 1950-2009, which ensures the linear independence between them. Additionally, the simulated P_o series have been compared with the Atlantic Multidecadal Oscillation (AMO, Kaplan et al. (1998); Enfield et al. (2001)), computed in terms of the Sea Surface Temperature (SST) at the Atlantic Ocean with a detrend of its long-term linear signal. This index has been included in this analysis, since the Atlantic oceanic circulation could be relevant to explain the low frequency of the wind variability.

Figure 5.2 shows the Pearson correlation coefficient (r) between wind power and the considered teleconnection indices. r was computed by using s 30 year running window on a seasonal basis, for the period where the three indices are available (1950-2009). This running window was chosen in order to investigate the temporal stationarity of the relationship. It can be observed that T1 holds higher and significant correlations within the four seasons compared to T3. The highest correlations at DJF (NAO) and SON (SCAND) show a stationary behavior, with maximum signals of -0.55 and 0.79 respectively. The NAO signal in DJF shows the typical pattern in Western Europe in winter (Hurrell et al., 2003), while the positive connection with SCAND is attributable to a low pressure system centered over the Iberian Peninsula (Barnston and Livezey (1987), SCAND referred there as Eurasia-1). A similarly high and stationary connection with an opposite signal is also observed at SON in T3 (average r=-0.63), representing by far the highest correlation at this location. This negative coupling with SCAND is explained by a marked anticyclone centered over northern Spain (Barnston and Livezey, 1987), which is consistent with the persistent easterly winds over the Gibraltar area.

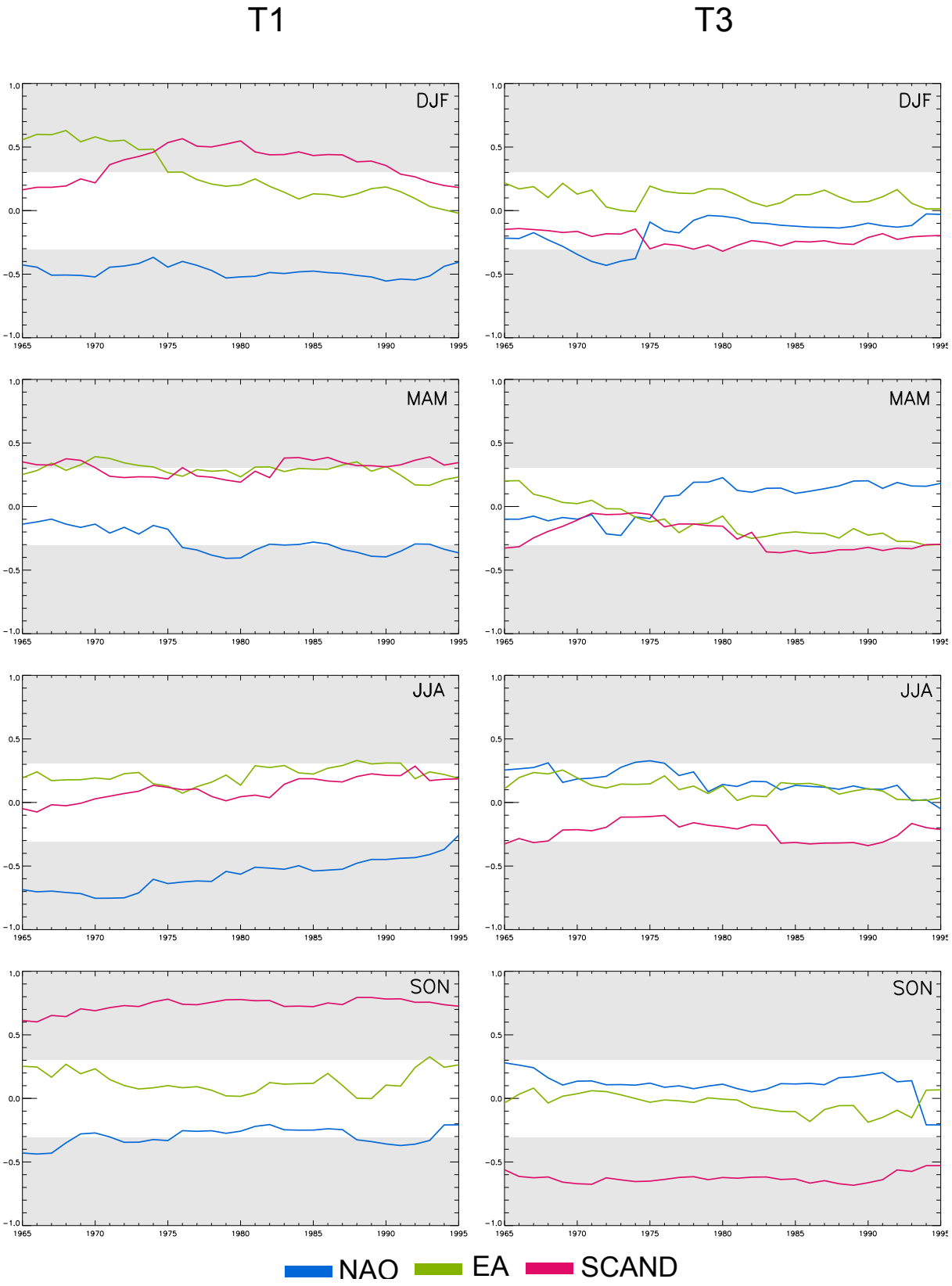

Figure 5.2: Seasonal evolution of the Pearson correlation coefficient (r) between the annual wind power output at T1 (left) and T3 (right) and each one of the three principal variability modes of the North Atlantic-European region. Years refer to the 15th year of a 30yr running correlation. Grey regions gather those statistically significant ($p<0.05$) r values.

Finally, the EA pattern shows a significant positive coupling with wind power in DJF at T1 during 1950-1990 (most of the time around 0.5), after which it vanishes. Although EA does not show any other significant seasonal connection with wind power, remarkable correlations are found if other intra-annual periods are considered. Thus, an homogenous and stationary (always significant) influence of EA over T1 during the extended summer (October-April, not shown) was detected, where peaks of r=0.6 are recorded along all the considered period. In turn, the impact of EA on T3 P_o is very strong in summer, specially in July, with significant correlations until 2006 (with peaks next to 0.6 during 1955-1990), presenting a decreasing influence in the last years (not shown).

The overall influence of the considered teleconnection indices on P_o can be represented by the cumulated explained variance (r^2) of the three patterns along the available periods. Figure 5.3 shows the contribution to r^2 of each variability mode. Results show that, in general terms, T1 P_o variability is more closely associated to large-scale circulation modes than T3. Thus, the overall explained variance is particularly high at T1 in winter (peaking at 69% in 1967) and autumn (highest in 1993, 79%), with a major contribution of NAO (negatively correlated) and SCAND (positively correlated) respectively. In JJA this tower shows a high signal with respect to NAO (also with a negative r), although it is highly non-stationary, with an abrupt decrease in the last years, even losing significance. Finally, MAM shows a stationary weak, but mostly significant contribution from each one of the three modes. Regarding T3, SON is the only season with significant results, with a high and stationary contribution from SCAND (r^2 peaking at 51%)

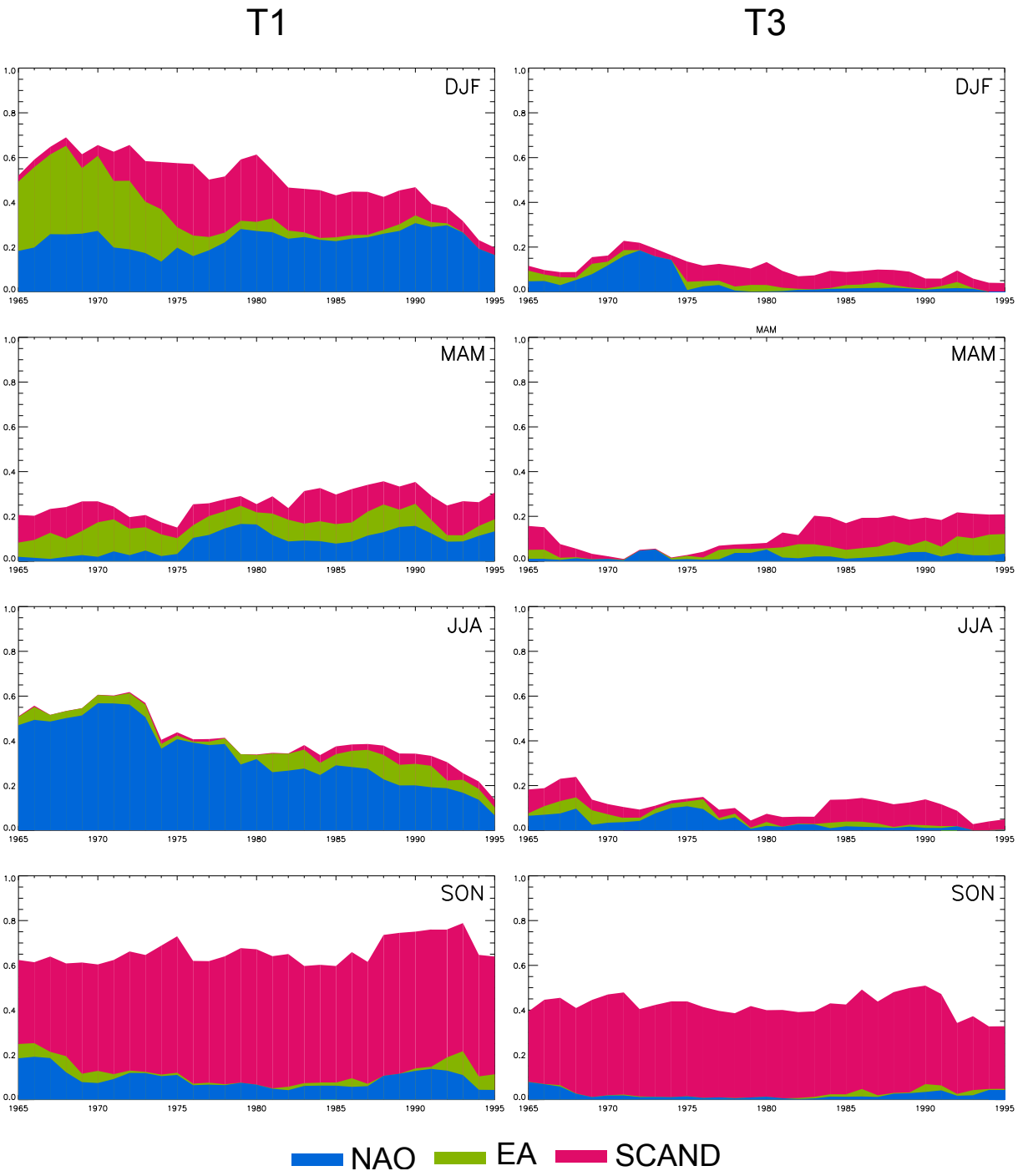

Figure 5.3: Seasonal evolution of the cumulated explained variance (r^2) of Output Power at T1 (left) and T3 (right) by the three principal variability modes of the North Atlantic-European region. Years refer to the 15th year of a 30yr running correlation.

Finally, AMO shows a non stationary (although slightly significant) correlation with both towers (Fig. 5.4), switching its sign during the considered period. The correlation with T1 shows an opposite behavior compared to that obtained with T3 throughout the entire considered period (1871-2009). It is significantly correlated with T3 since 1966 (with a maximum r=0.49), while T1 (negatively) does so since 1977. The season with the highest correlation between wind speed and AMO is MAM (T3), with r=0.48 during 1950-1980.

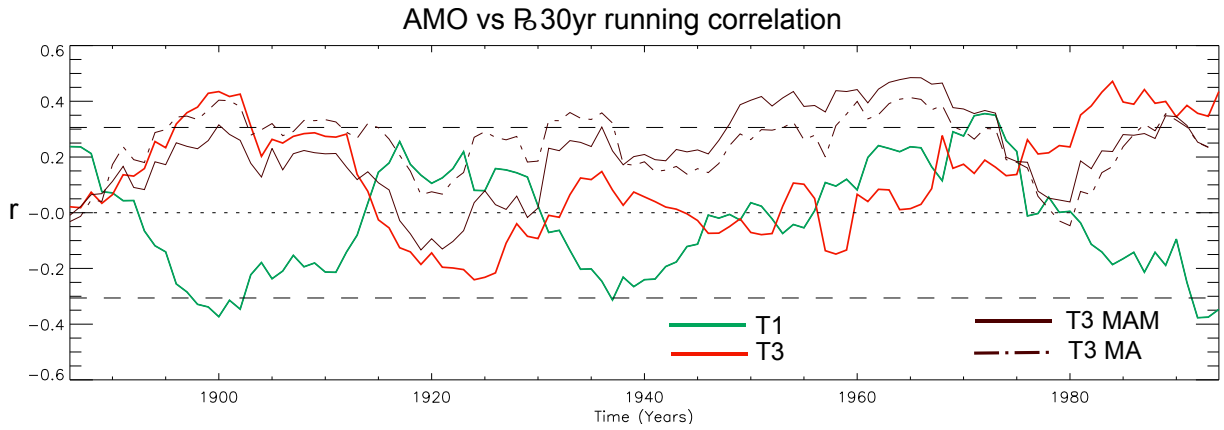

Figure 5.4: 30yr running Pearson correlation coefficient between AMO index and the annual power output series. The horizontal axis indicates the center of every 30 yr window considered. Dashed horizontal lines correspond to a p<(0.05) statistic significance level.

5.7 Incidence of the technical wind speed limits of a wind turbine on the Simulated Output Power

Every wind turbine has a limited operational wind speed interval where it generates usable power. This range is defined by an upper and a lower wind speed, called respectively *Cut-In* and *Cut-Out* speed. They are set by the manufacturer and usually range between 3-5 m/s and 20-25 m/s respectively. The *Cut-In* value refers to the minimum speed needed to generate power, and varies according to the turbine dimensions and aerodynamic features, while the *Cut-Out* responds to security and stability reasons to ensure the scheduled length of the device operational life. The empirical relationship of the variability of these technical limits on the annual P_o has been assessed.

We have reconstructed the annual frequency of wind speeds lower than the *Cut-In* speed and wind speeds higher than the *Cut-Out* value for T1 and T3. The variability of those wind speeds exceeding such technical limits can be computed through a so called *Point Over Threshold* (*POT*) methodology (Brabson and Palutikof, 2000), where a time series is computed from those values higher (or lower) than a certain threshold. In this way, two frequencies were computed: the upper POTs, which consist on those wind speeds higher than the *Cut-Out* value (hereafter, PCO), and the lower POTs, defined as those wind speeds lower than the device *Cut-In* speed (hereafter, PCI). The considered *Cut-In* and *Cut-Out* values are those of the Vestas V-82, 3.5 m/s and 20 m/s respectively.

As with the P_o simulation, the annual and seasonal averages of PCI and PCO were considered, so that a 139 year series was obtained for both POT frequencies at each tower. For the entire period, PCI corresponded to the 9 and 23 percentiles and the PCO to the 99 and 98 percentiles, for T1 and T3 respectively. Figures 5.5 and 5.6 show the annual series of reconstructed PCI and PCO series. Again there is an increasing spread in the early decades of the 20 CR dataset and overall, it is larger for T3 than for T1. As with P_o, the signal dispersion produced by having 56 different wind speed frequencies can be identified through the colour scale.

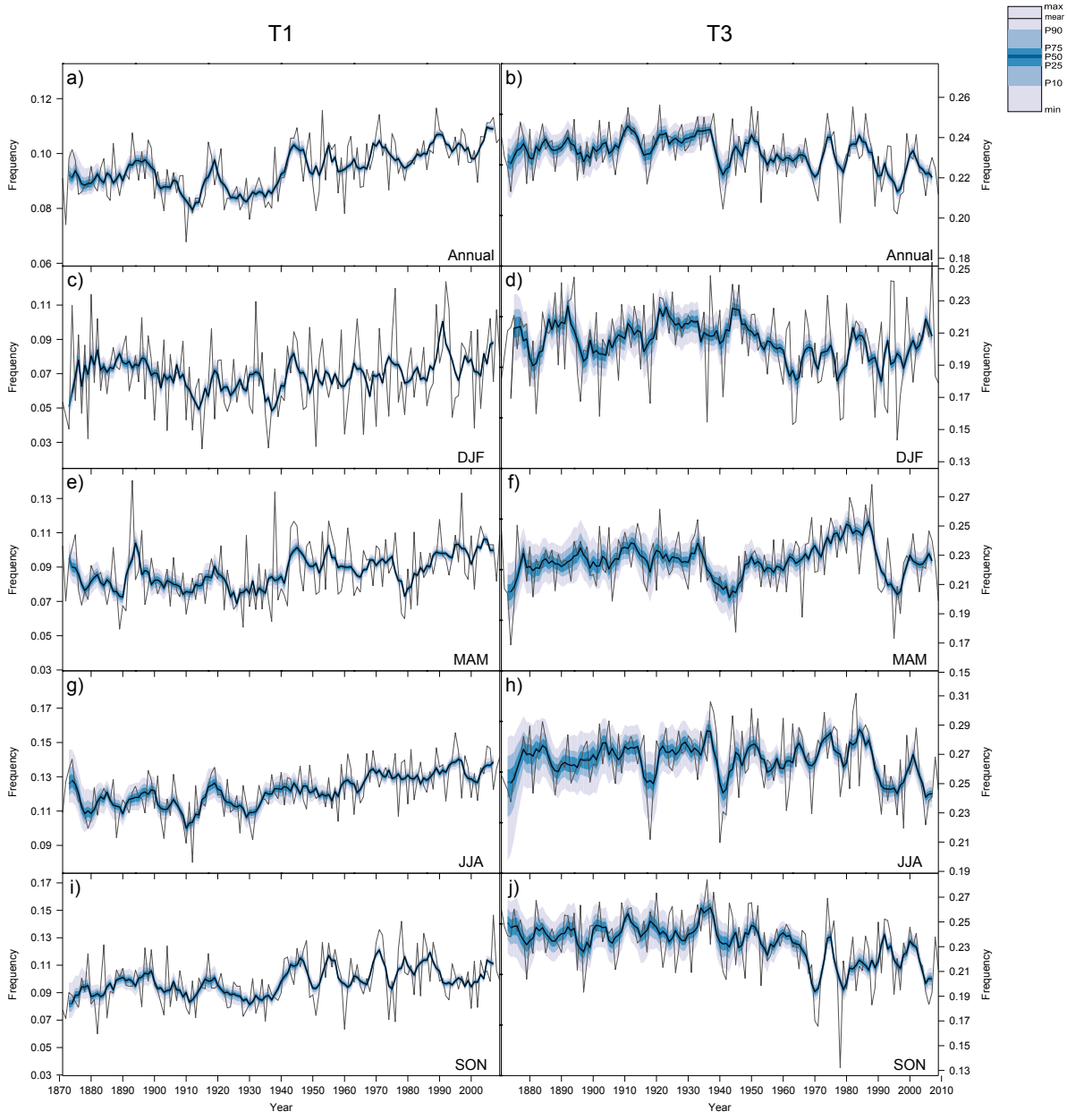

Figure 5.5: Annual(a) and seasonal ($b-e$) reconstruction of the wind speed frequency lower than the Vestas V-82 *Cut-In* value at T1 (left) and T3 (right). Black line stands for the ensemble annual mean, while blue tones represent the minimum value, the 10, 25, 50, 75, 90 percentiles and maximum value of the ensemble annual average frequency, represented with a smoothing of a 5-yr moving average.

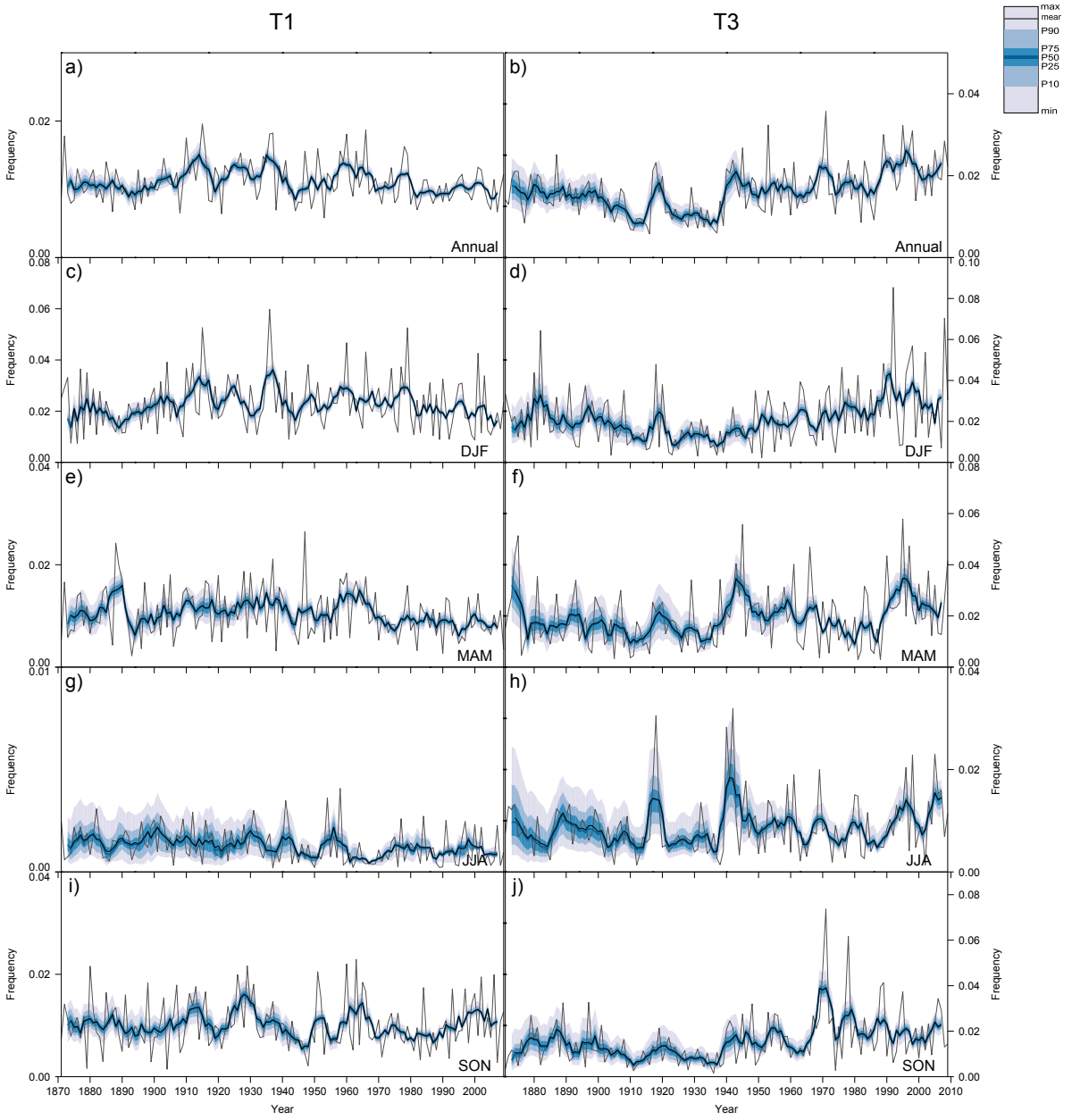

Figure 5.6: Same as Figure 5.5 but this time the wind speed frequency higher than the device $Cut - Out$ value.

The PCI and PCO long-term variability was found consistent with the results obtained for the of the annual P_o. The annual frequency of PCI values at T1 (Fig.5.5 *a*)) showed a significant increase expressed through a relative growth of 13.7% (as P_o, every 100 yr). In addition, all the seasonal series show a positive trend (statistically significant except

for DJF). The increase at the observational period with respect to the simulated one is also significant (specially at MAM and JJA), with a +8.21% increase. On the contrary, a negative trend (also significant at MAM) was detected for PCO (relative decrease of 5.4% -annual- and 14% -seasonal- every 100yr, Fig.5.5 $a), e)$).

As with P_o, the variability of T3 POTs behaved the opposite way. There, PCO values showed a significant increment (53%), while PCIs presented a significant decrease (-3.45%). Seasonally, frequency changes are mostly attributable to SON (89% at PCO, -10% at PCI) although DJF shows also a notable contribution (a 58% PCO growth).

Again, spectral Fourier analysis showed the existence of several statistically significant decadal and multi-decadal cyclic fluctuations in both towers, consistent with P_o results. Thus, T1 showed two annual periodicities of 27 yr and of 46 yr for PCI, the last one being observed also in MAM. Regarding T3, a 13 yr cycle was observed for the annual PCI frequency, also evidenced in the JJA, while cycles of 23, 27 and 17 years were shown by MAM, JJA and SON respectively.

The relationship of the cut-in and cut-out parameters with the annual output power has been computed. Results show that the negative relationship of PCI with the annual P_o shares an explained variance (r^2) of 95% (T1) and 81% (T3), while the positive coupling with PCO explains 53% (T1) and 33% (T3) of the annual power series variability. Hence, P_o changes respond more directly to PCI frequencies than to PCO, both of them higher at T1 than at T3.

5.8 Impact of the k and c Weibull parameters on the Simulated Output Power

Although as mentioned previously a fitting of a wind speed PDF to a Weibull (Weibull et al., 1951) distribution can present some troubles, it includes some advantages, as it allows to parameterize the shape (k, dimensionless) and scale (c, [m/s]) of a wind speed distribution into two single quantities. In this section the long-term interannual variability of these parameters is compared with the P_o multidecadal changes, in order to explore the empirical relationships existing among them. The motivation to investigate this arose from the fact that several works on either wind reconstruction (Curry et al., 2012) or climate projections (Pryor et al., 2005b; Bogardi and Matyasovzky, 1996) base their methodology on providing a value of c and k.

5.8. Impact of the k and c Weibull parameters on the Simulated Output Power

The Weibull fit of the 139 annual PDFs (each one averaged from the 56 ensemble members) allowed to compute the corresponding 139 annual values of k and c. This has been done, for every year, by considering the log-linear expression obtained for a certain probability P to have a wind speed v$< v_i$ (Miller et al., 1965). Since the Weibull PDF (F_W) is defined as

$$F_W(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} \mathrm{e}^{-\left(\frac{v}{c}\right)^k} \tag{5.3}$$

It can be verified that its Cumulative Distribution Function (CDF) in terms of the probability P_i is

$$P(v < v_i) = P_i = 1 - \mathrm{e}^{-\left(\frac{v}{c}\right)^k} \tag{5.4}$$

If a natural logarithm is applied twice, the following linear expression is obtained:

$$\ln[-\ln(1 - P_i)] = k\ln(v_i) - k\ln c \tag{5.5}$$

By considering $\ln[-\ln(1 - P_i)]$ as the dependent variable and $\ln(v_i)$ the independent one, k becomes the slope of the line and $-k\ln c$ the intercept with the y-axis. This way, for every year at T1 and T3 a linear fit was performed for 200 values of v_i obtained by defining wind speed intervals with an amplitude of 0.1 m/s along a speed range between 1.5 and 21.5 m/s. With them and their corresponding values of P_i the 139 annual values of k and c were obtained.

First of all, the long-term variability of these parameters was explored. Figure 5.7 shows the resulting wind speed distributions when adjusted to a Weibull distribution for T1 and T3. It can be seen that, through the whole period (1871-2009), the wind speed PDFs change differently in the two considered stations. Thus, at T1 the wind speed mode (the curve maximum) and the tail of the curve are reduced, which evidences a decrease in high and intermediate wind speed frequencies. On the contrary, in the T3 Weibull curve the frequency of intermediate (high) speeds is reduced (increased).

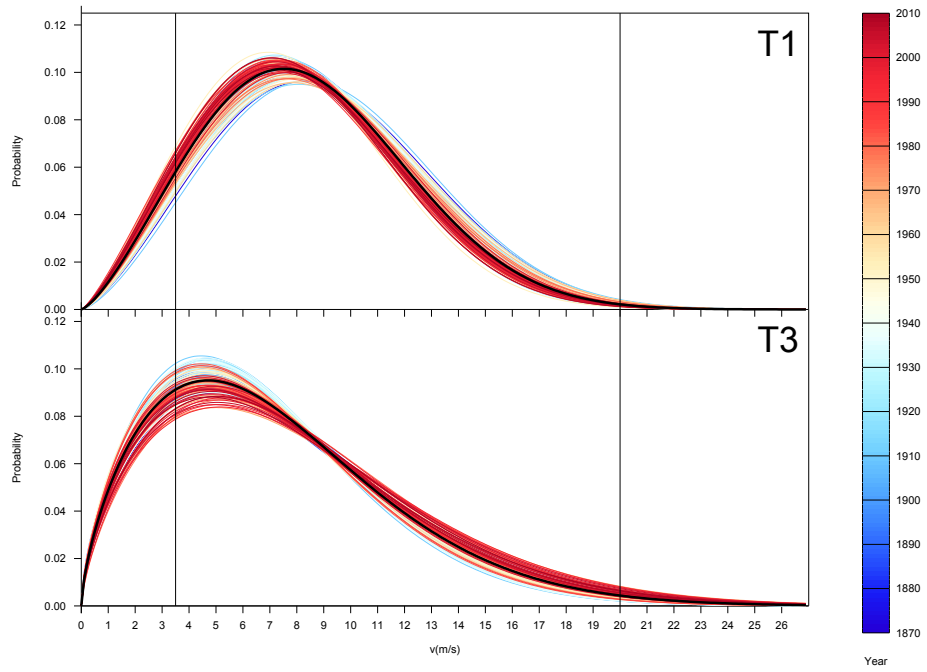

Figure 5.7: Set of reconstructed annual Weibull probability density functions of the wind speed distribution at T1 (top) and T3 (bottom), depicted according to time.

These morphological changes can be observed through the long-term variability of the Weibull parameters. In this sense, a statistically significant ($p<0.05$) negative (positive) trend was detected for c time series at T1 (T3). This implies a sharpening (flattering) of the corresponding distribution, as seen in the figure. A negative trend (also significant) is detected in the annual series of k for both towers. Usually, a decrease in k values should imply a translation to the left of the distribution maximum, which is evidenced at T1. However, at T3 the contribution of k appears to be overshadowed by the larger increment of c.

The difference between the observed period and the reconstructed one has also been calculated in order to assess the variability range between the two periods, showing results consistent with the linear trends, and obtaining statistically significant values within all differences. Finally, a spectral analysis through a Fourier transform revealed the existence of low frequency cycles, one of 27 years for k (T3) and 13 years for c (both towers). The origins of these cycles remain unknown and are matter of further work. All these results are summarized in Table 5.1.

Table 5.1: Statistics for the annual series of the Weibull parameters derived of the obtained wind speed annual PDFs, for T1 and T2.

	$T1$		$T3$	
	k	c	k	c
Mean	2.37	9.51	1.67	8.16
Std. Dev.	0.04	0.31	0.03	0.39
Lin. trend (%)	-2.31	-2.51	-1.56	+5.01
$Obs - Rec$ (%)	-1.37	-1.77	-1.13	+3.27
$Cycle(yrs)$	—	13	27	13

Once the multidecadal variability of the Weibull parameters was analyzed, its degree of coupling to the P_o signal was explored. The analytical relationship of k and c with P_o of a certain site is defined as:

$$P(k,c) = \frac{1}{2}\rho S c^3 \Gamma \left[1 + \frac{3}{k}\right] \tag{5.6}$$

where Γ stands for the Gamma function. If this expression is employed to compute P_o by considering the k and c values derived from the reconstructed annual PDFs, the obtained P_o values differ notably with respect to those observed at the simulation presented and analyzed before (RMSEs of 130 and 98 Kw for T1 and T3 resp.). Additionally, the annual P_o was estimated by employing the Weibull-fitted PDF and the turbine Power Curve, and the error diminished to half at T3 (53 Kw) although it did not substantially decreased at T1 (115 Kw).

In order to obtain a more efficient approach to P_o by only considering c and k, a multilinear fit was computed by regressing the 139 annual values of k and c on P_o. First, we computed linear regression coefficients between the P_o series and the respective c and k values. The contribution of c ($r^2 = 0.95$ at both towers) was clearly higher than that for k ($r^2 = 0.52$ at T1, $r^2 = 0.18$ at T3), which is consistent with Eq.5.6.

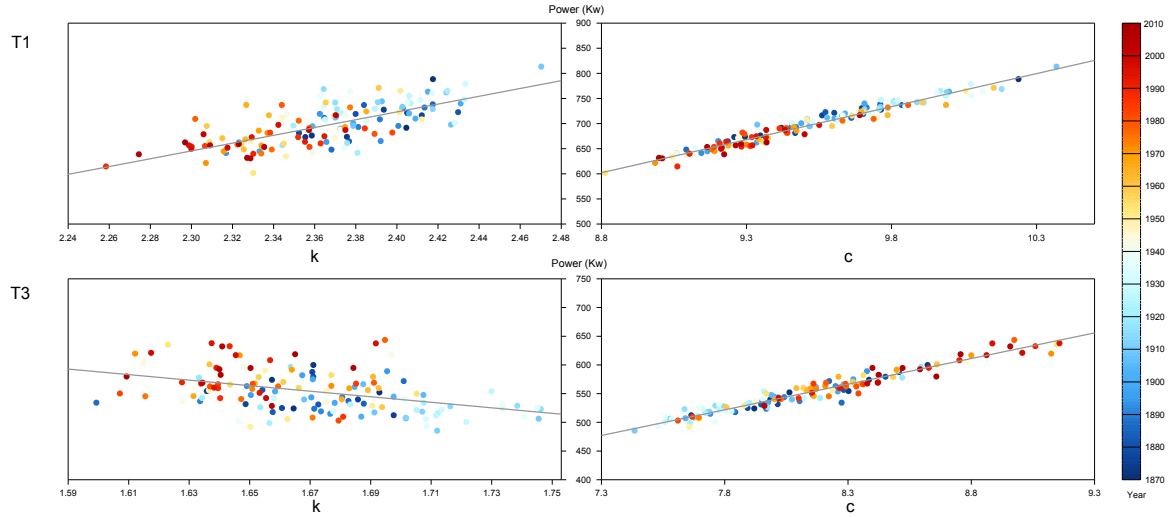

Figure 5.8: Linear fitting for the ratio between the obtained annual Weibull paramenters k (left) and c (right) and the annual P_o, for tower T1 (up) and T3 (down).

Figure 5.8 shows the obtained linear fits. There is a clear differentiation between remote (blue dots) and recent years (red dots), which are oppositely located in all cases, in correspondence with c and k trends. Both c and k are positively related to wind speed (and thus to wind power, Eq.5.6). However, the univariate linear correlation with wind power shows a negative relationship for k at T3. This happens because the impact of k is overshadowed by the high variability explained by c. Finally we computed a bivariate fit. The obtained expressions for the annual P_o were:

$$T1 : P_1(k, c) = -989 + 256k + 113c(Kw) \tag{5.7}$$

$$T3 : P_2(k, c) = -757 + 281k + 103c(Kw) \tag{5.8}$$

which denote a positive contribution of both parameters at both towers. This is consistent with the definition at Eq.5.6. The fitted expression showed r^2 values of 0.99 and 0.98 with the obtained P_o previously computed for T1 and T3 respectively, reflecting that this simple equation can describe the P_o variability in terms of c and k only.

The values of the RMSE between the annual simulated power and that estimated by these multilinear regressions (3.3 and 4.8 Kw for T1 and T3 resp.) was significantly smaller than those obtained through Eq.5.6 or through the PDF curves fitted to Weibull.

For this reason, this multiple linear fit represents a simple and accurate way to correct the estimations of the annual P_o of an hypothetical wind turbine for periods where wind data are available but wind power data are not. In addition, it can entail a useful solution at those sites where c and k values have been computed for a few years, but the wind speed PDF Weibull fitting is not a good approximation.

Chapter 6

Conclusions

The lack of an exhaustive knowledge on the multidecadal behavior of the wind power output has motivated the development of new techniques to reconstruct wind along past decades, for certain points with a wind power interest. In this thesis a series of tools for the classification and the statistical downscaling of wind has been built. The obtained results contribute to a better understanding of the surface wind and its associated wind power output variability from interannual to centennial scales. As case study the new methods have been tested and applied over six meteorological towers at wind farms spread over the Iberian Peninsula. The main conclusions raised in this work can be resumed in the following issues:

- By employing the daily SLP field from different reanalyses, a new statistical downscaling methodology based on clustering the synoptic configuration of the atmospheric circulation has been developed according to different geostrophic indices over a certain point of wind energy interest. The clustering problem has been addressed by employing different soft-computing algorithms (two different versions of a non-deterministic, evolutionary algorithm and one a deterministic, greedy algorithm). The introduced tools provided daily wind classifications from the synoptic circulation over a specific region, with a statistically significant smaller pressure field dispersion than that obtained by the Circulation Weather Types (Jenkinson and Collison, 1977) (WT).

- The three surface wind downscalings of the introduced methods performed better than approaches like WT and the NCEP/NCAR reanalysis (V.2) wind in terms of daily wind speed uncertainty reduction, with average errors 18 and 26% lower. The same happened with the correlation coefficients, with values 0.09 and 0.23 higher than NCEP/NCAR and WT respectively. At monthly scales the performance differences with the reference approaches (WT and reanalysis) became even higher, providing Pearson Correlation values up to 0.92. The methodology introduced shows an overall high performance compared to that shown by the existing bibliography on statistically downscaling wind.

- The introduced methods resulted appropriate to describe in detail the daily surface wind in terms of its associated large-scale circulation, providing a more clear wind classification than WT. Although surface wind and its associated geostrophic approach can present remarkable differences, these methods have demonstrated to retain similar wind conditions among elements within a same Wind Type.

- The new methods allowed the reconstruction of surface daily wind, at the considered locations throughout nearly 140 years (1871-2009). These obtained centennial wind series led to the following results:

    1. A negative significant trend from the second half of the 20th century is observed for all towers at the Central Iberian Plateau (CIP), while this trend is positive at the Gibraltar Strait Area (GSA). These evolutions are opposite at all towers during the first part of the time series (1871-1945).
    2. Statistically significant differences have been observed within the wind speed PDF and wind rose morphologies at interannual and interdecadal scales, being higher in GSA. The annual frequency of the most prevalent wind direction at CIP (SW) showed a significant correlation with wind speed. Regarding the Gibraltar Strait Area, the strong relationship between annual wind speed and SE direction appears to have become non-stationary, having diminished its correlation from 0.51 (until 1960) to 0.29 (non significant). This means that, although GSA wind speed appears to be rising in the last decades, this increase would be not associated to an increase of easterly winds.

- The wind speed statistical downscaling methods allowed to investigate the long-term variability of wind power output at the two representative regions. The main remarks on the annual variability of the wind power output for a market wind turbine are summarized as follows:

1. A wind speed significant periodicity in the 23-yr band was detected within five of the six considered locations. Additionally, cycles of 46 and 69 years (multiplicities of the 23-yr) were detected for GSA and in some CIP locations respectively. Similar results were obtained also for the annual series of the Weibull scale parameter c, wind speed frequency beyond the wind turbine technical operation limits (Cut-In and Cut-Out) and within the obtained Wind Power Output simulations. Although the origin of this pattern of wind variability is still unclear and subject of current research, interestingly a variability period in the 25y band has also been detected in the observed wind direction over the English Channel (Barriopedro et al., 2013), suggesting that this multidecadal variability may not result from an artifact derived from the reconstruction process.

2. The computed annual output power shows a different long-term variability depending on the location. In this way, a significant long-term power output decrease (increase) was observed in CIP (GSA). On a seasonal basis, although DJF is typically the season with the highest contribution to power output (and so occurs at CIP), no significant seasonality is detected at GSA.

3. CIP showed an overall higher and more non-stationary impact from teleconnection patterns than GSA. Wind power at CIP showed statistically significant negative correlations with NAO at all seasons, although no signal was detected at GSA. The opposite behavior of SCAND (positive at CIP, negative at GSA) is consistent with its spatial structure over Iberia, where an anticyclonic (cyclonic) system in its negative (positive) mode is located over Spain. The high variability in the relationship between the output power and large-scale climate should be taken into account when downscaling models at seasonal and larger time scales.

4. The reconstruction of the wind speed frequencies beyond the lower ($Cut\text{-}In$) and upper ($Cut\text{-}Out$) wind turbine technical limits showed consistency with the power output results. While in CIP the frequency is growing below cut-in and decreasing above cut-out, GSA behaves opposite. Although both technical limits are coupled to power output results, the variability of winds below Cut-In is shown to be more strongly related with power output.

5. The availability of a 139 year PDFs series and output power values has allowed the assessment of the empirical relationship between the annual generated Power and the Weibull parameters k and c, once the obtained annual

PDFs where adjusted to that function. A multilinear regression between annual power and Weibull parameters was computed for each tower, explaining up to 98% of the variance. This approach improved the estimation of the simulated output power as compared with classical ways, resulting an appropriate model for situations where k and c data are available but a an accurate fit of the speed distribution to a Weibull function can not be guaranteed.

References

Achberger, C., Ekström, M., and Bärring, L. 2002. Estimation of local near-surface wind conditions–a comparison of WASP and regression based techniques. *Meteorological Applications*, 9(2):211–221. p. 20, 37

Ackermann, T. et al. 2005. *Wind power in power systems*, volume 140. Wiley Online Library. p. 1

Allan, R. and Ansell, T. 2006. A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850-2004. *Journal of Climate*, 19(22):5816–5842. p. 15

Allen, J. T., Pezza, A. B., and Black, M. T. 2010. Explosive cyclogenesis: A global climatology comparing multiple reanalyses. *Journal of Climate*, 23(24):6468–6484. p. 67

Anderberg, M. R. 1973. Cluster analysis for applications. Technical report, DTIC Document. p. 20

Anderson, P. and Bose, A. 1983. Stability simulation of wind turbine systems. *Power Apparatus and Systems, IEEE transactions on*, (12):3791–3795. p. 2, 24, 94

Arakawa, A. 1966. Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. *Journal of Computational Physics*, 1(1):119–143. p. 22

Archer, C. L. and Jacobson, M. Z. 2003. Spatial and temporal distributions of US winds and wind power at 80 m derived from measurements. *Journal of Geophysical Research*, 108(D9):4289. p. 14

Aspray, W. 1990. *John von Neumann and the origins of modern computing*, volume 191. Mit Press Cambridge, MA. p. 15

AWS Truepower 2012. *Europe Wind Resource Map*. AWS Truepower. p. 13

Babyak, M. A. 2004. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine*, 66(3):411–421. p. 18

Back, T., Fogel, D. B., and Michalewicz, Z. 1997. *Handbook of evolutionary computation*. IOP Publishing Ltd. p. 21

Bäck, T. and Schwefel, H.-P. 1993. An overview of evolutionary algorithms for parameter optimization. *Evolutionary computation*, 1(1):1–23. p. 27

Badger, J., Larsén, X. G., Mortensen, N. G., Hahmann, A. N., Hansen, J. C., and Ejsing Jørgensen, H. 2010. A universal mesoscale to microscale modelling interface tool. p. 16

Baik, J.-J., Kim, J.-J., and Fernando, H. J. 2003. A CFD model for simulating urban flow and dispersion. *Journal of Applied Meteorology*, 42(11):1636–1648. p. 10

Baker, R. W., Hewson, E. W., Butler, N. G., and Warchol, E. J. 1978. Wind power potential in the Pacific Northwest. *Journal of Applied Meteorology*, 17(12):1814–1826. p. 14

Barnett, T. and Preisendorfer, R. 1987. Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Monthly Weather Review*, 115(9):1825–1850. p. 17

Barnston, A. G. and Livezey, R. E. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly weather review*, 115(6):1083–1126. p. 4, 99

Barriopedro, D., Gallego, D., Alvarez-Castro, M. C., García-Herrera, R., Wheeler, D., Peña-Ortiz, C., and Barbosa, S. M. 2013. Witnessing North Atlantic westerlies variability from ships logbooks (1685–2008). *Climate Dynamics*, pages 1–17. p. XXVI, 11, 115

Barth, N. H. 1992. Oceanographic experiment design II: Genetic algorithms. *Journal of Atmospheric and Oceanic Technology*, 9(4):434–443. p. 22

Bednorz, W. 2008. Advances in greedy algorithms. *Wienna: I-Tech Education and Publishing KG*. p. 21, 32, 36

Bell, G., Wei, S., L'Heureux, M., and M, H. Northern hemisphere teleconnection patterns. *Climate Diagnostics Bulletin*, (5). p. 99

Bellecci, C., Colacino, M., Gaudio, P., Aversa, P., Casella, L., Federico, S., and Feudo, T. L. 2001. Influence of orography on water vapor profiles over complex terrain in southern Italy. In *Europto Remote Sensing*, pages 64–74. International Society for Optics and Photonics. p. 9

Beranová, R. and Huth, R. 2008. Time variations of the effects of circulation variability modes on European temperature and precipitation in winter. *International Journal of Climatology*, 28(2):139–158. p. 19

Bergström, H. 1996. A Climatological Study of Boundary Layer Wind Speed Using a Meso-Scale High Order Closure Model. *Journal of Applied Meteorology*, 35:1291–1306. p. 16, 32, 33, 59

Bladé, I., Liebmann, B., Fortuny, D., and van Oldenborgh, G. J. 2012. Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region. *Climate dynamics*, 39(3-4):709–727. p. 7

Bogardi, I. and Matyasovzky, I. 1996. Estimating daily wind speed under climate change. *Solar Energy*, 57(3):239–248. p. 20, 33, 37, 59, 107

Bormans, M., Garrett, C., and Thompson, K. 1986. Seasonal variability of the surface inflow through the strait-of-gibraltar. *Oceanologica Acta*, 9(4):403–414. p. 7, 61

Bouzgou, H. and Benoudjit, N. 2011. Multiple architecture system for wind speed prediction. *Applied Energy*, 88(7):2463–2471. p. 21, 60

Brabson, B. and Palutikof, J. 2000. Tests of the generalized Pareto distribution for predicting extreme wind speeds. *Journal of applied meteorology*, 39(9):1627–1640. p. 104

Brayshaw, D. J., Troccoli, A., Fordham, R., and Methven, J. 2011. The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: A case study over the UK. *Renewable Energy*, 36(8):2087–2096. p. 12, 24, 25, 91

Breiman, L. 2001. Random forests. *Machine learning*, 45(1):5–32. p. 21

Bretherton, C. S., Smith, C., and Wallace, J. M. 1992. An intercomparison of methods for finding coupled patterns in climate data. *Journal of climate*, 5(6):541–560. p. 17

Bronnimann, S., Martius, O., Von Waldow, H., Welker, C., Luterbacher, J., Compo, G., Sardeshmukh, P., and Usbeck, T. 2012. Extreme winds at northern mid-latitudes since 1871. *Meteorologische Zeitschrift*, 21(1):13–27. p. 49

Buczak, A. L. and Barrett, J. J. 1998. Genetic algorithm method for determining temperature profiles. In *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on*, pages 142–147. IEEE. p. 22

Burlando, M. 2009. The synoptic-scale surface wind climate regimes of the Mediterranean Sea according to the cluster analysis of ERA-40 wind fields. *Theoretical and applied climatology*, 96(1-2):69–83. p. 19

Busuioc, A., Tomozeiu, R., and Cacciamani, C. 2008. Statistical downscaling model based on canonical correlation analysis for winter extreme precipitation events in the Emilia-Romagna region. *International Journal of Climatology*, 28(4):449–464. p. 17

Byrkjedal, Ø. and Berge, E. 2008. The use of WRF for wind resource mapping in Norway. In *The 9th WRF Users' Workshop, NCAR Green Campus. http://www. mmm. ucar. edu/wrf/users/workshops/WS2008/WorkshopProgram. pdf.* p. 12

Cadenas, E. and Rivera, W. 2009. Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. *Renewable Energy*, 34(1):274–278. p. 11

Cai, M. and Mak, M. 1990. Symbiotic relation between planetary and synoptic-scale waves. *Journal of the atmospheric sciences*, 47(24):2953–2968. p. 3

Candlish, G., Law, N., Bett, P., Clark, R., Thornton, H., and Wilson, C. 2012. Interannual wind variation from observations and numerical weather analyses. p. 49

Carro-Calvo, L., Salcedo-Sanz, S., Kirchner-Bossi, N., Portilla-Figueras, A., Prieto, L., Garcia-Herrera, R., and Hernández-Martín, E. 2011. Extraction of synoptic pressure patterns for long-term wind speed estimation in wind farms using evolutionary computing. *Energy*, 36(3):1571–1581. p. 38

Carro-Calvo, L., Salcedo-Sanz, S., Prieto, L., Kirchner-Bossi, N., Portilla-Figueras, A., and Jiménez-Fernández, S. 2012. Wind speed reconstruction from synoptic pressure patterns using an evolutionary algorithm. *Applied Energy*, 89(1):347–354. p. 60

Carta, J., Ramirez, P., and Velázquez, S. 2009. A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews*, 13(5):933–955. p. 94

Cavazos, T. 2000. Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans. *Journal of Climate*, 13(10):1718–1732. p. 17

Cebeci, T. and Bradshaw, P. 1977. Momentum transfer in boundary layers. *Washington, DC, Hemisphere Publishing Corp.; New York, McGraw-Hill Book Co., 1977. 407 p.*, 1. p. 3, 8

Celik, A. N. 2004. A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey. *Renewable energy*, 29(4):593–604. p. 2, 23, 92

Chang, E. K. 1993. Downstream development of baroclinic waves as inferred from regression analysis. *Journal of the atmospheric sciences*, 50(13):2038–2053. p. 5

Chang, E. K., Lee, S., and Swanson, K. L. 2002. Storm track dynamics. *Journal of climate*, 15(16):2163–2183. p. 5

Chen, Y.-W., Mendoza, N., Uehara, S., Nakao, Z., Adachi, T., and Masuda, Y. 2001. Estimating wind speed in lower atmosphere wind profiler based on genetic algorithms. In *Instrumentation and Measurement Technology Conference, 2001. IMTC 2001. Proceedings of the 18th IEEE*, volume 2, pages 1258–1263. IEEE. p. 22

Cliff, W. C. 1977. Effect of generalized wind characteristics on annual power estimates from wind turbine generators. Technical report, Battelle Pacific Northwest Labs., Richland, WA (USA). p. 24

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R., Yin, X., Gleason, B., Vose, R., Rutledge, G., Bessemoulin, P., et al. 2011. The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, 137(654):1–28. p. XXII, XXVII, 15, 48, 49, 96, 97

Corby, G. 1954. The airflow over mountains. A review of the state of current knowledge. *Quarterly Journal of the Royal Meteorological Society*, 80(346):491–521. p. 9

Corte-Real, J., Zhang, X., and Wang, X. 1995. Downscaling GCM information to regional scales: a non-parametric multivariate regression approach. *Climate Dynamics*, 11(7):413–424. p. 17

Counihan, J. 1975. Adiabatic atmospheric boundary layers: a review and analysis of data from the period 1880–1972. *Atmospheric Environment (1967)*, 9(10):871–905. p. 93

Crane, R. G. and Hewitson, B. C. 1998. Doubled co2 precipitation changes for the susquehanna basin: down-scaling from the genesis general circulation model. *International Journal of Climatology*, 18(1):65–76. p. 17

Crochet, P. 2004. Adaptive Kalman filtering of 2-metre temperature and 10-metre windspeed forecasts in Iceland. *Meteorological Applications*, 11(2):173–187. p. 11

Curry, C. L., van der Kamp, D., and Monahan, A. H. 2012. Statistical downscaling of historical monthly mean winds over a coastal region of complex terrain. I. Predicting wind speed. *Climate dynamics*, 38(7-8):1281–1299. p. 21, 37, 59, 107

Dafka, S., Xoplaki, E., Garcia-Bustamante, E., Toreti, A., Zanis, P., and Luterbacher, J. 2013. The Etesian wind system and wind energy potential over the Aegean Sea. In *EGU General Assembly Conference Abstracts*, volume 15, page 10133. p. 49

Davis, R. E., Hayden, B. P., Gay, D. A., Phillips, W. L., and Jones, G. V. 1997. The North Atlantic subtropical anticyclone. *Journal of Climate*, 10(4):728–744. p. 67

Davis, R. E. and Kalkstein, L. S. 1990. Development of an automated spatial synoptic climatological classification. *International Journal of Climatology*, 10(8):769–794. p. 19

Davy, R. J., Woods, M. J., Russell, C. J., and Coppin, P. A. 2010. Statistical downscaling of wind variability from meteorological fields. *Boundary-layer meteorology*, 135(1):161–175. p. 20, 21, 28, 34, 36, 59

de Pedraza, L. and Nacional, E. S. M. 1964. *La Predicción del tiempo en el Valle del Ebro*. Servicio Metereológico Nacional: Publicaciones, serie A. Ministerio del Aire, Servicio Metereológico Nacional. p. 7

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., et al. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597. p. 48

DeGaetano, A. T. 1997. A quality-control routine for hourly wind observations. *Journal of Atmospheric and Oceanic Technology*, 14(2):308–317. p. 14

DeVore, R. A. and Temlyakov, V. N. 1996. Some remarks on greedy algorithms. *Advances in computational Mathematics*, 5(1):173–187. p. 21

Donat, M., Renggli, D., Wild, S., Alexander, L., Leckebusch, G., and Ulbrich, U. 2011. Reanalysis suggests long-term upward trends in European storminess since 1871. *Geophysical Research Letters*, 38(14). p. 93

Dorman, C. E., Beardsley, R. C., and Limeburner, R. 1995. Winds in the Strait of Gibraltar. *Quarterly Journal of the Royal Meteorological Society*, 121(528):1903–1921. p. 61

Dunst, M. 1982. On the vertical structure of the eddy diffusion coefficient in the PBL. *Atmospheric Environment (1967)*, 16(9):2071–2074. p. 9

Earl, N., Dorling, S., Hewston, R., and von Glasow, R. 2013. 1980–2010 Variability in UK Surface Wind Climate. *Journal of Climate*, 26(4):1172–1191. p. 2, 12, 14, 23

Eiben, A. E. and Smith, J. E. 2010. *Introduction to evolutionary computing*, volume 2. Springer Berlin. p. 33

Eischeid, J. K., Bruce Baker, C., Karl, T. R., and Diaz, H. F. 1995. The quality control of long-term climatological data using objective data analysis. *Journal of Applied Meteorology*, 34(12):2787–2795. p. 14

Ekström, M. 2002. Estimating monthly surface winds for Scania, southern Sweden, using geostrophic wind (1899–1997). *Geografiska Annaler: Series A, Physical Geography*, 84(2):113–126. p. 12, 18, 20, 37, 48, 59

Enfield, D. B., Mestas-Nuñez, A. M., and Trimble, P. J. 2001. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophysical Research Letters*, 28(10):2077–2080. p. 99

Esteban, P., Martin-Vide, J., and Mases, M. 2006. Daily atmospheric circulation catalogue for Western Europe using multivariate techniques. *International journal of climatology*, 26(11):1501–1515. p. 19, 98

Etemadi, H., Samadi, S., and Sharifikia, M. 2013. Uncertainty analysis of statistical downscaling models using general circulation model over an international wetland. *Climate Dynamics*, pages 1–22. p. 23

Etienne, C., Lehmann, A., Goyette, S., Lopez-Moreno, J.-I., and Beniston, M. 2010. Spatial predictions of extreme wind speeds over Switzerland using generalized additive models. *Journal of Applied Meteorology and Climatology*, 49(9):1956–1970. p. 3

Fan, H. and Sailor, D. J. 2005. Modeling the impacts of anthropogenic heating on the urban climate of Philadelphia: a comparison of implementations in two PBL schemes. *Atmospheric Environment*, 39(1):73–84. p. 10

Fang, S., Wang, M., Qi, W., and Zheng, F. 2008. Hybrid genetic algorithms and support vector regression in forecasting atmospheric corrosion of metallic materials. *Computational Materials Science*, 44(2):647–655. p. 22

Faucher, M., Burrows, W. R., and Pandolfo, L. 1999. Empirical-statistical reconstruction of surface marine winds along the western coast of Canada. *Climate research*, 11(3):173–190. p. 21, 28, 34, 36, 59

Fernández-Montes, S., Rodrigo, F. S., Seubert, S., and Sousa, P. M. 2012. Spring and summer extreme temperatures in Iberia during last century in relation to circulation types. *Atmospheric Research*. p. 98

Fertig, E., Apt, J., Jaramillo, P., and Katzenstein, W. 2012. The effect of long-distance interconnection on wind power variability. *Environmental Research Letters*, 7(3):034017. p. 23

Fiedler, F. and Panofsky, H. A. 1970. Atmospheric scales and spectral gaps. *Bulletin of the American Meteorological Society*, 51:1114–1120. p. 3, 9

Fogel, D. B. 2006. *Evolutionary computation: toward a new philosophy of machine intelligence*, volume 1. Wiley. com. p. 21, 27

Frehlich, R. 2013. Scanning Doppler Lidar for Input into Short-Term Wind Power Forecasts. *Journal of Atmospheric and Oceanic Technology*, 30(2):230–244. p. 23

Frey-Buness, F., Heimann, D., and Sausen, R. 1995. A statistical-dynamical downscaling procedure for global climate simulations. *Theoretical and Applied Climatology*, 50(3-4):117–131. p. 16

Galanis, G., Louka, P., Katsafados, P., Pytharoulis, I., and Kallos, G. 2006. Applications of Kalman filters based on non-linear functions to numerical weather predictions. In *Annales geophysicae*, volume 24, pages 2451–2460. Copernicus GmbH. p. 11

García-Bustamante, E., González-Rouco, J., Jiménez, P., Navarro, J., and Montávez, J. 2008. The influence of the Weibull assumption in monthly wind energy estimation. *Wind Energy*, 11(5):483–502. p. 24, 94

García-Bustamante, E., González-Rouco, J., Navarro, J., Xoplaki, E., Jiménez, P., and Montávez, J. 2012. North Atlantic atmospheric circulation and surface wind in the Northeast of the Iberian Peninsula: uncertainty and long term downscaled variability. *Climate dynamics*, 38(1-2):141–160. p. 7, 12, 17, 18, 20, 25, 28, 37, 48, 60

García-Bustamante, E., González-Rouco, J., Navarro, J., Xoplaki, E., Luterbacher, J., Jiménez, P., Montávez, J., Hidalgo, A., and Lucio-Eceiza, E. 2013. Relationship between wind power production and North Atlantic atmospheric circulation over the northeastern Iberian Peninsula. *Climate Dynamics*, 40(3-4):935–949. p. 24, 91, 92

García-Herrera, R., Díaz, J., Trigo, R., and Hernández, E. 2005. Extreme summer temperatures in Iberia: health impacts and associated synoptic conditions. In *Annales Geophysicae*, volume 23, pages 239–251. Copernicus GmbH. p. 98

García-Valero, J. A., Montavez, J. P., Jerez, S., Gómez-Navarro, J. J., Lorente-Plazas, R., and Jiménez-Guerrero, P. 2012. A seasonal study of the atmospheric dynamics over the Iberian Peninsula based on circulation types. *Theoretical and Applied Climatology*, 110(1-2):291–310. p. 19

Garratt, J. R. 1994. *The atmospheric boundary layer*. Cambridge university press. p. 3

Gastón, M., Pascal, E., Frıas, L., Martı, I., Irigoyen, U., Cantero, E., Lozano, S., and Loureiro, Y. 2008. Wind resources map of Spain at mesoscale. Methodology and validation. In *Proceedings of European Wind Energy Conference*. p. 7

Gates, W. L., Henderson-Sellers, A., Boer, G. J., Folland, C. K., Kitoh, A., McAvaney, B. J., Semazzi, F., Smith, N., Weaver, A. J., and Zeng, Q. 1996. *Climate models–evaluation*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. p. 16

Gerstengarbe, F., Werner, P., Busold, W., Rüge, U., Wegener, K., and Deutscher Wetterdienst-Zentralamt, O. a. M. G. 1993. *Catalogue of'Grosswetterlagen'in Europe after Paul Hess and Helmuth Brezowski, 1881-1992*. Deutscher Wetterdienst. p. 32

Gerstengarbe, F.-W., Werner, P., and Fraedrich, K. 1999. Applying non-hierarchical cluster analysis algorithms to climate classification: some problems and their solution. *Theoretical and Applied Climatology*, 64(3-4):143–150. p. 32

Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., and Draxl, C. 2011. The state-of-the-art in short-term prediction of wind power: A literature overview. Technical report, ANEMOS. plus. p. 11

Glahn, H. R. and Lowry, D. A. 1972. The use of model output statistics (MOS) in objective weather forecasting. *Journal of applied meteorology*, 11(8):1203–1211. p. 16

Goubanova, K., Echevin, V., Dewitte, B., Codron, F., Takahashi, K., Terray, P., and Vrac, M. 2011. Statistical downscaling of sea-surface wind over the Peru–Chile upwelling region: diagnosing the impact of climate change from the IPSL-CM4 model. *Climate Dynamics*, 36(7-8):1365–1378. p. 20, 36, 48, 91

Goyette, S. 2011. Synoptic conditions of extreme windstorms over Switzerland in a changing climate. *Climate Dynamics*, 36(5-6):845–866. p. 67

Green, M. C., Flocchini, R. G., and Myrup, L. O. 1992. The relationship of the extinction coefficient distribution to wind field patterns in southern California. *Atmospheric Environment. Part A. General Topics*, 26(5):827–840. p. 19

Grotch, S. L. and MacCracken, M. C. 1991. The use of general circulation models to predict regional climatic change. *Journal of Climate*, 4(3):286–303. p. 16

Hahmann, A. N., Lange, J., Pena Diaz, A., and Bay Hasager, C. 2012. The NORSEWInD numerical wind atlas for the South Baltic. p. 16

Hanssen-Bauer, I., Achberger, C., Benestad, R., Chen, D., and Forland, E. 2005. Statistical downscaling of climate scenarios over Scandinavia. *Climate Research*, 29(3):255. p. 17

Harper, B. R., Katz, R. W., and Harriss, R. C. 2007. Statistical methods for quantifying the effect of the El Niño–Southern Oscillation on wind power in the northern Great Plains of the United States. *Wind Engineering*, 31(3):123–137. p. 12, 24, 25, 91

Hartigan, J. A. and Wong, M. A. 1979. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108. p. 19

Hau, E. 2000. *Windturbines*. Springer Berlin. p. 94

Hawkins, D. M. 2004. The problem of overfitting. *Journal of chemical information and computer sciences*, 44(1):1–12. p. 18

Hennessey Jr, J. P. 1978. A comparison of the Weibull and Rayleigh distributions for estimating wind power potential. *Wind Engineering*, 2:156–164. p. 24

Hess, P. and Brezowsky, H. 1952. Katalog der Großwetterlagen Europas (Catalog of the European Large Scale Weather Types). *Ber. Dt. Wetterd, in der US-Zone*, 33. p. 19

Hess, P. and Brezowsky, H. 1969. *Katalog der grosswetterlagen Europas*, volume 15. Deutscher Wetterdienst. p. 32

Hewitson, B. 1994. Regional climates in the GISS general circulation model: Surface air temperature. *Journal of climate*, 7(2). p. 17

Hewitson, B. C. and Crane, R. G. 1992. Regional-scale climate prediction from the GISS GCM. *Global and planetary change*, 5(3):249–267. p. 17

Heyen, H., Zorita, E., and Von Storch, H. 1996. Statistical downscaling of monthly mean North Atlantic air-pressure to sea level anomalies in the Baltic Sea. *Tellus A*, 48(2):312–323. p. 17

Hoinka, K. P. and Castro, M. D. 2003. The Iberian peninsula thermal low. *Quarterly Journal of the Royal Meteorological Society*, 129(590):1491–1511. p. 67

Holtslag, A. and De Bruin, H. 1988. Applied modeling of the nighttime surface energy balance over land. *Journal of Applied Meteorology*, 27(6):689–704. p. 9

Hong, S.-Y. and Lim, J.-O. J. 2006. The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc*, 42(2):129–151. p. 10

Hopfield, J. J. 1988. Artificial neural networks. *Circuits and Devices Magazine, IEEE*, 4(5):3–10. p. 21

Hopke, P. K. and Kaufman, L. 1990. The use of sampling to cluster large data sets. *Chemometrics and intelligent laboratory systems*, 8(2):195–204. p. 19

Hošek, J., Jež, J., Svoboda, J., and Štekl, J. 2004. Comparison of the mean wind speed fields computed by three models over the area of the Czech Republic. *DEWI Magazin*, 24:66–71. p. 14

Hoskins, B., McIntyre, M., and Robertson, A. W. 1985. On the use and significance of isentropic potential vorticity maps. *Quarterly Journal of the Royal Meteorological Society*, 111(470):877–946. p. 5

Houghton, E. and Carruthers, N. 1976a. *Wind Forces on Buildings and Structures: An Introduction*. Wiley. p. 10

Houghton, E. and Carruthers, N. 1976b. *Wind Forces on Buildings and Structures: An Introduction*. Wiley. p. 11

Hundecha, Y. and Bárdossy, A. 2008. Statistical downscaling of extremes of daily precipitation and temperature and construction of their future scenarios. *International Journal of Climatology*, 28(5):589–610. p. 17

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M. 2003. *The North Atlantic Oscillation: climatic significance and environmental impact*, volume 134. American Geophysical Union. p. 5, 99

Huth, R. 1999. Statistical downscaling in central Europe: Evaluation of methods and potential predictors. *Climate Research*, 13(2):91–101. p. 16

Jakob Themeßl, M., Gobiet, A., and Leuprecht, A. 2011. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, 31(10):1530–1544. p. 10

James, P. 2007. An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe. *Theoretical and Applied Climatology*, 88(1-2):17–42. p. 32

Jamil, M., Parsa, S., and Majidi, M. 1995. Wind power statistics and an evaluation of wind energy density. *Renewable energy*, 6(5):623–628. p. 24, 94

Jenkinson, A. and Collison, F. 1977. An initial climatology of gales over the North Sea. *Synoptic Climatology Branch Memorandum*, 62:18. p. XXI, XXVI, 19, 28, 30, 51, 113

Jerez, S. and Trigo, R. 2013. Time-scale and extent at which large-scale circulation modes determine the wind and solar potential in the Iberian Peninsula. *submitted*, (2013). p. 7, 16, 24, 25, 91

Jerez, S., Trigo, R., Vicente-Serrano, S., Pozo-Vázquez, D., Lorente-Plazas, R., Lorenzo-Lacruz, J., Santos-Alamillos, F., and Montávez, J. 2013. The impact of the North Atlantic Oscillation on the renewable energy resources in south-western Europe. *Journal of Applied Meteorology and Climatology*, 52:2204–2225. p. 5, 6, 7, 8, 16, 23, 24, 25, 61, 91

Jiménez, P. A., González-Rouco, J. F., Montávez, J. P., García-Bustamante, E., and Navarro, J. 2009. Climatology of wind patterns in the northeast of the Iberian Peninsula. *International Journal of Climatology*, 29(4):501–525. p. 19

Jiménez, P. A., González-Rouco, J. F., Montávez, J. P., García-Bustamante, E., Navarro, J., and Dudhia, J. 2012. Analysis of the long-term surface wind variability over complex terrain using a high spatial resolution WRF simulation. *Climate Dynamics*, pages 1–14. p. 16

Jones, C., Gregory, J., Thorpe, R., Cox, P., Murphy, J., Sexton, D., and Valdes, P. 2005. Systematic optimisation and climate simulation of FAMOUS, a fast version of HadCM3. *Climate dynamics*, 25(2-3):189–204. p. 15

Jones, P., Harpham, C., and Briffa, K. 2012. Lamb weather types derived from reanalysis products. *International Journal of Climatology*. p. 71

Joyce, T. M., Kwon, Y.-O., and Yu, L. 2009. On the relationship between synoptic wintertime atmospheric variability and path shifts in the Gulf Stream and the Kuroshio Extension. *Journal of Climate*, 22(12):3177–3192. p. 5

Jungo, P., Goyette, S., and Beniston, M. 2002. Daily wind gust speed probabilities over Switzerland according to three types of synoptic circulation. *International journal of climatology*, 22(4):485–499. p. 3

Jursa, R. and Rohrig, K. 2008. Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models. *International Journal of Forecasting*, 24(4):694–709. p. 22

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, 77(3):437–471. p. XXI, XXVI, 15, 48, 51, 99

Kanamitsu, M. 1989. Description of the NMC global data assimilation and forecast system. *Weather and Forecasting*, 4(3):335–342. p. 15

Kanamitsu, M. and Kanamaru, H. 2007. Fifty-seven-year California Reanalysis Downscaling at 10 km (CaRD10). Part I: System detail and validation with observations. *Journal of Climate*, 20(22):5553–5571. p. 16

Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., and Rajagopalan, B. 1998. Analyses of global sea surface temperature 1856–1991. *Journal of Geophysical Research*, 103(C9):18567–18. p. 99

Kariniotakis, G., Halliday, J., Brownsword, R., Marti, I., Palomares, A. M., Cruz, I., Madsen, H., Nielsen, T., Nielsen, H. A., Focken, U., et al. 2006. Next Generation Short-Term Forecasting of Wind Power–Overview of the ANEMOS Project. In *Proceedings of the European Wind Energy Conference, EWEC 2006*. p. 11

Katsoulis, B. D. 1996. The relationship between synoptic, mesoscale and microscale meteorological parameters during poor air quality events in Athens, Greece. *Science of the total environment*, 181(1):13–24. p. 3

Kaufmann, P. and Whiteman, C. D. 1999. Cluster-analysis classification of wintertime wind patterns in the Grand Canyon region. *Journal of Applied Meteorology*, 38(8):1131–1147. p. 32

Kirchhofer, W. 1974. *Classification of European 500 mb patterns*. Swiss Meteorological Institute. p. 19

Kirchner-Bossi, N., Prieto, L., García-Herrera, R., Carro-Calvo, L., and Salcedo-Sanz, S. 2013. Multi-decadal variability in a centennial reconstruction of daily wind. *Applied Energy*, 105:30–46. p. 37, 71, 96

Kirchner-Bossi, N., Prieto, L., García-Herrera, R., and Trigo, R. 2014. A Long-Term perspective of Wind Power Output Variability. *Submitted*. p. 91

Kjellin, J., Bülow, F., Eriksson, S., Deglaire, P., Leijon, M., and Bernhoff, H. 2011. Power coefficient measurement on a 12 kW straight bladed vertical axis wind turbine. *Renewable energy*, 36(11):3050–3053. p. 94

Klein, W. H. and Glahn, H. R. 1974. Forecasting local weather by means of model output statistics. *Bulletin of the American Meteorological Society*, 55(10):1217–1227. p. 16

Klink, K. 2007. Atmospheric circulation effects on wind speed variability at turbine height. *Journal of applied meteorology and climatology*, 46(4):445–456. p. 25, 91

Krekov, G. and Sukhanov, A. Y. 2012. Improved genetic algorithm for multiwave lidar sensing of atmospheric aerosol. *Atmospheric and Oceanic Optics*, 25(2):130–134. p. 22

Kubik, M., Coker, P., and Hunt, C. 2011. Using meteorological wind data to estimate turbine generation output: a sensitivity analysis. *Volume 15 Wind Energy Applications*, page 4074. p. 9

Küttel, M., Luterbacher, J., and Wanner, H. 2011. Multidecadal changes in winter circulation-climate relationship in Europe: frequency variations, within-type modifications, and long-term trends. *Climate Dynamics*, 36(5-6):957–972. p. 21

Lamb, H. H. 1972. *British Isles weather types and a register of the daily sequence of circulation patterns 1861-1971*. HM Stationery Office. p. 19, 32

Landberg, L., Myllerup, L., Rathmann, O., Petersen, E. L., Jørgensen, B. H., Badger, J., and Mortensen, N. G. 2003. Wind resource estimation - an overview. *Wind Energy*, 6(3):261–271. p. 14

Lanzafame, R. and Messina, M. 2010. Horizontal axis wind turbine working at maximum power coefficient continuously. *Renewable Energy*, 35(1):301–306. p. 94

Larsén, X. G., Badger, J., Hahmann, A. N., and Mortensen, N. G. 2012. The selective dynamical downscaling method for extreme-wind atlases. *Wind Energy*. p. 9

Launder, B., Reece, G. J., and Rodi, W. 1975. Progress in the development of a Reynolds-stress turbulence closure. *Journal of fluid mechanics*, 68(03):537–566. p. 15

Lee, D. and Baldick, R. 2012. Analyzing the variability of wind power output through the power spectral density. In *Power and Energy Society General Meeting, 2012 IEEE*, pages 1–8. IEEE. p. 23

Lee, S. 1995. Localized storm tracks in the absence of local instability. *Journal of Atmospheric Sciences*, 52:977–989. p. 5

Li, G. and Shi, J. 2010. Application of Bayesian model averaging in modeling long-term wind speed distributions. *Renewable Energy*, 35(6):1192–1202. p. 24

Linderson, M.-L. 2001. Objective classification of atmospheric circulation over southern Scandinavia. *International Journal of Climatology*, 21(2):155–169. p. 19, 31

Lionello, P., Malanotte-Rizzoli, P., and Boscolo, R. 2006a. *Mediterranean climate variability*, volume 4. Access Online via Elsevier. p. 5

Lionello, P., Malanotte-Rizzoli, P., Boscolo, R., Alpert, P., Artale, V., Li, L., Luterbacher, J., May, W., Trigo, R., Tsimplis, M., et al. 2006b. The Mediterranean climate: an overview of the main characteristics and issues. *Developments in earth and environmental sciences*, 4:1–26. p. 7

Lo, J. C.-F., Yang, Z.-L., and Pielke, R. A. 2008. Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D9). p. 12

Lopez-Bustins, J.-A., Martin-Vide, J., and Sanchez-Lorenzo, A. 2008. Iberia winter rainfall trends based upon changes in teleconnection and circulation patterns. *Global and Planetary Change*, 63(2):171–176. p. 67

Lorenz, E. N. 1965. A study of the predictability of a 28-variable atmospheric model. *Tellus*, 17(3):321–333. p. 22

Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., and Kallos, G. 2008. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(12):2348–2362. p. 11

Lund, I. A. 1963. Map-pattern classification by statistical methods. *Journal of Applied Meteorology*, 2(1):56–65. p. 19

Mahrt, L. and Ek, M. 1993. Spatial variability of turbulent fluxes and roughness lengths in HAPEX-MOBILHY. *Boundary-Layer Meteorology*, 65(4):381–400. p. 9

Makarieva, A., Gorshkov, V., Sheil, D., Nobre, A., and Li, B.-L. 2010. Where do winds come from? A new theory on how water vapor condensation influences atmospheric

pressure and dynamics. *Atmospheric Chemistry and Physics Discussions*, 10(10):24015–24052. p. 9

Mamdani, E. H. and Assilian, S. 1975. An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of man-machine studies*, 7(1):1–13. p. 21

Manwell, J., McGowan, J., and Rogers, A. 2002. Wind energy explained: theory, design and application. 2002. *John Wiley&Sons Ltd, UK*, page 577. p. 23, 96

Marsh, D. R., Mills, M. J., Kinnison, D. E., Lamarque, J.-F., Calvo, N., and Polvani, L. M. 2013. Climate change from 1850 to 2005 simulated in CESM1 (WACCM). *Journal of Climate*, (2013). p. 15

Martner, B. E. and Marwitz, J. D. 1982. Wind characteristics in Southern Wyoming. *Journal of Applied Meteorology*, 21:1815–1827. p. 14

Mathew, S. 2006. *Wind energy: fundamentals, resource analysis and economics*. Springer-Verlag New York, Inc. p. 1

Matulla, C., Scheifinger, H., Menzel, A., and Koch, E. 2003. Exploring two methods for statistical downscaling of Central European phenological time series. *International Journal of Biometeorology*, 48(2):56–64. p. 17

McFarlane, N. 1987. The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *Journal of the atmospheric sciences*, 44(14):1775–1800. p. 3

McIntyre, J. H., Lubitz, W. D., and Stiver, W. H. 2011. Local wind-energy potential for the city of Guelph, Ontario (Canada). *Renewable Energy*, 36(5):1437–1446. p. 9, 92

McVicar, T. R., Van Niel, T. G., Li, L. T., Roderick, M. L., Rayner, D. P., Ricciardulli, L., and Donohue, R. J. 2008. Wind speed climatology and trends for Australia, 1975–2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophysical Research Letters*, 35(20):L20403. p. 48

Mengelkamp, H.-T. 1999. Wind climate simulation over complex terrain and wind turbine energy output estimation. *Theoretical and applied climatology*, 63(3-4):129–139. p. 19, 33, 59

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E., Berbery, E. H., et al. 2006. North American regional reanalysis. *Bulletin of the American Meteorological Society*, 87(3):343–360. p. 15

Michelangeli, P.-A., Vautard, R., and Legras, B. 1995. Weather regimes: Recurrence and quasi stationarity. *Journal of the Atmospheric Sciences*, 52(8):1237–1256. p. 20, 37, 48, 59

Miller, I., Freund, J. E., and Johnson, R. A. 1965. *Probability and statistics for engineers*, volume 4. Prentice-Hall Englewood Cliffs, NJ. p. 108

Mohandes, M., Halawani, T., Rehman, S., and Hussain, A. A. 2004. Support vector machines for wind speed prediction. *Renewable Energy*, 29(6):939–947. p. 21, 60

Monahan, A. H. 2012. Can we see the wind? Statistical downscaling of historical sea surface winds in the subarctic northeast Pacific. *Journal of Climate*, 25(5):1511–1528. p. 18, 20, 21, 36, 38, 59

Monroy, A. and Alvarez-Icaza, L. 2006. Real-time identification of wind turbine rotor power coefficient. In *Decision and Control, 2006 45th IEEE Conference on*, pages 3690–3695. IEEE. p. 94

Mortensen, N. G., Hansen, J. C., Kelly, M. C., Szewczuk, S., Mabille, E., and Prinsloo, E. 2012. Wind Atlas for South Africa (WASA). p. 14

Mortensen, N. G., Landberg, L., Troen, I., and Lundtang Petersen, E. 1993. Wind atlas analysis and application program (WAsP). p. 12

Mortensen, N. G., Said, U. S., and Badger, J. 2006. *Wind Atlas of Egypt*. RISO National Laboratory, New and Renewable Energy Authority, Egyptian Meteorological Authority. p. 14

Mosmann, V., Castro, A., Fraile, R., Dessens, J., and Sanchez, J. 2004. Detection of statistically significant trends in the summer precipitation of mainland Spain. *Atmospheric Research*, 70(1):43–53. p. 67

Mott, R. and Lehning, M. 2010. Meteorological modeling of very high-resolution wind fields and snow deposition for mountains. *Journal of Hydrometeorology*, 11(4):934–949. p. 2

Najac, J., Boé, J., and Terray, L. 2009. A multi-model ensemble approach for assessment of climate change impact on surface winds in France. *Climate dynamics*, 32(5):615–634. p. 12, 20, 33, 36, 59

Noguer, M. 1994. Using statistical techniques to deduce local climate distributions. An application for model validation. *Meteorological Applications*, 1(3):277–287. p. 17

Ochotta, T., Gebhardt, C., Saupe, D., and Wergen, W. 2005. Adaptive thinning of atmospheric observations in data assimilation with vector quantization and filtering methods. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3427–3437. p. 22

Oehlert, G. W. 1996. Shrinking a wet deposition network. *Atmospheric Environment*, 30(8):1347–1357. p. 22

Orlanski, I. 1975. A rational subdivision of scales for atmospheric processes. *Bulletin American Meteorology Society*, (56):527–530. p. 3, 4

Osuna, E., Freund, R., and Girosit, F. 1997. Training support vector machines: an application to face detection. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 130–136. IEEE. p. 21

Palutikof, J., Kelly, P., Davies, T., and Halliday, J. 1987. Impacts of Spatial and Temporal Windspeed Variability on Wind Energy Output. *Journal of Applied Meteorology*, 26:1124–1133. p. 24

Palutikof, J. P., Guo, X., and Halliday, J. A. 1992. Climate variability and the UK wind resource. *Journal of Wind Engineering and Industrial Aerodynamics*, 39(1):243–249. p. 18, 19, 30, 36, 59

Paredes, D., Trigo, R. M., Garcia-Herrera, R., and Trigo, I. F. 2006. Understanding precipitation changes in Iberia in early spring: weather typing and storm-tracking approaches. *Journal of Hydrometeorology*, 7(1):101–113. p. 19

Pearl, J. 1984. Heuristics: intelligent search strategies for computer problem solving. p. 36

Pérez-Landa, G., Ciais, P., Sanz, M., Gioli, B., Miglietta, F., Palau, J., Gangoiti, G., Millán, M., et al. 2007. Mesoscale circulations over complex terrain in the Valencia coastal region, Spain–Part 1: Simulation of diurnal circulation regimes. *Atmospheric Chemistry and Physics*, 7(7). p. 2

Periera, R., Guedes, R., and Santos, C. S. 2010. Comparing WAsP and CFD wind resource estimates for the" regular" user. In *Proceedings from EWEC*. p. 15

Philipp, A., Della-Marta, P.-M., Jacobeit, J., Fereday, D. R., Jones, P. D., Moberg, A., and Wanner, H. 2007. Long-term variability of daily North Atlantic-European pressure patterns since 1850 classified by simulated annealing clustering. *Journal of Climate*, 20(16):4065–4095. p. 21

Phillips, N. A. 1956. The general circulation of the atmosphere: A numerical experiment. *Quarterly Journal of the Royal Meteorological Society*, 82(352):123–164. p. 15

Polya, G. 2008. *How to solve it: A new aspect of mathematical method*. Princeton University Press. p. 36

Prieto, L., Herrera, R. G., Díaz, J., Hernández, E., and Del Teso, T. 2004. Minimum extreme temperatures over Peninsular Spain. *Global and Planetary Change*, 44(1):59–71. p. 19

Pryor, S. and Barthelmie, R. 2003. Long-term trends in near-surface flow over the Baltic. *International Journal of Climatology*, 23(3):271–289. p. 12, 15

Pryor, S., Barthelmie, R., and Schoof, J. T. 2005a. The impact of non-stationarities in the climate system on the definition of a normal wind year: a case study from the Baltic. *International journal of climatology*, 25(6):735–752. p. 12, 48

Pryor, S., Schoof, J., and Barthelmie, R. 2005b. Climate change impacts on wind speeds and wind energy density in northern Europe: empirical downscaling of multiple AOGCMs. *Climate Research*, 29(3):183. p. 37, 59, 107

Pryor, S., Schoof, J. T., and Barthelmie, R. 2005c. Empirical downscaling of wind speed probability distributions. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D19). p. 21, 24, 37, 59

Qin, X., Zhang, S. J., and Yan, D. X. 2010. A New Circular Distribution and Its Application to Wind Data. *Journal of Mathematics Research*, 2(3):P12. p. 24

Queralt, S., Hernández, E., Barriopedro, D., Gallego, D., Ribera, P., and Casanova, C. 2009. North Atlantic Oscillation influence and weather types associated with winter total and extreme precipitation events in Spain. *Atmospheric Research*, 94(4):675–683. p. 19

REE 2012. Boletín mensual, Abril 2012. *Red Eléctrica de España, S.A.U.*, (64). p. 1

REE 2013. Avance del informe del Sistema Eléctrico español 2013. *Red Eléctrica de España, S.A.U.* p. 1

Rehman, S., Halawani, T., and Husain, T. 1994. Weibull parameters for wind speed distribution in Saudi Arabia. *Solar Energy*, 53(6):473–479. p. 24

Ribrant, J. and Bertling, L. 2007. Survey of failures in wind power systems with focus on Swedish wind power plants during 1997-2005. In *Power Engineering Society General Meeting, 2007. IEEE*, pages 1–8. IEEE. p. 23

Richardson, L. F. 1911. The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 210:307–357. p. 16

Rocha, A. 1999. Low-frequency variability of seasonal rainfall over the Iberian peninsula and ENSO. *International Journal of Climatology*, 19(8):889–901. p. 4

Rodriguez-Puebla, C., Encinas, A., Nieto, S., and Garmendia, J. 1998. Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula. *International Journal of Climatology*, 18(3):299–316. p. 5

Rogers, J. C. 1997. North Atlantic storm track variability and its association to the North Atlantic Oscillation and climate variability of northern Europe. *Journal of Climate*, 10(7):1635–1647. p. 5, 60

Romero, R., Sumner, G., Ramis, C., and Genovés, A. 1999. A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *International Journal of Climatology*, 19(7):765–785. p. 19

Ruiz-Arias, J., Terrados, J., Pérez-Higueras, P., Pozo-Vázquez, D., and Almonacid, G. 2012. Assessment of the renewable energies potential for intensive electricity production in the province of Jaén, southern Spain. *Renewable and Sustainable Energy Reviews*, 16(5):2994–3001. p. 24

Şahin, A. Z. and Aksakal, A. 1998. Wind power energy potential at the northeastern region of Saudi Arabia. *Renewable Energy*, 14(1):435–440. p. 2, 23

Sailor, D., Hu, T., Li, X., and Rosen, J. 2000. A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change. *Renewable Energy*, 19(3):359–378. p. 2, 12, 16, 20, 21, 23, 28, 37, 60

Salameh, T., Drobinski, P., Vrac, M., and Naveau, P. 2009. Statistical downscaling of near-surface wind over complex terrain in southern France. *Meteorology and Atmospheric Physics*, 103(1-4):253–265. p. 17, 20, 28, 33, 36, 59

Salcedo-Sanz, S., Fernández-Villacañas, J.-L., Segovia-Vargas, M. J., and Bousoño-Calzón, C. 2005. Genetic programming for the prediction of insolvency in non-life insurance companies. *Computers & Operations Research*, 32(4):749–765. p. 18

Salcedo-Sanz, S., Pérez-Bellido, Á. M., Ortiz-García, E. G., Portilla-Figueras, A., Prieto, L., and Paredes, D. 2009. Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction. *Renewable Energy*, 34(6):1451–1457. p. 11

Salcedo-Sanz, S., Pérez-Bellido, Á. M., Portilla-Figueras, A., Prieto, L., et al. 2011. Short term wind speed prediction based on evolutionary support vector regression algorithms. *Expert Systems with Applications*, 38(4):4052–4057. p. 11

Sandström, S. 1997. Simulations of the climatological wind field in the Baltic Sea area using a mesoscale higher-order closure model. *Journal of applied meteorology*, 36(11):1541–1552. p. 16, 32, 33, 59

Sandwell, D. T. and Agreen, R. W. 1984. Seasonal variation in wind speed and sea state from global satellite measurements. *Journal of Geophysical Research*, 89(C2):2041–2051. p. 12

Saunders, I. and Byrne, J. 1996. Generating regional precipitation from observed and GCM synoptic-scale pressure fields, southern Alberta, Canada. *Climate Research*, 6(3):237–249. p. 17

Schlichting, H., Kestin, J., Schlichting, H., and Schlichting, H. 1968. *Boundary-layer theory*, volume 539. McGraw-Hill New York. p. 8

Schubel, P. J. and Crossley, R. J. 2012. Wind Turbine Blade Design. *Energies*, 5(9):3425–3449. p. 92

Schubert, S. and Henderson-Sellers, A. 1997. A statistical model to downscale local daily temperature extremes from synoptic-scale atmospheric circulation patterns in the Australian region. *Climate Dynamics*, 13(3):223–234. p. 17

Seguro, J. and Lambert, T. 2000. Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis. *Journal of Wind Engineering and Industrial Aerodynamics*, 85(1):75–84. p. 24

Seierstad, I., Stephenson, D., and Kvamstø, N. 2007. How useful are teleconnection patterns for explaining variability in extratropical storminess? *Tellus A*, 59(2):170–181. p. 5

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G. 2005. A description of the advanced research WRF version 2. Technical report, DTIC Document. p. 10

Spears, B. M. and Jones, I. D. 2010. The long-term (1979–2005) effects of the North Atlantic Oscillation on wind-induced wave mixing in Loch Leven (Scotland). *Hydrobiologia*, 646(1):49–59. p. 25, 91

Spellman, G. 2000. The application of an objective weather-typing system to the Iberian peninsula. *Weather*, 55(10):375–385. p. 19, 31

Stevens, M. and Smulders, P. 1979. The estimation of the parameters of the Weibull wind speed distribution for wind energy utilization purposes. *Wind engineering*, 3:132–145. p. 24

Tenguria, N., Mittal, N., and Ahmed, S. 2011. EVALUATION OF PERFORMANCE OF HORIZONTAL AXIS WIND TURBINE BLADES BASED ON OPTIMAL ROTOR THEORY. *Journal of Urban and Environmental Engineering*, (1):15–23. p. 92

Torres, J. L., Garcia, A., De Blas, M., and De Francisco, A. 2005. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy*, 79(1):65–77. p. 60

Tribbia, J. and Baumhefner, D. 2004. Scale interactions and atmospheric predictability: An updated perspective. *Monthly weather review*, 132(3):703–713. p. 3

Trigo, I. F. 2006. Climatology and interannual variability of storm-tracks in the Euro-Atlantic sector: a comparison between ERA-40 and NCEP/NCAR reanalyses. *Climate Dynamics*, 26(2-3):127–143. p. 67

Trigo, I. F., Bigg, G. R., and Davies, T. D. 2002. Climatology of cyclogenesis mechanisms in the Mediterranean. *Monthly Weather Review*, 130(3):549–569. p. 5

Trigo, I. F., Davies, T. D., and Bigg, G. R. 2000. Decline in Mediterranean rainfall caused by weakening of Mediterranean cyclones. *Geophysical Research Letters*, 27(18):2913–2916. p. 5

Trigo, R. M. and DaCamara, C. C. 2000. Circulation weather types and their influence on the precipitation regime in Portugal. *International Journal of Climatology*, 20(13):1559–1581. p. 19, 31

Trigo, R. M. and Palutikof, J. P. 1999. Simulation of daily temperatures for climate change scenarios over Portugal: a neural network model approach. *Climate Research*, 13(1):45–59. p. 17

Trigo, R. M., Valente, M. A., Trigo, I. F., Miranda, P., Ramos, A. M., Paredes, D., and García-Herrera, R. 2008. The impact of North Atlantic wind and cyclone trends on European precipitation and significant wave height in the Atlantic. *Annals of the New York Academy of Sciences*, 1146(1):212–234. p. 5

Troen, I. and Lundtang Petersen, E. 1989. European wind atlas. p. 14

Tuller, S. E. and Brett, A. C. 1984. The characteristics of wind velocity that favor the fitting of a Weibull distribution in wind speed analysis. *Journal of Applied Meteorology*, 23:124–134. p. 24, 94

Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., et al. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612):2961–3012. p. 6, 15

Van der Hoven, I. 1957. Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour. *Journal of Meteorology*, 14(2):160–164. p. 11

van der Kamp, D., Curry, C. L., and Monahan, A. H. 2012. Statistical downscaling of historical monthly mean winds over a coastal region of complex terrain. II. Predicting wind components. *Climate dynamics*, 38(7-8):1301–1311. p. 17, 20, 36, 59

Van Wachem, B., Schouten, J., Van den Bleek, C., Krishna, R., and Sinclair, J. 2001. Comparative analysis of CFD models of dense gas–solid systems. *AIChE Journal*, 47(5):1035–1051. p. 15

Veigas, M. and Iglesias, G. 2012. Evaluation of the wind resource and power performance of a turbine in Tenerife. *Journal of Renewable and Sustainable Energy*, 4:053106. p. 2, 14, 24

Verseghy, D. L. 1991. CLASS-A Canadian land surface scheme for GCMs. I. Soil model. *International Journal of Climatology*, 11(2):111–133. p. 10

Vestas 2005. *General Specifications V82-1.65 MW MK II*. Vestas. p. 92

von Storch, H., Zorita, E., and Cubasch, U. 1993. Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *Journal of Climate*, 6(6):1161–1171. p. 16

Von Storch, H. and Zwiers, F. W. 2001. *Statistical analysis in climate research*. Cambridge University Press. p. 17

Voronoi, G. 1908. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les parallélloèdres primitifs. *Journal für die reine und angewandte Mathematik*, 134:198–287. p. 40

Vries, O. d. 1979. Fluid dynamic aspects of wind energy conversion. Technical report, DTIC Document. p. 24, 94

Walter, A., Keuler, K., Jacob, D., Knoche, R., Block, A., Kotlarski, S., Muller-Westermeier, G., Rechid, D., and Ahrens, W. 2006. A high resolution reference data set of German wind velocity 19512001 and comparison with regional climate model results. *Meteorologische Zeitschrift*, 15(6):585–596. p. 16

Wang, X. L., Feng, Y., Compo, G., Swail, V., Zwiers, F., Allan, R., and Sardeshmukh, P. 2012. Trends and low frequency variability of extra-tropical cyclone activity in the ensemble of twentieth century reanalysis. *Climate Dynamics*, pages 1–26. p. 49, 93

Ward Jr, J. H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244. p. 19

Weber, R. O. and Kaufmann, P. 1995. Automated classification scheme for wind fields. *Journal of Applied Meteorology*, 34(5):1133–1141. p. 32

Weibull, W. et al. 1951. A statistical distribution function of wide applicability. *Journal of applied mechanics*, 18(3):293–297. p. 24, 107

Weisser, D. and Foxon, T. 2003. Implications of seasonal and diurnal variations of wind velocity for power output estimation of a turbine: a case study of Grenada. *International journal of energy research*, 27(13):1165–1179. p. 12

Wendland, W. M. 1982. Wind power as an electrical power source in Illinois. *J. Appl. Meteor.*, 21:423–428. p. 14

Whiteman, C. D. and Doran, J. C. 1993. The relationship between overlying synoptic-scale flows and winds within a valley. *Journal of Applied Meteorology*, 32(11):1669–1682. p. 3

Wilby, R. L. and Wigley, T. 1997. Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, 21(4):530–548. p. 16

Wilson, R. E., Lissaman, P. B., and Walker, S. N. 1976. Aerodynamic performance of wind turbines. Technical report, Oregon State Univ., Corvallis (USA). p. 24, 94

Wood, N. 2000. Wind flow over complex terrain: a historical perspective and the prospect for large-eddy modelling. *Boundary-Layer Meteorology*, 96(1-2):11–32. p. 9

Yao, X., Liu, Y., and Lin, G. 1999. Evolutionary programming made faster. *Evolutionary Computation, IEEE Transactions on*, 3(2):82–102. p. 32, 42, 60

Yeter, P., Güler, Ö., and Akdağ, S. 2012. The Impact of Wind Speed Variability on Wind Power Potential and Estimated Generation Cost. *Energy Sources, Part B: Economics, Planning, and Policy*, 7(4):339–347. p. 94

Yim, S. H., Fung, J. C., Lau, A. K., and Kot, S. 2007. Developing a high-resolution wind map for a complex terrain with a coupled MM5/CALMET system. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 112(D5). p. 12

Zhang, Y. and Held, I. M. 1999. A linear stochastic model of a GCM's midlatitude storm tracks. *Journal of the atmospheric sciences*, 56(19):3416–3435. p. 5

Zorita, E. and Von Storch, H. 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of climate*, 12(8):2474–2489. p. 10