

# UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

Departamento de Física Atómica, Molecular y Nuclear



## TESIS DOCTORAL

Análisis y predicción de series temporales de irradiancia solar global  
mediante modelos estadísticos

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Luis Martín Pomares

Directores

Luis Fernando Zarzalejo Tirado  
Carlos Armenta Deu

Madrid, 2012

**Universidad Complutense de Madrid**

**DEPARTAMENTO DE FÍSICA ATÓMICA, MOLECULAR Y NUCLEAR**

---



**ANÁLISIS Y PREDICCIÓN DE SERIES TEMPORALES  
DE IRRADIANCIA SOLAR GLOBAL MEDIANTE  
MODELOS ESTADÍSTICOS**

Memoria para optar al grado de Doctor

presentada por

**LUIS MARTÍN POMARES**

Dirigida por:

LUIS FERNANDO ZARZALEJO TIRADO

CARLOS ARMENTA DEU

MADRID, DICIEMBRE DE 2011



## **AGRADECIMIENTOS**

El periodo de mi vida que ha transcurrido durante la elaboración de este trabajo ha sido una fase que me ha ayudado a enriquecerme y formarme personalmente. Este documento no habría visto la luz sin la colaboración y el apoyo de tantísimas personas, a las que quiero agradecer todo el apoyo recibido. Son tantos los nombres que me vienen a mi memoria en estos instantes, que necesitaría días y días para no olvidarme de ninguna de ellas. En especial me gustaría destacar a:

Luis Zarzalejo, por todo su esfuerzo y paciencia. Además de los conocimientos científicos que he adquirido, este periodo me ha permitido enriquecerme personalmente. A mis compañeros (y amigos) de CIEMAT y la PSA: Jesús, Ana, Antonio Campos,...A Domingo, Elia, Pablo, María, Miguel, Bea, Pablo, Esther, Jorge, Edu, Juan, Andrés, Paco, Oscar, Antonio... y a los que se fueron: Lourdes, Bella, Marco, Irene, Cristina, Joao, Elena, Jaime, Verónica... A mis compañeros de IrSOLaV: Marco, Gonzalo, Ruth y Manuel. A todas las personas que fueron mi familia durante mis días en Madrid: David, Roberto, Carola, Estela, Julio, José, Manuel, Carmen, Pilar,... A Ángel Fornieles por su paciente revisión. A Mari Carmen. A mis compañeros de la Universidad de Almería. A mis amigos de la infancia por estar siempre a mi lado: Salva, Fran, Victor, José Manuel, Tania, Saray...A mis hermanos, a mis primos hermanos y a mis tíos y tías.

A Camino, por sus ánimos y todo su apoyo, por darme el empujoncito necesario para terminar este trabajo, gracias.



*A mis padres,  
a mi abuela.  
Gracias.*



## **PREFACIO**

La memoria aquí presentada ha sido realizada en el Departamento de Energía, División de energías Renovables del Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Ministerio de Educación y Ciencia.

El desarrollo de los modelos de predicción de la irradiancia global semidiaria se ha llevado a cabo utilizando datos piranométricos horarios procedentes de la Agencia Estatal de Meteorología (AEMet).



# ÍNDICE

<b>1</b>	<b>INTRODUCCIÓN.....</b>	<b>17</b>
1.1	OBJETIVOS.....	21
1.2	ESTRUCTURA DE LA MEMORIA.....	22
<b>2</b>	<b>PREDICCIÓN DE LA IRRADIANCIA SOLAR: ANTECEDENTES Y ESTADO ACTUAL</b>	<b>25</b>
2.1	PREDICCIÓN DE LA IRRADIANCIA SOLAR: CONCEPTOS PREVIOS.....	26
2.2	MODELOS DE PREDICCIÓN DE LA IRRADIANCIA SOLAR.....	30
2.2.1	<i>Modelos dinámicos aplicados a la predicción de la irradiancia solar.....</i>	<i>33</i>
2.2.2	<i>Modelos basados en downscaling estadístico aplicados a la predicción de la irradiancia solar 41</i>	
2.2.3	<i>Modelos estadísticos clásicos aplicados a la predicción de la irradiancia solar.....</i>	<i>45</i>
2.2.4	<i>Modelos basados en imágenes de satélite.....</i>	<i>48</i>
2.2.5	<i>Comparación de modelos de predicción de la irradiancia solar.....</i>	<i>50</i>
2.3	RESUMEN DE LOS MODELOS DE PREDICCIÓN DE LA IRRADIANCIA SOLAR.....	58
2.4	CONCLUSIONES.....	61
<b>3</b>	<b>ELECCIÓN DEL MODELO Y METODOLOGÍA EXPERIMENTAL.....</b>	<b>65</b>
3.1	APLICACIONES DE LAS PREDICCIÓNES DE IRRADIANCIA SOLAR PARA LA PRODUCCIÓN DE ENERGÍA ELÉCTRICA.....	66
3.1.1	<i>Requisitos de predicción en las plantas termosolares y fotovoltaicas.....</i>	<i>67</i>
3.1.2	<i>Predicciones de la irradiancia solar para plantas termosolares y fotovoltaicas.....</i>	<i>68</i>
3.2	DETERMINACIÓN DEL TIPO DE MODELO METEOROLÓGICO: MODELOS FÍSICOS Y MODELOS ESTADÍSTICOS.....	72
3.2.1	<i>Selección del conjunto inicial de modelos estadísticos a ensayar en base a su linealidad</i>	<i>75</i>
3.3	ESQUEMA DE LA METODOLOGÍA EXPERIMENTAL DE PREDICCIÓN.....	77
3.4	PREPROCESADO DE DATOS EXPERIMENTALES.....	79
3.4.1	<i>Transformación de series temporales no-estacionarias.....</i>	<i>79</i>
3.4.2	<i>Identificación de la linealidad de la serie temporal estacionaria.....</i>	<i>82</i>
3.5	MODELOS ESTADÍSTICOS DE PREDICCIÓN PROPUESTOS: AJUSTE Y GENERACIÓN DE LAS PREDICCIÓNES.....	87
3.5.1	<i>Modelos de predicción.....</i>	<i>87</i>
3.5.2	<i>Generación de las predicciones de irradiancia solar semidiaria.....</i>	<i>92</i>
3.6	VALIDACIÓN DE LOS MODELOS DE PREDICCIÓN.....	95
3.7	CONCLUSIONES.....	96
<b>4</b>	<b>PREPROCESADO.....</b>	<b>99</b>
4.1	CONJUNTO DE DATOS RADIOMÉTRICOS.....	99
4.2	APLICACIÓN DE LA METODOLOGÍA DE PREPROCESADO DE LOS DATOS RADIOMÉTRICOS.....	103
4.3	PROCEDIMIENTOS DE CONTROL DE CALIDAD.....	103
4.4	SELECCIÓN DE LOS PERIODOS DE DATOS DISPONIBLES.....	106
4.5	TRANSFORMACIÓN DE LA SERIE TEMPORAL DE IRRADIANCIA GLOBAL SEMIDIARIA EN UNA SERIE ESTACIONARIA.....	109
4.6	ANÁLISIS DE LAS PROPIEDADES ESTADÍSTICAS DE LAS SERIES TEMPORALES ESTACIONARIAS.....	111
4.7	APLICACIÓN DEL TEST DE LINEALIDAD A LAS SERIES TEMPORALES ESTACIONARIAS.....	115
4.8	CONCLUSIONES.....	119
<b>5</b>	<b>ANÁLISIS Y VALIDACIÓN DE LOS RESULTADOS.....</b>	<b>121</b>
5.1	VALIDACIÓN DE MODELOS AUTORREGRESIVOS.....	122
5.2	VALIDACIÓN DE MODELOS BASADOS EN REDES NEURONALES.....	127
5.3	VALIDACIÓN DE MODELOS ANFIS.....	131
5.4	ANÁLISIS DE MODELOS ÓPTIMOS PARA CADA ESTACIÓN RADIOMÉTRICA.....	135
5.5	CONCLUSIONES.....	140

<b>6</b>	<b>SELECCIÓN DEL MODELO FINAL DE PREDICCIÓN.....</b>	<b>145</b>
6.1	ANÁLISIS DE LOS MODELOS FINALES PARA CADA ESTACIÓN RADIOMÉTRICA .....	145
6.2	CONCLUSIONES.....	171
<b>7</b>	<b>RESUMEN.....</b>	<b>173</b>
<b>8</b>	<b>CONCLUSIONES.....</b>	<b>181</b>
	<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>187</b>
	<b>APÉNDICE A – FUNDAMENTOS DE IRRADIANCIA SOLAR .....</b>	<b>197</b>
A.1	SISTEMA TEMPORAL DE REFERENCIA .....	197
A.2	SISTEMA ESPACIAL DE REFERENCIA. GEOMETRÍA SOLAR .....	199
A.3	VARIACIÓN DE LA IRRADIANCIA EXTRATERRESTRE.....	202
A.4	ÍNDICE DE CLARIDAD ( $K_T$ ) .....	204
A.4	COMPONENTE SOLAR ATENUADA ( $LC$ ).....	205
A.6	INTERACCIÓN RADIACIÓN-ATMÓSFERA .....	205
A.7	MODELO DE CIELO CLARO DEL EUROPEAN SOLAR RADIATION ATLAS (ESRA) .....	212
A.8	BASELINE SURFACE RADIATION NETWORK (BSRN).....	216
	<b>APÉNDICE B – ANÁLISIS GRÁFICO DE LOS MODELOS FINALES.....</b>	<b>225</b>
	<b>APÉNDICE C – ANÁLISIS MENSUAL DE LOS RESULTADOS .....</b>	<b>251</b>
	<b>APÉNDICE D – ANÁLISIS DE LA FUNCIÓN DE DISTRIBUCIÓN ACUMULADA Y TEST DE KOLMOGOROV-SMIRNOV.....</b>	<b>277</b>

## NOMENCLATURA

A continuación se presenta una descripción detallada de los símbolos utilizados en este trabajo así como sus unidades y rango. Ordenados alfabéticamente se transcriben los símbolos que con más frecuencia aparecen en esta memoria, omitiendo aquellos que solo son utilizados como apoyo para la definición de otra variable y/o parámetro dentro de una ecuación. En cualquier caso todas las variables y símbolos que las representan son definidas, salvo error u omisión, la primera vez que aparecen en el texto.

Se ha intentado respetar las recomendaciones aceptadas internacionalmente en el campo de la energía solar [Beckman et al., 1978]; en caso de duplicidad en la simbología recomendada, se han adoptado símbolos alternativos.

Se incluye también una lista con los acrónimos y abreviaturas que aparecen en este trabajo. Como en el caso de los símbolos, son descritos la primera vez que aparecen en el texto. Se ha respetado la grafía correspondiente a acrónimos procedentes de la lengua inglesa en el caso de no existir equivalencia en castellano o ser internacionalmente reconocibles.

<b>Símbolo</b>	<b>Magnitud</b>	<b>Rango, signo y unidades</b>
$k_t$	Índice de transparencia atmosférica	
$K_T$	Índice de transparencia atmosférica semidiario	
lc	Componente solar atenuada o <i>lost component</i>	
LC	Componente solar atenuada o <i>lost component</i> semidiario	
$S$	Variación estacional de una serie temporal	
$C$	Cambios cíclicos de una serie temporal	
$T$	Tendencia de una serie temporal	
$a$	Componente aleatoria de una serie temporal	
$r_i$	Rango de los discriminantes estadísticos no-lineales	
$y(t_i)$	Observación de una serie temporal estacionaria	
$t_i$	Instante temporal	
$\delta_{AR}$	Parámetro constante de un modelo autorregresivo	
$\phi_n$	Parámetro de un modelo autorregresivo	
$\varepsilon(t_i)$	Errores aleatorios de predicción para un instante temporal	
$w_{ij}$	Efecto de la célula j-ésima sobre la célula i-ésima	
$g$	Modelos estadísticos de predicción	
$\hat{y}(t_i)$	Valor estimado para el instante temporal $t_i$	
$\bar{y}$	Esperanza de un conjunto de valores	
$S_y$	Desviación típica de un conjunto de valores	
$\hat{y}(t_{i+1}   t_k)$	Predicción del valor $t_{i+1}$ a partir del valor en el instante temporal $t_k$	
$I_{CS}$	Constante solar	1367 Wm <sup>-2</sup>

---

$G$	Irradiancia solar global	$\text{Wm}^{-2}$
$I_0$	Irradiancia solar extraterrestre	$\text{Wm}^{-2}$
$\Lambda$	Longitud geográfica	$-180^\circ \sim 180^\circ$ / Oeste $\sim$ Este
$\Lambda_{of}$	Longitud del meridiano de referencia horaria oficial	$-180^\circ \sim 180^\circ$ / Oeste $\sim$ Este
$\Lambda_{loc}$	Longitud local	$-180^\circ \sim 180^\circ$ / Oeste $\sim$ Este
$\Phi$	Latitud geográfica	$-180^\circ \sim 180^\circ$ / Sur $\sim$ Norte
$J_d$	Día juliano	1~366
$\delta$	Declinación	$-0.41 \sim 0.41$ radianes
$\omega$	Ángulo horario	$-\pi \sim \pi$ radianes ( $\omega > 0$ AM)
$\psi$	Azimut solar	$-\pi \sim \pi$ radianes ( $\omega > 0$ AM)
$\theta_z$	Ángulo cenital solar	$0 \sim \pi/2$ radianes
$\alpha$	Altura o elevación solar	$0 \sim \pi/2$ radianes
$\varepsilon$	Excentricidad de la órbita terrestre	m
$I_{0n}$	Irradiancia solar extraterrestre normal	$\text{Wm}^{-2}$
$\lambda$	Longitud de onda incidente	m
$\Theta$	Tamaño relativo de las partículas responsables del scattering	
$q$	Tamaño medio de una partícula	
$m$	Masa óptica relativa del aire	
$\alpha^*$	Altura solar verdadera	
$p$	Presión atmosférica	hPa
$p_0$	Presión atmosférica al nivel del mar	hPa
$z$	Altitud sobre el nivel del mar	m
$\delta_r$	Profundidad óptica de Rayleigh	
$T_L$	Factor de turbidez de Linke	

---

**Lista de acrónimos, siglas y abreviaturas**

AEMet	Agencia Estatal de Meteorología
AMF	Atmospheric Motion Vector Field
ANFIS	Adaptative Network based Fuzzy Inference System (Redes Adaptativas basadas en sistemas de inferencia difusa)
ANN	Artificial Neural Network
AU	Unidad Astronómica
BSRN	Baseline Surface Radiation Network
CIEMAT	Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
CM-SAF	Climate Monitoring Satellite Applications Facility
CSP	Concentrated Solar Power
DVV	Delay Vector Variance
DWD	Deutscher Wetterdienst
ECMWF	European Centre for Medium-Range Weather Forecasts Centro Europeo de Predicción a Plazo Medio
EPIA	Asociación de la Industria Solar Fotovoltaica
ESRA	European Solar Radiation Atlas (Atlas Europeo de Radiación Solar)
ESTELA	Asociación Europea de la Energía Solar Termoeléctrica
EURAD	EUROpean Air Pollution Disperison
FT	Fourier Transform
GEM	Canadian Meteorological Center's Global Enviromental Multiscale Model
GFS	Global Forecast System
HIRLAM	High Resolution Limited Area Model
iAAFT	Iterative Amplitude Adjusted Fourier Transform
IEA	International Energy Agency
LM	Lokal-Modell
MBD	Mean Bias Deviation
MCGA	Modelos Globales de Circulación Atmosférica

---

MM5	Mesoscale Model Version 5
MODIS	Moderate Resolution Imaging Spectro Radiometer
MOS	Model Output Statistics
MRE	Mean Relative Error (Error Relativo Medio)
MVF	Motion Vector Field
NARX	Red Neuronal Autorregresiva No-Lineal
NCAR	National Center for Atmospheric Research
NCEP	Nacional Centers for Enviromental Prediction
NDFD	National Digital Forecast Database
NOAA	National Oceanic and Atmospheric Administration (USA)
NWS	Nacional Weather Service
OMEL	Operador del Mercado Eléctrico
REV	Reversibilidad Temporal
RMSD	Root Mean Square Deviation
RNA	Red Neuronal Artificial
RRN	Red Radiométrica Nacional
SHC	Solar Heating and Cooling
TLM	Tiempo Local Medio
TSI	Total Sky Imager
TSV	Tiempo Solar Verdadero
WMO	World Meteorological Center (Centro Meteorológico Mundial)
WRC	World Radiometric Network (Centro Radiométrico Mundial)
WRF	Weather Research & Forecasting Model
WSC	Whole Sky Camera

---



# 1 INTRODUCCIÓN

El establecimiento de un sistema energético sostenible es uno de los grandes desafíos a los que debe enfrentarse la humanidad en el siglo XXI. El modelo actual, que ha permitido alcanzar el bienestar y desarrollo económico de nuestra sociedad desde la revolución industrial, está basado en el uso de combustibles fósiles con gran impacto sobre el medioambiente. Otro inconveniente del uso de los combustibles fósiles es la creciente inseguridad internacional que desemboca en conflictos bélicos de escala mundial y catástrofes humanitarias por el control de dichos recursos.

Las principales ventajas derivadas del uso de las energías renovables son la diversificación del suministro energético nacional, el fomento de los recursos autóctonos, la reducción de la dependencia energética exterior, el incremento del nivel de autoabastecimiento y autosuficiencia energética, la creación de empleo local, el menor impacto medioambiental y la generación distribuida de la energía que evita problemas de suministro ante cualquier fallo. Además, los compromisos internacionales de los países en relación a la lucha contra el cambio climático (y, en consecuencia, la reducción de las emisiones de CO<sub>2</sub> a la atmósfera) provocan el fomento y aumento de peso de las energías renovables en el conjunto de fuentes energéticas o *mix de la energía*.

Entre las energías renovables, la energía solar presenta la mayor cantidad absoluta de potencial energético; la energía solar que llega a la superficie terrestre es más de tres órdenes de magnitud superior a la demanda energética mundial actual. Sin embargo, existe una serie de factores que disminuyen su potencial y que se enumeran a continuación:

- La variación temporal de la radiación solar en un emplazamiento debido a los ciclos diarios y estacionales.
- La variación geográfica de la radiación solar dependiente de la latitud. Las regiones situadas más próximas al Ecuador reciben un mayor nivel potencial de radiación solar que las regiones más próximas a los polos terrestres.

- Las condiciones atmosféricas de nubosidad y turbidez atmosférica debido a los aerosoles que tienen un gran efecto sobre la radiación solar.
- El uso de la energía solar necesita grandes extensiones de terreno debido a su baja densidad espacial.

La energía solar fotovoltaica y termoeléctrica son las dos principales tecnologías de aprovechamiento del recurso solar como fuente energética para la producción de electricidad. Las perspectivas de desarrollo futuro de la energía solar fotovoltaica y termoeléctrica para abastecer de energía eléctrica a gran parte de la población del planeta han sido plasmadas en varios estudios.

De entre ellos, un reciente estudio promovido por la Asociación de la Industria Solar Fotovoltaica (EPIA), asegura que la energía solar fotovoltaica será la fuente de energía mayoritaria antes de 2030 en los países del llamado “Cinturón Solar”, que engloba a los estados entre las latitudes  $\pm 35^\circ$ . Esta zona del planeta representa en la actualidad el 75% de la población mundial y el 40% de la demanda global de electricidad.

Por otro lado, el informe “Energía Solar Térmica de Concentración – Perspectiva Mundial 2009”, en cuya elaboración han participado la Asociación Europea de la Energía Solar Termoeléctrica (ESTELA), SolarPACES y Greenpeace, destaca que la energía solar térmica podría cubrir el 7% de la demanda eléctrica mundial en el año 2030 y más del 25% en 2050.

En consecuencia, ambas tecnologías disponen de un mercado de amplio potencial cuyo desarrollo está siendo impulsado por el progresivo apoyo regulatorio que están recibiendo en los diversos países. En concreto, España establece un sistema de apoyo a la instalación de energías renovables con un marco regulatorio basado en un sistema de tarificación *feed-in* y un sistema de primas regulado en el Real Decreto 661/2007 [BOE núm.126, 2007].

En la tarifa regulada (sistema *feed-in tariff*) el productor vierte la energía al sistema a través de la red de distribución eléctrica y obtiene una cantidad económica fija que

---

depende de la tecnología que se utilice. Las plantas en este régimen económico deben tener un tamaño inferior a los 50MW. Existe además una garantía a los operadores de que toda la energía que produzcan e inyecten a la red es remunerada.

En la tarifa basada en primas, la producción de la planta se vende libremente en el mercado de la energía eléctrica (también conocido como *pool* de la energía o mercado organizado), gestionado por el Operador del Mercado Eléctrico (OMEL), más un complemento adicional, o prima de referencia según el tipo de tecnología que se utilice. Las plantas con potencia nominal inferior a 50MW no están obligadas a acogerse a este sistema retributivo; no así las que tienen potencia superior a 50MW que deben vender su producción dentro del mercado libre de la energía y recibir como complemento la prima tarifaria.

En la tarifa de primas, los productores de energías renovables pueden competir de igual a igual con los productores convencionales de energía eléctrica, aunque existen ciertas desventajas con respecto al régimen de tarifa regulada. Por un lado, no existe una obligación de compra de toda la energía que se produce, ya que sólo es posible verter a la red aquella energía que se consigue vender en el mercado libre de la energía. Por otro lado, las instalaciones que se acojan a este sistema regulatorio se ven en la obligación de soportar al operador del sistema en caso de requerir apoyo en la estabilización del sistema y control de tensión [BOE núm.184, 2008].

En un futuro cercano, todas las plantas fotovoltaicas y termosolares tendrán que acogerse a la tarifa basada en primas y participar en el mercado libre de la energía, compitiendo directamente y a igualdad de condiciones con el resto de tecnologías que utilizan fuentes convencionales para la producción de energía eléctrica. El objetivo es que la energía solar sea competitiva con el resto de energías y se integre en el *mix de la energía* [Ruiz, 2006].

En ese contexto, la gestión de la producción de las tecnologías solares necesita un conocimiento preciso de las predicciones del recurso solar y la disponibilidad de energía

---

mediante un sistema de almacenamiento. Además, la maximización del beneficio económico de las plantas solares requiere la utilización de estrategias de optimización que utilicen tres módulos compuestos por las predicciones del recurso solar y de los precios de la electricidad en el mercado libre de la energía y un sistema de almacenamiento [Wittmann et al., 2008].

Uno de los principales inconvenientes para la entrada de la energía solar en el mercado libre de la energía, y por ende, para su competitividad en igualdad de condiciones con el resto de energías convencionales para la producción de electricidad es la alta incertidumbre asociada a la disponibilidad del recurso solar debido al gran efecto que tienen las condiciones atmosféricas de nubosidad y de turbidez por los aerosoles.

El estudio, la profundización y el perfeccionamiento de los modelos de predicción de la irradiancia solar se convierte pues en pilar fundamental del desarrollo y apuesta definitiva por las tecnologías solares para la producción de electricidad. Cuanto más precisas son las predicciones, las fuentes productivas de electricidad a partir del recurso solar son más fiables. Partiendo de tal necesidad, en este trabajo se plantea la propuesta y validación de modelos para la mejora de las predicciones de los valores de irradiancia solar para su posterior uso en plantas fotovoltaicas y termosolares que dispongan de un sistema de almacenamiento.

Las líneas actuales de investigación combinan la salida de los modelos globales o de mesoescala con un post-procesado o *downscaling* estadístico para la obtención de predicciones con frecuencia temporal horaria. Los modelos estadísticos son una parte fundamental que integran las predicciones de los modelos numéricos de predicción, y corrigen posibles desviaciones respecto a las condiciones meteorológicas locales. Sin embargo, como se muestra en este trabajo, los resultados que se obtienen para el paso temporal horario no tienen la precisión suficiente para las aplicaciones que hacen uso del recurso solar como principal fuente energética.

En este trabajo, se opta por realizar las predicciones con resolución temporal semidiaria, como complemento a las predicciones horarias, debido a que es el paso temporal que demandan las plantas termosolares que disponen de un sistema de almacenamiento.

Se desarrolla una metodología de predicción para la predicción de la irradiancia solar semidiaria mediante modelos estadísticos de predicción aplicados a series temporales y su validación en estaciones repartidas por diferentes zonas de la Península Ibérica. Los valores semidiarios de las predicciones de la energía solar se definen como los valores de la irradiancia solar integrados para los intervalos temporales entre la salida del Sol y el cenit Solar, y entre este último y la puesta del Sol. La predicción de la irradiancia solar se realiza utilizando dos variables intermedias, el índice de claridad y la componente solar atenuada. Se aplican diferentes modelos de predicción en base a sus propiedades de linealidad y se obtiene como resultado que la bondad en la predicción de dichos modelos depende de la dinámica de la propia serie temporal, índice de claridad o componente solar atenuada.

## **1.1 Objetivos**

Este trabajo persigue desarrollar una metodología de predicción que permita establecer una relación directa con la mayor precisión y exactitud posible, entre los valores de irradiancia solar y la energía producida a partir de sistemas solares termoeléctricos y fotovoltaicos, lo cual obliga a establecer ciertas condiciones de operación, reflejadas en los siguientes objetivos específicos:

- Proposición de una metodología para la predicción de la irradiancia solar semidiaria.
- Evaluación de los modelos propuestos para la predicción de la irradiancia solar semidiaria.
- Obtención de un modelo final de predicción de la irradiancia solar semidiaria basado en la combinación de los modelos que se ensayan en este trabajo.

Los objetivos arriba mencionados suponen la principal aportación de este trabajo. No obstante, el desarrollo del mismo ha supuesto una serie de estudios en paralelo que se pueden considerar como aportaciones adicionales de este trabajo:

1. Filtrado de datos piranométricos horarios medidos en estaciones radiométricas para la obtención de series temporales de irradiancia solar semidiaria.
2. Obtención de series temporales estacionarias a partir de la serie temporal de irradiancia solar semidiaria.
3. Estudio de la linealidad de las series temporales estacionarias.
4. Evaluación de la mejora de las predicciones con variantes más complejas a los modelos propuestos.

Los distintos estudios aquí propuestos abren la base de nuevas investigaciones encaminadas al desarrollo de metodologías y modelos que permitan la mejora de las predicciones de la irradiancia solar en un horizonte predictivo del corto plazo.

## **1.2 Estructura de la memoria**

El contenido del presente trabajo se ha estructurado de la siguiente forma:

- El capítulo 2 presenta una revisión del estado del arte de la predicción de la irradiancia solar que nos sirve como capítulo introductorio para plantear la necesidad de predecir la irradiancia solar.
- El capítulo 3 muestra la metodología de predicción propuesta en este trabajo. Se describen las diferentes fases que se llevan a cabo para la obtención de las predicciones de la irradiancia solar semidiaria y que se enumeran a continuación:
  - Preprocesado de datos experimentales.
  - Selección de los modelos de predicción.
  - Generación de las predicciones de irradiancia solar semidiaria y postprocesado.
  - Validación de los modelos de predicción de la irradiancia solar semidiaria.

- El capítulo 4 presenta los resultados tras realizar el preprocesado. Se filtran los datos de radiación solar y se construye la serie temporal semidiaria. Se eliminan las posibles tendencias que pueda presentar la serie temporal de irradiancia solar semidiaria para que tenga un comportamiento estacionario. Como resultado de este proceso se obtienen dos series temporales intermedias estacionarias que se utilizan como entrada de los modelos de predicción. Además, se identifican las propiedades lineales de las series temporales intermedias.
- El capítulo 5 muestra los resultados de las tres últimas fases: selección de los modelos predictivos, generación y validación de las predicciones. Para realizar las predicciones se utiliza un modelo lineal autorregresivo y dos modelos no lineales basados en redes neuronales y lógica difusa. Las predicciones se realiza sobre las series temporales intermedias y es necesario realizar un postprocesado a partir de dichas series que se basa en la introducción de las componentes deterministas que se eliminaron en el paso de preprocesado, para la obtención de los valores futuros de irradiancia solar. La validación de las predicciones se realiza analizando la calidad de las predicciones a partir de los estadísticos error medio relativo (MBD), error cuadrático medio relativo (RMSD) y mejora en términos de RMSD relativo frente a las predicciones de un modelo basado en la persistencia.
- El capítulo 6 presenta los modelos finales de predicción seleccionados para cada estación radiométrica. El criterio de selección de los modelos se basa en la elección de los modelos con menor orden y menor RMSD relativo.
- Finalmente, el capítulo 7 muestra las conclusiones fundamentales y aportaciones de los estudios desarrollados así como las principales líneas identificadas como futuros trabajos.



## **2 PREDICCIÓN DE LA IRRADIANCIA SOLAR: ANTECEDENTES Y ESTADO ACTUAL**

Tradicionalmente, la irradiancia solar se mide a intervalos muy reducidos, de forma prácticamente instantánea, si bien con objeto de hacer manejable la serie de datos, los valores se promedian a intervalos de 10 segundos y luego estos promedios lo hacen a su vez en series diezminutales. Las series diezminutales constituyen una buena base para el análisis y estudio estadístico, pero no son prácticas para el manejo en aplicaciones de orden práctico, como son el diseño y dimensionado de sistemas solares. Por ello, los valores más habitualmente utilizados se proporcionan en bases horarios o diarias. La ventaja que tiene una base horaria es que proporcionan una mayor precisión, aunque genera un error más elevado a la hora de predecir el comportamiento a lo largo de un día debido a la variabilidad del recurso solar. Por el contrario, mediante la base horaria se obtiene un error cuadrático medio inferior, pero se pierde información sobre lo ocurrido a lo largo del día, ya que considera un proceso de homogenización que no responde a la propia variabilidad diaria del recurso solar. Es posible encontrar bases temporales diferentes, por ejemplo semidiaria, que sin perder demasiada información mejoran las características, esto es, tenga una precisión adecuada y un error cuadrático medio aceptable.

La predicción de la irradiancia solar para usos energéticos requiere, como paso previo, la introducción de una serie de conceptos fundamentales para comprender la importancia y los múltiples esfuerzos que se están realizando actualmente en este campo en todo el mundo.

En este capítulo se presentan los conceptos fundamentales relacionados con la predicción meteorológica que se utilizarán a lo largo de esta memoria. Posteriormente, se realiza una revisión de los principales trabajos previos en el ámbito de la predicción de la irradiancia solar cuyas resoluciones temporales (horaria y diaria) son las más destacables a nuestros fines.

## **2.1 Predicción de la irradiancia solar: conceptos previos**

La energía que llega a la superficie de la atmósfera terrestre procedente del Sol es la principal fuente de energía a partir de la cual se desarrollan todos los procesos en el sistema Tierra-Atmósfera-Océano, siendo la principal responsable, por ejemplo, del ciclo del agua, los cambios de presión y de vital importancia en el proceso de fotosíntesis. La proporción de energía que este sistema recibe del Sol es superior al 99.9%. Además, la energía solar es el origen del resto de fuentes energéticas de nuestro planeta, a excepción de la energía geotérmica y la energía nuclear.

El clima se define como el conjunto de fenómenos meteorológicos que caracterizan las condiciones habituales o más probables de un punto determinado de la superficie terrestre. La climatología es la ciencia que estudia el clima y sus variaciones a lo largo del tiempo. Aunque utiliza los mismos parámetros que la meteorología, su objetivo es distinto, ya que no pretende hacer previsiones inmediatas, si no estudiar las características ambientales a largo plazo.

La meteorología es la ciencia que se encarga del estudio de los fenómenos, las leyes y el estado de la atmósfera en las capas bajas, donde se desarrolla la vida de plantas y animales. Uno de sus objetivos es el conocimiento de dichos fenómenos orientado a su predicción.

A pesar de que la irradiancia solar es una variable de suma importancia, como se ha comentado anteriormente, la meteorología ha prestado mayor importancia al estudio y predicción de otras variables meteorológicas, entre ellas la presión, la temperatura y la humedad. De hecho, es común ver en los partes meteorológicos la cuantificación de dichas variables con magnitudes exactas y precisas. Sin embargo, en el caso de la irradiancia solar la variable con mayor relación directa y de la que históricamente se dan estimaciones es el nivel de nubosidad que establece únicamente seis valores discretos como se puede observar en la Figura 2.1.



Figura 2.1 Clasificación del estado del cielo realizada por AEMet (*Agencia Estatal de Meteorología*) en sus predicciones

Esto se debe principalmente a que para realizar la predicción meteorológica se utilizan modelos matemáticos que simulan el estado futuro de la atmósfera. La capacidad de dichos modelos está limitada por la potencia de procesamiento de las computadoras y a pesar de los avances realizados en los últimos años en el campo del procesamiento de la información, en la actualidad, no es posible disponer de un único modelo que represente el flujo atmosférico en todas sus escalas de movimiento existentes.

Los modelos de predicción meteorológicos pueden clasificarse en función de la resolución temporal y espacial, o bien en función de las herramientas matemáticas que se utilicen. La primera clasificación requiere la definición de los conceptos relacionados con la resolución espacial y temporal de los modelos meteorológicos de predicción y definen el tipo de predicciones que éste producirá. Así, se distingue:

- Resolución espacial. Tamaño que representa el área espacial a la que corresponde la predicción que genera el modelo.
- Resolución temporal. Está relacionado con dos conceptos. Por un lado, el *paso temporal*, o frecuencia temporal de la predicción, se define como el intervalo de tiempo entre cada valor predicho por el modelo; el *horizonte de predicción* hace referencia al alcance temporal de las predicciones.

Dependiendo del paso temporal y del horizonte de predicción, existen cuatro categorías en las que se pueden dividir los modelos meteorológicos y que se presentan en la Tabla

2.1 [Orlanski, 1975]. Los modelos de predicción con escala de nowcasting tienen un horizonte de predicción de una hora, también se considera esta escala como aquella en la que se utilizan técnicas que permiten tener un conocimiento del estado actual de la atmósfera. La escala a corto plazo tiene un horizonte que va desde una hora hasta una semana y la meteorología es la ciencia que se encarga de su estudio. Los modelos con escalas de medio y largo plazo tienen un horizonte de predicción mayor a una semana y la climatología es la ciencia que se ocupa del estudio de esta escala.

Tabla 2.1 Escalas temporales de la predicción según su horizonte de predicción y su paso temporal

Tipo	Horizonte	Paso temporal
<b>Muy corto plazo o nowcasting</b>	Inferior a una hora	Segundos a minutos
<b>Corto plazo</b>	Entre una hora y una semana	Minutos, horas o días
<b>Medio plazo</b>	Semana a un año	Días
<b>Largo plazo</b>	Mayor de un año	Meses

En función de las herramientas matemáticas utilizadas para modelar las variables meteorológicas los modelos se pueden clasificar en los siguientes [Wilks, 2005]:

- Modelos numéricos de predicción (NWP). Se conocen también como modelos dinámicos y se basan en la resolución de las ecuaciones primitivas de la atmósfera utilizando técnicas numéricas que discretizan el espacio y el tiempo. La principal ventaja de estos modelos es que su aplicación es –o se considera– universal. Como inconveniente, necesitan conocer con la mayor fidelidad posible el estado inicial de la atmósfera [Lorenz, 1963].
- Modelos estadísticos. Se aplican directamente sobre las variables meteorológicas medidas o hacen uso de las salidas de los modelos dinámicos para mejorar su resolución. Se caracterizan por su sencillez y el hecho de que sólo se pueden aplicar en aquellos emplazamientos donde se tienen identificados los parámetros del modelo previamente [Gutiérrez et al., 2006].

Entre los diferentes modelos de predicción numérica, y atendiendo a su escala temporal y espacial, podemos encontrar los *Modelos Globales de Circulación Atmosférica (MCGA)*. Estos modelos integran sobre todo el globo terrestre con una baja resolución

temporal (3-6 horas) y espacial ( $0.5^{\circ}$ - $1^{\circ}$ ) y capturan únicamente la dinámica sinóptica de la atmósfera. La salida de estos modelos se puede utilizar posteriormente en modelos de mesoescala o modelos estadísticos para aumentar la resolución espacial y temporal original de las predicciones para una zona geográfica determinada.

El proceso de aumentar la resolución de las predicciones de los MCGA mediante modelos numéricos se conoce como *downscaling dinámico*. Para este fin se utilizan modelos de mesoescala que se caracterizan por tener una resolución espacial de entre 1 y 100 km.<sup>2</sup> y una mejor parametrización de los procesos físicos (evaporación, turbulencia, irradiancia,...) que los modelos globales. De esta forma, al aumentar la resolución, consiguen una mayor precisión que los modelos globales para reproducir fenómenos de pequeña escala, debido al dominio espacial y temporal de estos últimos. Una de las ventajas de estos modelos es que permiten tanto controlar la resolución como seleccionar el esquema de parametrizaciones más apropiado con el objetivo de encontrar la mejor combinación posible para simular los fenómenos meteorológicos particulares de un entorno concreto. Los modelos de mesoescala más utilizados son MM5 (Mesoscale Model version 5) -de *Pennsylvania State University /National Center for Atmospheric Research-* y WRF (Weather Research and Forecast Model), desarrollado en colaboración entre diferentes organismos, como el *National Center for Atmospheric Research* (NCAR) y el *National Oceanic and Atmospheric Administration* (NOAA), entre otros.

Dentro de los modelos estadísticos es posible diferenciar dos grupos dependiendo de si éstos hacen uso de la salida de los modelos numéricos para realizar las predicciones o no:

- Los *modelos estadísticos clásicos* se basan en el uso exclusivo de técnicas matemáticas, principalmente el análisis de series temporales. El calificativo de clásicos se debe a que fueron los primeros modelos que se desarrollaron, cuando no se disponía de la información procedente de los modelos de predicción numérica [Weigend y Gershenfeld, 1993]. En la actualidad, se utilizan para los horizontes temporales de predicción de corto y medio plazo, donde la

información procedente de los modelos numéricos no presenta la suficiente resolución.

- Los modelos de *downscaling estadístico*, también conocidos como técnicas MOS (*Model Output Statistics*), se basan en la aplicación de los métodos estadísticos en conjunción con los modelos de predicción numérica. Éstos se utilizan para mejorar las salidas de los modelos dinámicos de predicción sobre variables o emplazamientos no cubiertos por la resolución del modelo de predicción numérica. Además, la aplicación de este tipo de técnicas permite simular efectos locales que los modelos numéricos de predicción no son capaces de capturar [Glahn y Lowry, 1972].

A partir de estos conceptos, se han ido construyendo modelos diferenciados de predicción de irradiancia solar para un horizonte temporal de predicción del corto plazo realizados y adaptados al contexto y objetivos propios del trabajo en el que se desarrollaron.

## 2.2 Modelos de predicción de la irradiancia solar

La complejidad de los procesos físicos relacionados con la dinámica atmosférica, y la interacción de la irradiancia solar con la atmósfera, ha condicionado que la ciencia no haya concebido un modelo único y universal de predicción de la irradiancia solar. Existe una serie de factores físicos que, en su influencia sobre la irradiancia solar recibida, obliga a diferenciar los modelos: la nubosidad.

La clasificación del estado de cielo se realiza habitualmente de forma paramétrica, caracterizando a la atmósfera como un todo a partir de un parámetro fácilmente accesible que cuantifica la situación de nubosidad (despejado, poco nuboso, intervalos nubosos, muy nuboso, cubierto). Anteriormente, en la Figura 2.1, se presentó una clasificación del estado de cielo realizada por AEMet en sus predicciones. Una magnitud habitualmente utilizada para realizar la parametrización del tipo de cielo es el índice de claridad ( $k_t$ ). El  $k_t$  se define como el cociente entre la irradiancia solar, que se recibe en la superficie terrestre y el valor correspondiente de la irradiancia solar

extraterrestre. En el apéndice A se una descripción detallada de los efectos atmosféricos, pero no se relaciona el valor de  $k_t$  con dichos efectos. Los valores de  $k_t$  se obtienen a partir de la irradiancia solar, aplicando la siguiente expresión:

$$k_t = \frac{G}{I_0}$$

donde  $G$  es la radiación global que llega a la superficie terrestre sobre superficie horizontal e  $I_0$  es la irradiancia extraterrestre sobre superficie horizontal, ambas para el mismo periodo temporal.

En situaciones de cielo claro o despejado, la irradiancia solar a su paso por la atmósfera sufre procesos de absorción y difusión, provocados por los diferentes componentes de la atmósfera (entre los más importantes, vapor de agua ( $H_2O$ ), dióxido de carbono ( $CO_2$ ) y ozono ( $O_3$ )) y los aerosoles formados por partículas sólidas y líquidas en suspensión. Los modelos que estiman la radiación que llega a la superficie de la Tierra, para este tipo de situaciones de cielo, se conocen como *modelos de cielo despejado* y utilizan diversas aproximaciones matemáticas para su resolución. Existen diferentes aproximaciones como los modelos de Bourges [Bourges, 1979], SOLIS [Mueller et al., 2004], Rigollier [Rigollier et al., 2000] adoptado en el Atlas Europeo de Radiación Solar (ESRA) [ESRA, 2000] y modificaciones para la estimación de la irradiancia directa para cielos despejados [Polo et al., 2008].

La aplicación de los modelos de cielo claro está limitada a las situaciones atmosféricas con ausencia de nubosidad. Ya que dichas condiciones para un tiempo futuro no son conocidas de antemano, es necesaria la aplicación previa de los modelos meteorológicos de predicción. Estos modelos son capaces de resolver la dinámica atmosférica relacionada con las masas nubosas, que son las principales causantes de la atenuación de la irradiancia solar a su paso por la atmósfera, y que los modelos de cielo claro obvian en sus cálculos debido a su simplicidad.

Los diferentes autores han abordado la resolución del problema de la predicción de la irradiancia solar que llega a la superficie terrestre mediante diversas perspectivas que deben ser analizadas detenidamente. A tal fin, se ha realizado una clasificación de los diferentes tipos de modelos recogidos en la bibliografía, en función del tipo de modelo empleado (dinámico o estadístico) y del origen de los datos con los que se realiza la predicción:

- Modelos dinámicos.
- Modelos basados en *downscaling* estadístico.
- Modelos estadísticos clásicos a partir de medidas radiométricas.
- Modelos estadísticos que hacen uso de las imágenes de satélite.

Los trabajos aplicados a la predicción de la irradiancia solar, realizados mediante modelos dinámicos se basan en la obtención de las predicciones mediante modelos de mesoescala o aplicando un postprocesado simple a la salida de los modelos globales de predicción (ver 2.2.1).

Los modelos de *downscaling estadístico* utilizan como entrada datos de los modelos numéricos de predicción y la utilización de métodos estadísticos para mejorar las predicciones teniendo en cuenta los efectos locales del emplazamiento donde se pretende realizar las predicciones (ver 2.2.2).

Los *modelos estadísticos clásicos* se basan en el uso exclusivo de técnicas matemáticas y las observaciones recientes de las variables meteorológicas (ver 2.2.3).

Dentro de los modelos estadísticos y atendiendo al tipo de datos sobre el que se realizan las predicciones, se ha realizado una nueva división que da lugar a los *modelos estadísticos aplicados sobre imágenes procedentes de satélite*. El tipo de técnicas que se aplican a los datos procedentes de satélite no sólo tienen en cuenta la evolución temporal de los predictores, sino también su evolución espacial. Una de las características principales de los satélites es que registran de forma simultánea grandes

áreas, lo que permite conocer la distribución espacial de la información y de esta forma es posible modelar la evolución espacial de la misma (ver 2.2.4).

En las dos últimas secciones se mostrará una comparativa de modelos y un resumen de modelos predictivos de la irradiancia solar realizada en el marco de la Tarea 36 (*Solar Resource Knowledge Management*) de la IEA/SHC (*International Energy Agency/Solar Heating and Cooling*). Cabe señalar, que el principal objetivo de la Tarea 36 es el desarrollo, validación y acceso a la información del recurso solar a partir de medidas en superficie y de satélite [Renne, 2009]. Finalmente, se presentará un resumen de todos los modelos expuestos que nos ayudará a obtener una serie de conclusiones en conjunto sobre la precisión de los modelos aplicados a la predicción de la irradiancia solar para usos energéticos.

### *2.2.1 Modelos dinámicos aplicados a la predicción de la irradiancia solar*

Como se ha comentado anteriormente, una forma de mejorar la modelización de las diferentes variables meteorológicas es aplicando técnicas de *downscaling* dinámico o aumento dinámico de la resolución. Estas técnicas se basan en utilizar la salida de los modelos globales en un modelo de mesoescala con mayor resolución espacial y temporal y una mejor modelización física de los procesos de capa límite. En esta sección se presentan los principales trabajos que hacen uso del *downscaling* dinámico para obtener predicciones de la irradiancia solar.

#### Modelo MM5 Heinemann

Heinemann realiza una selección de la mejor parametrización para el modelo de mesoescala de área limitada MM5 en una región del sur de Alemania [Heinemann et al., 2006]. El modelo MM5 se inicializa con datos de reanálisis de NCEP (National Centers for Environmental Prediction) con una resolución temporal de 6 horas y espacial de 1°x1°. Los resultados se calculan como medias espaciales para una región de 150x150 km. a partir de la salida del modelo de mesoescala con resolución espacial de 3x3 km., resolución temporal de 1 hora y horizonte de predicción de 48 horas.

Para realizar la validación se utilizan siete configuraciones diferentes del modelo MM5 y se identifican tres situaciones de nubosidad: *cielo despejado*, *intervalos nubosos* y *cielo cubierto*. La comparativa muestra diferencias significativas en los errores para cada una de las configuraciones ensayadas. En el caso de las situaciones de cielos despejados los errores varían entre 0.25% y 14.7% en términos del RMSD relativo. Para las situaciones de intervalos nubosos los errores van desde el 2.3% al 64.4%, mientras que para los cielos cubiertos los errores varían entre -7.8% y 129%. La mejor configuración posible presenta un error relativo en términos de RMSD para cada una de las situaciones de cielo definidas anteriormente de 1.85%, 6.75% y 25.8% respectivamente.

Posteriormente, una vez obtenida la mejor configuración del modelo MM5, Heinemann realiza nuevas simulaciones con datos iniciales procedentes de dos fuentes diferentes: por un lado, el modelo de mesoescala LM (*Lokal-Modell*), utilizado por el Servicio Meteorológico Alemán DWD (*Deutscher Wetterdienst*), con una resolución espacial de 7km. y una resolución temporal de 1 hora; por otro lado, el modelo operacional del Centro de Predicción Europeo a Plazo Medio ECMWF (*European Centre for Medium-Range Weather Forecasts*) con una resolución temporal de 6 horas y resolución espacial de 60 km.

La verificación de las predicciones del modelo MM5 con datos de entrada procedentes del ECMWF (MM5-ECMWF) y DWD (MM5-LM) se lleva a cabo para un periodo de 40 días durante el verano de 2003 en una región del sur de Alemania de 200x200 km. y se comparan con medidas radiométricas en superficie.

Para el modelo MM5-ECMWF y todo tipo de situaciones de cielo, los errores en términos de RMSD son de 51% y del 56% para el primer y segundo día de predicción, respectivamente. Para el modelo MM5-LM, los errores en términos de RMSD son de 54% y del 57% para el primer y segundo día de predicción, respectivamente

Ambas configuraciones del modelo MM5 presentan mejores resultados para situaciones de cielos despejados. En el caso de las situaciones de intervalos nubosos y cielos cubiertos, el modelo con datos iniciales del ECMWF ofrece mejores resultados que el modelo MM5 con datos iniciales del modelo DWD. Por otro lado, el modelo ECMWF tiende a sobreestimar las predicciones de irradiancia solar en términos del sesgo.

#### Modelo MM5 Universidad de Jaén

Otro estudio relacionado con el modelo MM5 ha sido realizado por la Universidad de Jaén; en él se compara el rendimiento de dos parametrizaciones de capa límite atmosférica del modelo MM5, *Blackadar* y *MRF* [Pozo et al., 2009]. Las simulaciones de irradiancia solar global horaria se llevan a cabo en el sureste de la Península Ibérica; con datos correspondientes al año 2005. La validación se realiza con datos radiométricos medidos sobre superficie en el emplazamiento donde se sitúa la planta termosolar ANDASOL. El horizonte de predicción es de 72 horas y la resolución espacial del modelo para la zona bajo estudio es de 4 km<sup>2</sup>. Los datos iniciales del modelo de mesoescala se obtienen a partir de los análisis del modelo global del centro europeo de predicción (ECMWF) con una resolución temporal de 6 horas.

Los resultados de la validación se clasifican en tres conjuntos dependiendo del tipo de condición de cielo clasificado a partir del índice de claridad diario ( $k_t$ ) (ver APÉNDICE A). Los tipos de condición de cielo claro se dividen en: *cielo despejado* para valores de  $k_t$  mayores a 0.7, *intervalos nubosos* para valores de  $k_t$  menores o iguales a 0.7 y mayores o iguales a 0.4 y *cielo cubierto* para valores de  $k_t$  menores a 0.4.

En este trabajo el autor realiza la validación únicamente para las estaciones de invierno y primavera y los estadísticos de error se dan en valor absoluto. Para situaciones de cielo claro, la media de la incertidumbre es de 30Wm<sup>-2</sup> para el verano y de 130Wm<sup>-2</sup> en primavera. Existe una tendencia a infraestimar la irradiancia solar para las estaciones de otoño y primavera y a sobreestimar en verano e invierno. Las dos parametrizaciones utilizadas tienen un resultado similar en términos de RMSD y MBD, de modo que la

selección de la mejor parametrización bajo condiciones de cielo claro no es primordial a la hora de mejorar los resultados de predicción.

La utilización de modelos de cielo despejado y componentes atmosféricos que afectan a la irradiancia solar bajo este tipo de condiciones de cielo se antoja mucho más importante que la selección del tipo de parametrización de los diferentes procesos físicos para reducir el error de predicción, como queda demostrado por Wittmann en el trabajo que se presenta más adelante.

En el caso de las situaciones de intervalos nubosos, el modelo no es capaz de simular los continuos cambios en el índice de nubosidad relacionado con los tránsitos o pasos de nubes para una determinada situación general atmosférica. Los errores asociados en términos de RMSD van desde los  $200\text{Wm}^{-2}$  en invierno hasta los  $300\text{Wm}^{-2}$  en verano.

En situaciones de cielo cubierto, el modelo MM5 ofrece peores resultados que para situaciones de cielo despejado, aunque estos resultados son mejores comparándolos con las situaciones de intervalos nubosos.

Por lo tanto, los mayores errores del modelo se producen para las situaciones intermedias de nubosidad donde tiene que simular el paso o tránsito de nubes. Para las situaciones de cielo cubierto, la configuración del modelo con la parametrización *Blackadar* presenta menores errores durante invierno, mientras que la parametrización MRF obtiene mejores resultados durante el otoño.

### Modelo HIRLAM AEMet

El modelo regional operativo HIRLAM (*High Resolution Limited Area Model*) de AEMet (*Agencia Estatal de Meteorología*) se analiza comparando las predicciones de la irradiancia global con datos radiométricos medidos en superficie y con estimaciones a partir de imágenes de satélite [Geijo y Orfila, 2006]. La resolución temporal del modelo es horaria y se dispone de predicciones con un horizonte de predicción de 72 horas, con resolución espacial de  $20\text{ km}^2$ .

El ejercicio de verificación comprende varios meses en los que las propiedades de la variable irradiancia solar y su distribución de probabilidad sean totalmente diferentes (agosto, octubre y diciembre de 2005 y enero de 2006). En ellos se utilizan dos fuentes diferentes y complementarias de datos de irradiancia solar: mediciones piranométricas de estaciones de la Red Radiométrica Nacional (RRN), perteneciente a la Agencia Estatal de Meteorología (AEMet), y estimaciones derivadas de imágenes de satélite generadas por el consorcio europeo CM-SAF (*Climate Monitoring Satellite Applications Facility*).

Los datos obtenidos tras la aplicación del modelo HIRLAM presenta un coeficiente de correlación entre 0.8 y 0.9 para las siguientes 48 horas de predicción. En su trabajo, Geijo no informa del error cuadrático medio relativo, por lo tanto, no es posible realizar la comparación de los resultados de validación con el resto de trabajos presentados.

Adicionalmente, la validación del modelo HIRLAM se ha llevado a cabo durante la realización de esta tesis, cuyos resultados se han presentado en la Tarea 36 de la IEA [Martín et al., 2009a]. En dicho trabajo se analizan las predicciones de la irradiancia solar horaria para las estaciones de Granada y Huelva. El periodo de datos va desde el 16 de diciembre de 2005 hasta el 29 de febrero de 2008. En el caso de Huelva, el RMSD relativo obtenido ha sido del 31.06% y en el caso de Granada de 37.37%. Por otro lado, el sesgo relativo para cada estación ha sido de 18.74% y 22.81%, respectivamente.

Los resultados anteriores sugieren que, podría incrementarse la precisión del modelo aplicando un post-procesado mediante métodos de *downscaling* estadístico para corregir el bias del modelo que, teóricamente, se debe a los efectos microclimáticos locales que se producen sobre la radiación solar.

### Modelo AFSOL

Wittmann propone la combinación de un modelo de predicción numérica y un modelo de calidad del aire, con el fin de obtener predicciones de la irradiancia solar horaria con horizonte de predicción de 72 horas para la entrada en el mercado libre de la energía

---

[Wittmann et al., 2008]. Las predicciones meteorológicas y de diferentes componentes atmosféricos se combinan utilizando asimilación de datos procedente de registros de satélite y medidas de tierra. El objetivo de su trabajo es analizar el impacto de las predicciones de irradiancia solar en las estrategias de operación de las plantas termosolares para su participación en el mercado libre de la energía. Las predicciones se realizan en la comarca del Marquesado del Zenete, próximo a Guadix (Granada), donde se sitúa ANDASOL [Kistner et al., 2004], la primera planta termosolar en Europa con tecnología cilindro-parabólica y almacenamiento térmico mediante sales fundidas.

La zona es de gran interés para el estudio de la energía solar debido a que ANDASOL tendrá capacidad para producir 150MW cuando finalicen las tres fases constructivas en las que se divide el proyecto [EFE, 2008]. Actualmente está en funcionamiento ANDASOL-1 y ANDASOL-2, con 50MW cada una de ellas, mientras que ANDASOL-3 tiene programada su finalización en 2011.

Wittmann propone obtener predicciones de la irradiancia solar para aplicaciones energéticas basadas en el modelo de predicción de aerosoles AFSOL [Breitkreuz et al., 2007]. Este modelo operativo está especializado en la obtención de predicciones de irradiancia solar exclusivamente para condiciones de cielo claro, cubriendo Europa y el Mediterráneo. Se emplea el modelo de transferencia radiativa *libRadtran* (<http://www.libradtran.org>) para realizar los cálculos de irradiancia dentro del sistema AFSOL. El modelado de aerosoles es de gran relevancia en este sistema debido a su influencia sobre la irradiancia solar en condiciones de cielo despejado.

Las estimaciones de aerosoles, vapor de agua e información de nubes se obtienen a partir de imágenes de satélite y utilizando como base el modelo EURAD (*EUROpean Air Pollution Dispersion*). Los parámetros de las nubes que se modelan son: altura y base de la nube, contenido de agua de la nube e índice de nubosidad (*cloud fraction*). Además, se obtienen estimaciones del albedo terrestre y del ozono en las diferentes capas atmosféricas a partir de imágenes de satélite. La sonda *Earth Probe Satellite* de la NASA emplea el instrumento TOMS (*Total Ozone Mapping Spectrometer*) para obtener

el contenido de la columna de ozono atmosférico. En el caso de los valores de albedo terrestre, se utiliza el sensor MODIS (*Moderate Resolution Imaging Spectro Radiometer*) de los satélites *Terra* y *Aqua* de la NASA. El resto de variables meteorológicas del sistema AFSOL se obtienen a partir del modelo MM5.

Combinando las diferentes partes que forman el sistema AFSOL se obtienen predicciones de irradiancia solar cada hora con un horizonte temporal de 72 horas vista y una resolución espacial de 50 km<sup>2</sup>. *LibRadtran* es el encargado de integrar las predicciones de irradiancia global del modelo MM5 con medidas de ozono, aerosoles, vapor de agua y parámetros de las nubes para obtener las predicciones de irradiancia solar directa.

Las predicciones que se obtienen se comparan con las del modelo global operativo del Centro Europeo de Predicción a Medio Plazo (ECMWF), que tiene una resolución temporal de 3 horas, espacial de 25 km<sup>2</sup> y un horizonte temporal de predicción de 240 horas. Los valores de irradiancia solar global sobre superficie del modelo del ECMWF se presentan como valores acumulados cada 3 horas y se transforman en valores de irradiancia solar horarios. Para realizar el aumento de la resolución temporal a valores horarios se emplea un método de interpolación que tiene en cuenta el índice de claridad para evitar valores sobreestimados en condiciones de elevaciones solares bajas [Lorenz, 2009]. La principal diferencia entre el modelo AFSOL y ECMWF radica en que éste último utiliza valores climatológicos (de aerosoles, dióxido de carbono y contenido de ozono), que, por lo tanto, no se modifican en tiempo real.

El conjunto de datos de validación corresponde a medidas horarias de radiación solar sobre la superficie terrestre durante el mes julio de 2003. En la Tabla 2.2 se resumen los resultados de la validación de las predicciones horarias obtenidos por Wittmann considerando únicamente los días con condiciones de cielo claro.

Tabla 2.2 MBD y RMSD predicciones horarias modelos AFSOL y ECMWF situaciones de cielo claro

Modelo	MBD ( $Wm^{-2}$ )	MBD (%)	RMSD ( $Wm^{-2}$ )	RMSD (%)
<b>AFSOL</b>	28.6	3.9%	37.9	5.2%
<b>ECMWF</b>	-86.1	-11.9%	89.8	12.4%

El modelo AFSOL ofrece mejores resultados que el modelo ECWFM, siendo el valor del RMSD relativo del 5.2%, mucho menor que el 12.4% que ofrece el modelo de la Universidad de Oldenburg. El modelo AFSOL presenta una pequeña sobreestimación, mientras que en el caso del modelo modificado del ECMWF se produce una infraestimación significativa en torno al 12% de los valores de irradiación global horaria.

La Tabla 2.3 muestra los resultados de ambos modelos para todas las situaciones de cielo. Los errores relativos en términos del RMSD son considerablemente altos para ambos modelos: 47% para el caso del modelo AFSOL y 42% para el caso del modelo ECMWF. Se observa también que existe una infraestimación por parte de ambos modelos, aunque en el caso del modelo ECMWF sigue siendo bastante considerable.

Tabla 2.3 MBD y RMSD predicciones horarias modelos AFSOL y ECMWF para todas las situaciones de cielo

Modelo	MBD ( $Wm^{-2}$ )	MBD (%)	RMSD ( $Wm^{-2}$ )	RMSD (%)
<b>AFSOL</b>	-12.7	-2.2%	148.0	25.1%
<b>ECMWF</b>	-65.6	-11.1%	109.0	18.5%

En el trabajo de Wittmann se pone de manifiesto que el modelo ECMWF obtiene mejores resultados que el modelo AFSOL, a excepción de las situaciones de cielo claro, donde la mayor información contenida en el sistema AFSOL sobre parámetros atmosféricos mejora significativamente las predicciones. Otro hecho destacable es que el modelo ECMWF infraestima considerablemente las predicciones de irradiación solar del orden del 10% para el emplazamiento estudiado.

### 2.2.2 Modelos basados en *downscaling* estadístico aplicados a la predicción de la irradiancia solar

El uso de técnicas de *downscaling* estadístico se utiliza principalmente en la predicción de la precipitación y la temperatura, siendo actualmente escaso el número de trabajos dedicados exclusivamente a la irradiancia solar. A continuación, se presentan las principales aportaciones realizadas hasta la fecha que utilizan este tipo de técnicas aplicadas al recurso solar para las resoluciones temporales horaria y diaria con distintos horizontes de predicción.

#### Modelo Jensenius

Uno de los primeros trabajos en el ámbito de la predicción de la irradiancia solar fue desarrollado por Jensenius en 1981 [Jensenius, 1981]. En este trabajo se realiza la predicción de la irradiancia solar diaria para un día vista utilizando diversos predictores relacionados con la nubosidad: humedad relativa media, velocidad vertical del viento a 750 hPa, vorticidad relativa a 850 hPa, velocidad del viento a 200 hPa, punto de rocío a 700 hPa, etc. La validación se lleva a cabo en diferentes emplazamientos de los Estados Unidos; el error relativo medio en términos del MBD es del 2% y en términos del RMSD del 25% para un horizonte de predicción de un día vista.

#### Modelo Baker y Casper

Baker y Casper [Baker y Casper, 1981] desarrollan tres tipos de modelos de regresión para predecir el porcentaje de radiación extraterrestre diaria recibida en tres localizaciones diferentes de los Estados Unidos para el próximo día: Bismark (ND), Madison (WI) y Chicago (IL):

- En el primer tipo de modelo se establece una relación entre la predicción probabilística de la precipitación para el día siguiente y la irradiancia observada para obtener la predicción diaria de la irradiancia solar.
- En el segundo tipo de modelo se utiliza también una aproximación MOS (Model Output Statistics) para relacionar la predicción de humedad relativa (procedente de un modelo de predicción numérica) con la irradiancia observada.

- En el tercer tipo de modelo se establece una relación entre las observaciones de humedad relativa e irradiancia solar. Una vez obtenidos los parámetros de la ecuación MOS se utilizan las predicciones de humedad relativa del modelo operacional como predictores.

La validación se realiza para un horizonte de predicción de 24 horas. El modelo que presenta mayor precisión es el tercero cuyo RMSD en términos porcentuales es del 18%.

### Modelo NOAA

En 1983, *National Oceanic and Atmospheric Administration* (NOAA) propone un sistema para la predicción de la energía solar en los Estados Unidos en el que se utiliza la técnica MOS para obtener la predicción de la irradiancia solar a 1 y 2 días vista [Jensenius, 1983]. Para el desarrollo del modelo se utiliza un conjunto de datos de un año, procedente de 34 estaciones de *Solar Radiation Network* de NOAA. Se emplean dos métodos MOS para realizar la predicción de la irradiancia solar:

- En el primero, se deriva el índice de claridad utilizando la técnica MOS para posteriormente obtener la cantidad de irradiancia solar.
- En el segundo método, se calcula directamente la irradiancia solar aplicando la técnica MOS.

Además, se desarrollaran tres tipos de modelos para cada método MOS:

- Ecuaciones independientes para cada estación.
- Ecuaciones regionalizadas dividiendo los Estados Unidos en 6 regiones.
- Finalmente, combinando los datos para todas las estaciones dentro de una única región.

El proceso de validación demuestra que los modelos regionalizados obtenidos a partir de la primera aproximación proporcionan los mejores resultados. Lamentablemente, la calidad de las predicciones no se puede comparar con el resto de trabajos que se presentan en esta Tesis ya que las medidas de error que se utilizan en el trabajo

realizado por Jensenius es diferente; se utiliza el erro cuadrático medio relativo en términos de la irradiancia extraterrestre diaria.

### Modelo Meteocontrol

La empresa Meteocontrol propone dos modelos MOS basados en datos radiométricos medidos en Tierra y estimaciones de satélite. Ambos modelos utilizan la salida del modelo operativo del ECMWF [Bofinger y Heilscher, 2004], según el cual los predictores más relevantes son: índice de nubosidad en toda la atmósfera, punto de rocío, humedad relativa a 500 hPa, índice de nubosidad por debajo de 2000m y probabilidad de precipitación.

Los resultados de validación para un periodo de un año y 32 estaciones en Alemania utilizando datos radiométricos de Tierra y estimaciones de satélite. En concreto el RMSD para valores horarios es del 32.1% para la técnica MOS basada en datos de Tierra y del 34.9% para la técnica MOS basada en datos de satélite.

En la Tabla 2.4 se muestran los resultados de la validación utilizando datos medidos sobre superficie terrestre comparados con la validación del modelo basado en la persistencia. Como se puede apreciar, la predicción MOS basada en datos radiométricos medidos sobre superficie terrestre ofrece mejores resultados que el modelo basado en la persistencia.

Tabla 2.4 Validación de resultados para las predicciones MOS basadas en el modelo global del ECMWF y datos sobre superficie.

Modelo	RMSD horario	RMSD diario	MBD horario	MBD diario
<b>MOS, datos superficie</b>	32.1%	19.0%	2.9%	2.8%
<b>Persistencia superficie</b>	54.5%	48.1%	-0.2%	-0.6%

### Modelo Perez

Richard Perez propone un modelo para la predicción de la irradiancia solar horaria haciendo uso de las predicciones del índice de nubosidad (*sky cover*) ofrecidas por el Servicio Meteorológico de los Estados Unidos en su base de datos NDFD (*National*

*Digital Forecast Database*) [Perez et al., 2007]. El horizonte temporal de las predicciones que obtiene Perez es de 76 horas.

Perez propone una nueva relación no-lineal entre las predicciones del coeficiente de nubosidad del NDFD, como modelo base, y la irradiancia solar medida en superficie. Para establecer esta relación se utilizan medidas de irradiancia solar en superficie y valores de predicción del coeficiente de nubosidad del NDFD para la localidad de Albany en los Estados Unidos para un periodo comprendido entre abril y septiembre de 2004.

Las predicciones del NDFD se ofrecen en dos conjuntos. El primero tiene una resolución temporal de 3 horas y un horizonte de predicción de 72 horas y el segundo tiene una resolución temporal de 6 horas y un horizonte de predicción de 168 horas. La resolución espacial de las predicciones es de  $0.05^{\circ} \times 0.05^{\circ}$ . En su trabajo, Perez no informa de la metodología aplicada para obtener las predicciones horarias a partir de las predicciones originales del NDFD cada 3 y 6 horas.

La validación de los modelos se realiza en la localidad de Albany, utilizando medidas radiométricas de Tierra para un horizonte temporal de predicción de 76 horas (ver Tabla 2.5). Como se puede observar en ella, la precisión de las predicciones disminuye al aumentar el horizonte de predicción, aunque el error permanece bastante estabilizado dentro de unos márgenes:  $RMSD=30\%-40\%$  al validar con datos de tierra.

Tabla 2.5 Errores de la técnica MOS propuesta por Perez realizando la validación con datos de tierra

	Rango de Predicción	MBD (%)	RMSD ( $Wm^{-2}$ )
<b>Tierra</b>	< 4 horas	-10%	32%
	4 – 8 horas	-3%	34%
	8 – 26 horas	-1%	38%
	26 – 76 horas	-4%	40%

### 2.2.3 Modelos estadísticos clásicos aplicados a la predicción de la irradiancia solar

En las siguientes secciones se analizan algunos de los trabajos desarrollados en el ámbito de la predicción de la irradiancia solar utilizando métodos estadísticos clásicos: lineales y no lineales.

#### Modelo Red Neuronal Multivariable

Uno de los primeros trabajos relevantes aplicados en la predicción de la irradiancia solar diaria mediante redes neuronales artificiales (RNA) se publicó en 1994 [Elizondo et al., 1994]. El objetivo de esta investigación es el desarrollo de un modelo de red neuronal con el cual predecir la irradiancia solar diaria utilizando como valores de entrada las variables meteorológicas observadas (temperatura máxima y mínima del aire y precipitación diaria acumulada) y variables calculadas analíticamente (longitud del día y radiación diaria total de cielo claro).

En este trabajo se utiliza un algoritmo de retropropagación para el entrenamiento de las redes neuronales. El entrenamiento y evaluación de la red neuronal se realiza con 20 años de datos procedentes de 4 emplazamientos en el sureste de los Estados Unidos. Los datos se dividen en un conjunto de datos de entrenamiento constituido por los 11 primeros años y un conjunto de validación formado por los años restantes.

La evaluación del modelo de red neuronal se realiza comparando la irradiancia solar diaria predicha (salida de la red neuronal) con las observaciones de radiación solar. El modelo de red neuronal propuesto presenta una tendencia a sobreestimar la irradiancia solar para valores bajos e infraestimar la irradiancia para valores altos. El autor informa que el error cuadrático medio (RMSD) es de  $2.92\text{MJ/m}^2\text{día}$ .

#### Modelo Red Neuronal Multietapa

Otro modelo más complejo basado en redes neuronales en múltiples etapas fue desarrollado por Kemmoku para obtener la predicción de la irradiancia global total del próximo día (resolución temporal y horizonte de predicción de un día) [Kemmoku et al., 1999]. El autor utiliza datos meteorológicos diarios de 6 años (1988-1993) procedentes

de la localidad de Omaezaki en Japón; la validación del modelo se realiza sobre un conjunto de datos de irradiancia solar diaria para cuatro meses de 1994 (Abril, Agosto, Octubre y Diciembre), independiente del conjunto de datos utilizado para la estimación de los parámetros de la red neuronal.

El conjunto de parámetros meteorológicos usados en la entrada de la red neuronal está formado por la diferencia entre el primer y segundo máximo y el primer y segundo mínimo local de la presión atmosférica, su valor medio y el valor registrado a las 18:00 horas, el índice de transparencia atmosférica, la temperatura y la irradiancia solar para la hora previa a la puesta del Sol.

La predicción se realiza utilizando redes neuronales independientes en tres etapas con las siguientes características:

- La primera etapa se basa en la consideración de que existe una correlación entre la irradiancia del próximo día y la diferencia entre las presiones atmosféricas medias entre el día anterior y el día siguiente. Esto se debe a que la irradiancia solar presenta una dependencia con las condiciones meteorológicas relacionadas estrechamente con el paso de masas de aire de alta y baja presión atmosférica. De esta forma, en esta primera fase la presión atmosférica media se puede utilizar como entrada a una red neuronal para predecir tres niveles de irradiancia solar.
- La segunda etapa utiliza una red neuronal para predecir la irradiancia solar del próximo día a partir de la presión atmosférica media y de los datos meteorológicos.
- La tercera etapa refina el valor de la predicción de la irradiancia solar de la segunda etapa. En esta etapa final se consideran niveles de irradiancia solar altos, medios y bajos y se realiza la predicción final utilizando una red neuronal, entre tres independientes, de acuerdo al nivel de irradiancia solar predichos en la primera etapa.

El RMSD en términos porcentuales es del 20% para los cuatro meses que componen el conjunto de datos de validación.

### Modelo Safi

Safi [Safi et al., 2002] propone dos nuevos procedimientos para la modelización de la irradiancia global diaria basados en la transformación de la serie temporal para que cumpla las condiciones de estacionariedad y normalidad.

Para realizar la validación de las predicciones, Safi utiliza medidas radiométricas realizadas en la Universidad Cadi Ayyad, situada en la ciudad de Marrakech (Marruecos), durante un periodo de un año. Los resultados de la validación muestran que, para el caso del procedimiento basado en el índice de claridad diario, el error en términos de MBD es del 11.87% y en términos del RMSD es del 28%, mientras que para la *componente solar atenuada* (ver APÉNDICE A) los errores en términos del MBD son del 2.28% y en términos del RMSD del 16%. En este trabajo se concluye que la predicción basada en el predictor componente solar atenuada presenta mayor precisión que el índice de claridad aplicando un modelo autorregresivo del mismo orden sobre ambas series temporales.

### Modelo Preprocesado Wavelet + Red Neuronal

El uso combinado del análisis wavelet y redes neuronales da lugar a un nuevo modelo de predicción que se conoce como *wavelets-networks* y se caracteriza porque la función de activación de las redes neuronales son funciones *wavelets*. Esta técnica permite predecir valores futuros teniendo en cuenta la variabilidad en el dominio de la frecuencia de los datos recientes [Zhang y Benveniste, 1992].

Mellit utiliza las redes neuronales wavelets para predecir la irradiancia solar diaria del próximo día (resolución temporal y horizonte de predicción de un día) [Mellit et al., 2006]. Los datos radiométricos que utiliza Mellit corresponden a una estación situada en Argelia durante el periodo comprendido entre 1981 y 2001. Se ensayan diferentes

estructuras de redes neuronales, siendo la configuración que ofrece mejores resultados la constituida por 5 valores de entrada y 12 neuronas en la capa oculta.

La validación se realiza comparando los datos estimados por la red neuronal frente a un conjunto independiente de datos observados. El autor informa únicamente que el error relativo medio (MRE) del modelo es del 5.6%.

### Modelos CIEMAT

Durante la realización de esta tesis, se ha realizado diversos trabajos en los que se utilizan técnicas predictivas lineales y no-lineales basadas en la serie temporal índice de claridad semidiario dividido en el cenit solar [Martín et al., 2006; Martín et al., 2007; Martín et al., 2008a; Martín et al., 2008b; Martín et al., 2008c; Martín et al., 2009b]. En dichos trabajos, la salida del índice de claridad de los modelos se transforma en valores de irradiancia solar semidiaria acumulados con un horizonte de predicción de tres días. El modelo lineal ensayado es el modelo autorregresivo y los modelos no-lineales son los modelos basados en redes neuronales y lógica difusa [Martín et al., 2010].

La medida de error que se ofrece en este trabajo es la mejora de los modelos en términos de RMSD con respecto al modelo basado en la persistencia. Los resultados de validación que se obtienen muestran que los modelos no-lineales mejoran al modelo basado en la persistencia y al modelo lineal para todos los horizontes de predicción. Entre los dos modelos no-lineales ensayados para el primer semidía de predicción el mejor modelo es el modelo basado en lógica difusa mientras que para el resto de horizontes de predicción el modelo que mejores resultados ofrece es la red neuronal. Los valores de error, fruto del proceso de investigación que se ha realizado se presentan a lo largo de esta memoria.

#### *2.2.4 Modelos basados en imágenes de satélite*

Las imágenes de satélite, gracias a su resolución espacial y temporal, son una excelente fuente de datos para realizar predicciones en el muy corto plazo o *nowcasting*. Su utilidad se encuentra principalmente en el efecto sustancial que tienen las nubes en la

irradiancia solar que llega a la superficie terrestre [Zarzalejo et al., 2007; Zarzalejo et al., 2009; Espinar et al., 2009].

La variación temporal de las estructuras nubosas se debe fundamentalmente al movimiento propio de las nubes, causa ésta por la que los campos de vectores de movimiento (*Motion Vector Field* MVF o *Atmospheric Motion Vector Field* AMF) son una herramienta apropiada para la predicción del movimiento de las masas nubosas [Lorenz, 2004; Holmlund et al., 2008].

La aplicación de los campos de vectores de movimiento directamente sobre el coeficiente de cobertura nubosa permite obtener predicciones de la irradiancia solar a partir de imágenes de satélite para un horizonte de predicción del muy corto plazo o *nowcasting* [Lorenz, 2004]. El procedimiento para derivar los MVFs a partir del coeficiente de cobertura nubosa consiste en identificar regiones de dos imágenes consecutivas y calcular un vector de desplazamiento óptimo entre las dos regiones a partir de la diferencia cuadrática media en cada píxel. Los vectores de campo de movimiento que se obtienen se aplican a la imagen actual para derivar la posición de las nubes en el siguiente instante temporal. Finalmente, se aplica un filtro de suavizado para la eliminación de pequeñas estructuras que varían aleatoriamente.

Situaciones con nubes no homogéneas corresponden a imágenes con alta variabilidad del coeficiente de cobertura nubosa y son más difíciles de predecir, por lo que presentan mayor incertidumbre que las situaciones de cielos despejados. En los casos en los que existen valores altos de irradiancia solar con baja variabilidad, el error es del 10% en términos de RMSD para un horizonte de predicción de 30 minutos mientras que si éste aumenta hasta 6 horas es del 25% en términos del RMSD. En el proceso de validación no se ha tenido en cuenta el efecto que tiene sobre el valor del índice de cobertura nubosa las elevaciones solares bajas o los ángulos de *scattering* elevados debido a que el punto donde se realiza la predicción está excesivamente distante del nadir del satélite y se produce una distorsión en la imagen.

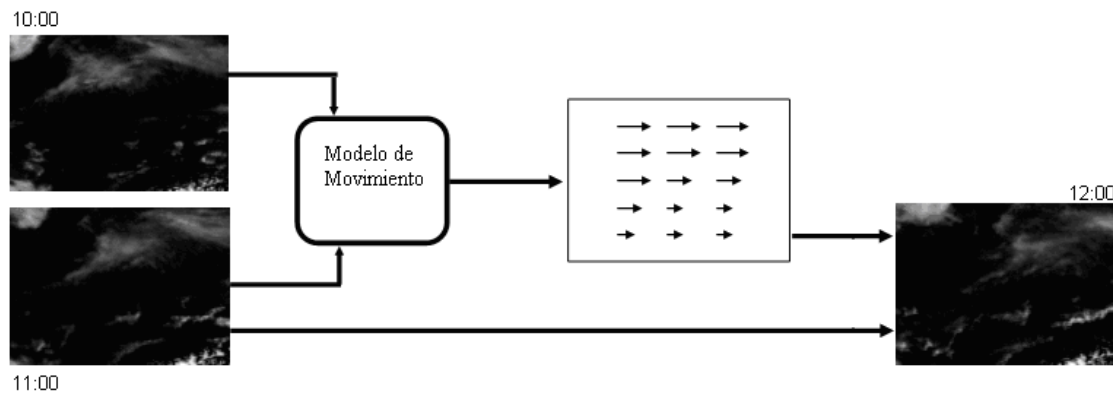


Figura 2.2 Campo de vectores de movimiento calculados en un esquema de predicción de corto plazo [Hammer et al., 1999]

Otra metodología propuesta por Lorenz para la predicción del coeficiente de cobertura nubosa está basada en el empleo de redes neuronales, aunque el error obtenido no mejora al de la predicción mediante MVFs.

La aplicación de los campos de vectores de movimiento sobre las imágenes de satélite es una herramienta fundamental para la obtención de predicciones de la irradiancia solar global en un horizonte temporal que va desde la propia resolución temporal del satélite hasta las 6 horas [Heinemann et al., 2006]. La obtención de predicciones con mayor resolución temporal a la propia del satélite debe realizarse mediante la aplicación de técnicas estadísticas, como el análisis de series temporales o redes neuronales, sobre medidas radiométricas en superficie o utilizando la técnica de campos de vectores de movimiento sobre medidas de cámaras de cielo, ya que para estos horizontes temporales no está disponible la información de los modelos globales [Casa-Nova et al., 2005].

### 2.2.5 Comparación de modelos de predicción de la irradiancia solar

En este apartado se presenta una comparativa de modelos realizada recientemente dentro de la Tarea 36 (*Solar Resource Knowledge Management*) de la IEA/SHC. Una de sus líneas de trabajo se centra de forma exclusiva en la predicción de la irradiancia solar. Dentro de esta línea, se ha realizado recientemente una comparativa de diferentes

modelos de predicción de la irradiancia solar horaria; todos ellos basados en modelos dinámicos y técnicas de *downscaling* estadístico. El horizonte de predicción es de 72 horas. El estudio comparativo se divide en diferentes zonas de la Tierra que se enumeran y detallan a continuación:

- *Estados Unidos*, cuyo periodo de validación va desde el 1 de abril de 2007 hasta el 30 de septiembre de 2007.
- *Canadá*, cuyo periodo de validación va desde el 1 de octubre de 2008 hasta el 30 de septiembre de 2009.
- *Europa*, donde se incluyen Alemania, Austria, Suiza y España. El periodo de validación de los modelos va desde el 1 de julio de 2007 hasta el 30 de junio de 2008.

Las instituciones participantes y los modelos que utilizan cada una de ellas son:

- SUNY/Albany. El modelo utilizado hace uso de las predicciones del Servicio Meteorológico de los Estados Unidos (NDFD). Este modelo se presentó anteriormente [Perez et al., 2007].
- MeteoTest. El modelo utilizado emplea como entrada los datos del modelo global de predicción GFS (*Global Forecasting System*) de NOAA y realiza un *downscaling* dinámico utilizando el modelo de mesoescala WRF.
- Universidad de Oldenburg. Realiza las predicciones mediante un *downscaling* estadístico a partir de la salida del modelo de predicción operativo del centro Europeo de Predicción (ECMWF).
- CanMet. El modelo numérico utilizando es el *Canadian Meteorological Centre's Global Environmental Multiscale Model* (GEM) [Pelland, 2009].
- Blue Sky. Realiza un *downscaling* estadístico sobre las predicciones del modelo global GFS.
- CENER. El modelo implementado utiliza los datos del modelo GFS y realiza un *downscaling* dinámico utilizando el modelo de mesoescala Skiron (GFS-Skiron). El modelo Skiron no da directamente la salida de la irradiancia solar por lo que se aplica un post-procesado para obtener esta variable.

- Universidad de Jaén. Realiza un *downscaling* dinámico utilizando el modelo WRF a partir de la salida del modelo global GFS.

Los resultados de validación de las predicciones se presentan para cada región, donde se presenta cada uno de los equipos participantes y los resultados de predicción para cada emplazamiento.

En general se aplican los modelos para zonas geográficas concretas dado que existe una fuerte dependencia de la radiación solar con la climatología, que es característica de la zona de estudio. Cabe destacar que cada institución participante en la validación ha seleccionado por propia voluntad las zonas en las que realizar las predicciones.

#### Estados Unidos

Los emplazamientos donde existen estaciones radiométricas de medida en la superficie con las que comparar las predicciones que se utilizan en el estudio son Desert Rock (NV), Boulder (CO) y Goodwin Creek (MS). Los resultados que se obtienen tras realizar la validación sobre las predicciones se muestra en la Tabla 2.6, donde se presenta la institución participante en la validación, la estación donde se realiza la validación y los errores relativos en términos porcentuales del MBD y RMSD.

El modelo ECMWF de la Universidad de Oldenburg es el que mejores resultados ofrece de los tres modelos analizados, seguido del modelo de SUNNY/Albany y finalmente el modelo de Meteotest. Para todos los modelos (excepto para el modelo SUNNY/Albany en la estación de Goodwin) existe una sobreestimación las predicciones ( $MBD > 0$ ).

El mejor ajuste se produce para la estación de Desert Rock con valores de predicción muy ajustados ( $MBD < 3\%$ ) y errores relativos aceptables ( $RMSD = 18\%$ ).

El resto de emplazamientos presentan errores considerables (30-50% en términos de RMSD), debido a que existen efectos locales que no se tienen en cuenta en las parametrizaciones de los modelos físicos y que los modelos tienen una baja resolución

espacial de salida (11-25km.), para poder realizar extrapolaciones entre los puntos de rejilla del modelo y la estación bajo estudio. De esta forma, la aplicación de los modelos de *downscaling* estadístico puede ser una forma de mejorar la precisión de los resultados que se obtienen.

Tabla 2.6 Errores de predicción de la irradiancia solar global horaria para los EE.UU.

<b>Institución</b>	<b>Estación</b>	<b>MBD (%)</b>	<b>RMSD (%)</b>
<i>SUNNY/Albany</i>	Desert Rock	2%	18%
<i>Meteotest</i>		2%	18%
<i>U. Oldenburg</i>		3%	18%
<i>SUNNY/Albany</i>	Boulder	3%	41%
<i>Meteotest</i>		19%	50%
<i>U. Oldenburg</i>		11%	40%
<i>SUNNY/Albany</i>	Goodwin	-4%	36%
<i>Meteotest</i>		18%	41%
<i>U. Oldenburg</i>		6%	32%

El error cuadrático medio de todos los modelos es de 33%. El RMSD medio de SUNNY/Albany es 32%, el de Meteotest es 36% y el de la Universidad de Oldenburg del 30%. El MBD medio de todos los modelos es 7% y el de cada modelo es: SUNNY/Albany es 0%, el de Meteotest es 13% y el de la Universidad de Oldenburg del 7%.

### Canadá

Se estudian cinco localidades de Canadá: Bondville, Fort Peck, Sioux Falls, Penn State y Varennes; obteniéndose los resultados que se muestran en la Tabla 2.7 tras realizar la validación sobre las predicciones.

Tabla 2.7 Errores de predicción de la irradiancia solar global horaria para Canadá

<b>Institución</b>	<b>Estación</b>	<b>MBD (%)</b>	<b>RMSD (%)</b>
<i>CanMet</i>	Bondville	8%	28%
	Fort Peck	5%	28%
	Sioux Falls	7%	32%
	Penn State	10%	29%
	Varennes	4%	33%

El error cuadrático medio de CanMet es 30% y el MBD medio es 7%.

Del análisis de los resultados anteriores se pueden destacar las siguientes apreciaciones:

- Los errores relativos en términos de RMSD son menores al comparar con las estaciones de Goodwin y Boulder, estudiadas anteriormente. Si bien si comparamos los valores medios, los resultados son muy parecidos, 7% de MBD para ambos casos y 33% de RMSD para EEUU frente al 30% para CanMet. El modelo ECMWF es el único que se aproxima al nivel de precisión, aunque los resultados de predicción no se pueden comparar, ya que el conjunto de datos es totalmente diferente.
- A partir del MBD relativo puede comprobarse que existe sobreestimación de la predicción de la irradiancia solar en todas las estaciones.

### Suiza

El periodo de validación de los datos va desde el 1 de julio de 2007 hasta el 31 de junio de 2008. Los resultados de la validación de los resultados de predicción se muestran en la Tabla 2.8.

Tabla 2.8 Errores de predicción de la irradiancia global horaria para Suiza

Institución	Estación	Horizonte Predicción	MBD (%)	RMSD (%)
<i>Meteotest</i>	Suiza	1 <sup>er</sup> día	1.6%	51.2%
<i>U. Oldenburg</i>		1 <sup>er</sup> día	-0.3%	39.6%
<i>Meteocontrol</i>		1 <sup>er</sup> día	-6.6%	45.0%
<i>Meteotest</i>		2 <sup>o</sup> día	1.0%	53.2%
<i>U. Oldenburg</i>		2 <sup>o</sup> día	-0.4%	41.8%
<i>Meteocontrol</i>		2 <sup>o</sup> día	-6.7%	46.3%
<i>Meteotest</i>		3 <sup>er</sup> día	-0.8%	57.8%
<i>U. Oldenburg</i>		3 <sup>er</sup> día	-0.6%	42.7%
<i>Meteocontrol</i>		3 <sup>er</sup> día	-7.2%	48.1%

El error cuadrático medio de Meteotest es 54.07%, el de U. Oldenburg es 41.37% y el de la Meteocontrol es del 46.47%. El MBD medio de Meteotest es 0.6%, el de U. Oldenburg es -0.43% y el de la Meteocontrol es del -4.3%

Todos los modelos analizados ofrecen mejores resultados que el modelo basado en la persistencia del día anterior para un horizonte de predicción a 3 días vista. El modelo de persistencia que se utiliza consiste en considerar la irradiancia solar para los próximos tres días igual a la medida en el día actual. El modelo con mejor rendimiento de los analizados es el modelo utilizado en la Universidad de Oldenburg que utiliza las predicciones del modelo global del Centro Europeo de predicción (ECMWF). Reseñar los resultados obtenidos por Meteocontrol, que tras realizar un post-procesado a los datos del ECMWF obtiene peores resultados que el modelo más simple utilizado por la Universidad de Oldenburg. Es de señalar, por otro lado, que los valores de RMSD se mantienen considerablemente elevados en todos los casos, si bien el valor no crece de forma apreciable cuando se extiende el horizonte de predicción de 1 a 2 y 3 días.

Austria

La Tabla 2.9 muestra los resultados de la validación de las predicciones de irradiancia global con resolución horaria para Austria llevadas a cabo por las diferentes instituciones participantes para las estaciones situadas en Viena y Linz.

Tabla 2.9 Errores de predicción de la irradiancia solar global horaria para Austria

Institución	Estación	MBD (%)	RMSD (%)
<i>Blue Sky</i>	Viena	3.1%	45.9%
<i>U. Oldenburg</i>		2.7%	44.1%
<i>Meteotest</i>		5.9%	66.8%
<i>CENER</i>		4.5%	56.3%
<i>Blue Sky</i>	Linz	5.2%	57.8%
<i>U. Oldenburg</i>		12.9%	51.8%
<i>Meteotest</i>		28.0%	81.5%
<i>CENER</i>		18.8%	66.4%

El único modelo que presenta valores de predicción bastante ajustados para ambas estaciones es el modelo de Blue Sky en tanto que el resto sólo muestra un ajuste para Viena (MBD<6%) mientras que la sobreestimación para Linz suepra el 13% en todos los casos. Los errores cuadráticos medios siguen siendo altos para cualquiera de los modelos, entre el 47.9% y el 74.3% en media siendo el modelo de la Universidad de Oldenburg el que mejores resultados ofrece (47.9%). Se puede observar, también, que todos los modelo sobreestiman el valor de la irradiancia solar.

El error cuadrático medio de todos los modelos es de 58.9%. El RMSD medio de Blue Sky es 51.3%, el de U. Oldenburg es 47.9%, el de Meteotest es del 74.3% y el de CENER es del 61.3%. El MBD medio de Blue Sky es 4.15%, el de U. Oldenburg es 7.8%, el de Meteotest es del 16.95% y el de CENER es del 11.65%.

En el caso de las predicciones de irradiancia global con resolución diaria los resultados de validación se muestran en la Tabla 2.10.

Tabla 2.10 Errores de predicción de la irradiancia solar global diaria para Austria

<b>Institución</b>	<b>Estación</b>	<b>MBD (%)</b>	<b>RMSD (%)</b>
<i>Blue Sky</i>	Viena	1.2%	24.6%
<i>U. Oldenburg</i>		0.0%	26.6%
<i>Meteotest</i>		-0.7%	41.3%
<i>CENER</i>		-0.1%	34.2%
<i>Blue Sky</i>	Linz	1.2%	26.4%
<i>U. Oldenburg</i>		9.7%	28.5%
<i>Meteotest</i>		20.7%	48.8%
<i>CENER</i>		12.5%	36.7%

En el caso de la radiación solar diaria, el modelo que mejores resultados ofrece en términos de RMSD (%) es el modelo de Blue Sky. Los resultados de ambos modelos que realizan *downscaling* estadístico (Blue Sky y U. Oldenburg) son mucho mejores que los de los dos modelos que realizan *downscaling* dinámico (Meteotest y CENER). El modelo de la Universidad de Oldenburg (basado en ECMWF) ofrece resultados algo peores a los del modelo de Blue Sky (basado en GFS).

Como en el caso de la irradiancia horaria, el modelo Blue Sky presenta un buen ajuste (MBD=1.2%) mientras que el resto de los modelos ajusta bien para Viena aunque no para Linz, lo que indica que la precisión de dichos modelos y dependencia de la localización geográfica, esto es las condiciones climáticas y meteorológicas. En el caso de los errores cuadráticos medios, los valores se reducen notablemente, en torno al 25%, lo que indica que la dispersión en datos es mucho menor para las predicciones diarias que para las horarias.

El error cuadrático medio de todos los modelos es de 33.3%. El RMSD medio de Blue Sky es 25.5%, el de U. Oldenburg es 27.6%, el de Meteotest es de 45.1% y el de CENER es del 35%. El MBD medio de Blue Sky es 1.2%, el de U. Oldenburg es 4.85%, el de Meteotest es del 10% y el de CENER es del 6.2%.

España

La Tabla 2.11 muestra los resultados de la validación de las predicciones de irradiancia global con resolución horaria para España llevadas a cabo por las diferentes instituciones participantes para las estaciones situadas en Andasol y Córdoba.

Tabla 2.11 Errores de predicción de la irradiancia solar global horaria para España

Institución	Estación	MBD (%)	RMSD (%)
<i>Universidad de Jaén</i>	Andasol	17.7%	37.3%
<i>Universidad de Jaén</i>	Córdoba	7.3%	19.7%
<i>U. Oldenburg</i>		-1.3%	16.7%
<i>CENER</i>		-2.5%	21%

Al analizar los resultados que se presentan en la Tabla 2.11 se puede destacar:

- Comparando los errores que se obtienen para la estación de ANDASOL con los presentados para la estación de Córdoba, se observa que son considerablemente peores para la primera estación. Esta gran diferencia en la incertidumbre puede deberse a la baja resolución de los datos de entrada de los modelos globales (0.5 grados o unos 50km. de resolución espacial) que no tienen en cuenta la gran complejidad orográfica, ni los efectos meteorológicos locales en la zona donde se sitúa ANDASOL.
- Al comparar los resultados del emplazamiento ANDASOL para el modelo de la Universidad de Jaén con los que obtiene Wittmann [Wittmann et al., 2008], se puede comprobar que el modelo ECMWF de la Universidad de Oldenburg ofrece mejores resultados (RMSD=16.7%) que el modelo AFSOL (RMSD=21%) y que el modelo de la Universidad de Jaén (RMSD=37.3%). No obstante, se debe resaltar que los periodos temporales de ambos trabajos son diferentes: Wittmann analiza únicamente el periodo de julio de 2003, donde existe un gran predominio de días de cielo despejado, mientras que la

Universidad de Jaén analiza el periodo que va desde agosto de 2007 hasta marzo de 2008, que contiene un mayor número de días de cielo cubierto o semicubierto al incluir en este periodo las estaciones de otoño, invierno y primavera, donde se concentra la época de lluvias y nevadas en la zona donde se localiza ANDASOL.

- En el caso de la estación de Córdoba, se puede comprobar que el modelo ECMWF de la Universidad de Oldenburg ofrece mejores resultados, mientras que el modelo de la Universidad de Jaén tiene menor incertidumbre que el modelo Skiron-MOS de CENER.
- Los altos valores del error cuadrático medio indican que la influencia de la serie temporal sobre la que se extiende el estudio; 1 mes para Andasol y 8 meses para Córdoba, es notable, por lo que un mayor rango en el periodo probablemente reduciría el valor de RMSD.

### **2.3 Resumen de los modelos de predicción de la irradiancia solar**

Un resumen de los resultados de validación, de los modelos presentados en el apartado anterior, se muestra en la Tabla 2.12 para la predicción horaria y en la Tabla 2.13 para la predicción diaria. En estas tablas se presenta el autor o equipo del trabajo, el emplazamiento del estudio, el periodo de validación, el modelo empleado y los errores MBD y RMSD en términos porcentuales.

Al analizar los resultados de la Tabla 2.12 se obtienen las siguientes conclusiones:

- Los errores de predicción varían considerablemente dependiendo de la latitud geográfica y de las condiciones de cielo claro predominantes en la zona de estudio. Para estaciones con función de distribución unimodal del índice de claridad -que se caracterizan por tener un predominio de días despejados-, los errores en términos de RMSD se sitúan en torno al 20%; mientras que, si la función de distribución del índice de claridad es bimodal, los errores están por encima del 40%.

- El modelo basado en campos de vectores de movimiento (MVF-Meteosat) ofrece los menores errores para la predicción en el muy corto plazo o *nowcasting*.
- El modelo que presenta menores errores, para todos los emplazamientos y periodos, es el utilizado por la Universidad de Oldenburg, que está basado en el modelo ECMWF.

Tabla 2.12 Resumen de resultados trabajos predicción de la irradiancia global horaria

Autor/Institución	Región	Periodo	Modelo	Min/Max MBD(%)	Min/Max RMSD(%)
Meteocontrol	Europa	-	ECMWF-MOS Tierra		32.1%
Meteocontrol		-	ECMWF-MOS Satélite	2.3%	34.9%
[Lorenz, 2004]			MVF-Meteosat		25%
[Heinemann et al., 2006]		-	MM5-ECMWF		51%
[Heinemann et al., 2006]		-	MM5-LM		54%
Sunny/Albany	EEUU	-	NDFDMOS	-2/5%	35%/48%
Wittmann	Guadix	Julio 2003	ECMWF-MOS	-11.1%	18.5%
Wittmann		Julio 2003	AFSOL-MOS	-2.2%	25.1%
SUNNY/Albany	EEUU	4/2007-9/2007	NDFD-MOS	-4/3%	18/41%
Meteotest		4/2007-9/2007	GFS-WRF	2/19%	18/50%
U. Oldenburg		4/2007-9/2007	ECMWF-OL	3/11%	18/40%
CanMet		Canadá	10/2008-9/2009	GEM	4/10%
Meteotest	Suiza	7/2007-6/2008	GFS-WRF	-0.8/1.6%	51/58%
U. Oldenburg		7/2007-6/2008	ECMWF-OL	-0.3/-0.6%	39/43%
Meteocontrol		7/2007-6/2008	ECMWF-MC	-6/-7%	45/48%
Blue Sky	Austria	4/2007-9/2007	GFS-MOS	3/5%	46/58%
U. Oldenburg		4/2007-9/2007	ECMWF-OL	2.7/13%	44/52%
Meteotest		4/2007-9/2007	GFS-WRF	6/28%	67/82%
CENER		4/2007-9/2007	GFS-Skiron	4.5/19%	56/66%
Universidad Jaén	Andasol	7/2007-6/2008	GFS WRF	17%	34/41%
Universidad Jaén	Córdoba	7/2007-6/2008	GFS WRF	7/8%	19/21%
U. Oldenburg		7/2007-6/2008	ECMWF-OL	-3/0%	16/18%
CENER		7/2007-6/2008	GFS-Skiron	-3/-2%	19/23%
AEMet-CIEMAT	Huelva	12/2005-2/2008	HIRLAM	18.74%	31.06%
AEMet-CIEMAT	Granada	12/2005-2/2008	HIRLAM	22.81%	37.37%

A partir de los resultados anteriores se puede concluir que los errores de las predicciones de la irradiancia solar horaria tienen una gran incertidumbre, por lo que no se puede considerar su utilización como apta en cualquier emplazamiento para aplicaciones energéticas.

Tabla 2.13 Resumen de resultados trabajos predicción de la irradiancia global diaria

Autor/Institución	Región	Periodo	Modelo	Min/Max MBD(%)	Min/Max RMSD(%)
<b>Safi</b>	Marrakesh	-	AR $K_T$	11.87%	28%
<b>Safi</b>		-	AR LC	2.28%	16%
<b>Jensenius</b>	EEUU	-	MOS	2%	25%
Baker y Casper		-	MOS		18%
<b>Red Neuronal Multietapa</b>	Japón	1988-1994	Red Neuronal		20%
<b>Meteocontrol</b>	Europa	-	ECMWF-MOS Tierra	2.8%	19.0%
<b>Meteocontrol</b>		-	ECMWF-MOS Satélite	2.1%	22.9%
<b>Blue Sky</b>	Austria	4/2007-9/2007	GFS-MOS	1.2%	25/27%
<b>U. Oldenburg</b>		4/2007-9/2007	ECMWF-OL	0/9.7%	27/28.5%
<b>Meteotest</b>		4/2007-9/2007	GFS-WRF	-0.7/21%	41/49%
<b>CENER</b>		4/2007-9/2007	GFS-Skiron	0/12.5%	34/37%

Al analizar los resultados de la Tabla 2.13 se obtienen las siguientes conclusiones:

- Los errores de la predicción de la irradiancia solar diaria son inferiores a los de la predicción horaria, estando el MBD de los modelos estadísticos en torno al 2%.
- Destacan los bajos errores del modelo autorregresivo que utiliza el predictor componente solar atenuada frente a las predicciones a partir del índice de claridad para obtener las predicciones de irradiancia solar diaria para el día siguiente. Este modelo no utiliza ningún tipo de información de los modelos numéricos.
- Los modelos de Baker y Casper y Red Neuronal Multietapa, que utilizan información meteorológica de irradiancia solar, humedad, presión y temperatura; obtienen errores en torno al 20% en términos del RMSD relativo [Baker y Casper, 1981; Kemmoku et al., 1999].
- Los errores que se obtienen para los modelos empleados en Austria son bastante superiores al resto de modelos, aunque existe una gran diferencia (entorno al 10-20% en términos de RMSD) entre los modelos puramente físicos (GFS-WRF y GFS-Skiron) y los modelos que realizan un postprocesado estadístico a la salida de los modelos globales de predicción (GFS-MOS y ECMWF-OL).
- Los resultados que se obtienen con el modelo *ECMWF-MOS-Tierra* para emplazamientos de Alemania presenta errores por debajo del 20% en términos

de RMSD. Este valor es bastante menor al obtenido en Austria, que se sitúa en latitudes próximas [Bofinger y Heilscher, 2007].

- Ninguno de los trabajos presentados para la predicción de la irradiancia solar diaria presenta resultados específicos para emplazamientos localizados en la Península Ibérica.

De todo lo anterior, se puede obtener como reflexión principal que los modelos que obtienen mejores resultados son el modelo, que utiliza la variable componente solar atenuada como predictor, y las técnicas de *downscaling* estadístico aplicadas a la salida de los modelos globales. De esta forma, se puede postular que la mejora de los resultados de predicción diaria puede tener un camino a recorrer en el estudio y conocimiento de las propiedades estadísticas y de linealidad de la serie temporal irradiancia solar, así como la selección de predictores para su utilización en el *downscaling* estadístico.

## 2.4 Conclusiones

En este capítulo se han presentado los principales conceptos relacionados con la predicción meteorológica que son de utilidad para la comprensión de esta memoria. Además, se ha realizado una revisión de los trabajos desarrollados hasta la fecha en el ámbito de la predicción de la irradiancia solar para un horizonte temporal de predicción de *nowcasting* y corto plazo.

La predicción de la irradiancia solar se puede realizar utilizando diferentes técnicas, que se pueden agrupar de la siguiente forma:

- Predicción mediante modelos numéricos de predicción y *downscaling* dinámico.
- Predicción mediante métodos de *downscaling* estadístico aplicado a la salida de los modelos numéricos de predicción.
- Predicción de series temporales basada en la aplicación de técnicas estadísticas lineales y no lineales.
- Predicción del coeficiente de cobertura nubosa a partir de imágenes de satélite.

La predicción de la irradiancia solar utilizando únicamente la salida de los modelos numéricos presenta un sesgo alto y un elevado error cuadrático medio, que depende de la función de distribución de los datos de irradiancia para la estación bajo estudio (en emplazamientos con predominio de condiciones de cielo claro los errores son menores). Debido a la complejidad de la dinámica atmosférica, los modelos de predicción numérica no son capaces de resolver las situaciones de cambios continuos de nubosidad relacionados con los pasos de nubes para las situaciones de cielo cubierto e intervalos nubosos. Estos modelos predicen únicamente las situaciones sinópticas o de mesoescala. Por estas razones, los modelos físicos presentan grandes errores en la predicción de la irradiancia solar para emplazamientos donde no existe un predominio de situaciones de cielo claro.

La mejora de las predicciones de los modelos globales y de mesoescala se puede lograr mediante un post-procesado utilizando los modelos estadísticos, dando lugar a los modelos de *downscaling* estadístico o MOS. Este tipo de modelo se utiliza para corregir la salida de los modelos dinámicos teniendo en cuenta las condiciones locales de un emplazamiento.

Por otro lado, los modelos puramente estadísticos no utilizan ningún tipo de información dinámica referente a la evolución de las condiciones meteorológicas. Debido a esta razón, se aplican fundamentalmente en situaciones donde no se dispone de información de los modelos de predicción numérica; su utilización queda además restringida para un número limitado de pasos temporales, ya que la radiación solar horaria y diaria presenta una correlación significativa únicamente con el valor anterior [Aguar y Collares-Pereira, 1992a]. Además, los modelos estadísticos clásicos presentan menores errores que los modelos numéricos para los primeros pasos temporales de predicción; debido a la alta correlación con el valor anterior de irradiancia solar y a que los modelos numéricos necesitan un tiempo inicial para generar los campos de nubosidad que puede variar entre 6 y 24 horas, lo que se conoce como *spin-up* del modelo. El *spin-up* inicial se debe a que en algunos modelos la nubosidad no es una

variable que se asimile directamente, por lo que, es necesario generar su estado inicial a partir del resto de variables para lograr que el modelo se estabilice estadísticamente.

El horizonte temporal de predicción límite en el que los modelos estadísticos clásicos mejoran a los modelos de numéricos no es conocido, ni la existencia de relación con el emplazamiento. Queda pues una línea de trabajo abierta a explorar.

Se debe señalar que la predicción de la irradiancia solar a partir de imágenes de satélite es una herramienta fundamental en los horizontes temporales que van desde la propia resolución temporal del satélite hasta las 6 horas [Heinemann et al., 2006]. La obtención de predicciones con mayor resolución temporal a la propia del satélite debe realizarse mediante la aplicación de técnicas estadísticas, sobre medidas radiométricas en superficie o utilizando la imágenes a partir de cámaras de cielo (Total Sky Imager TSI o Whole Sky Cameras WSC) [Casa-Nova et al., 2005].

Las líneas actuales de investigación combinan la salida de los modelos globales o de mesoescala con un post-procesado o *downscaling* estadístico para mejorar la calidad de las predicciones. Los modelos estadísticos son una parte fundamental que integran las predicciones de los modelos numéricos y corrigen posibles desviaciones con respecto a las condiciones meteorológicas locales, a partir de las observaciones en tierra (o estimaciones indirectas a partir a partir de imágenes de satélite, como en el caso de la irradiancia solar). Sin embargo, los resultados que se obtienen para el paso temporal horario no tienen la precisión suficiente para las aplicaciones que hacen uso del recurso solar como principal fuente energética; sobre todo en aquellos sistemas donde no se realiza una gestión del recurso energético, como en las plantas con almacenamiento de energía o sistemas híbridos basados en gas natural o biomasa.



### **3 ELECCIÓN DEL MODELO Y METODOLOGÍA EXPERIMENTAL**

Los trabajos realizados hasta la fecha en el ámbito de la predicción de la irradiancia solar han abordado el problema planteando un único modelo, o varios modelos a partir de un modelo base, que los diferentes autores no han contrastado mediante una comparativa o selección previa. De este modo, se ha partido de la suposición de que el modelo planteado era el más idóneo para la obtención de las predicciones con aplicación a un contexto y objetivos concretos.

Por lo tanto, debe considerarse que el planteamiento más correcto para la mejora de los resultados de predicción de la irradiancia solar, necesita de la realización de una propuesta de modelos previa. Los requerimientos de los modelos se definen según el tipo de tecnología que haga uso de las predicciones; cada uno de los modelos planteados presentará una serie de propiedades concretas, que serán de utilidad para reproducir el comportamiento dinámico de la irradiancia solar. De esta forma, la comparativa de los modelos, mediante el ensayo y validación, permitirá la selección de los mismos en base a una serie de criterios.

A partir del planteamiento anterior y dentro del contexto en el que se enmarca este trabajo, se pretende mejorar el conocimiento actual sobre los modelos de predicción de la irradiancia solar, orientados a su aplicación a las tecnologías solares (fotovoltaicas y fototérmicas) de producción eléctrica. Así se presentan las necesidades de predicción de dichas tecnologías en términos de los horizontes temporales de predicción y resolución espacial. A partir del planteamiento anterior, se identifican las necesidades específicas con objeto de cubrir los requerimientos de las tareas de operación y programación de la producción para la entrada en el mercado libre de la energía. De esta forma, la disponibilidad de predicciones fiables permitirá lograr la gestionabilidad de dichas plantas para competir en igualdad de condiciones frente a las tecnologías convencionales y, de otra parte, conseguir una maximización del beneficio económico.

Siguiendo el iter procedimental, se requiere realizar un estudio previo de las ventajas e inconvenientes de los diferentes tipos de modelos de predicción meteorológica (estadísticos y físicos). La elección del tipo de modelos predictivos que se aplicarán es un paso previo importante ya que determinará la metodología experimental de predicción.

En base a la elección del tipo de modelo, se plantean un conjunto de modelos iniciales que se utilizan para simular las propiedades de la irradiancia solar; para cada emplazamiento, el ensayo y validación de los modelos propuestos permite realizar una comparativa de los mismos y la selección del modelo óptimo para cada emplazamiento en función de una serie de criterios (simulación de ciertas propiedades estadísticas, bondad del modelo, menores errores, etc).

El objetivo principal de la comparativa es la comprobación de que los modelos que están especializados en la simulación de ciertas propiedades presentes en la dinámica de la irradiancia solar obtendrán mejores resultados. De esta forma, se pretende demostrar que un modelo que presente previamente menores errores de predicción no será necesariamente el más idóneo para un nuevo emplazamiento. Así, la bondad de los modelos de predicción dependerá de las propiedades en las que se especialice dicho modelo y de que éstas estén presentes localmente en la irradiancia solar semidiaria para un determinado lugar.

### **3.1 Aplicaciones de las predicciones de irradiancia solar para la producción de energía eléctrica**

Los usos que hace el hombre de la energía solar son diversos: iluminación natural, calentamiento de agua sanitaria, refrigeración, calefacción producción de energía térmica para procesos industriales, detoxificación de agua y aire, producción de energía eléctrica, desalación, secado de alimentos, etc...

Entre los diversos usos, el aprovechamiento de la energía solar como fuente para la producción de energía eléctrica está cobrando una gran importancia, debido

fundamentalmente a la sobreexplotación de las reservas energéticas de combustibles fósiles. Este hecho está provocando que el precio del petróleo, el gas natural y sus derivados aumente hasta unos márgenes que se consideran excesivos para mantener el nivel de crecimiento y desarrollo de las naciones industrializadas. Una de las razones del aumento de los precios es el fuerte incremento en la demanda energética de los países en vías de desarrollo que está poniendo de manifiesto, aún más si cabe, la insostenibilidad del modelo energético actual.

Por el contrario, una de las principales ventajas en el uso de la energía solar es la estabilidad en los precios de la energía eléctrica a largo plazo frente a las variaciones impredecibles en el caso de los combustibles fósiles, debido a que los precios de producción de las tecnologías solares son fijos. La mayor inversión que se debe llevar a cabo en una planta de energía solar se realiza al inicio, como consecuencia de la adquisición e instalación de los equipos; a lo largo de la vida útil de la planta, únicamente se deberán considerar los costes de mantenimiento y operación, ya que el recurso solar es gratuito.

Entre las tecnologías solares existentes para la producción de electricidad, la energía solar fotovoltaica y solar termoeléctrica son las dos principales fuentes. Por esta razón, este trabajo se centra en la mejora de las predicciones para la operación y planificación de la producción de estos dos tipos de tecnologías.

### *3.1.1 Requisitos de predicción en las plantas termosolares y fotovoltaicas*

La disponibilidad, rentabilidad económica y competitividad de las instalaciones solares frente a las plantas convencionales depende en gran medida de su buena gestión. El control y la operación -de las plantas termosolares- y la planificación de la producción y el mantenimiento -de las plantas termosolares y fotovoltaicas- son actividades críticas en la explotación de dichas instalaciones. Cada una de estas actividades requiere predicciones de la irradiancia solar con resoluciones temporales y espaciales específicas y un horizonte de predicción de *nowcasting* o de corto plazo, como puede apreciarse en la Figura 3.1.

En el caso de las plantas termosolares (CSP – Concentrated Solar Power), el sistema de control del campo solar precisa de predicciones con una resolución temporal entre 5 y 30 minutos, un horizonte de predicción de una hora y una resolución espacial menor a un kilómetro. Por otro lado, la operación de las plantas CSP -con o sin sistema de almacenamiento- necesita de predicciones con una resolución temporal de 10 minutos y un horizonte temporal de predicción entre 4 y 12 horas vista. Finalmente, la planificación de la entrega de la producción precisa de predicciones del recurso solar con una resolución temporal horaria y un horizonte temporal de uno a tres días vista.

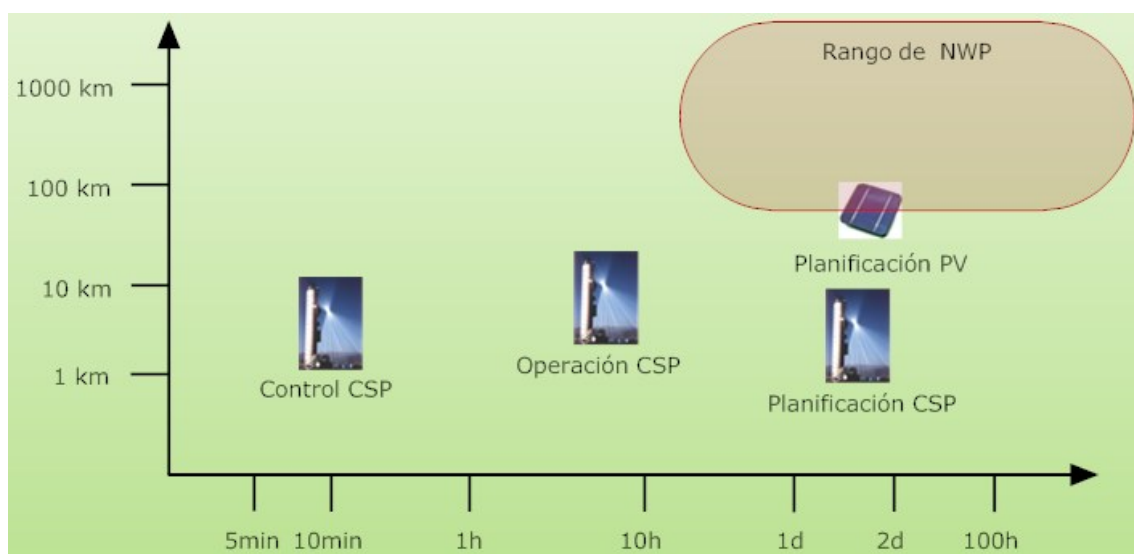


Figura 3.1 Actividades relacionadas con las plantas termosolares y fotovoltaicas y sus requerimientos de las predicciones de la irradiación solar

Por otro lado, la integración de la producción de las plantas fotovoltaicas en la red eléctrica requiere, al igual que las plantas termosolares, de predicciones horarias a varios días vista, aunque con una resolución espacial diferente a la de las plantas termosolares.

### 3.1.2 Predicciones de la irradiación solar para plantas termosolares y fotovoltaicas

El principal inconveniente en el uso de la energía solar es su alta variabilidad e incertidumbre asociadas a la disponibilidad del recurso, debido fundamentalmente a los

ciclos día/noche y estacionales y al efecto que tienen las condiciones atmosféricas de nubosidad y turbidez sobre la irradiancia solar que se recibe en la superficie terrestre.

La gestión de las tecnologías solares para la producción de electricidad se define como la posibilidad de programar arranques/ paradas de la planta y producir incluso durante la noche, cuando el recurso solar no está disponible. Éste es uno de los grandes retos de futuro que se presentan ante las tecnologías solares para lograr su competitividad en igualdad de condiciones frente al resto de energías convencionales.

Para lograr la *gestionabilidad* y fiabilidad de las fuentes productivas de electricidad a partir del recurso solar, será necesario la disponibilidad de predicciones de la irradiancia solar lo más precisas posibles, así como de sistemas de almacenamiento de energía.

Una consecuencia derivada de la gestión óptima de las plantas termosolares y fotovoltaicas es la de poder realizar una planificación de la producción que permita maximizar el beneficio económico. La Figura 3.2 presenta las componentes que deben estar involucradas en la planificación y operación de las plantas que hacen uso de la energía solar [Wittmann et al., 2008]. En dicha figura se puede observar como la planificación se debe realizar mediante técnicas de optimización que utilicen tres módulos principales compuestos por las predicciones del recurso solar, las predicciones de los precios de la electricidad en el mercado libre de la energía y la disponibilidad de energía en el sistema de almacenamiento.

En función de las predicciones del recurso solar, las técnicas de optimización programarán cuándo se debe entregar la producción de la electricidad a la red en base a las predicciones de los precios más altos de la energía eléctrica en el mercado libre. El sistema de almacenamiento de la planta permitirá desplazar el instante en el que se hará entrega de la producción eléctrica en función de la programación y, además, permitirá corregir las posibles desviaciones del recurso solar previsto.

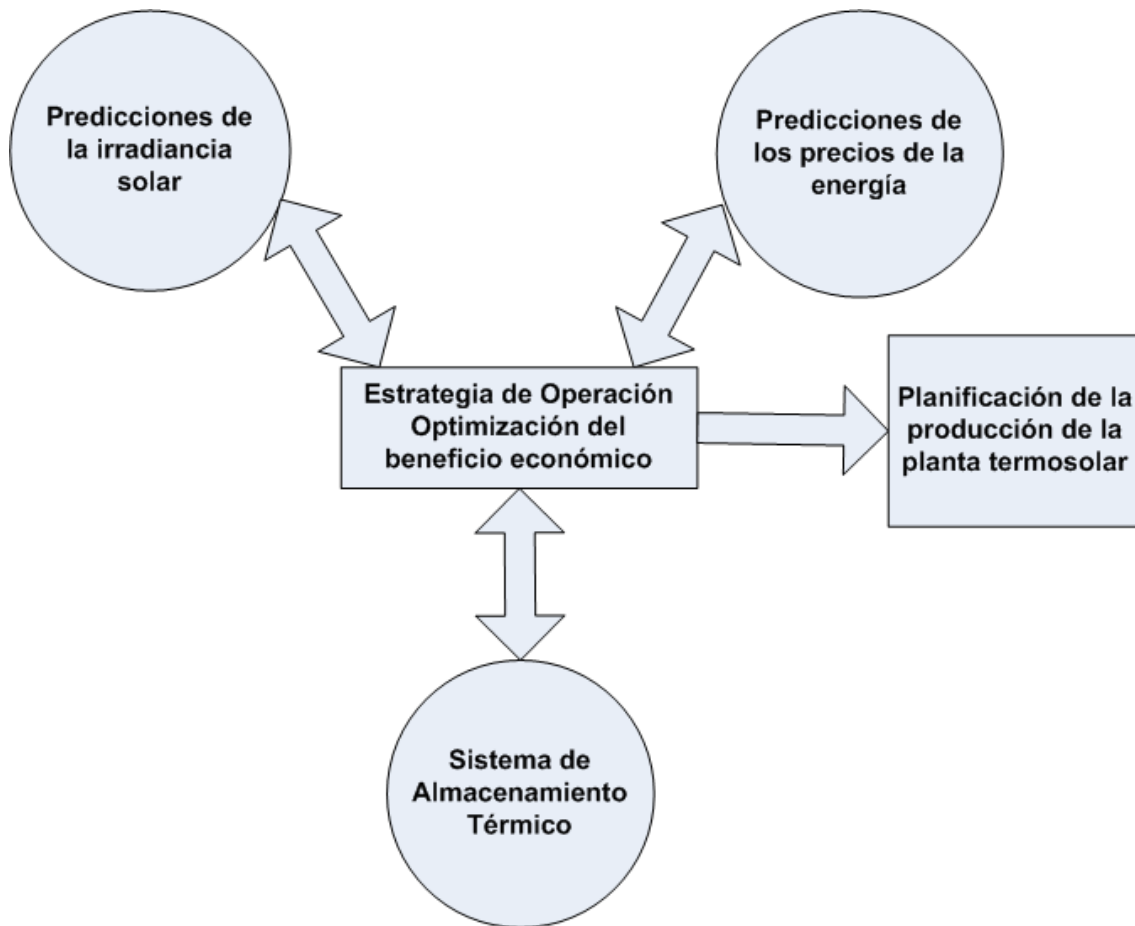


Figura 3.2 Esquema de un sistema de optimización para la planificación de la producción en plantas termosolares

Una de las principales ventajas de la energía solar termoeléctrica frente a la fotovoltaica es la disponibilidad de sistemas de acumulación de energía plenamente comerciales y con precio razonable en comparación a los costes de producción. En una planta termosolar, la energía térmica del campo solar se puede almacenar en tanques de sales fundidas para su utilización posterior. La acumulación incrementa notablemente la inversión económica inicial; sin embargo, permite reducir el coste de producción de la electricidad, debido al incremento de capacidad de generación.

Las plantas termosolares que disponen de un sistema de almacenamiento térmico suelen tener una capacidad temporal de almacenamiento entre seis y nueve horas de la potencia nominal de la planta. Este valor es el óptimo para minimizar el coste de los

componentes y el precio de producción de la energía eléctrica a partir de sistemas de aprovechamiento de la energía solar [Herrmann, 2009].

En el caso de la energía solar fotovoltaica se prevé que en un futuro próximo estén disponibles sistemas de almacenamiento comerciales y a un coste competitivo. Los sistemas de almacenamiento en los que se está investigando actualmente para su aplicación en diversas tecnologías renovables son los sistemas de almacenamiento mediante condensadores, producción de hidrógeno, volantes de inercia y baterías [Armenta-Deu, 1998; Armenta-Deu, 2003; Kazempour et al., 2009; Beaudin et al., 2010].

En este trabajo se propone la validación de modelos para la mejora de las predicciones de la irradiancia solar global para su uso en plantas solares que dispongan de un sistema de almacenamiento. Un tipo de predicciones de la irradiancia solar útiles para la planificación de la entrega de energía eléctrica en plantas termosolares son las predicciones con resolución temporal de predicción similar a la capacidad del sistema de almacenamiento térmico. Los valores semidiarios de las predicciones de la energía solar se definen como los valores de la irradiancia solar integrados para los intervalos temporales entre la salida del Sol y el cenit Solar, y entre este último y la puesta del Sol, valores que oscilan entre 6 y 9 horas por término medio para la mayoría de las localidades donde se utilizan sistemas solares de producción termoeléctrica con almacenamiento.

Las predicciones de la irradiancia solar semidiaria y la gestión del almacenamiento térmico mediante estrategias óptimas de operación permitirán realizar la programación de la entrega de la producción de las plantas solares maximizando el beneficio de las mismas. Además, la disponibilidad de energía térmica de forma continua, a partir del campo solar o del sistema de almacenamiento en base a las predicciones de energía solar, posibilitará optimizar la operación de la planta, obteniendo el máximo rendimiento de la misma e incurriendo en un menor número de paradas de los diferentes

componentes (principalmente en el ciclo de potencia) minimizando el alto coste de operación de las arrancadas.

Las predicciones de la irradiancia solar global semidiaria pretenden ser de utilidad para las plantas termosolares que concentran la irradiancia solar directa. La estimación de la componente directa de la irradiancia a partir de la global es una tarea que no se llevará a cabo en este trabajo, porque es digno de un estudio individualizado. Por lo tanto, se deja como propuesta futura a partir de la utilización de modelos  $k_d-k_t$  [Orgill J.F. y Hollands K.G.T., 1977; Erbs D.G. et al., 1982; Reindl D.T. et al., 1990] y modelos  $k_b-k_t$  [Maxwell A.L., 1987; Louche A. et al., 1991].

### **3.2 Determinación del tipo de modelo meteorológico: modelos físicos y modelos estadísticos**

Los modelos meteorológicos que se pueden aplicar para realizar las predicciones se dividen en modelos físicos y modelos estadísticos, como se ha mencionado con anterioridad. Los primeros, también conocidos como modelos numéricos de predicción, se basan en la resolución de las ecuaciones primitivas de la atmósfera y tienen una aplicación universal. Precisan conocer el estado inicial de la atmósfera con la mayor fiabilidad posible, ya que dos estados iniciales similares tendrán una evolución de las trayectorias de sus estados completamente diferentes [Lorenz, 1963]. En este tipo de modelos la solución de las variables meteorológicas tiene una escala global o de mesoescala, de forma que necesitan un postprocesado o *downscaling* para tener en cuenta el efecto de las condiciones locales de un emplazamiento.

Por otro lado, los modelos estadísticos tienen una gran importancia dentro de la predicción meteorológica ya que para obtener predicciones fiables en un emplazamiento concreto es necesario su uso, ya sea en solitario o combinados a la salida de los modelos numéricos de predicción, para establecer una relación entre las variables meteorológicas a nivel global y las condiciones locales.

Los modelos estadísticos de predicción tienen diversas ventajas frente a los modelos físicos. Las predicciones que generan los modelos estadísticos representan las condiciones locales de un emplazamiento concreto. Además, la aplicación de los modelos estadísticos de predicción se caracteriza por su sencillez, ya que no es necesario un conocimiento previo detallado del proceso físico que se pretende predecir. Además, desde el punto de vista de la operación, los modelos estadísticos no necesitan grandes recursos computacionales. Desde el punto de vista del desarrollo requieren un bajo tiempo de desarrollo y son fáciles de implementar.

Ello no obstante, los modelos estadísticos presentan también desventajas frente a los modelos dinámicos. La aplicación de los modelos estadísticos necesita disponer de una base de datos histórica para el emplazamiento donde se pretende predecir la variable meteorológica. Por otro lado, la precisión de las predicciones que se obtienen mediante los modelos estadísticos depende del tamaño y de la calidad de la base de datos, ya que es necesario entrenar los modelos con todas las posibles situaciones que pueden presentarse para que los modelos estadísticos capturen las relaciones existentes entre predictandos y predictores. En el caso particular de la irradiancia solar, es posible disponer de bases de datos radiométricas registradas o estimadas (por ejemplo, en cualquier punto de la Península Ibérica, estimadas a partir de imágenes de satélite [Zarzalejo, 2005]) que pueden ser utilizadas para entrenar los modelos estadísticos de predicción.

Otra de las desventajas de los modelos estadísticos es que la interpretación de las predicciones generadas por ellos es complicada ya que éstos se comportan como una caja negra donde el sistema se estudia desde el punto de vista de las entradas que recibe y sus salidas sin tener en cuenta su funcionamiento interno. Además, los modelos estadísticos de predicción necesitan a la entrada datos estacionarios con una función de distribución normal. Las diferentes variables meteorológicas tienen un comportamiento no estacionario, por lo tanto, es necesario realizar una serie de transformaciones previas a la aplicación de los modelos estadísticos de predicción.

Finalmente, destacar que los modelos estadísticos solo establecen relaciones entre los patrones pasados y futuros sin tener en cuenta la evolución temporal de las variables meteorológicas. Este hecho provoca que, a partir de un cierto horizonte temporal de predicción, los modelos estadísticos no sean capaces de simular la evolución dinámica de las variables meteorológicas a predecir, por lo que es preciso combinarlos con los modelos numéricos de predicción [Araujo da Costa, 2005].

En este trabajo se ha optado por realizar las predicciones exclusivamente mediante modelos estadísticos de predicción. Más adelante se presentarán las razones de esta elección. Pero como quiera que la elección requiere de la atención a las principales aplicaciones de los modelos estadísticos en la predicción meteorológica y su utilidad para obtener y mejorar las predicciones de la irradiancia solar; el estudio se hace previo a la determinación del tipo de modelo. Entre las diversas aplicaciones de los modelos estadísticos se pueden destacar las siguientes:

- Predicción de series temporales. Los modelos que se utilizan en la predicción de series temporales se conocen también como modelos clásicos. Los modelos estadísticos se utilizan únicamente sobre las medidas observadas para obtener valores futuros en base a las relaciones existentes entre ambas. Se aplica en horizontes temporales de *nowcasting* donde los modelos físicos no sean capaces de generar predicciones fiables, ya que necesitan un tiempo inicial (*spin-up*) para generar los campos de nubosidad [Huang, 1995].
- Downscaling temporal. Los modelos estadísticos se aplican a la salida de los modelos de predicción numérica para aumentar la resolución temporal de las predicciones y obtener los estados de la atmósfera con una mayor frecuencia temporal.
- Downscaling espacial. Se utilizan para conocer las condiciones locales de un emplazamiento que no está dentro de los puntos de rejilla que forman la salida del modelo de predicción numérica [Marzban et al., 2006].
- Post-procesado. Se utilizan para corregir los posibles errores de las predicciones, como valores fuera de los límites físicos y la corrección del bias mensual o estacional de las predicciones.

- Generación de predicciones óptimas. Se aplican para establecer una ponderación entre las predicciones de los diferentes modelos.

De entre las aplicaciones anteriores, se ha elegido realizar las predicciones de la irradiancia solar semidiaria mediante modelos estadísticos clásicos basados en el análisis de series temporales, debido a las siguientes razones:

- Simplicidad de los modelos estadísticos basados en el análisis de series temporales. Son una alternativa simple frente a los modelos numéricos, que tienen grandes requisitos computacionales.
- Disponibilidad de una extensa base de datos piranométricos en emplazamientos repartidos por toda la geografía de la Península Ibérica.
- Mayor experiencia previa en la aplicación de los modelos estadísticos en otros campos de la ciencia.
- Necesidad de un aumento en el conocimiento de las propiedades estadísticas de la irradiancia solar y la aplicación de los modelos estadísticos, a fin de que la obtención de predicciones tenga en cuenta los efectos locales.
- Menores costes desde el punto de vista del desarrollo de los modelos y operación en plantas.
- Desde el punto de vista del desarrollo y la operación en plantas que hacen uso de la energía solar tienen menores costes, ofreciendo predicciones fiables en el corto plazo.

### *3.2.1 Selección del conjunto inicial de modelos estadísticos a ensayar en base a su linealidad*

Una vez identificado el tipo de modelo que se va a utilizar, el siguiente paso es determinar el criterio de selección del conjunto de modelos iniciales sobre los que desplegar la metodología experimental de predicción. El criterio de selección de los modelos estadísticos de predicción que se ha elegido está basado en la linealidad, ya que la bondad de los modelos de predicción dependerá del determinismo (lineal o no-lineal) que presente la serie temporal de irradiancia solar. De esta forma, en caso que la irradiancia solar muestre un comportamiento lineal, la utilización de los modelos

lineales será más conveniente, ya que, previsiblemente, los resultados de predicción tendrán menores errores y los modelos serán más sencillos en cuanto a su uso y complejidad.

Un sistema se considera lineal si cumple las propiedades de homogeneidad y aditividad; la propiedad de homogeneidad señala que la salida de un sistema es proporcional a su entrada, por su parte, la propiedad de aditividad se cumple cuando el resultado de una entrada compleja sobre el sistema tiene el mismo efecto que la suma de los efectos simples que forman la entrada compleja. La combinación de las propiedades de homogeneidad y aditividad se conoce como principio de superposición. La aplicación del concepto de linealidad a una serie temporal permite definir una serie temporal lineal como aquella en la que sus valores futuros se pueden expresar como combinación lineal de los  $n$  valores anteriores.

Los modelos estadísticos de predicción que se van a aplicar a las series temporales son un tipo de modelo lineal y dos tipos de modelos no-lineales:

- Modelo autorregresivo (modelo lineal).
- Modelo basado en redes neuronales (modelo no-lineal).
- Modelo de lógica difusa (modelo no-lineal).

Los modelos autorregresivos tienen como principal característica la generación de predicciones que conservan las propiedades lineales de las series temporales que se utilizan a su entrada. La elección de este tipo de modelos para realizar las predicciones se debe a los buenos resultados previos obtenidos dentro del campo de la irradiancia solar en la modelización [Aguiar y Collares-Pereira, 1992b] y en la generación de datos sintéticos [Amato et al., 1989; Aguiar y Collares-Pereira, 1992b].

El otro tipo de modelos que se utilizará es un modelo no-lineal basado en las redes neuronales artificiales (RNA). La aplicación de este tipo de modelos a la irradiancia solar ha demostrado buenos resultados en trabajos previos [Dorvlo et al., 2002]; las RNA se han aplicado anteriormente a la irradiancia solar en la generación de mapas

[Hontoria et al., 2005], en la generación de datos sobre terreno complejo [Bosch et al., 2008] y en la predicción de la irradiancia solar [Elizondo et al., 1994; Cao et al., 2009], entre otros.

El tercer tipo de modelo de predicción es un modelo no-lineal que se basa en los sistemas neuro-difusos y emplea reglas de lógica borrosa del tipo *if-then* para modelar el conocimiento humano y los procesos de razonamiento sin emplear análisis precisos; los modelos difusos se han empleado previamente en la estimación [Ramirez Santigosa et al., 2003] y la predicción de la radiación solar [Cao y Lin, 2008], entre otros.

Una vez seleccionado el conjunto inicial de modelos de predicción que se van a utilizar, a continuación se presenta la metodología experimental de predicción que se va a seguir en este trabajo para la obtención de las predicciones de irradiancia solar semidiaria.

### **3.3 Esquema de la metodología experimental de predicción**

La metodología de predicción que se va a aplicar en este trabajo se basa en el análisis de series temporales. Ésta metodología permite el estudio de las propiedades estadísticas y ayuda a la interpretación de los datos con el objetivo de poder extrapolar los valores futuros en base a los pasados [Box y Jenkins, 1976].

Una serie temporal se define como una secuencia de datos ordenados cronológicamente con respecto a su referencia temporal. En el mundo real, los procesos o sistemas presentan comportamientos complejos que no pueden modelarse mediante un único modelo. Así, la metodología de predicción basada en el análisis de series temporales consta de un conjunto de técnicas y modelos que se aplican en diferentes fases, con la intención de predecir de forma independiente las diferentes variabilidades de la serie temporal [Schreiber, 1999]. El análisis que se aplica se basa en la identificación de las componentes determinísticas y no-determinísticas (o aleatorias) de la serie temporal, de forma que se aplica un modelo de predicción diferente a cada una de las componentes.

Las diferentes etapas de la metodología experimental de predicción son:

1. Preprocesado de los datos experimentales. En esta fase se filtran los datos y se construye la serie temporal semidiaria. Se eliminan las posibles tendencias de la serie para que presente un comportamiento estacionario. Posteriormente, se identificarán las propiedades lineales de la serie temporal estacionaria tras su transformación. La identificación de las propiedades de linealidad nos permitirá interpretar posteriormente los resultados de los diferentes modelos de predicción.
2. Ajuste de los modelos estadísticos propuestos y generación de las predicciones. Para realizar la predicción se utilizará un modelo lineal autorregresivo y dos modelos no-lineales basados en redes neuronales y lógica difusa. La serie temporal estacionaria intermedia se dividirá en dos conjuntos independientes: un primer conjunto para realizar el ajuste de los modelos de predicción y un segundo conjunto para realizar la validación de los modelos de predicción. En un primer paso, se utilizará el conjunto de datos de entrenamiento para ajustar los parámetros de los diferentes modelos de predicción. Posteriormente, una vez ajustados los parámetros de los modelos de predicción, se obtendrán las predicciones de las series temporales intermedias a partir del conjunto de datos de validación. El paso final de esta fase consistirá en restaurar las tendencias eliminadas en la fase de preprocesado para obtener las predicciones de la irradiancia solar.
3. Validación de los modelos de predicción. En esta fase, a partir de las predicciones de la irradiancia solar obtenidas en el paso anterior, se realizará la validación de los diferentes modelos planteados para comprobar la adecuación y representatividad de los mismos a la dinámica de la serie temporal de irradiancia solar semidiaria. Se analiza la calidad de las predicciones a partir de los estadísticos error medio relativo (MBD), error cuadrático medio relativo (RMSD) y mejora en términos de RMSD frente a las predicciones de un modelo basado en la persistencia.

La Figura 3.3 muestra el diagrama de flujo que representa la metodología de predicción que se va a seguir en este trabajo. En dicho gráfico, se muestra cada uno de los subprocesos que se van a realizar y la etapa a la que corresponde cada uno.

Una vez presentado el esquema general de la metodología experimental de predicción, se está en disposición de exponer los conceptos teóricos generales relacionados con cada una de las etapas que se van a seguir para la obtención de las predicciones. Los resultados particulares relacionados con la *toma de datos*, el *acondicionamiento*, el *filtrado* y la *preparación de las series temporales* se mostrarán en los siguientes capítulos.

### **3.4 Preprocesado de datos experimentales**

La transformación de las series temporales originales (no-estacionarias) para que presenten un comportamiento estacionario y la identificación de las propiedades de linealidad toma como base unos fundamentos teóricos generales que requieren de unas mínimas aclaraciones previas.

#### *3.4.1 Transformación de series temporales no-estacionarias*

Un proceso estrictamente estacionario se basa en la suposición de que los datos presentan un estado particular de *equilibrio estadístico*, siendo su media, varianza y distribución de probabilidad la misma para todos los instantes temporales. Una forma menos restrictiva de estacionariedad es la denominada *estacionariedad débil*, en la que sólo se requiere que la media y la varianza permanezcan constantes con respecto al tiempo, excluyendo la distribución de probabilidad [Hamilton, 1994].

La teoría probabilística para la predicción de series temporales se aplica sobre datos con un comportamiento estacionario en media y varianza, dicho en otros términos, es preciso que la serie temporal no tenga tendencia alguna y que presente un grado de dispersión similar en cualquier intervalo temporal.

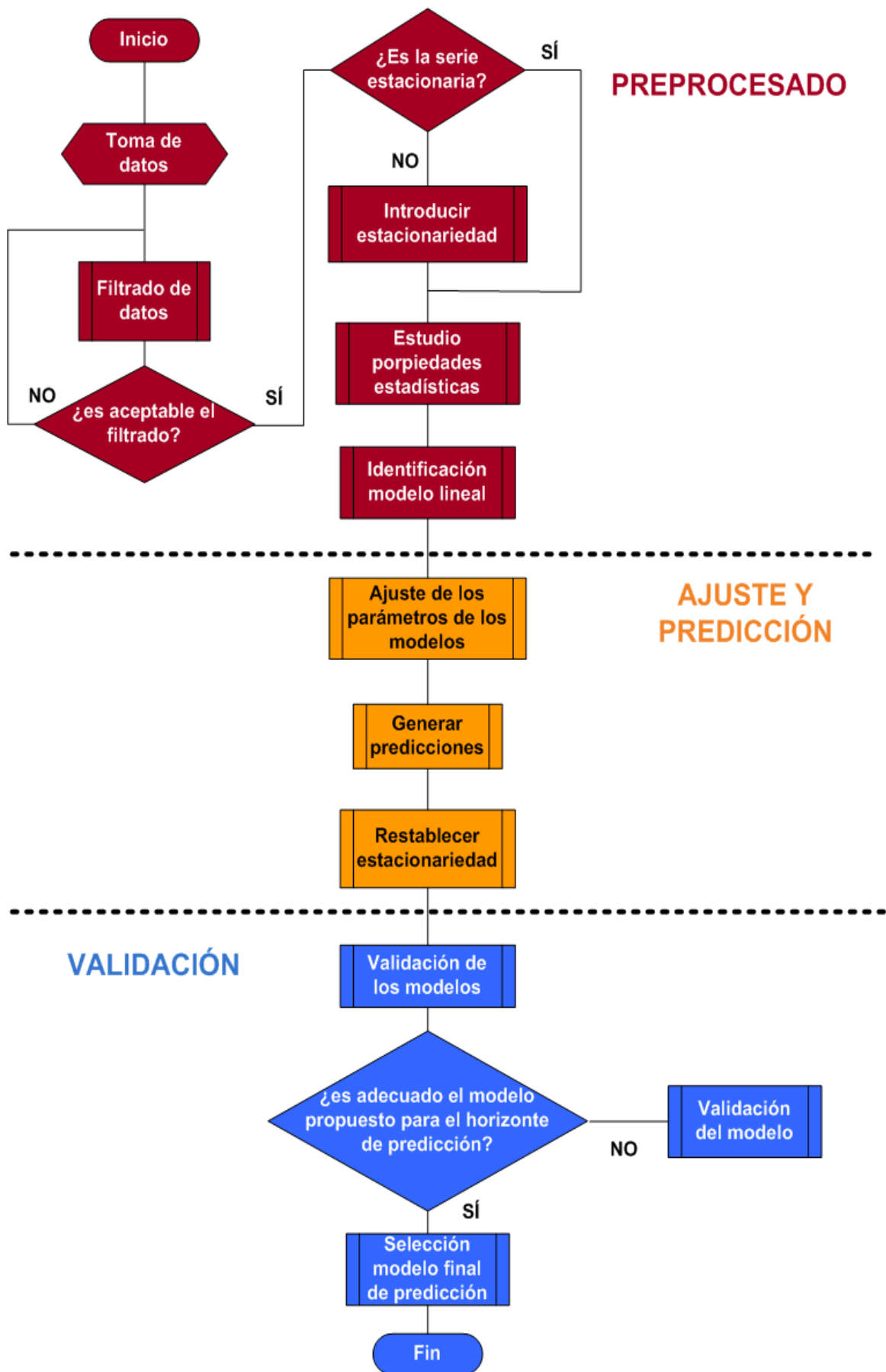


Figura 3.3 Metodología experimental de predicción

Una serie temporal no-estacionaria puede presentar diferentes modos de variabilidad (determinística y aleatoria) que dependen de las propiedades del proceso físico al que pertenecen los datos observados. Para extraer los diferentes modos de variabilidad que puede presentar una serie temporal no-estacionaria es necesario aplicar una serie de transformaciones que nos permita obtener valores estacionarios.

### Tipos de variación de una serie temporal

Los métodos tradicionales de análisis de series temporales están principalmente relacionados con la descomposición de la variación determinística de los datos en tendencia, variación estacional, otros cambios cíclicos y resto de fluctuaciones irregulares. Sea  $y$  una variable determinada, es posible realizar su descomposición de la siguiente forma:

$$y = S + C + T + a \quad (5.1)$$

donde  $S$  es la variación estacional,  $C$  son los cambios cíclicos,  $T$  es la tendencia de la serie y  $a$  es la componente aleatoria o residual que corresponde a una señal para la cual no existe un modelo de predicción único. La característica fundamental de la serie temporal de residuos es que cualquier segmento de elementos consecutivos tiene la misma forma que cualquier otro segmento de similar tamaño. Si el residuo que se obtiene presenta una tendencia o manifiesta un patrón más o menos regular, entonces contiene propiedades que deben ser atribuidas a las otras componentes determinísticas y será necesario volver a descomponer la serie temporal.

El objetivo principal de la descomposición es el de desacoplar las diferentes componentes que forman la serie temporal para simplificar el proceso de predicción. De esta forma, se tendrá una información más detallada y precisa sobre los diferentes modos de variabilidad de la serie temporal. La predicción de los valores futuros se podrá realizar mediante el empleo de un método de predicción apropiado a cada componente de la serie temporal para, posteriormente, combinarlas y obtener una única predicción. Es de esperar que, a priori, esta aproximación tenga un mayor potencial que

la utilización de un único modelo de predicción, que posiblemente no prestará atención a la estructura oculta de la serie temporal.

La predicción de las componentes estacional ( $S$ ), cíclica ( $C$ ) y tendencia ( $T$ ) se realizará mediante modelos de predicción determinísticos. Por otro lado, la componente aleatoria, que tendrá una media y varianza constante, será la serie temporal estacionaria intermedia sobre la que se aplicarán los modelos de predicción estadísticos lineales y no-lineales.

En el siguiente capítulo se mostrarán los resultados, tras aplicar los conceptos que aquí se exponen, relacionados con la identificación de los modos de variabilidad de la serie temporal irradiancia solar semidiaria.

### *3.4.2 Identificación de la linealidad de la serie temporal estacionaria*

Existen determinados conceptos cuyo tratamiento se hace necesario para valorar qué modelos, dependiendo de sus propiedades de linealidad, serán capaces de reproducir mejor la dinámica de la serie temporal de irradiancia solar; de esta forma, se podrán contrastar posteriormente las hipótesis planteadas a partir de la aplicación de los test de linealidad.

La identificación de las propiedades de linealidad de la serie temporal estacionaria se realizará mediante la aplicación del test estadístico de no-linealidad basado en el método de generación de *series subrogadas* [Theiler et al., 1992]. La precisión de éste método para identificar la linealidad dependerá de la naturaleza de los datos y de la longitud de la serie temporal muestral disponible [Schreiber y Schmitz, 1996].

Los datos subrogados son series temporales lineales que se construyen a partir de una serie temporal (que aquí llamamos original) y compartirán ciertas propiedades estadísticas con ésta. La generación de las series subrogadas se realiza bajo una hipótesis nula que debe ser lo más general posible, ya que no se pretende crear un

modelo que extraiga las propiedades de los datos originales sino que se trata de obtener datos subrogados con las propiedades lineales generales de los datos originales.

Posteriormente, se calcula una serie de discriminantes estadísticos no-lineales que caractericen las series temporales original y subrogadas utilizando un único escalar. El estadístico discriminante no-lineal de los datos originales se compara con la distribución empírica del conjunto de datos subrogados generados bajo la hipótesis nula. En caso de que el discriminante estadístico de los datos originales no pertenezca a la distribución experimental de los datos subrogados, la hipótesis nula de linealidad bajo la que se generaron los datos subrogados será rechazada y se podrá afirmar que la serie temporal original presenta una dinámica no-lineal.

La idea de este método es comprobar que la serie temporal original es lineal, planteando de antemano dicha hipótesis; de esta forma, se crean nuevas series temporales con las mismas propiedades lineales que la serie temporal original y se calculan una serie de discriminantes estadísticos no-lineales para todas las series. Ya que no conocemos de antemano el valor teórico de dichos estadísticos no-lineales para las series temporales original y subrogadas, es de esperar que las series temporales tengan la misma distribución de probabilidad del estadístico de no-linealidad, ya que suponemos que estas pertenecen el mismo proceso físico.

Los pasos que se van a seguir en la aplicación del método de datos subrogados para la identificación de la linealidad de la serie temporal estacionaria son los que se enumeran a continuación:

1. Generación de los datos subrogados a partir de la serie temporal original bajo la hipótesis nula de linealidad. Se utilizarán dos métodos, el primer método basado en la transformada de Fourier (FT) y el segundo conocido como iAAFT (*iterative Amplitude Adjusted Fourier Transform*).
2. Selección de los discriminantes estadísticos no-lineales que se calcularán para la serie temporal original y sus versiones lineales.

3. Aplicación de un test estadístico paramétrico para comprobar si el discriminante estadístico no-lineal de la serie temporal original sigue la misma función de distribución de probabilidad que las series temporales subrogadas.

#### Generación de las series temporales subrogadas

El proceso de generación de los datos subrogados se ha realizado utilizando el método basado en la Transformada de Fourier (FT) (también conocido como el método de datos subrogados con fase aleatoria) y el método iAAFT (*iterative Amplitude Adjusted Fourier Transform*) [Walpole et al., 2002].

La hipótesis nula que se establece en el método de la transformada de Fourier (FT) es que los datos originales se caracterizan por un ruido gaussiano con autocorrelación lineal. El teorema de Wiener-Khinchin establece que la función de autocorrelación es la transformada inversa de Fourier del espectro de densidad de potencia. De la misma forma, el espectro de densidades se relaciona con la transformada de Fourier de la función de autocorrelación. La consecuencia de este teorema es que una señal puede expresarse de forma indistinta en el dominio del tiempo o de la frecuencia a partir de la correspondencia que existe entre ambas.

De esta forma, el método de la Transformada de Fourier (FT) se basa en el cálculo de las amplitudes del espectro de Fourier de la serie temporal original y la generación de las series subrogadas, mediante la transformada inversa de Fourier y la distribución aleatoria del espectro de ángulos de fase. Como resultado de aplicar el método FT se obtienen series subrogadas con ángulo de fase del espectro aleatorio que conservan el espectro de Fourier y las mismas propiedades lineales (media, varianza y función de autocorrelación) de la serie temporal original.

El principal problema en el uso del método anterior (FT) es que puede darse el caso en que la función distribución de amplitudes de los valores de la serie original no coincida con la de los datos subrogados. Este hecho puede dar lugar al falso rechazo de la hipótesis nula, ya que, para aceptar la hipótesis nula y asegurar que la serie temporal

original tiene un comportamiento lineal, es necesario que las series subrogadas mantengan las propiedades lineales de la serie temporal original. Por lo tanto, es necesario establecer una hipótesis nula aún más general, que considere la posibilidad de que la serie temporal sea lineal pero que exista un ruido no-lineal en dicha serie temporal introducido durante el proceso de medida de la señal. Por lo tanto, la hipótesis nula que se establezca debe mantener la misma función de distribución de amplitudes y espectro de potencia que la serie original.

El método iAAFT (*iterative Amplitude Adjusted Fourier Transform*) evita los problemas que pueden surgir de la utilización del método FT y genera series subrogadas que conservan las mismas propiedades lineales (media, varianza y función de autocorrelación) y distribución de amplitudes que la serie temporal original [Schreiber y Schmitz, 1996]. La hipótesis nula que se establece en este método es que la serie temporal original es el resultado de la observación a través de una función no-lineal, monótona e invertible que actúa sobre una serie temporal que representa un proceso linealmente estocástico.

El procedimiento que se sigue en el método iAAFT es el siguiente: en primer lugar, se transforma la serie temporal original, para que tenga una distribución gaussiana y conserve la evolución temporal de los valores medidos, mediante un cambio de orden de rango de sus amplitudes de acuerdo a una distribución normal. A partir de la serie temporal obtenida en el paso anterior, se generan los datos subrogados con fase aleatoria a partir del método FT. Finalmente, se realiza un cambio de escala de las series subrogadas con distribución normal, obtenidas en el paso anterior, aplicando un cambio de orden de acuerdo a la distribución de amplitudes de la serie temporal original. El espectro de Fourier de las series subrogadas que se obtienen tras este último paso no coincidirá con el espectro de la serie original. Por lo tanto, los pasos anteriores se realizarán de forma iterativa hasta que las series subrogadas tengan una distribución y un espectro de Fourier similar a la de la serie temporal original, con un determinado nivel de tolerancia.

### Discriminantes estadísticos de no-linealidad

La selección de los estadísticos no-lineales discriminantes es de suma importancia, ya que una serie temporal puede ser muy sensible a las características de no-linealidad que representa un determinado estadístico y dar lugar al falso rechazo de la hipótesis nula. En este trabajo se han elegido las medidas de no-linealidad que se describen a continuación:

- Autocovarianza de tercer orden (C3). Mide la simetría de la función de distribución de probabilidad [Swami et al., 1998].
- Reversibilidad Temporal (REV). Una serie temporal se dice que es reversible si sus propiedades de probabilidad son invariantes al aplicar una inversión en el orden temporal de la serie temporal original [Diks et al., 1995].
- Varianza del Vector de Retrasos (*Delay Vector Variance DVV*). El último estadístico que se utiliza para medir la no-linealidad se basa en el cálculo de los vectores de retrasos (*Delay Vectors DVs*), que tiene sus fundamentos en el paradigma de la Teoría del Caos [Gautama et al., 2003]. La aplicación de la Teoría del Caos a una serie temporal se define a partir del Teorema de Takens [Takens, 1981].

### Verificación de la hipótesis nula mediante un test paramétrico

En este apartado, se presenta la aplicación de un test estadístico paramétrico para comprobar si el discriminante estadístico no-lineal de las series temporales (originales y subrogadas) pertenece a la misma función de distribución de probabilidad. En caso de que no se cumpla la suposición anterior, se rechazará la hipótesis nula de linealidad bajo la que se generaron los datos subrogados y se podrá afirmar que la serie temporal original presenta una dinámica no lineal.

El procedimiento que se sigue es el siguiente: para cada serie temporal original, se generan  $N_s$  series subrogadas y se calculan los discriminantes estadísticos no-lineales (C3, REV, DVV) de las series temporales (original y subrogadas). El rechazo de la hipótesis nula se basa en comprobar si el estadístico no-lineal de la serie original pertenece a la misma función de distribución que las series temporales linealizadas.

Debido a que la forma analítica de las funciones de distribución de probabilidad de los discriminantes estadísticos no-lineales no se conoce, es necesario utilizar un test no-paramétrico basado en el orden de los resultados [Theiler y Prichard, 1996]. El criterio no paramétrico se basa en examinar el rango ( $r_i$ ) de los discriminantes estadísticos no-lineales de la serie temporal original y sus versiones lineales. El conjunto de estadísticos ( $r_i$ ) se ordena y se determina el índice de la posición ( $r_o$ ) del estadístico de la serie temporal original. Un test paramétrico de cola derecha (ó cola izquierda) rechaza la hipótesis nula ( $H_0$ ) si el rango  $r_o$  es mayor a 90 ( $r_o \leq 10$ ), mientras que un test de dos colas o bilateral rechaza  $H_0$  si el rango  $r_o$  es mayor que 95 o menor o igual que 5. El test paramétrico sobre el estadístico discriminante DVV se basa en un test de cola derecha, mientras que para los discriminantes C3 y REV se basan en un test de dos colas [Mandic et al., 2008].

### **3.5 Modelos estadísticos de predicción propuestos: ajuste y generación de las predicciones**

Una vez presentada la metodología de preprocesado, en este apartado se describen los modelos de predicción estadísticos y las técnicas que se utilizarán en la generación de las predicciones para los diferentes horizontes temporales de predicción.

Sería lógico presuponer que se seleccionará la utilización de los modelos lineales o no-lineales en función de los resultados que se hayan obtenido en la fase anterior de identificación de la linealidad de la serie temporal estacionaria. Sin embargo, la metodología experimental de predicción que se propone plantea ensayar todos los modelos estadísticos de predicción, para verificar las hipótesis planteadas en el test de linealidad de la serie temporal estacionaria.

#### *3.5.1 Modelos de predicción*

En los siguientes apartados se presentan los modelos estadísticos que se van a utilizar en la predicción de la serie temporal estacionaria intermedia; un modelo lineal

autorregresivo y dos modelos no-lineales basados en las redes neuronales y lógica difusa tal y como se expuso en 3.2.1.

#### Predicción mediante modelos autorregresivos

El modelo autorregresivo realiza la predicción de los valores futuros en base a una suma ponderada de observaciones previas de la serie temporal y la adición de la influencia de los sistemas externos en forma de ruido.

La serie temporal estacionaria  $y(t_i)$  se puede considerar como una combinación lineal de observaciones pasadas  $y(t_{i-1}), y(t_{i-2}), \dots$ ; de un proceso así definido se dice que es invertible y su representación se denomina autorregresiva:

$$y(t_i) = \delta + \phi_1 y(t_{i-1}) + \phi_2 y(t_{i-2}) + \dots + \varepsilon(t_i) \quad (5.2)$$

donde  $y(t_i)$  representa la observación estacionaria actual;  $y(t_{i-1}), y(t_{i-2})$ , las observaciones pasadas;  $\delta, \phi_1, \phi_2$ , los parámetros (constante y coeficientes) del modelo autorregresivo y  $\varepsilon(t_i)$ , los errores aleatorios de predicción para el instante temporal actual.

#### Predicción mediante redes neuronales artificiales

Las redes neuronales artificiales (RNA o *Artificial Neural Networks ANN*) vienen demostrando buenos resultados en los últimos años. Las RNA son un nuevo paradigma de aprendizaje y procesamiento automático inspirado en los sistemas nerviosos biológicos, que tratan de mimetizar sus funciones con modelos matemáticos aunque, evidentemente, lo hacen de una forma extremadamente simplificada [Haykin, 2005].

Las partes que constituyen una neurona artificial son las siguientes (ver Figura 3.4):

- Entrada. Está formada por un vector o patrón de entrada.
- Función de red. Calcula el valor de base o entrada total a la unidad, generalmente como simple suma ponderada de todas las entradas recibidas; es decir, de las entradas multiplicadas por el peso sináptico.

- Función de activación. Se encarga de calcular el nivel o estado de activación de la neurona a partir de la entrada que proviene de la función de red. Las funciones de activación más comunes son la función escalón, rampa ó sigmoide y gaussiana.
- Pesos sinápticos ó conexiones. El valor de los pesos y el signo de los mismos definen el tipo (excitatorio/inhibitorio) y la intensidad de la influencia de las conexiones entre neuronas. Cada peso sinóptico ( $w_{ij}$ ) determina el efecto de la célula  $j$ -ésima sobre la célula  $i$ -ésima.
- Salida. Se expresa como función de la entrada y los pesos.

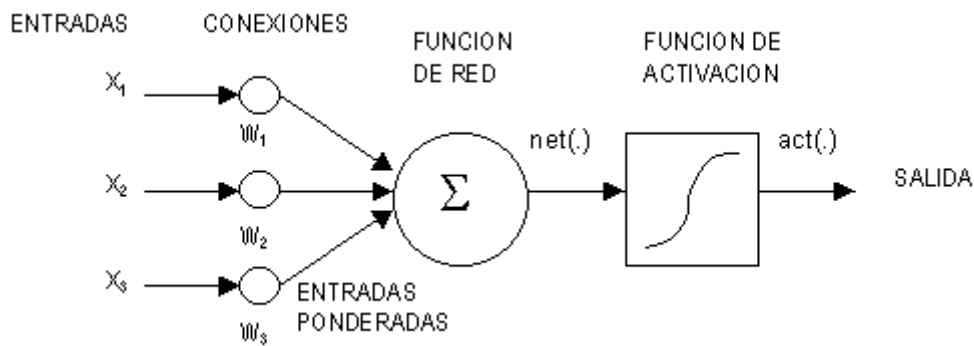


Figura 3.4 Modelización de las redes neuronales

Las neuronas se interconectan entre ellas en capas sucesivas para formar una red neuronal. Las conexiones entre neuronas pueden estar total o parcialmente interconectadas.

Existen diferentes tipos de redes neuronales que pueden ser empleadas para la predicción de la irradiancia solar. En este trabajo se utilizará el tipo de red neuronal que se conoce como redes neuronales supervisadas y que se caracteriza por ser el modelo más ampliamente utilizado y con mayor desarrollo desde los inicios de estos diseños.

En una red neuronal supervisada, el conjunto de datos de entrenamiento está formado por vectores o patrones que se presentan a la entrada y salida de la red neuronal. El error entre la salida de la red neuronal y la salida deseada se utiliza para modificar los pesos

de la red neuronal, por lo tanto, el aprendizaje se beneficia de la supervisión controlada de un maestro. Este tipo de aprendizaje se conoce como aprendizaje por retropropagación (backpropagation) y está basado en el método del gradiente descendiente, que calcula el error a la salida de la red neuronal y lo retropropaga desde las últimas capas hasta la capa de entrada para modificar los pesos.

### Predicción mediante redes adaptativas basadas en sistemas de inferencia difusa (ANFIS)

El tercer modelo de predicción que se utilizará se basa en los sistemas neuro-difusos, que emplean reglas de lógica borrosa del tipo *if-then* para modelar el conocimiento humano y los procesos de razonamiento sin emplear análisis precisos. El modelado difuso fue empleado por primera vez por Takagi y Sugeno [Takagi y Sugeno, 1985], y actualmente tiene numerosas aplicaciones prácticas en control [Sugeno, 1985; Pedrycz, 1989], inferencia y predicción [Kandel, 1991].

Un caso especial de sistemas neuro-difusos son los modelos ANFIS (*Adaptive Network based Fuzzy Inference System*), que están basados en los modelos de redes neuronales e intentan simular el modo de razonamiento de los humanos mediante reglas de lógica borrosa *if-then* [Palit y Popovic, 2005].

Un sistema de inferencia difusa está compuesto por cinco bloques funcionales, que se presentan a continuación (ver Figura 3.5):

- Base de reglas. Contiene las reglas difusas *if-then*.
- Base de datos. Define las funciones de pertenencia del conjunto difuso que se utilizan en las reglas difusas *if-then*.
- Unidad de decisión. Realiza las operaciones de inferencia sobre las reglas.
- Interfaz de fuzzificación. Transforma las entradas en valores difusos para que concuerden con los valores lingüísticos.
- Interfaz de defuzzificación. Transforma los resultados difusos del sistema de inferencia en valores de salida.

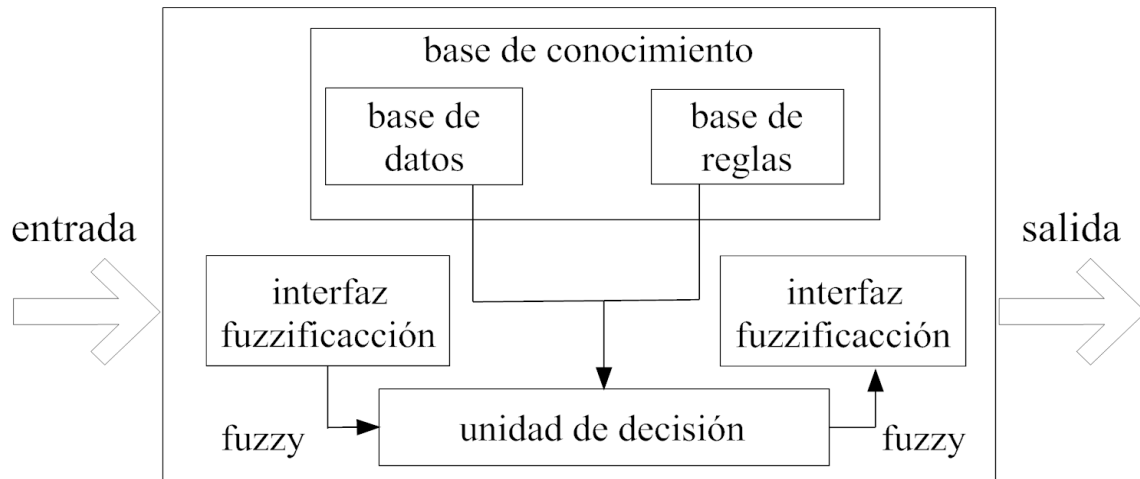


Figura 3.5 Sistema de inferencia difusa

La arquitectura ANFIS consta de cinco capas secuenciales, que se describen a continuación:

- **Capa 1 (Fuzzificación).** En esta capa se utilizan los valores de las variables de entrada, en la parte de la premisa, para adaptar las funciones de pertenencia.
- **Capa 2 (Conjunción).** Se combinan los valores de pertenencia en la parte de la premisa para obtener el *peso* de cada regla.
- **Capa 3 (Normalización).** En esta capa se genera el consecuente de cada regla en función de su peso.
- **Capa 4 (Defuzzificación).** En esta capa se agregan los consecuentes para producir la salida del sistema de lógica difusa.
- **Capa 5 (Agregación).** El único nodo en esta capa es un nodo que calcula la salida del sistema de inferencia difusa como suma de todas las señales de entrada a este nodo.

La Figura 3.6 muestra un ejemplo de sistema ANFIS con tres funciones de pertenencia asociadas a dos entradas, por lo que el espacio de entrada está particionado en nueve subespacios difusos (SD). La premisa de cada regla delimita SD, mientras que la parte consecuente especifica la salida dentro de SD.

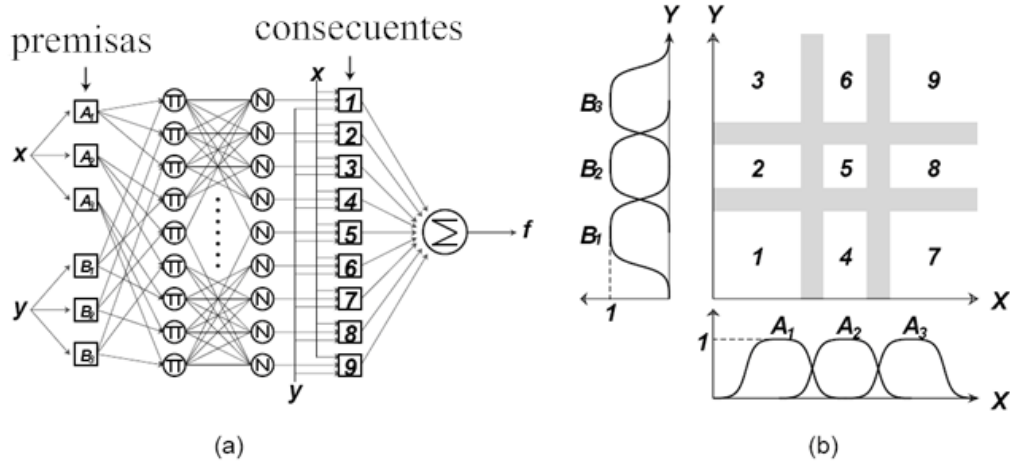


Figura 3.6 ANFIS tipo 3 con dos entradas y 9 reglas (b) subespacios fuzzy correspondientes

El entrenamiento del sistema ANFIS se realiza utilizando un algoritmo de aprendizaje híbrido que combina el método del gradiente descendente (*backpropagation*) y una estimación de los parámetros de la función de pertenencia basada en el método de mínimos cuadrados [Jang, 1993].

El siguiente paso que se va a mostrar es la forma en la que se van a emplear los modelos estadísticos para obtener las predicciones en los diferentes horizontes temporales de predicción y el procedimiento que se va a seguir para realizar la cuantificación de la calidad de las predicciones.

### 3.5.2 Generación de las predicciones de irradiancia solar semidiaria

Sea  $y(t_i)$  una serie temporal estacionaria; formado por los  $k$  valores observados anteriores al instante temporal  $t_i$ :

$$y(t_k) = (y(t_{i-k}), \dots, y(t_{i-1}), y(t_i)) \tag{5.3}$$

La predicción ( $\hat{y}$ ) para el instante temporal  $t_{i+1}$  mediante el vector formado por los  $k$  valores anteriores observados  $y(t_k)$  utilizados como entrada de los modelos estadísticos de predicción ( $g$ ) se puede expresar mediante la siguiente ecuación:

$$\hat{y}(t_{i+1} | t_k) = g(y(t_k)) \quad (5.4)$$

donde  $\hat{y}(t_{i+1} | t_k)$  se define como la predicción para el instante temporal  $t_{i+1}$  a partir del vector formado por los  $k$  valores observados previos al instante temporal  $t_i$  y se denomina  $y(t_k)$ .

De esta forma, las predicciones hasta el horizonte temporal de predicción  $t_{i+h}$  desde el instante temporal  $t_i$  pueden representarse como:

$$\{\hat{y}(t_{i+1}), \hat{y}(t_{i+2}), \dots\} \quad (5.5)$$

El cálculo de las predicciones para los sucesivos horizontes temporales de predicción ( $t_{i+1}, \dots, t_{i+h}$ ) requiere conocer de antemano los valores observados anteriores a cada instante temporal de predicción, es decir, para realizar la predicción del valor  $t_{i+2}$  es necesario conocer el valor observado  $y(t_{i+1})$  y anteriores  $y(t_k)$ , que se puede expresar de la siguiente forma:

$$\begin{aligned} \hat{y}(t_{i+2} | t_{i+1}, t_k) &= g(y(t_{i+1}, t_k)) \\ \hat{y}(t_{i+3} | t_{i+2}, t_{i+1}, t_k) &= g(y(t_{i+2}, t_{i+1}, t_k)) \\ &\dots \\ \hat{y}(t_{i+h} | t_{i+h-1}, \dots, t_{i+2}, t_{i+1}, t_k) &= g(y(t_{i+h-1}, \dots, t_{i+2}, t_{i+1}, t_k)) \end{aligned} \quad (5.6)$$

Si queremos realizar predicciones mayores a un paso temporal (desde  $t_{i+1}$  hasta  $t_{i+h}$ ), los valores observados ( $y$ ) en los instantes temporales ( $t_{i+1}, \dots, t_{i+h-1}$ ) no son conocidos de antemano para el instante temporal  $t_i$ .

El problema puede simplificarse hallando en primer lugar  $\hat{y}(t_{i+1} | t_k)$ , para posteriormente utilizar  $\hat{y}(t_{i+1})$  en el regresor; aplicando de nuevo la ecuación (5.4) para obtener  $\hat{y}(t_{i+2} | t_k)$ , ya que el valor correcto de  $y(t_{i+1})$  no es conocido *a priori*. Posteriormente, se puede aplicar el mismo procedimiento para obtener los valores de  $\hat{y}(t_{i+3} | t_k)$  a partir de  $\hat{y}(t_{i+2})$  y así sucesivamente hasta obtener todas las predicciones hasta llegar al horizonte temporal  $t_{i+h}$ . Esta técnica de predicción se conoce como *predicción recurrente*, ya que utiliza de forma recurrente las predicciones que se obtienen en los pasos anteriores para generar los nuevos valores.

Otra técnica para generar predicciones se conoce como *predicción multipaso* y se basa en la utilización de un modelo de predicción independiente ( $g_i$ ) para cada horizonte de predicción, ( $t_i$ ), que utiliza como entrada únicamente el vector observado,  $y(t_k)$ . De esta forma, las predicciones se generan de la siguiente forma:

$$\begin{aligned}\hat{y}(t_{i+1} | t_k) &= g_1(y(t_k)) \\ \hat{y}(t_{i+2} | t_k) &= g_2(y(t_k)) \\ &\vdots \\ \hat{y}(t_{i+h} | t_k) &= g_h(y(t_k))\end{aligned}\tag{5.7}$$

En este trabajo se utilizarán ambas técnicas (recurrente y multipaso) para realizar las predicciones.

Una vez obtenidas las predicciones de las series temporales intermedias, el siguiente paso que se realizará es el de post-procesado, rehaciendo las descomposiciones efectuadas previamente (ver 3.4.1) de la magnitud de trabajo analizada; en nuestro caso, la irradiancia solar semidiaria para los diferentes horizontes temporales de predicción. A partir de dichas predicciones, se realizará la validación de los resultados obtenidos mediante los diferentes modelos ensayados.

### 3.6 Validación de los modelos de predicción

El proceso de validación tiene como principal objetivo establecer la credibilidad de un modelo para un propósito específico, en nuestro caso el resultado final de los modelos son las predicciones de irradiancia global semidiaria.

Las medidas de error elegidas son las siguientes [Pielke, 2002]:

- Sesgo ó Bias, *Mean Bias Deviation* (MBD), se define como:

$$\text{MBD} = \frac{1}{N} \sum_{i=1}^N (\hat{y}(t_i) - y(t_i)) \quad (5.8)$$

- Error cuadrático medio, *Root Mean Squared Deviation* (RMSD), se define como:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}(t_i) - y(t_i))^2} \quad (5.9)$$

siendo  $N$  la longitud del conjunto,  $y(t_i)$  el valor observado e  $\hat{y}(t_i)$  el valor predicho mediante el modelo.

Para poder comparar el error entre diferentes modelos y estaciones es necesario estandarizar los estimadores anteriores, expresando su valor con respecto al valor medio del conjunto de datos observados de la siguiente forma:

$$\begin{aligned} \text{MBD}(\%) &= \left( \frac{\text{MBD}}{\bar{y}} \right) \cdot 100 \\ \text{RMSD}(\%) &= \left( \frac{\text{RMSD}}{\bar{y}} \right) \cdot 100 \end{aligned} \quad / \quad \bar{y} = \sum_{i=1}^N \frac{y(t_i)}{N} \quad (5.10)$$

Siendo  $\bar{y}$  e  $S_y$  la esperanza y la desviación típica del conjunto de valores, respectivamente.

Otro estimador utilizado en la validación de los resultados se basa en el porcentaje de mejora de cada modelo en términos del error cuadrático medio con respecto a un modelo base o de referencia, que suele ser el modelo más sencillo posible de aplicar. El modelo base elegido en este estudio es el modelo basado en la persistencia, ya que es el más extendido para contrastar propuestas de nuevos modelos. La persistencia se define como:

$$\hat{y}(t_{i+1}) = y(t_i) \quad (5.11)$$

Donde  $y(t_i)$  es el valor observado en el instante temporal  $t_i$  e  $\hat{y}(t_{i+1})$  es la predicción para el siguiente paso temporal ( $t_{i+1}$ ). El modelo basado en la persistencia se comporta como un modelo lineal de orden uno, con ruido gaussiano de media igual a cero.

La mejora porcentual sobre la persistencia en base al RMSD se calcula a partir de la siguiente expresión:

$$MSP(\%) = \left( 1 - \frac{RMSD_m}{RMSD_p} \right) \cdot 100 \quad (5.12)$$

donde  $RMSD_m$  es el error cuadrático medio del modelo propuesto y  $RMSD_p$  es el error cuadrático medio de la persistencia.

### 3.7 Conclusiones

Al comienzo de este capítulo se han presentado las necesidades de las plantas termosolares y fotovoltaicas con respecto a las predicciones de la irradiancia solar. Este trabajo pretende ser de utilidad en la operación y planificación de plantas termosolares y

fotovoltaicas, que precisan de una gestión que debe estar encaminada a la maximización del beneficio económico. A partir de las necesidades que se han presentado, se ha optado por realizar las predicciones de la irradiancia solar con una resolución temporal semidiaria. La elección del paso temporal semidiario para las predicciones se debe a que es el paso temporal que demandan las plantas termosolares, que disponen de un sistema de almacenamiento y contienen mayor información de la dinámica de la irradiancia solar que los valores diarios. Además, este paso temporal presenta una menor incertidumbre comparado con las predicciones horarias de la irradiancia solar. El mercado libre de la energía requiere que la planificación de la producción se realice con un horizonte temporal de predicción de 6 semidías. Por lo tanto, éste será el horizonte temporal de las predicciones que se van a generar.

Los modelos que se pueden utilizar para realizar las predicciones de la irradiancia solar global semidiaria se dividen en modelos físicos y modelos estadísticos. Se han presentado las ventajas y desventajas de ambos tipos de modelos. En este trabajo se ha optado por realizar la predicción únicamente mediante modelos estadísticos, conocidos también como modelos clásicos.

La metodología experimental de predicción que se ha propuesto se basa en el estudio y validación de modelos en función de su linealidad para la mejora de las predicciones de los valores semidiarios de la radiación solar. Las etapas en las que se divide la metodología experimental de predicción que se va a aplicar son las siguientes:

1. Preprocesado de los datos experimentales.
2. Ajuste de los modelos estadísticos de predicción y generación de las predicciones.
3. Validación de los modelos de predicción.

Los siguientes capítulos muestran los principales resultados y el análisis de los mismos tras aplicar la metodología de predicción que se acaba de presentar. En el capítulo 4 se presentarán los resultados de la etapa de preprocesado. Los capítulos 5 y 6 presentarán

respectivamente los resultados de la aplicación de los modelos de predicción y la selección de un modelo final de predicción para cada estación radiométrica.

## 4 PREPROCESADO

En este capítulo, se presentan los resultados que se han obtenido, tras la aplicación de la metodología de preprocesado propuesta en el capítulo 3. En un primer paso, se describe el origen de los datos radiométricos que se van a utilizar en este trabajo. Posteriormente, se muestran los resultados particulares de la metodología de preprocesado aplicada a la serie temporal irradiancia solar global, que consta de los siguientes pasos intermedios:

- Filtrado de los datos piranométricos, a partir de las recomendaciones de la red radiométrica BSRN (*Baseline Surface Radiation Network*).
- Estudio y selección de los periodos de datos radiométricos disponibles con mayor longitud temporal en el registro.
- Construcción de la serie temporal de irradiancia solar semidiaria a partir de los valores disponibles para cada estación radiométrica.
- Transformación de la serie temporal irradiancia solar semidiaria en las series temporales estacionarias intermedias.
- Construcción de un fichero único de datos para cada estación radiométrica.
- Análisis de las propiedades estadísticas de las series temporales estacionarias intermedias semidiarias para cada estación radiométrica.
- Aplicación del test de linealidad sobre las series temporales semidiarias intermedias para conocer la naturaleza de los datos.

### 4.1 Conjunto de datos radiométricos

En este trabajo, se utilizan datos de irradiancia solar global horaria correspondientes a cuatro estaciones de medida de la España peninsular pertenecientes a la Red Radiométrica de AEMet. La elección de la red radiométrica Nacional (RRN) de AEMet como fuente de datos se basa en diversas razones. La RRN es la única red de datos radiométricos dentro de la Península Ibérica con un número extenso de estaciones repartidas por diferentes zonas climáticas dentro de la Península Ibérica. Estos es

importante dada la elevada influencia que la climatología tiene sobre la modelización y predicción de la irradiancia solar.

Adicionalmente, otras de las razones es la calidad de los registros radiométricos, ya que se lleva a cabo un programa de mantenimiento y calibración de los piranómetros. Finalmente, la selección se justifica por la disponibilidad de series radiométricas temporales con varios años de medida para el ajuste de los modelos predictivos. Como normal no se puede admitir una base de datos para análisis y predicción si no se ajustan a unos niveles de calidad y si no tiene una extensión temporal determinada (mínimo 3 años y recomendable 11 años).

Entre todas las estaciones pertenecientes a AEMet disponibles, se ha hecho una selección previa de emplazamientos en los que se realizarán los diversos estudios que se desarrollan en este trabajo. Los criterios se basan en la calidad de los datos radiométricos registrados y en la continuidad del registro temporal de los datos piranométricos. El hecho de garantizar la continuidad de los datos radiométricos es fundamental para asegurar el ajuste de los parámetros de los modelos de predicción de forma óptima.

La Figura 4.1 presenta las diferentes regiones climáticas de la península ibérica. Con la intención de comprobar la bondad en la predicción de cada modelo que se va a ensayar se han elegido estaciones radiométricas representativas de las diferentes zonas climáticas donde existe un mayor número de plantas fotovoltaicas y termosolares instaladas. Las diferentes zonas climáticas seleccionadas son la zona mediterránea árida, mediterránea noroeste, continental y continental extrema.

Entre las estaciones radiométricas de AEMet disponibles para las diferentes zonas climáticas, los datos de las estaciones situadas en Lérida, Madrid, Albacete y Murcia son las que presentan una mejor calidad, al aplicar los procedimientos de filtrado que se muestran más adelante. Además, estas estaciones tienen el menor número de *huecos* en el registro temporal de los datos. En la Figura 4.2, se puede observar la localización de

las estaciones seleccionadas dentro de la Península Ibérica. En dicha figura se incluye el nombre de cada estación, su latitud, longitud, altura, así como el periodo de registro de datos.



Figura 4.1 Regiones climáticas de la Península Ibérica

Las estaciones radiométricas de AEMet emplean piranómetros modelo CM11 de la firma holandesa Kipp&Zonen. Este instrumento utiliza sensores termoeléctricos (termopilas) formados por termopares en serie asentados sobre un substrato metálico ennegrecido [Kipp and Zonen, 2009]. Las principales características técnicas del piranómetro CM11 instalado en las estaciones radiométricas se presenta en la Tabla 4.1. Este instrumento se considera de 1ª clase y las medidas que se obtienen son de alta

calidad, exceptuando los posibles errores en la operación o el mantenimiento del radiómetro.

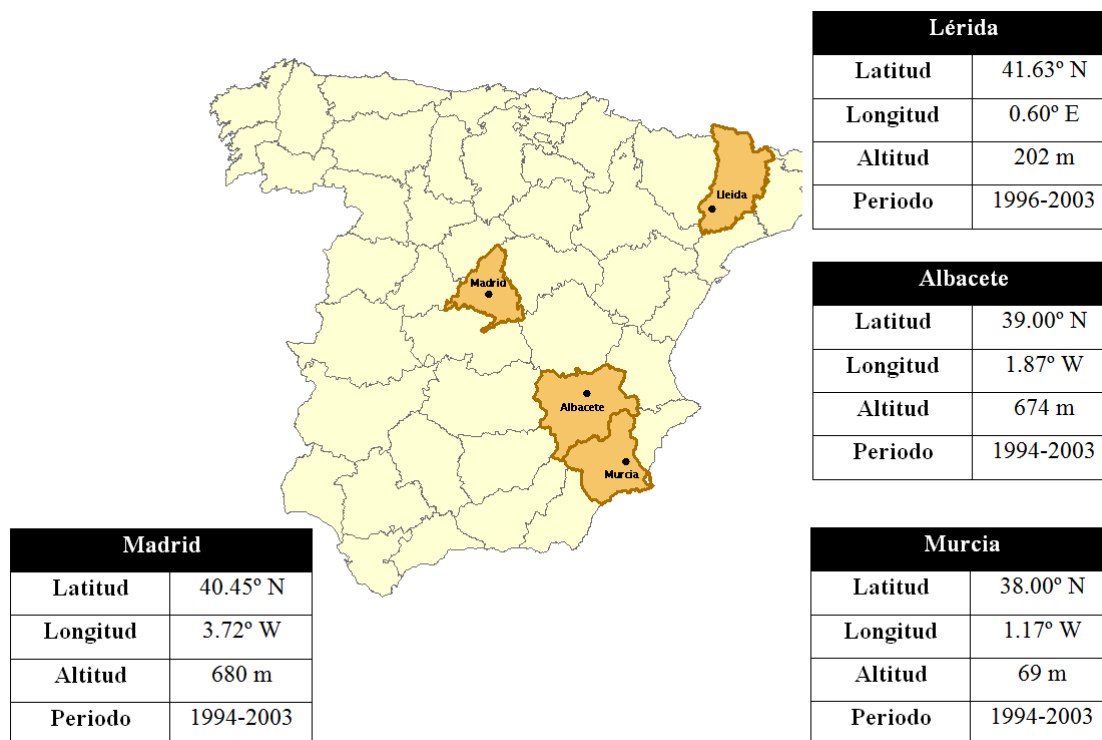


Figura 4.2 Ubicación de las estaciones radiométricas seleccionadas

Tabla 4.1 Características del piranómetro CM11

CM 11	
Rango espectral	0.305 ~ 2.8
Tiempo de respuesta	< 15 s (95 %)
Resolución	< ± 1 Wm <sup>-2</sup>
Estabilidad	< ± 0.5 %
No-linealidad	< ± 0.2 %
Sensibilidad térmica	< ± 1 %
Sensibilidad espectral	< ± 2 %
Respuesta cosenoidal	< ± 3 %
Respuesta acimutal	< ± 3 %
<b>Error máximo en medida integrada</b>	
Radiación horaria	2 %
Radiación diaria	1 %

## 4.2 Aplicación de la metodología de preprocesado de los datos radiométricos

Una vez recopilados los datos piranométricos horarios de las estaciones de medida seleccionadas, es preciso realizar un preprocesado previo a su utilización en el ajuste y validación de los modelos predictivos. El tratamiento de los datos consta de los siguientes pasos:

- Aplicación de los procedimientos de control de calidad recomendados por la BSRN a los datos piranométricos horarios.
- Estudio y selección de los periodos temporales de datos radiométricos disponibles.
- Construcción de la serie temporal de irradiancia solar semidiaria a partir de los valores horarios disponibles para cada estación radiométrica.
- Transformación de la serie temporal irradiancia solar semidiaria en una serie estacionaria.
- Construcción de un fichero único para cada estación radiométrica, a utilizar para realizar el análisis sobre los datos y como entrada a los modelos de predicción.
- Análisis de las propiedades estadísticas de las series temporales estacionarias para cada estación radiométrica.

Todos los fundamentos teóricos relacionados con la radiación solar, que se utilizan en esta memoria, se exponen en el APÉNDICE A.

## 4.3 Procedimientos de control de calidad

En este trabajo se ha realizado un análisis exhaustivo en dos fases de las series temporales de radiación solar horaria. El estudio realizado se caracteriza por la aplicación de una serie de procedimientos para la obtención de datos radiométricos de la más alta calidad.

En una primera fase se han aplicado los procedimientos de filtrado recomendados por la red radiométrica BSRN (ver APÉNDICE A). En concreto, se ha aplicado el nivel de filtros físicos consistente en comprobar que los valores registrados no superen los valores físicamente posibles: como valor inferior y la *constante solar* ( $I_{CS}$ ) como valor

superior. Se define la *constante solar* como la irradiancia solar sobre superficie plana normal al vector posición del Sol ubicada en el límite superior de la atmósfera terrestre (ver APÉNDICE A).

Además, se ha realizado una comparación de la irradiancia solar medida en la superficie terrestre con la irradiancia solar extraterrestre recibida en la superficie de la atmósfera. Este último filtro, se aplica como un filtro paso bajo, presuponiendo que la hora de registro de los datos sea la correcta. La aplicación de esta primera fase en el proceso de filtrado de datos permite:

- Identificar la información relativa a la disponibilidad de datos para cada mes y cada año.
- Conocer la distribución horaria de los valores erróneos detectados, en consecuencia, identificar la distribución de errores en el registro temporal.

En una segunda fase se ha realizado un análisis gráfico de los valores del índice de claridad horario. Este índice es un parámetro adimensional normalizado y su representación gráfica permite identificar comportamientos sospechosos. El análisis gráfico se ha efectuado mediante la representación de la serie temporal índice de claridad horaria ( $k_t$ ) y el diagrama de Box-Whisker.

La representación gráfica de la serie temporal  $k_t$  permite identificar comportamientos anómalos durante periodos concretos. La visualización de la serie  $k_t$  habitualmente presenta valores por debajo de 0.8 debido a la atenuación de la irradiancia solar, por los componentes atmosféricos para condiciones de cielo claro.  $k_t$  representa habitualmente un valor mínimo de 0.2 [Aguiar y Collares-Pereira, 1992b]. Los valores que se han eliminado de la serie temporal de irradiancia solar horaria se representan en la figura de índice de claridad en color rojo (Figura 4.3, Figura 4.5, Figura 4.7 y Figura 4.9).

En el diagrama de Box-Whisker (para cada hora del día) se dibuja una caja con el 50% de los valores centrales que se extienden entre el primer y el tercer cuartil; la línea central que atraviesa la caja corresponde a la mediana y a partir de los extremos de la

caja se extienden los valores que se consideran “normales”. Los valores extremos que difieren del cuartil más próximo en más de 1.5 veces el intervalo intercuartílico se grafican como puntos aislados, por considerarse que pueden corresponder a datos anómalos (*outliers*). El diagrama Box-Whisker permite identificar posibles asimetrías en los valores horarios con respecto al mediodía solar, debidas a errores en el registro de la hora.

Las Figura 4.4, 4.5, 4.7 y 4.9 muestran los diagramas de Box-Whisker de  $k_t$  para cada hora del día. Los *outliers* aparecen marcados en color rojo con el símbolo (+) y la línea que marca la mediana horaria dentro de la caja representa el rango intercuartílico. Las cuatro estaciones analizadas presentan una buena simetría con respecto al cenit solar. Este hecho es de especial importancia para la construcción de las series temporales semidiarias, cuyos valores representan la irradiación solar acumulada en las horas anteriores y posteriores al mediodía solar.

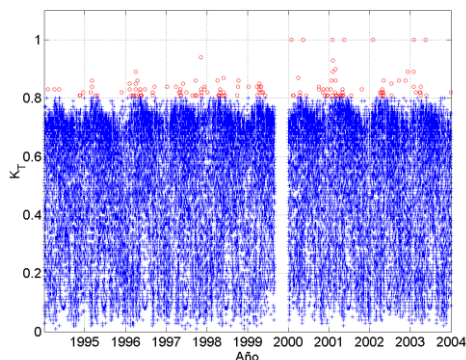


Figura 4.3 Murcia:  $k_t$  (datos filtrados)

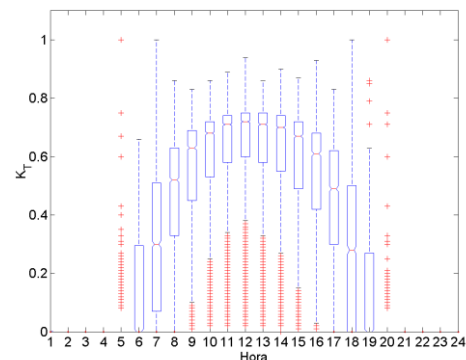


Figura 4.4 Distribución horaria  $k_t$  para Murcia

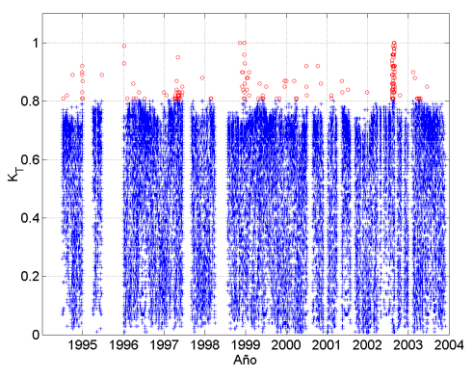


Figura 4.5 Albacete:  $k_t$  (datos filtrados)

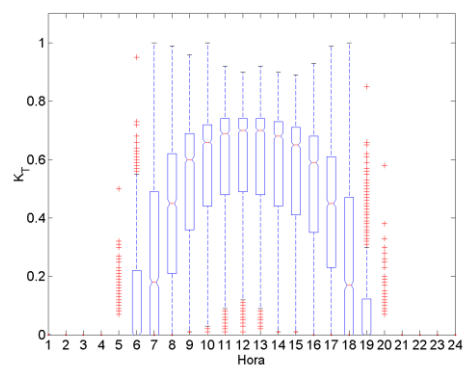
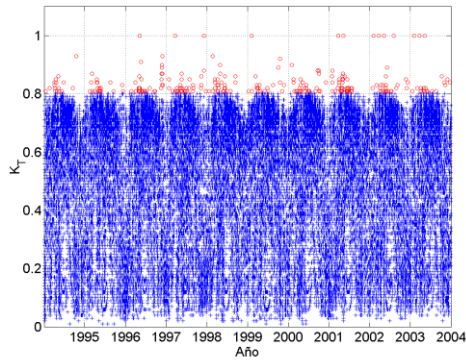
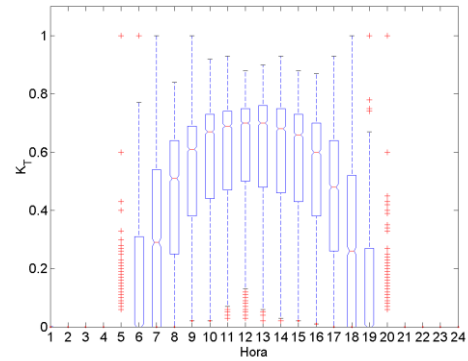
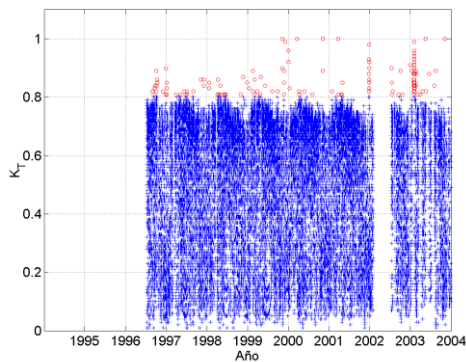
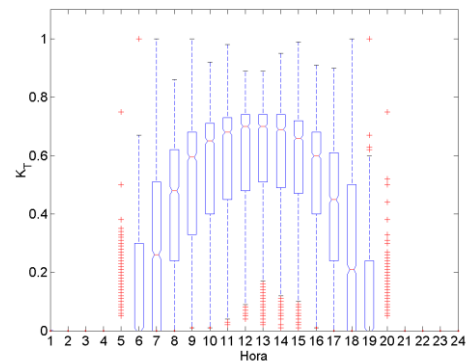


Figura 4.6 Distribución horaria  $k_t$  para Albacete

Figura 4.7 Madrid:  $k_t$  (datos filtrados)Figura 4.8 Distribución horaria  $k_t$  para MadridFigura 4.9 Lérida:  $k_t$  (datos filtrados)Figura 4.10 Distribución horaria  $k_t$  para Lérida

#### 4.4 Selección de los periodos de datos disponibles

El filtrado de los datos de radiación solar horaria, presentado anteriormente, nos ha permitido identificar los registros radiométricos anómalos. El siguiente paso es el estudio de la disponibilidad de periodos de datos radiométricos horarios registrados con continuidad temporal. La intención de este estudio previo es identificar periodos en los que existen averías en la medida piranométrica o falta de mantenimiento en la instrumentación; para asegurar la continuidad de las series temporales que se utilizan en el ajuste y validación de los modelos de predicción. Al disponer de series de datos continuas se consigue mantener la correlación (ya sea lineal o no-lineal) original de las series temporales, por lo tanto, no se introducen nuevas perturbaciones en la dinámica de la irradiancia solar.

La Tabla 4.2 representa, para todas las estaciones, los datos de disponibilidad temporal de registros con el número de días por mes para cada año en los que se dispone de datos

---

de irradiancia global horaria. Cada color de fondo de la tabla indica los diferentes periodos de datos seleccionados. Se muestra, además, el número de registros seleccionados por periodo y el % de datos no disponibles para dicho periodo. En la última fila, se presenta el número total de registros diarios disponibles y el porcentaje de datos no disponibles.

De la Tabla 4.2, para cada estación radiométrica, se puede extraer la siguiente información:

- Murcia. En esta estación, destaca la existencia de un periodo en el que se produce una interrupción en la medida de datos radiométricos de más de cuatro meses. En consecuencia, se tienen dos periodos de datos bien diferenciados que van desde 1 de enero de 1994 hasta 31 de agosto de 1999 y desde 1 de enero de 2000 hasta 31 de diciembre de 2003.
- Albacete. La principal característica de esta estación es que presenta continuas interrupciones temporales en la medida de la serie temporal de irradiancia solar. Se pueden identificar dos periodos de datos, que presentan el mínimo número de discontinuidades y van desde 1 de enero de 1996 hasta 31 de mayo de 1997 y desde 9 de agosto de 1998 hasta 30 de junio de 2000. A pesar de la discontinuidad en el registro de las medidas radiométricas, esta estación se ha seleccionado para comprobar el efecto y adecuación del uso de la metodología propuesta a estaciones con falta de continuidad en los datos.
- Madrid. La estación de Madrid presenta una alta calidad en cuanto al registro temporal de los datos. En esta estación, se dispone de un único periodo de datos de la serie temporal de irradiancia solar, que va desde 1 de enero de 1994 hasta 31 de diciembre de 2003.
- Lérida. Esta estación presenta una alta continuidad temporal, destacando únicamente una interrupción de más de cinco meses en la medida de los datos radiométricos. En consecuencia, se obtienen dos periodos de datos bien diferenciados, que van desde 1 de mayo de 1997 hasta 31 de marzo de 1999 y desde 1 de junio de 1999 hasta 31 de enero de 2002.

Tabla 4.2 Número de días disponibles por mes para cada estación.

	E	F	M	A	M	J	J	A	S	O	N	D	# Registros	%Disponibilidad	
	<b>MURCIA</b>	31	28	31	30	31	30	31	31	30	31	30			30
1994	31	28	31	30	31	30	31	31	30	31	30	31			
1995	31	28	31	30	31	30	31	31	30	31	30	31			
1996	31	29	31	30	31	30	31	31	30	31	30	31			
1997	31	28	31	30	31	30	31	31	30	31	30	31			
1998	31	28	31	30	31	30	31	30	30	31	30	31			
1999	31	28	31	30	31	28	31	31	0	0	0	0	1442	99.987	
2000	30	29	31	30	31	30	31	31	30	30	30	26			
2001	31	27	31	29	31	30	31	31	30	31	30	31			
2002	31	28	31	30	25	30	31	31	30	31	30	31			
2003	31	26	31	30	31	30	31	31	27	31	30	31	<b>TOTAL</b>	3507	99.994
<b>ALBACETE</b>	E	F	M	A	M	J	J	A	S	O	N	D	# Registros	%Disponibilidad	
1994	0	0	0	0	0	0	31	31	30	31	30	31	515	99.996	
1995	0	0	0	30	25	15	0	0	0	0	0	0			
1996	29	29	31	30	31	30	31	31	30	31	30	31			
1997	31	28	31	30	31	15	0	0	30	31	30	31			
1998	31	28	31	0	0	0	9	23	30	31	30	31			
1999	31	28	31	30	31	29	31	31	30	31	30	31			668
2000	31	29	31	30	31	30	6	7	30	31	30	0			
2001	24	28	27	0	20	30	31	0	23	31	20	29			
2002	31	28	31	13	9	14	30	25	8	31	8	31			
2003	0	21	31	30	31	30	31	31	30	31	21	0	<b>TOTAL</b>	1183	99.998
<b>MADRID</b>	E	F	M	A	M	J	J	A	S	O	N	D	# Registros	%Disponibilidad	
1994	31	28	31	30	31	30	31	31	30	31	30	31	3652	100	
1995	31	28	31	30	31	30	31	31	30	31	30	31			
1996	31	29	31	30	31	30	31	31	30	31	30	31			
1997	31	28	31	30	31	30	31	31	30	31	30	31			
1998	31	28	31	30	31	30	31	31	30	31	30	31			
1999	31	28	31	30	31	30	31	31	30	31	30	31			
2000	31	29	31	30	31	30	31	31	30	31	30	31			
2001	31	28	31	30	31	30	31	31	30	31	30	31			
2002	31	28	31	30	31	30	31	31	30	31	30	31			
2003	31	28	31	30	31	30	31	31	30	31	30	31			<b>TOTAL</b>
<b>LÉRIDA</b>	E	F	M	A	M	J	J	A	S	O	N	D	# Registros	%Disponibilidad	
1994	0	0	0	0	0	0	0	0	0	0	0	0	692	99.989	
1995	0	0	0	0	0	0	0	0	0	0	0	0			
1996	0	0	0	0	0	0	22	31	30	23	30	31			
1997	31	28	17	24	31	30	31	31	30	31	30	31			
1998	31	28	23	30	31	30	31	31	30	31	30	31			
1999	31	28	31	24	27	30	31	31	30	31	30	31			970
2000	31	29	30	30	31	29	31	31	30	28	30	31			
2001	31	28	29	30	31	30	31	31	30	31	30	31			
2002	31	0	0	0	0	0	13	14	30	27	30	19			
2003	12	20	17	12	10	7	5	19	22	31	23	14	<b>TOTAL</b>	1662	99.992

---

Una vez obtenidos los diferentes periodos que forman las series temporales de irradiancia solar horaria, se construyen las series temporales de irradiancia solar semidiaria para cada estación radiométrica.

#### **4.5 Transformación de la serie temporal de irradiancia global semidiaria en una serie estacionaria**

Los diferentes modos de variabilidad de la serie temporal de radiación solar semidiaria pueden descomponerse, a partir de los conceptos expuestos en 3.2.1 y de la naturaleza de la irradiancia solar que se describe en el APÉNDICE A, de la siguiente forma:

- Efecto estacional (*S*). Se debe a la variación anual de la irradiancia solar por los cambios de distancia entre la Tierra y el Sol ocasionados por la excentricidad de la órbita terrestre y el efecto de la declinación solar [Iqbal, 1983].
- Otros cambios cíclicos (*C*). Además de los efectos estacionales, la serie temporal de irradiancia solar con frecuencia temporal inferior a un día, presenta un ciclo día/noche debido al movimiento de rotación de la Tierra sobre su propio eje.
- Tendencia Global (*T*). El régimen de variación de la irradiancia solar de largo plazo se rige de acuerdo a la actividad magnética en el interior del astro y se repite cada 22 años (ciclo de Hale). La polaridad del Sol se invierte cada once años, esto es, el norte magnético cambia hacia el sur y viceversa. Este tipo de variación, ocasionadas por las fluctuaciones de las manchas solares, se omite en el presente estudio, ya que se asume un valor constante de la irradiancia solar que llega al borde superior de la atmósfera; esta magnitud se conoce como *Constante Solar* y se asume un valor de  $1367 \text{ Wm}^{-2}$ , a partir de estimaciones realizadas por el Centro Radiológico Mundial (WRC) [WMO, 1981]. Sin embargo, en los cálculos que se realicen se tendrá en cuenta la variación de la distancia Sol-Tierra, debido a la excentricidad de la órbita terrestre, que provoca una variación estacional de la irradiancia solar que llega al borde superior de la atmósfera de  $\pm 3.3\%$ . En el APÉNDICE A se realiza una descripción más detallada sobre estos conceptos relacionados con la geometría Sol-Tierra.
- Otras fluctuaciones (*a*). Una vez eliminadas la tendencia y la variación cíclica de la serie temporal, se obtiene una serie de residuos que pueden ser o no aleatorios.

La parte aleatoria (o no-determinista) se debe a la influencia cuasi-aleatoria que tiene la nubosidad (movimiento y estructuras de las altas y bajas presiones) y el contenido de aerosoles en la atmósfera (vapor de agua, ozono, aerosoles,...).

Existen diferentes métodos para separar las componentes determinísticas y aleatoria de la serie temporal de irradiancia solar semidiaria, siendo los más importantes los que se presentan a continuación:

- Índice de claridad semidiario ( $K_T$ ). Se define como el cociente entre los valores semidiarios de irradiancia solar registrados en la superficie terrestre y los correspondientes valores de irradiancia solar extraterrestre [Kreider y Kreith, 1978; Bouhadou et al., 1997; Safi S. y Zeroual A., 1999]. Esta transformación es similar a la ya realizada por otros autores para la generación sintética de series de datos de radiación solar [Brinkworth, 1977; Bendt et al., 1981; Aguiar et al., 1988; Aguiar y Collares-Pereira, 1992b; Mora-Lopez y Sidrach-de-Cardona, 1997; Mora-Lopez y Sidrach-de-Cardona, 2003]. Los valores de  $K_T$  se obtienen a partir de la irradiancia solar, aplicando la siguiente expresión:

$$K_T(t) = \frac{G(t)}{I_0(t)} \quad (6.1)$$

donde  $G(t)$  es la radiación global semidiaria que llega a la superficie terrestre sobre superficie horizontal y  $I_0(t)$  es la irradiancia extraterrestre semidiaria sobre superficie horizontal.

- Componente solar atenuada o *lost component* semidiario (LC). Se obtiene mediante la diferencia entre el valor de irradiancia solar extraterrestre semidiario y el valor correspondiente de la irradiancia global solar recibida en la superficie terrestre sobre superficie horizontal para el periodo de un semidía [Safi et al., 2002]. Se obtiene utilizando la siguiente ecuación:

$$LC(t) = I_0(t) - G(t) \quad (6.2)$$

donde la serie temporal LC representa la energía solar semidiaria, absorbida o reflejada a su paso por la atmósfera.

- Otro método propuesto, se basa en el análisis de la serie temporal en el dominio espectral mediante la transformada de Fourier para la identificación de la media, la desviación estándar y la parte aleatoria de la serie temporal de irradiancia solar semidiaria [Amato et al., 1986; Amato et al., 1989].

En este trabajo, se han seleccionado los dos primeros métodos, debido a que, el tercero es redundante al primero y presenta una mayor complejidad. Por lo tanto, la serie temporal de irradiancia solar semidiaria se transformará en una señal estacionaria utilizando los valores de  $K_T$  y LC, que se utilizarán como predictores en los modelos estadísticos de predicción.

El siguiente paso que se ha realizado, es la construcción de un fichero único para cada estación radiométrica con un registro por cada semidía y ocho columnas por registro (año, mes, día, periodo con respecto al meridiano local, irradiancia solar semidiaria acumulada, radiación solar extraterrestre semidiaria,  $K_T$  y LC). Los datos de irradiancia solar se almacenan en  $Wm^{-2}semidía^{-1}$  y se utiliza el valor -999 como identificador de dato erróneo.

#### **4.6 Análisis de las propiedades estadísticas de las series temporales estacionarias**

En esta sección se presentan los resultados del análisis exploratorio que se ha realizado sobre las series temporales semidiarias índice de claridad ( $K_T$ ) y componente solar atenuada (LC). El objetivo es conocer las propiedades estadísticas de las series temporales  $K_T$  y LC a partir de sus funciones de densidad de probabilidad empírica y de autocorrelación parcial. Las conclusiones que aquí se obtienen nos ayudarán a interpretar los resultados de los modelos de predicción.

El estudio de la *función de densidad de probabilidad* empírica de las series temporales  $K_T$  y LC nos permite conocer la adecuación de sus distribuciones muestrales a la

---

función de distribución gaussiana; ya que los modelos de predicción estadísticos que se van a utilizar, requieren que los datos tengan una distribución normal. En caso de que los datos no presenten una distribución gaussiana, los valores con mayor frecuencia en la muestra, tendrán mayor peso en el ajuste de los parámetros de los modelos de predicción, lo que dará lugar a que, los errores de predicción no sigan una función de distribución normal.

La función de densidad de probabilidad muestra el rango de valores de la serie temporal ( $K_T$  o LC) y el número de muestras para dicho valor. El máximo valor de la función de densidad de probabilidad empírica puede ser mayor a uno, ya que el único requisito es que la integral de la función de distribución de probabilidad sobre la línea real sea igual a uno.

Por otro lado, la *función de autocorrelación parcial* es una guía importante de las propiedades estadísticas de una serie temporal, ya que mide la correlación entre dos observaciones separadas por sucesivos retardos temporales, aislando los posibles efectos que puedan tener el resto de retardos temporales. Los coeficientes de la autocorrelación nos dan información sobre la dinámica y la memoria del proceso que genera los datos porque recogen la influencia del pasado en el presente.

El coeficiente de autocorrelación parcial representa el retardo temporal de la serie temporal ( $K_T$  o LC) y su valor de autocorrelación parcial. En el eje X, se exhiben los 20 primeros retardos temporales; y, en el eje Y, el valor de la función de autocorrelación parcial para dicho retardo. Además, en las gráficas se muestra la banda de incertidumbre del 95% de los valores de autocorrelación para la serie temporal ( $K_T$  o LC).

#### Propiedades estadísticas de la serie temporal índice de claridad semidiaria ( $K_T$ )

Las Figuras 4.10, 4.12, 4.14 y 4.16 muestran la función de densidad de la serie temporal  $K_T$  para cada estación radiométrica. Las Figuras 4.11, 4.13, 4.15 y 4.17 presentan la función de autocorrelación de la serie temporal  $K_T$  para cada estación radiométrica. A partir de estas figuras, se puede extraer la información que a continuación se expone.

---

La función de distribución de probabilidad de los valores  $K_T$  semidiarios tiene una forma similar a la de los valores diarios y horarios. En los resultados obtenidos se observa una función unimodal, ésta se puede aproximar a una función bimodal o según algunos autores a una función de distribución biexponencial, dependiente del índice de claridad medio mensual y de los valores máximo y mínimos de cada mes [Ibañez et al., 2002]. Por otro lado, la función de densidad de probabilidad muestra una alta predominancia de condiciones de cielo claro para las cuatro estaciones radiométricas, aunque es algo menor para el caso de Madrid. A partir de estos resultados es de esperar que los modelos de predicción se comporten mejor para las condiciones de cielo claro.

En los estudios realizados por Aguiar, se obtiene un valor medio de la autocorrelación parcial de  $K_T$  diario para el primer día de  $r=0.29$  [Aguiar y Collares-Pereira, 1992b]. En el caso de los datos  $K_T$  analizados,  $r$  es bastante más alto, en comparación con los resultados obtenidos por Aguiar; en torno a 0.6. Como se puede observar en la figuras, es evidente que existe una alta autocorrelación parcial exclusivamente para el primer retardo temporal semidiario para todas las estaciones radiométricas seleccionadas en este trabajo. Por lo tanto, se puede afirmar que al ensayar los modelos de predicción lineal, se obtendrán buenos resultados con modelos de orden uno; o expresado de otra forma, aquellos modelos que utilicen exclusivamente el primer retardo temporal de la serie temporal para realizar las predicciones. Por lo tanto, al aumentar el orden de los modelos lineales, por lo tanto, es previsible que no se reduzca en gran medida el error de las predicciones.

#### Propiedades estadísticas de la componente solar atenuada semidiaria (LC)

Al igual que en el apartado anterior, para cada estación se ha realizado un análisis exploratorio sobre la serie temporal LC semidiaria en forma gráfica. Se calcula su función de densidad de probabilidad empírica (Figura 4.19, Figura 4.21, Figura 4.23 y Figura 4.25) y su función de autocorrelación parcial (Figura 4.20, Figura 4.22, Figura 4.24 y Figura 4.26).

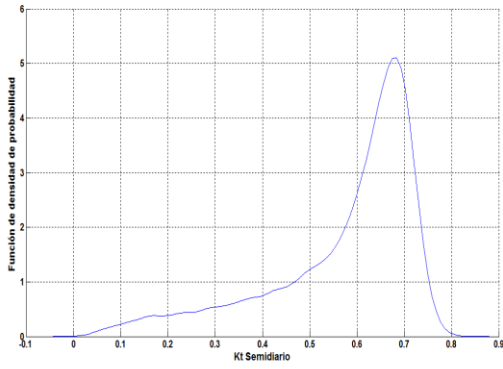


Figura 4.11 Función de densidad  $K_T$  (Murcia)

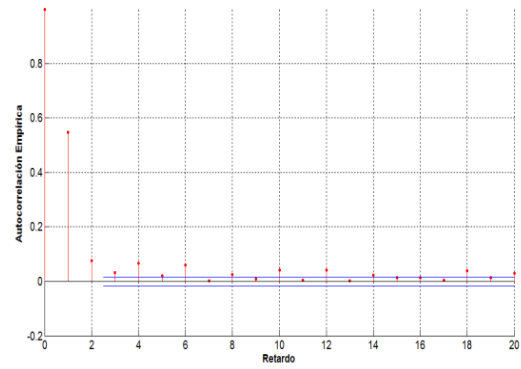


Figura 4.12 Autocorrelación parcial  $K_T$  (Murcia)

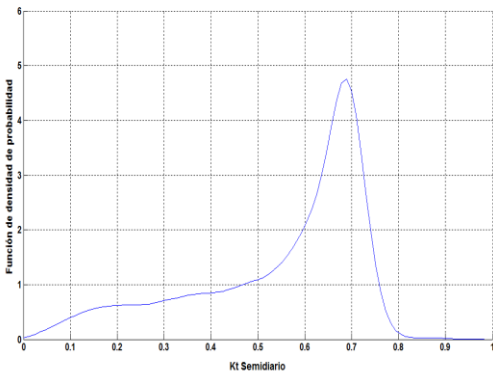


Figura 4.13 Función densidad  $K_T$  (Albacete)

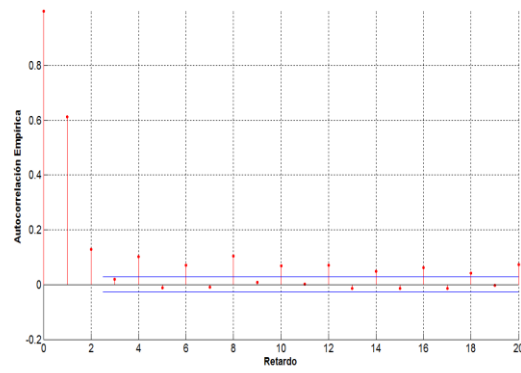


Figura 4.14 Autocorrelación parcial  $K_T$  (Albacete)

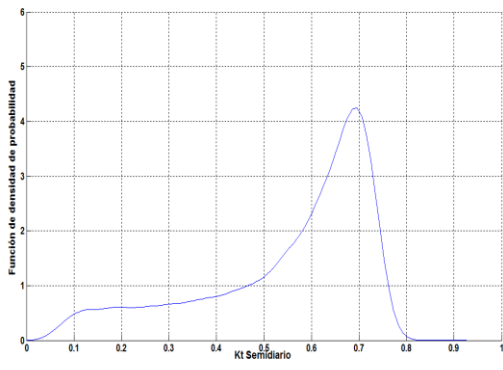


Figura 4.15 Función densidad  $K_T$  (Madrid)

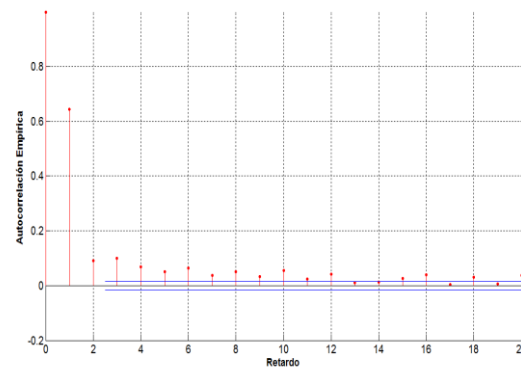


Figura 4.16 Autocorrelación parcial  $K_T$  (Madrid)

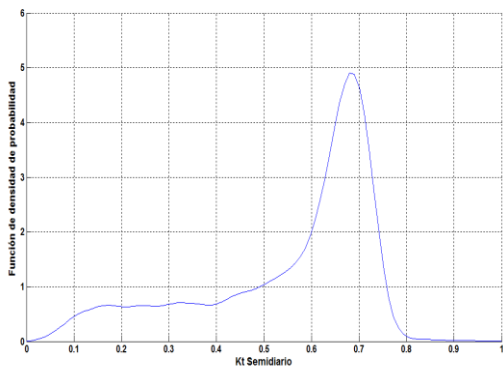


Figura 4.17 Función de densidad  $K_T$  (Lérida)

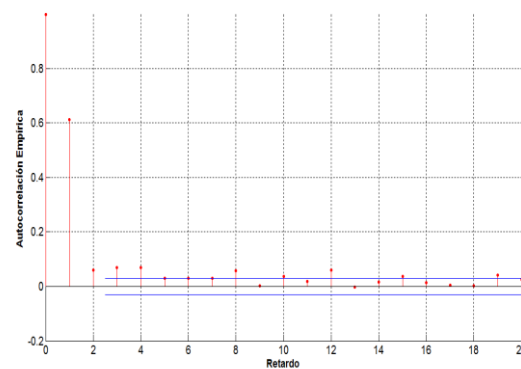


Figura 4.18 Autocorrelación parcial  $K_T$  (Lérida)

Como se puede observar en las figuras, la función de distribución de la serie temporal LC presenta una asimetría. Por lo tanto, se pueden obtener las mismas conclusiones que se comentaron anteriormente para la serie temporal  $K_T$ . De esta forma, a partir de los resultados (y al igual que se afirmó para la serie temporal  $K_T$ ) es de esperar que los modelos de predicción (a partir de la serie temporal LC) se comporten mejor para las condiciones de cielo claro.

El valor de la función de autocorrelación parcial de la serie temporal LC para el primer retardo temporal, muestra un valor en torno a 0.6 para las cuatro estaciones. Para el resto de retardos temporales, el valor de autocorrelación está por debajo de 0.1. De esta forma, se puede afirmar que la función de autocorrelación parcial, que muestra la serie temporal LC es similar a la de la serie temporal  $K_T$ . Por lo tanto, se pueden extraer las mismas conclusiones, que se comentaron previamente para la serie temporal  $K_T$ . Al ensayar los modelos de predicción lineal, es de esperar, que se obtengan buenos resultados con aquellos modelos que utilicen exclusivamente el primer retardo temporal de la serie temporal para realizar las predicciones.

#### **4.7 Aplicación del test de linealidad a las series temporales estacionarias**

El análisis de la linealidad de las series temporales  $K_T$  y LC nos permite conocer la naturaleza del proceso que generan ambas series temporales. Esta información será de utilidad a la hora de interpretar los resultados de los modelos estadísticos en función de su linealidad.

El análisis de linealidad que se ha aplicado está basado en el método de datos subrogados y Delay Vector Variance, que es una técnica novedosa para detectar la no linealidad de una serie temporal.

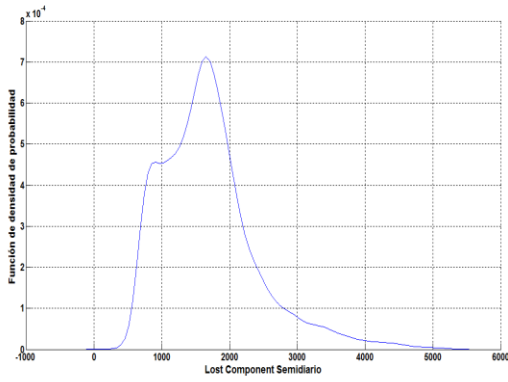


Figura 4.19 Función densidad LC (Murcia)

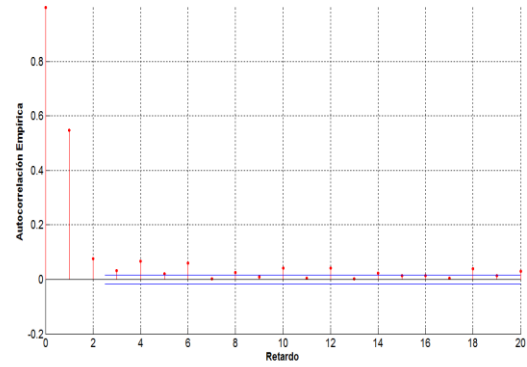


Figura 4.20 Autocorrelación parcial LC (Murcia)

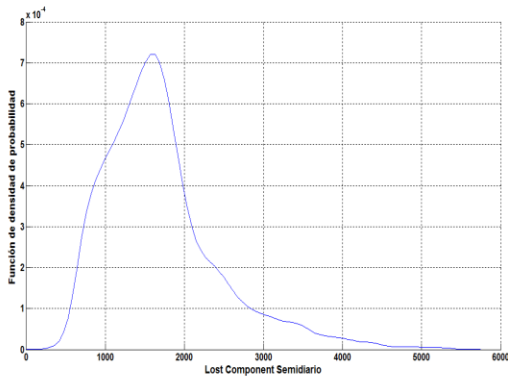


Figura 4.21 Función densidad LC (Albacete)

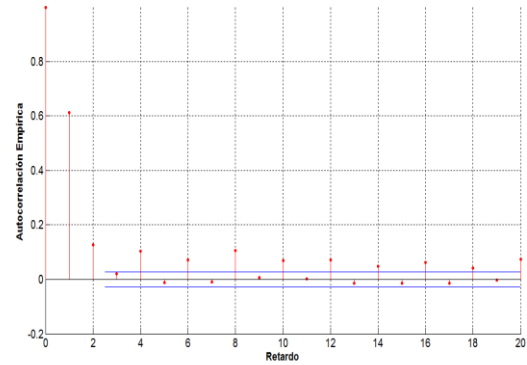


Figura 4.22 Autocorrelación parcial LC (Albacete)

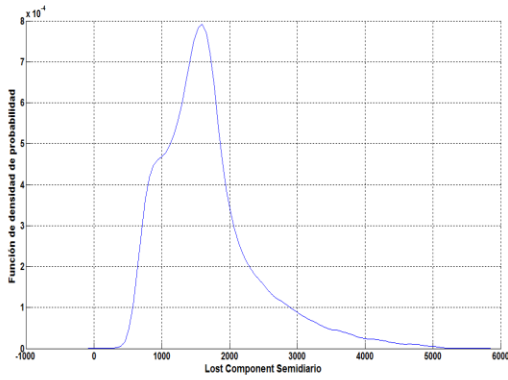


Figura 4.23 Función densidad LC (Madrid)

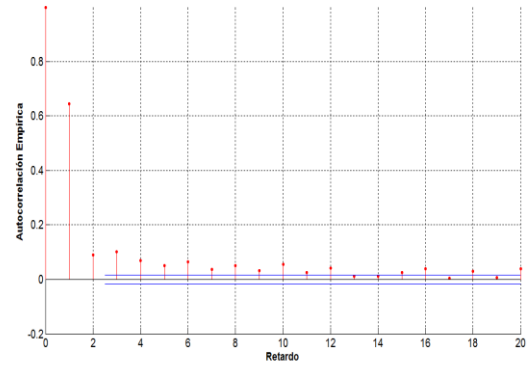


Figura 4.24 Autocorrelación parcial LC (Madrid)

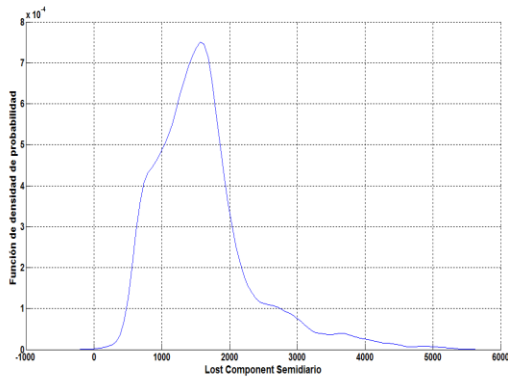


Figura 4.25 Función densidad LC (Lérida)

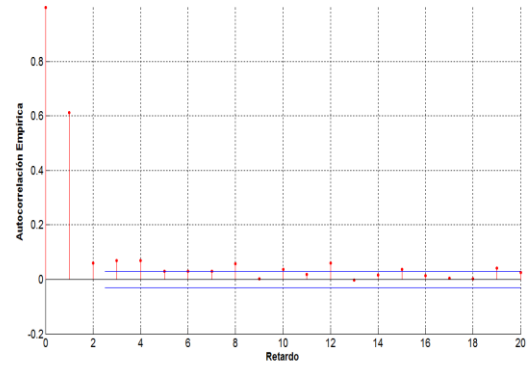


Figura 4.26 Autocorrelación parcial LC (Lérida)

El proceso de generación de datos subrogados, o versiones lineales, se ha realizado utilizando los métodos basados en la Transformada de Fourier (FT) y iAAFT (*iterative Amplitude Adjusted Fourier Transform*). El método FT garantiza que los datos subrogados tendrán la misma amplitud espectral y propiedades lineales que las series originales. Por otro lado, el método iAAFT garantiza que los datos subrogados tengan la misma función de distribución, amplitud espectral y propiedades lineales que las series temporales  $K_T$  y LC originales.

Para medir la linealidad de la serie original y series subrogadas, se utilizan los discriminantes estadísticos no-lineales: autocovarianza de tercer orden (C3), reversibilidad Temporal (REV) y varianza del vector de retrasos (DVV).

#### Verificación del test de hipótesis

La verificación de la hipótesis nula o de linealidad se ha realizado mediante la aplicación de un test no-paramétrico, basado en el orden de los resultados de los estadísticos discriminantes, para cada serie temporal original y subrogadas [Theiler y Prichard, 1996].

Para cada serie temporal original, se han generado 99 series subrogadas y se han calculado los valores de los discriminantes estadísticos no-lineales para la serie original y las series temporales subrogadas. Una vez calculado el conjunto de estadísticos para cada serie temporal, éstos se ordenan, determinando el índice de la posición  $r_0$  de la serie temporal original.

Para la aplicación del test paramétrico, se ha utilizado un intervalo de confianza del 10%. El estadístico discriminante DVV se basa en un test de cola derecha, que rechazará la hipótesis nula si el rango  $r_0$  de la serie original es mayor a 90, o menor o igual a 10. Por otro lado, los discriminantes C3 y REV se basan en un test de dos colas o bilateral, que rechaza la hipótesis nula si el rango  $r_0$  es mayor que 95, o menor o igual a 5.

La Tabla 4.3 presenta los resultados de verificación del test de hipótesis para la serie temporal  $K_T$  y la Tabla 4.4, para la serie temporal LC. En cursiva, subrayado y negrita se presentan los casos en los que se acepta la hipótesis nula; es decir, la serie temporal original presenta un comportamiento lineal. Así, el rechazo de la hipótesis nula puede ser atribuido a la presencia de no linealidad en la serie temporal original. En tal caso, las versiones lineales y la serie temporal original no tendrán las mismas propiedades de no-linealidad y ambas series temporales no compartirán la misma función de distribución para los discriminantes estadísticos no-lineales que se han calculado.

Tabla 4.3 Porcentaje de rechazo (%) de la hipótesis nula para la serie temporal  $K_T$

Estación	Algoritmo	C3	REV	DDV
Murcia	FT	1/100	100/100	100/100
	iAAFT	<b><u>80/100</u></b>	100/100	91/100
Albacete	FT	1/100	100/100	100/100
	iAAFT	99/100	96/100	95/100
Madrid	FT	1/100	100/100	100/100
	iAAFT	100/100	99/100	<b><u>30/100</u></b>
Lérida	FT	1/100	94/100	100/100
	iAAFT	<b><u>81/100</u></b>	<b><u>89/100</u></b>	<b><u>58/100</u></b>

Tabla 4.4 Porcentaje de rechazo (%) de la hipótesis nula para la serie temporal LC

Estación	Algoritmo	C3	REV	DDV
Murcia	FT	100/100	1/100	100/100
	iAAFT	1/100	1/100	<b><u>59/100</u></b>
Albacete	FT	100/100	1/100	100/100
	iAAFT	9/100	1/100	<b><u>63/100</u></b>
Madrid	FT	100/100	1/100	100/100
	iAAFT	9/100	1/100	97/100
Lérida	FT	100/100	1/100	100/100
	iAAFT	2/100	1/100	94/100

Los resultados, que se obtienen tras la aplicación del test de linealidad paramétrico, nos permiten afirmar que existen indicios de que la serie temporal  $K_T$  tiene un comportamiento lineal (o una menor no-linealidad) en la estación de Lérida; ya que, para el caso del algoritmo iAAFT, todos los estadísticos aceptan la hipótesis nula. De esta forma, es de esperar que se obtengan buenos resultados a partir de la aplicación de

los modelos de predicción lineal a la serie temporal  $K_T$  en Lérida. Para el resto de estaciones, se concluye que ambas series temporales ( $K_T$  y LC) presentan un comportamiento no-lineal, por lo que se puede presuponer, que la aplicación de modelos de predicción no-lineales obtendrá mejores resultados que el modelo lineal para ambas series temporales  $K_T$  y LC.

#### **4.8 Conclusiones**

La aplicación de la metodología de preprocesado a la serie temporal de irradiancia global solar horaria ha permitido obtener las series temporales estacionarias ( $K_T$  y LC) y un conjunto de conclusiones, que serán de utilidad para interpretar los resultados de predicción a partir de la aplicación de los diferentes modelos.

Los pasos intermedios que se han realizado en esta fase son la recopilación de datos, la aplicación de filtros a los datos horarios, el estudio de los periodos temporales disponibles y la construcción de la serie temporal de irradiancia solar semidiaria.

Debido a que la irradiancia solar semidiaria presenta un comportamiento no estacionario, se ha realizado la transformación de dicha serie en dos series temporales estacionarias intermedias:  $K_T$  y LC. Ambas series se utilizarán como predictores en los modelos estadísticos de predicción, propuestos en la metodología experimental (Capítulo 3) y cuyos resultados, se presentará en los próximos capítulos.

El análisis de las propiedades estadísticas de ambas series temporales ( $K_T$  y LC), ha permitido comprobar que sus funciones de densidad de probabilidad muestran una alta predominancia de condiciones de cielo claro para las estaciones radiométricas estudiadas. Estos resultados permiten presuponer que los modelos de predicción se comportarán mejor bajo situaciones de cielo claro.

Por otro lado, a partir de las figuras de autocorrelación parcial de los diferentes emplazamientos analizados se ha podido comprobar que existe una alta autocorrelación parcial exclusivamente para el primer retardo temporal semidiario. De esta forma, se

puede afirmar que, al ensayar los modelos de predicción lineal, se obtendrán resultados *óptimos* con modelos que utilicen exclusivamente el primer retardo temporal de la serie temporal para realizar las predicciones.

El estudio de la linealidad de las series temporales  $K_T$  y  $LC$  nos ha permitido identificar que todas las series temporales tienen un comportamiento no-lineal, a excepción de la serie temporal  $K_T$  en la estación de Lérída.

Estos resultados nos permiten conocer de antemano qué modelos, en función de sus propiedades lineales, ofrecerán mejores resultados en la predicción de la irradiancia solar semidiaria y en función de la linealidad de las series temporales  $K_T$  y  $LC$ .

El paso lógico siguiente es el ajuste y validación de los modelos estadísticos de predicción propuestos en el Capítulo 3. El horizonte temporal de predicción es de seis semidías y cada modelo se ensaya utilizando un vector de entrada, de tamaño variable, formado por los valores de  $K_T$  y  $LC$  observados. A pesar de que, se ha realizado un estudio previo de la linealidad de las series temporales estacionarias, se ensayarán todos los modelos estadísticos de predicción para contrastar las hipótesis planteadas en el test de linealidad. De esta forma, se comprobará qué modelos, en función de su linealidad, tienen una mayor adecuación a las propiedades de las series temporales  $K_T$  y  $LC$ .

## 5 ANÁLISIS Y VALIDACIÓN DE LOS RESULTADOS

La aplicación de la metodología de preprocesado ha permitido obtener dos series temporales estacionarias ( $K_T$  y LC) a partir de la serie temporal de irradiancia solar semidiaria. Los valores de  $K_T$  y LC se utilizarán como predictores para realizar el ajuste y validación de los resultados de predicción.

Los modelos de predicción que se han utilizado para realizar la predicción son los siguientes (ver capítulo 3):

- Modelos autorregresivos. Tienen como principal característica la generación de predicciones que conservan las propiedades lineales de las series temporales.
- Modelos basados en redes neuronales. Están basados en los sistemas nerviosos biológicos y se caracterizan por realizar un procesamiento no-lineal.
- Modelos basados en redes neuronales y lógica difusa. El otro tipo de modelos no-lineales ensayados, se basan en los modelos de redes neuronales y lógica difusa. Estos modelos tratan de mimetizar el conocimiento humano y los procesos de razonamiento.

El ajuste y ensayo de los modelos de predicción permite obtener los valores futuros de  $K_T$  y LC, a partir de cuales la obtención de los valores predichos de irradiancia solar semidiaria requiere un post-procesado, consistente en la aplicación de las transformaciones inversas a las aplicadas en la fase de preprocesado (ver sección 3.5.2).

Una vez obtenidas las predicciones de la irradiancia solar semidiaria, se realizará la validación de los diferentes modelos estadísticos ensayados. Los estadísticos de error que se utilizan para realizar la validación de los modelos son MBD(%), RMSD(%) y MSP(%) presentados en la sección 3.6.

Finalmente, se realiza la selección de los mejores modelos predictivos para cada horizonte de predicción y cada estación radiométrica, teniendo en cuenta la sencillez de

los modelos y el menor RMSD(%) que presenten dichos modelos. El MBD(%) no se utiliza como criterio de selección, ya que dicho estadístico de error presenta, como se verá en este capítulo, un valor próximo a cero para la mayoría de modelos ensayados.

### **5.1 Validación de modelos autorregresivos**

El modelo autorregresivo (AR) presentado en la sección 3.5.1 utiliza la serie temporal índice de claridad ( $K_T$ ) y componente solar atenuada (LC) como entrada, utilizando como vector de entrada un número de semidías variable que va desde 1 hasta 10 semidías previos. Así, para cada serie temporal  $K_T$  o LC perteneciente a una estación radiométrica se ensayarán 10 modelos autorregresivos de orden variable. Los diferentes modelos autorregresivos se expresan como  $AR(X)$ , siendo  $X$  el orden del modelo o valores previos que se utilizan para realizar la predicción. La estimación de los parámetros de estos modelos  $K_T AR(X)$  y  $LC AR(X)$  se realiza aplicando un procedimiento de máxima similitud [Harvey y Phipips, 1979].

La obtención de valores con horizontes de predicción mayores a un semidía, se realiza utilizando un esquema de predicción recurrente; así la salida del modelo del paso anterior se utiliza como vector de entrada del modelo para predecir el siguiente valor temporal.

Los datos disponibles para cada estación radiométrica se dividen en un conjunto de datos de entrenamiento y un conjunto de validación, ambos consecutivos e independientes. El primero corresponde con el 75% de los primeros datos de la serie temporal y se utiliza para la identificación y estimación de los parámetros de los modelos. El segundo bloque comprende el 25% restante y se utiliza para realizar el proceso de validación. En total se ensayan 80 modelos autorregresivos (20 para cada estación radiométrica) y, en cada una de ellas, 10 para cada serie temporal de entrada ( $K_T$  y LC).

Una vez obtenidas las predicciones de  $K_T$  y LC, se aplica el post-procesado descrito en la sección 3.5.2 sobre dichas series para obtener las predicciones de irradiancia solar semidiaria. El postprocesado se basa en la aplicación de las ecuaciones determinísticas que describen la geometría del sistema del sistema Sol-Tierra (ver APÉNDICE A), según:

- En el caso de las predicciones de índice de claridad semidiario, para el paso temporal  $t_{k+i}$  ( $\hat{K}_T(t_{k+i} | t_k)$ ), la irradiancia solar predicha ( $\hat{G}(t_{k+i} | t_k)$ ) se puede obtener utilizando la siguiente expresión:

$$\hat{G}(t_{k+i} | t_k) = \hat{K}_T(t_{k+i} | t_k) I_0(t_{k+i}) \quad (7.1)$$

donde  $I_0(t_{k+i})$  es la irradiancia extraterrestre.

- A partir de la serie temporal LC, para el paso temporal  $t_{k+i}$  ( $\hat{K}_T(t_{k+i} | t_k)$ ), la irradiancia solar predicha ( $\hat{G}(t_{k+i} | t_k)$ ) se puede obtener mediante la siguiente expresión:

$$\hat{G}(t_{k+i} | t_k) = L\hat{C}(t_{k+i} | t_k) I_0(t_{k+i}) \quad (7.2)$$

Una vez obtenidas las predicciones de irradiancia solar semidiaria a partir del postprocesado el último paso que se realiza para los modelos autorregresivos es su validación.

La Tabla 5.1 muestra para cada estación radiométrica los resultados de la validación. Debido al elevado número de modelos autorregresivos ensayados, se muestran únicamente, para cada horizonte temporal de predicción, el modelo que presenta menores errores en términos de RMSD. La fuente de texto elegida para destacar los resultados de dichos modelos para un cierto horizonte temporal de predicción en cursiva y subrayado. Adicionalmente, en el formato de texto por defecto se incluye el RMSD(%) para el resto de horizontes de cada modelo seleccionado, aunque dichos

resultados no sean los mejores. En los resultados que aquí se presentan, se omite mostrar los errores en términos de MBD relativo ya que su magnitud es baja, inferior 2% en términos de MBD, y por lo tanto, dicho estadístico de error no se utiliza como criterio de selección de los modelos, como se mencionó anteriormente.

Tabla 5.1 RMSD en términos porcentuales modelos autorregresivos con menor error para cada estación radiométrica

Estación	Modelos	Horizonte Predicción (Semidías)					
		1	2	3	4	5	6
Murcia	K <sub>T</sub> AR(1)	20.99	23.84	24.84	<b><u>25.31</u></b>	<b><u>25.57</u></b>	<b><u>25.71</u></b>
	K <sub>T</sub> AR(6)	20.71	23.90	25.15	25.76	26.08	26.29
	K <sub>T</sub> AR(9)	20.69	<b><u>23.67</u></b>	<b><u>24.83</u></b>	25.44	25.72	25.96
	K <sub>T</sub> AR(10)	<b><u>20.65</u></b>	23.78	24.97	25.59	25.87	26.12
	LC AR(1)	21.55	24.77	26.09	26.89	27.47	27.86
	LC AR(2)	21.40	24.91	26.40	27.02	27.43	27.54
	LC AR(9)	20.91	24.17	25.42	26.09	26.41	26.67
	LC AR(10)	20.86	24.20	25.50	26.17	26.50	26.75
Albacete	K <sub>T</sub> AR(1)	27.94	28.44	31.67	31.30	32.31	32.23
	K <sub>T</sub> AR(2)	27.26	28.33	33.58	31.43	35.05	32.29
	K <sub>T</sub> AR(3)	27.63	28.26	33.23	31.55	34.80	32.40
	K <sub>T</sub> AR(4)	26.99	28.15	33.36	31.55	34.92	32.40
	K <sub>T</sub> AR(5)	27.12	28.06	32.85	31.34	34.37	32.23
	K <sub>T</sub> AR(6)	26.93	28.03	32.86	31.29	34.39	32.20
	K <sub>T</sub> AR(7)	27.04	28.02	32.76	31.28	34.31	32.20
	K <sub>T</sub> AR(8)	26.55	28.27	33.11	31.62	34.71	32.58
	K <sub>T</sub> AR(9)	26.95	28.27	32.81	31.56	34.36	32.47
	K <sub>T</sub> AR(10)	26.66	28.26	32.93	31.53	34.45	32.39
	LC AR(1)	28.24	<b><u>27.91</u></b>	<b><u>31.44</u></b>	<b><u>30.27</u></b>	<b><u>31.56</u></b>	<b><u>31.06</u></b>
	LC AR(8)	<b><u>26.54</u></b>	28.44	33.60	31.93	35.32	33.00
Madrid	K <sub>T</sub> AR(9)	23.16	26.82	28.32	29.24	29.75	<b><u>30.10</u></b>
	K <sub>T</sub> AR(10)	<b><u>23.11</u></b>	26.87	28.39	29.33	29.85	30.19
	LC AR(9)	23.21	<b><u>26.81</u></b>	<b><u>28.29</u></b>	<b><u>29.21</u></b>	<b><u>29.73</u></b>	30.12
Lérida	K <sub>T</sub> AR(1)	23.51	<b><u>27.27</u></b>	<b><u>28.72</u></b>	<b><u>29.14</u></b>	<b><u>29.43</u></b>	<b><u>29.69</u></b>
	K <sub>T</sub> AR(4)	<b><u>23.45</u></b>	27.98	30.35	30.79	30.94	31.19

A partir de los resultados anteriores, se pueden extraer la siguiente información para cada estación radiométrica:

- Murcia. En el caso de la predicción a partir de la serie temporal  $K_T$  semidiaria los mejores modelos, en función al error cuadrático medio, son el modelo autorregresivo  $K_T$  de orden 10 para el primer paso temporal, el modelo  $K_T$  de orden 9 para el segundo, tercer paso temporal y el modelo  $K_T$  de orden 1 para el resto de pasos temporales. Los modelos que presentan peores resultados son: el modelo  $K_T$  de orden 1 desde el primer y segundo horizonte temporal de predicción y el modelo  $K_T$  de orden 6 para los horizontes quinto y sexto. Sin embargo, la diferencia de error, para un determinado horizonte de predicción, entre los diversos modelos  $K_T$  es muy pequeña, inferior al 0.3% en términos de RMSD.

La predicción a partir de la serie temporal LC semidiaria presenta peores resultados. Al igual que en el caso anterior el mejor modelo para el primer horizonte temporal de predicción es el modelo LC de orden 10 y para el resto de horizontes predictivos es el modelo LC de orden 9. Los modelos que peores resultados ofrecen son los modelos LC de orden 1 y 2. La diferencia entre los diversos modelos LC ensayados es mayor que en el caso del índice de claridad.

- Albacete. Al analizar los modelos de predicción a partir  $K_T$ , no existe ningún modelo que sobresalga sobre el resto como ocurre en el caso de la estación de Murcia.

La predicción a partir de la serie temporal LC presenta mejores resultados que los modelos basados en  $K_T$ . Para el primer horizonte de predicción, el mejor modelo es el modelo autorregresivo LC de orden 8 y para el resto de horizontes de predicción el modelo LC de orden 1.

- Madrid. La estación de Madrid es la estación con mejor calidad en el registro de datos piranométricos. Para el primer horizonte de predicción, el mejor modelo es el modelo  $K_T$  de orden 10 y para el resto de horizontes temporales, el modelo con mejores resultados es el modelo  $K_T$ /LC de orden 9.
- Lérida. En esta estación, los modelos que menor incertidumbre presentan son los modelos aplicados a la serie temporal  $K_T$  semidiaria; y en concreto, el modelo  $K_T$  de orden 4 para el primer horizonte de predicción y el modelo  $K_T$  de orden 1 para el resto de horizontes temporales de predicción.

La información general que se puede destacar a partir de la tabla anterior es la siguiente:

- A partir del análisis de los resultados para todas las estaciones radiométricas, se puede observar que a medida que aumenta el horizonte temporal de predicción se va reduciendo la relación lineal entre el valor a predecir y el tamaño del vector de entrada de valores observados. Se comprueba que para el primer paso temporal el orden de los modelos es mayor y éste se va reduciendo conforme va aumentando el horizonte temporal de las predicciones.
- En general, los modelos autorregresivos estimados a partir de la serie temporal índice de claridad ofrecen mejores resultados; a partir de estos resultados y a los de linealidad presentados en el capítulo anterior, se puede afirmar que la serie temporal  $K_T$  tiene una menor no-linealidad que la serie temporal LC.
- La estación que presenta menor incertidumbre es la estación de Murcia, que es la situada más al sur, la situada a una menor elevación y la más próxima a la costa.
- La estación de Albacete presenta una mayor complejidad en su comportamiento dinámico, muy probablemente, debido a la baja continuidad temporal de la serie temporal medida. De esta forma, se pone de manifiesto que es de vital importancia disponer de datos de irradiancia solar (ya sean datos radiométricos o estimaciones a partir de satélite) de calidad y continuidad suficiente para realizar el entrenamiento de los modelos estadísticos.
- Las estaciones de Madrid y Lérida presentan un orden de magnitud del error similar en términos de RMSD relativo.
- Los modelos aplicados a las estaciones de Murcia y Madrid muestran similares resultados. La diferencia entre los diversos modelos ensayados para ambas estaciones es menor a 0.3% del error en términos de RMSD para un mismo horizonte de predicción. Por lo tanto, aumentar la complejidad de los modelos en términos del tamaño del vector de entrada no supone una reducción significativa en los errores de predicción en los términos del análisis realizado.
- Para la estación de Albacete, los errores son bastantes elevados y no existe un aumento gradual del mismo, en forma logarítmica, al aumentar el horizonte de predicción como ocurre en el resto de estaciones radiométricas. Al analizar detenidamente los resultados, se comprueba que no existe un aumento del error

de predicción entre el primer y segundo horizonte temporal de predicción, como sucede en el resto de estaciones. En Albacete, este incremento está en torno al 1.5%, en términos de RMSD, y en el resto de estaciones está en torno al 3%. El *gran salto* en el error para la estación de Albacete se produce en el paso del segundo al tercer horizonte temporal de predicción, en torno al 3.5% en términos de RMSD. A partir del tercer horizonte temporal de predicción se produce una fluctuación en torno al valor medio de error 30.5% en términos de RMSD relativo.

- Para las estaciones de Murcia, Madrid y Lérida, los errores de predicción aumentan significativamente a partir del segundo horizonte de predicción y se estabilizan en un valor máximo que está en torno al 4-6%, mayor que el error del primer horizonte temporal de predicción en términos de RMSD porcentual.

## 5.2 Validación de modelos basados en redes neuronales

Los modelos basados en las redes neuronales *backpropagation* que se han ensayado utilizan las series temporales  $K_T$  y LC para realizar las predicciones. El horizonte temporal de predicción es de seis semidías. Las diferentes configuraciones de los modelos de redes neuronales ensayadas son las siguientes:

- Modelo 1 (NN1):
  - Número de Capas: 1.
  - Función de activación: sigmoide (SG).
- Modelo 2 (NN2):
  - Número de Capas: 2.
  - Funciones de activación: tangente sigmoide (TG) y SG.
  - Número de neuronas por capa: 3 y 1.
- Modelo 3 (NN3):
  - Número de Capas: 3.
  - Funciones de activación: TG, TG, TG y SG.
  - Número de neuronas por capa: 5, 3 y 1.
- Modelo 4 (NN4):
  - Número de Capas: 4.

- Funciones de activación: TG, TG, TG, TG y SG.
- Número de neuronas por capa: 8, 5, 3 y 1.
- Modelo 5 (NN5):
  - Número de Capas: 5.
  - Funciones de activación: TG, TG, TG, TG y SG.
  - Número de neuronas por capa: 15, 8, 5, 3 y 1.

Para cada configuración, se utilizará un vector de entrada con tamaño variable. El modelo de RNA empleado se representa como  $NNI(X)$ , donde  $I$  es la configuración del modelo RNA ( $I=1..5$ ) y  $X$  es la longitud del vector de datos de entrada utilizado para predecir los valores futuros, considerando el valor actual y los  $X-1$  valores anteriores. El tamaño del vector de entrada que se ensaya tiene un tamaño variable desde 1 hasta 10.

El software utilizado es la *Toolbox de Matlab Neural Network 6* [Demuth et al., 2010]. La implementación de los algoritmos en la *Toolbox* emplea una única red neuronal para cada configuración, que dispone de seis salidas correspondientes a los seis semidías de predicción (horizontes temporales de predicción). Este tipo de esquema de predicción es similar a la predicción multipaso (ver sección 3.5.2) aunque más simple ya que utiliza una única red neuronal para todos los horizontes temporales de predicción en lugar de usar un modelo independiente para cada horizonte temporal de predicción.

Al igual que en el caso de los modelos autorregresivos, a partir del conjunto de datos disponible para cada estación radiométrica, se utiliza una muestra correspondiente al 75% de los primeros datos de la serie temporal de cada estación radiométrica para realizar la identificación y estimación de los modelos (conjunto de entrenamiento) y una muestra correspondiente al 25% restante, independiente del anterior, para la validación del ajuste del modelo estimado (conjunto de validación). En cada estación radiométrica se han ensayado 100 modelos de redes neuronales diferentes; 50 por cada serie temporal de entrada ( $K_T$  y LC).

Una vez obtenidas las predicciones de  $K_T$  y LC, se aplica el post-procesado descrito anteriormente (ver 3.5.2) sobre dichas series para obtener las predicciones de irradiancia solar semidiario.

El último paso que se realiza es la validación de las predicciones de irradiancia solar semidiaria, para las diferentes configuraciones de modelos basados en redes neuronales que se han ensayado. A continuación, se presentan en forma de tablas los resultados de la validación de las predicciones de irradiancia solar semidiaria en términos de RMSD relativo para cada estación radiométrica de forma independiente. En cada estación y para cada configuración (NNI), que se ha ensayado a partir de los vectores de entrada de tamaño variable desde 1 hasta 10, se muestra únicamente aquella que tiene los menores errores de predicción. Los modelos con menor RMSD relativo, para cada horizonte temporal de predicción, se señalan en color rojo y subrayado.

La siguiente tabla presenta para cada estación radiométrica los resultados de la validación. Debido al elevado número de modelos basado en redes neuronales ensayados, se muestran únicamente para cada horizonte temporal de predicción, el modelo que presenta menores errores en términos de RMSD. La fuente de texto elegida para destacar los resultados de dichos modelos para un cierto horizonte temporal de predicción en cursiva y subrayado. Adicionalmente, en el formato de texto por defecto se incluye el RMSD(%) para el resto de horizontes de cada modelo seleccionado.

Tabla 5.2 RMSD(%) de los modelos basados en redes neuronales

Estación	Modelos	Horizonte Predicción (Semidías)					
		1	2	3	4	5	6
Murcia	$K_T$ NN3(8)	<b><u>20.58</u></b>	23.62	24.66	25.30	25.54	25.70
	LC NN2(10)	21.34	<b><u>23.38</u></b>	24.16	24.55	24.76	24.90
	LC NN4(9)	21.55	23.39	<b><u>24.05</u></b>	<b><u>24.44</u></b>	<b><u>24.62</u></b>	<b><u>24.68</u></b>
Albacete	LC NN1(6)	26.55	<b><u>27.78</u></b>	<b><u>30.24</u></b>	<b><u>29.80</u></b>	<b><u>31.25</u></b>	30.39
	LC NN2(8)	26.81	27.92	30.58	29.91	31.52	<b><u>30.38</u></b>
	LC NN3(2)	<b><u>26.37</u></b>	28.35	30.60	30.26	31.45	30.63
Madrid	LC NN3(3)	<b><u>23.35</u></b>	26.40	27.42	28.06	28.44	28.77
	LC NN4(9)	23.59	<b><u>26.07</u></b>	<b><u>27.08</u></b>	<b><u>27.68</u></b>	<b><u>28.06</u></b>	<b><u>28.31</u></b>
Lérida	$K_T$ NN1(9)	<b><u>23.65</u></b>	27.09	<b><u>28.54</u></b>	<b><u>28.88</u></b>	<b><u>29.12</u></b>	29.33
	$K_T$ NN2(9)	23.94	<b><u>26.98</u></b>	28.58	28.91	29.15	<b><u>29.31</u></b>

A partir de la tabla anterior, se puede extraer la siguiente información:

- Murcia. Para todos los horizontes temporales de predicción, exceptuando el primero, los mejores modelos son los modelos basados en la serie temporal LC. El modelo de red neuronal LC NN2(10) obtiene mejores resultados para los horizontes de predicción segundo. Para los horizontes de predicción tercero a sexto, el mejor modelo es el modelo LC NN4(9).
- Albacete. En esta estación, los modelos basados en LC son los que tienen menor incertidumbre. El mejor modelo para el primer horizonte de predicción es el modelo de red neuronal LC NN3(2). Para los horizontes segundo hasta el cuarto, el modelo con menor incertidumbre es el modelo LC NN1(6) y, para el último horizonte de predicción, el modelo LC NN2(8).
- Madrid. Al igual que sucede en la estación de Albacete, los modelos con mejores resultados son los modelos basados en la componente solar atenuada. Para el primer horizonte de predicción, el modelo con menor error es el modelo LC NN3(3) y para el resto de horizontes de predicción, el modelo óptimo es el modelo LC NN4(9).
- Lérida. En esta estación, los modelos que mejores resultados presentan son los modelos basados en el índice de claridad. Para los horizontes de predicción primero, tercero, cuarto y quinto el mejor modelo es el de red neuronal  $K_T$  NN1(9). Para los horizontes de predicción segundo y sexto, el mejor modelo es el modelo  $K_T$  NN2(9).

La información general que se puede destacar es la siguiente:

- Los modelos de redes neuronales obtienen mejores resultados al utilizar la serie temporal LC como vector de entrada, excepto en el caso de la estación de Lérida donde los modelos que presentan menor incertidumbre son los modelos basados en la serie temporal  $K_T$ .
- En el caso de la estación de Albacete, los errores en términos de RMSD relativo son mayores al del resto de estaciones. Esto es posible que sea debido a la falta de continuidad en la base de datos. Por otra parte, no existe un aumento

exponencial del mismo conforme aumenta el horizonte de predicción, como se comentó anteriormente para el caso de los modelos autorregresivos.

- Como sucedía para los modelos autorregresivos, en las estaciones de Murcia, Madrid y Lérida los errores de predicción aumentan significativamente a partir del segundo horizonte de predicción y se estabilizan en un valor máximo. Dicho error está en torno al 4-6%, mayor que el error del primer horizonte temporal de predicción en términos de RMSD porcentual.
- La precisión de los modelos basados en redes neuronales depende del tipo de condiciones de cielo predominante en el emplazamiento bajo estudio; para un lugar donde las condiciones de cielo predominantes son de cielo claro (función de distribución de  $K_T$  unimodal), los errores predictivos son menores para este tipo de días debido a que los parámetros de los modelos adquieren un mayor peso. Este hecho se debe a que la función de distribución de probabilidad de  $K_T$  y LC no tiene una forma gaussiana. Las estaciones radiométricas cuyos datos de radiación solar semidiaria tienen una función de distribución bi-modal, presentan mayores errores predictivos, ya que la existencia de diferentes condiciones de tipos de cielo (cielos claros y nubosos) hace que los parámetros de la red neuronal no se especialicen en un único tipo de condición de cielo. Una posible mejora de los modelos planteados podría ser la utilización de diferentes modelos de redes neuronales (modelos multietapa) aplicados en diferentes fases, de acuerdo a las predicciones del tipo de cielos para el próximo día.

### 5.3 Validación de modelos ANFIS

En esta sección se presentan los principales resultados tras el ajuste y validación de los modelos ANFIS propuestos en la sección 3.5.1. Se han ensayado diferentes modelos ANFIS basados en la serie temporal índice de claridad ( $K_T$  ANFIS( $X$ )) y componente solar atenuada (LC ANFIS( $X$ )) donde  $X$  es el orden o valores previos que se utilizan para realizar la predicción; siendo  $X=1..6$ . El horizonte temporal de predicción es de seis semidías.

Para el desarrollo de los algoritmos de predicción basados en los modelos ANFIS se ha utilizado la *Toolbox Fuzzy Logic* de *Matlab* [The MathWorks, 2010]. Las estructuras ANFIS que se aplican sobre las series temporales  $K_T$  y  $LC$ , se definen bajo las siguientes condiciones:

- Se crean dos funciones de pertenencia para cada entrada a la estructura ANFIS.
- El tipo de función de pertenencia que se utiliza es una curva basada en *splines* (ver Figura 5.1). Esta función se utiliza como método de “difusión” de la variable de entrada dándole un nuevo valor borroso.
- El algoritmo de entrenamiento que se utiliza se basa en un método híbrido, que combina el método *backpropagation* de errores y la estimación de los parámetros mediante un método de mínimos cuadrados.
- El tipo de salida asociada a la función de pertenencia es lineal.

Los modelos ANFIS se representan como  $ANFIS(X)$ , donde  $X$  representa la longitud del vector de datos de entrada utilizado en los modelos, considerando el valor observado actual y los  $X-1$  valores pasados.

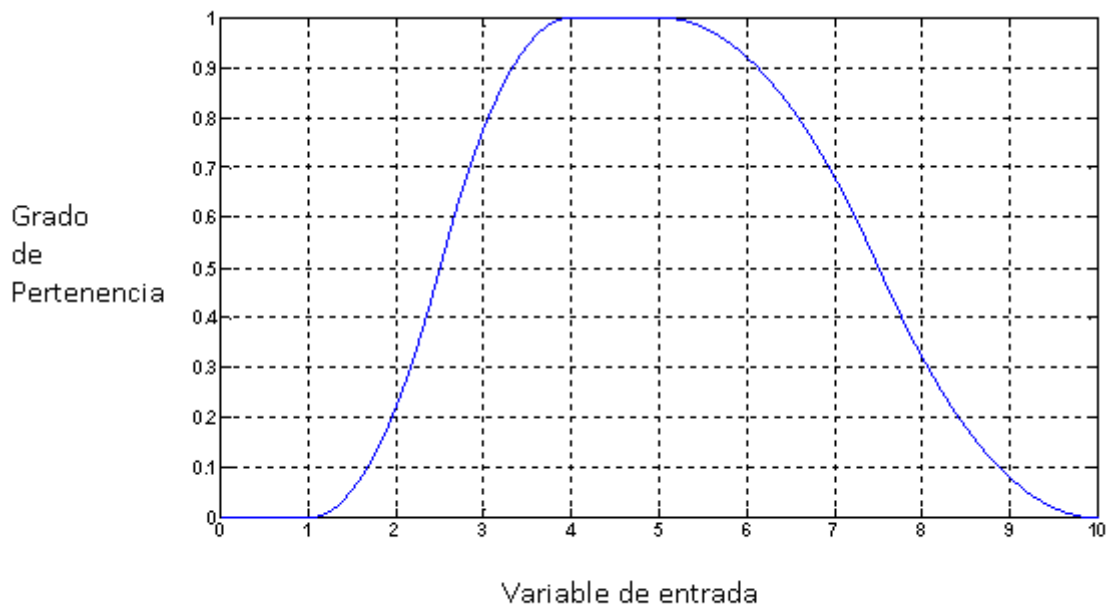


Figura 5.1 Función de pertenencia con forma de número  $\pi$  empleada en la estructura ANFIS

Para cada estación radiométrica se han ensayado 12 modelos ANFIS, 6 por cada serie temporal de entrada ( $K_T$  y LC). Se debe destacar que no se ha podido aumentar el tamaño del vector de entrada hasta 10 semidía previos como en el resto de modelos debido a los elevados requisitos computacionales de los modelos ANFIS, y a que, los errores que se obtienen con órdenes mayores a seis del vector de entrada están por encima del 100% en términos de RMSD relativo.

Al igual que se ha realizado en los ensayos anteriores, a partir del conjunto de datos disponible para cada estación radiométrica, se utiliza un porcentaje correspondiente al 75% para la identificación y estimación de los parámetros de los modelos y el 25% restante consecutivo para la validación del ajuste del modelo estimado. Las predicciones de la radiación solar semidiaria para un horizonte temporal mayor a un semidía se obtienen utilizando un esquema de predicción recurrente (ver sección 3.5.2).

Una vez obtenidas las predicciones de  $K_T$  y LC, se aplica el post-procesado, descrito anteriormente (ver sección 3.5.2), para obtener las predicciones de irradiancia solar semidiaria para los diferentes horizontes temporales de predicción. El último paso realizado es la validación de los resultados para los diferentes modelos ANFIS ensayados. En la siguiente tabla se presentan los resultados de validación de los modelos con menores errores en cada horizonte temporal de predicción para cada estación radiométrica de forma independiente en términos de RMSD(%). Dichos resultados se destacan en cursiva y subrayado.

Los resultados de la Tabla 5.3 permiten destacar la siguiente información:

- Murcia. El modelo ANFIS óptimo es el modelo LC ANFIS (4), para todos los horizontes temporales de predicción, excepto en el segundo horizonte, donde el modelo que presenta una menor incertidumbre es el modelo  $K_T$  ANFIS (5).
- Albacete. Los modelos que menor error predictivo presentan en esta estación son los modelos basados en la serie temporal componente solar atenuada, en concreto, el modelo LC ANFIS (2) para los horizontes predictivos uno, tres y cinco y el modelo LC ANFIS (1) para los horizontes predictivos dos, cuatro y cinco.

- Madrid. Para el primer horizonte temporal de predicción se selecciona el modelo  $K_T$  ANFIS (4) y para el resto de horizontes temporales el modelo LC ANFIS (4).
- Lérida. En esta estación, los modelos ANFIS basados en  $K_T$  presentan una menor incertidumbre. Para los cuatro primeros horizontes temporales de predicción se selecciona el modelo  $K_T$  ANFIS (3) y para el resto de horizontes el modelo  $K_T$  ANFIS (2).

Tabla 5.3 RMSD (%) modelos ANFIS

Estación	Modelos	Horizonte Predicción (Semidías)					
		1	2	3	4	5	6
Murcia	$K_T$ ANFIS (3)	20.91	23.79	25.00	<b><u>25.36</u></b>	25.76	25.84
	$K_T$ ANFIS (4)	<b><u>20.86</u></b>	23.81	24.88	25.38	25.73	25.82
	$K_T$ ANFIS (5)	21.00	<b><u>23.77</u></b>	25.04	25.56	25.81	25.87
	LC ANFIS (4)	21.07	23.94	<b><u>24.83</u></b>	25.27	<b><u>25.53</u></b>	<b><u>25.69</u></b>
Albacete	LC ANFIS (1)	28.17	<b><u>27.51</u></b>	31.80	<b><u>29.77</u></b>	32.37	<b><u>30.42</u></b>
	LC ANFIS (2)	<b><u>27.36</u></b>	28.10	<b><u>30.55</u></b>	30.17	<b><u>31.37</u></b>	30.91
Madrid	$K_T$ ANFIS (4)	<b><u>23.40</u></b>	26.77	27.92	28.54	28.96	29.27
	LC ANFIS (3)	23.52	<b><u>26.71</u></b>	27.72	28.32	<b><u>28.81</u></b>	29.05
	LC ANFIS (4)	23.50	<b><u>26.71</u></b>	<b><u>27.62</u></b>	<b><u>28.27</u></b>	<b><u>28.81</u></b>	<b><u>28.96</u></b>
Lérida	$K_T$ ANFIS (2)	23.63	27.36	28.83	29.21	<b><u>29.37</u></b>	<b><u>29.66</u></b>
	$K_T$ ANFIS (3)	<b><u>23.52</u></b>	<b><u>27.22</u></b>	<b><u>28.81</u></b>	<b><u>29.21</u></b>	29.45	29.69

La siguiente tabla se presenta la diferencia en términos de RSMD (%) entre los mejores modelos NN y lo mejores modelos ANFIS para cada estación radiométrica y cada horizonte de predicción.

Tabla 5.4 Diferencia del RMSD (%) entre los mejores modelos NN y ANFIS

Estación	Horizonte Predicción (Semidías)					
	1	2	3	4	5	6
<b>Murcia</b>	1.3	1.6	3.1	3.6	3.6	3.9
<b>Albacete</b>	3.6	-1.0	1.0	-0.1	0.4	0.1
<b>Madrid</b>	0.2	2.4	2.0	2.1	2.6	2.2
<b>Lérida</b>	-0.6	0.9	0.9	1.1	0.9	1.2

La tabla anterior permite señalar la información general:

- Los modelos ANFIS generan mayores errores RMSD que los modelos lineales para el mejor modelo en cada uno de los horizontes de predicción, si bien las diferencias no son significativas (menor al 4% en términos de RMSD). En el caso de Lérida esta diferencia está en torno al 1% para todos los horizontes de predicción. Lo mismo que para Albacete en el caso de 1 semidía de predicción. Por el contrario, para Madrid para un semidía de predicción no hay diferencia apreciable, pero para horizontes mayores aumenta y se mantiene en alrededor del 2% para el resto. Murcia presenta las mayores diferencias con una tendencia creciente según aumenta el horizonte de predicción.
- No existe una clara definición sobre la mejor forma de predicción para los modelos ANFIS, ya que para algunas estaciones, Lérida, el mejor modelo es el que realiza la predicción a partir de la serie  $K_T$ , mientras que en otras, Albacete y Madrid, la mejor serie de partida para realizar la predicción es LC, en tanto que para Murcia no hay una serie de partida claramente definida.
- La serie temporal índice de claridad presenta una dinámica mucho más sencilla que la componente solar atenuada en la estación de Lérida, por lo tanto, los modelos ANFIS  $K_T$  tienen una menor incertidumbre.
- Al igual que ocurre con los modelos anteriores, la estación que presenta menor error es la estación de Murcia.
- Los errores de los modelos LC ANFIS (5) y LC ANFIS (6) son varios órdenes de magnitud mayores que los del resto de modelos ANFIS ensayados.

#### 5.4 Análisis de modelos óptimos para cada estación radiométrica

A partir de los resultados anteriores y con la intención de comparar los diferentes tipos de modelos (autorregresivo, RNA y ANFIS), en la Tabla 5.5 se presenta para cada tipo de modelo y cada horizonte temporal de predicción los que presentan menores errores en términos de RMSD(%) para los diferentes emplazamientos. Se destaca en cursiva y subrayado los modelos que presentan menor error predictivo en términos de RMSD para cada horizonte temporal de predicción. Además, se incluyen los errores del modelo basado en la persistencia en términos de RMSD(%).

Tabla 5.5 Menores errores en términos de RMSD relativo para cada estación radiométrica y cada tipo de modelo

Estación	Modelos	Horizonte Predicción (Semidías)					
		1	2	3	4	5	6
Murcia	AR	20.65	23.67	24.83	25.31	25.57	25.71
	NN	<b><u>20.58</u></b>	<b><u>23.38</u></b>	<b><u>24.05</u></b>	<b><u>24.44</u></b>	<b><u>24.62</u></b>	<b><u>24.68</u></b>
	ANFIS	20.86	23.77	24.83	25.27	25.53	25.69
	PERSISTENCIA	25.22	29.36	31.94	32.64	33.47	33.46
Albacete	AR	26.54	27.91	31.44	30.27	31.56	31.06
	NN	<b><u>26.37</u></b>	27.78	<b><u>30.24</u></b>	29.80	<b><u>31.25</u></b>	<b><u>30.38</u></b>
	ANFIS	27.36	<b><u>27.51</u></b>	30.55	<b><u>29.77</u></b>	31.37	30.42
	PERSISTENCIA	27.90	32.06	35.91	36.30	38.39	38.19
Madrid	AR	<b><u>23.11</u></b>	26.81	28.29	29.21	29.73	30.10
	NN	23.35	<b><u>26.07</u></b>	<b><u>27.08</u></b>	<b><u>27.68</u></b>	<b><u>28.06</u></b>	<b><u>28.31</u></b>
	ANFIS	23.40	26.71	27.62	28.27	28.81	28.96
	PERSISTENCIA	26.17	31.05	33.43	34.75	35.79	36.20
Lérida	AR	<b><u>23.45</u></b>	27.27	28.72	29.14	29.43	29.69
	NN	23.65	<b><u>26.98</u></b>	<b><u>28.54</u></b>	<b><u>28.88</u></b>	<b><u>29.12</u></b>	<b><u>29.31</u></b>
	ANFIS	23.52	27.22	28.81	29.21	29.37	29.66
	PERSISTENCIA	27.45	33.78	36.50	37.45	38.24	38.99

A partir de los resultados de la Tabla 5.5 se puede destacar que, en líneas generales, los modelos ensayados que presentan menores errores son los modelos no-lineales basados en redes neuronales, que utilizan la serie temporal LC semidiaria como entrada excepto para la estación de Lérida. La estación de Lérida supone una excepción, debido a que los modelos NN  $K_T$  presentan una menor incertidumbre por la dinámica lineal de la serie temporal  $K_T$ .

En la Figura 5.2, Figura 5.3, Figura 5.4 y Figura 5.5 se muestra el grado de mejora que se obtiene con cada tipo de modelo (autorregresivo, RNA y ANFIS), con respecto al modelo basado en la persistencia en términos de RMSD relativo para cada estación radiométrica.

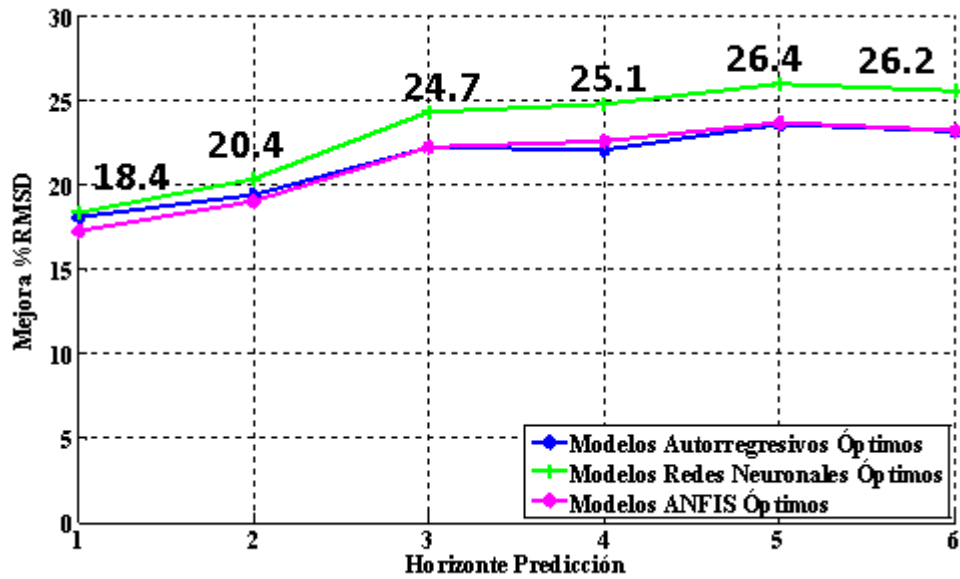


Figura 5.2 Murcia: mejora de los modelos ensayados frente al modelo de persistencia

Los modelos que menor incertidumbre presentan en la estación de Murcia son los modelos de redes neuronales basados en  $K_T$  para el primer horizonte temporal de predicción y en LC para el resto de horizontes. Los modelos autorregresivos tienen un buen desempeño siendo mejores que los modelos ANFIS para los tres primeros horizontes de predicción. En el caso de los horizontes de predicción cuarto, quinto y sexto, los modelos ANFIS tienen menor incertidumbre que los modelos autorregresivos. Los modelos de redes neuronales que mejor comportamiento presentan son los que utilizan entre ocho y diez semidías previos para generar las predicciones. El valor de MSP(%) está por encima del 18% para todos los horizontes de predicción, y su valor medio es de 23.5% para el conjunto del horizonte de predicción.

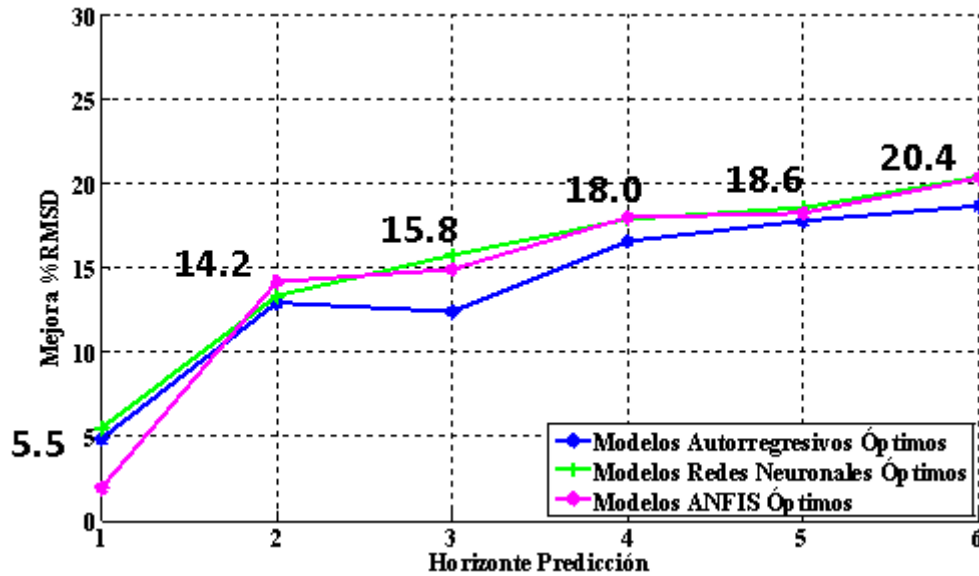


Figura 5.3 Albacete: mejora de los modelos ensayados frente al modelo de persistencia

En la estación de Albacete, para el primer horizonte temporal de predicción, el mejor modelo es el modelo de red neuronal, mientras que, para el resto de horizontes existe una alternancia entre los modelos de redes neuronales y los modelos ANFIS. Los modelos autorregresivos son los que peor comportamiento presentan para todos los horizontes temporales de predicción, esto se debe a que, la baja continuidad de las series temporales en esta estación hace que su dinámica sea mucho más compleja, debido a la mayor no-linealidad que presenta la serie temporal de irradiancia solar en dicha estación. El predictor con el que mejores resultados se obtiene es la serie temporal LC. Otro hecho destacable es que, los modelos ANFIS ensayados son mucho más sencillos que los modelos basados en redes neuronales. Los primeros utilizan entre uno y dos días previos y los segundos obtienen mejores resultados al utilizar entre seis y ocho valores previos. La mejora de los modelos óptimos ensayados con respecto al modelo basado en la persistencia está por encima del 5%, en términos de RMSD relativo, para todos los horizontes de predicción, con un valor medio de 15.4% para el conjunto de los horizontes.

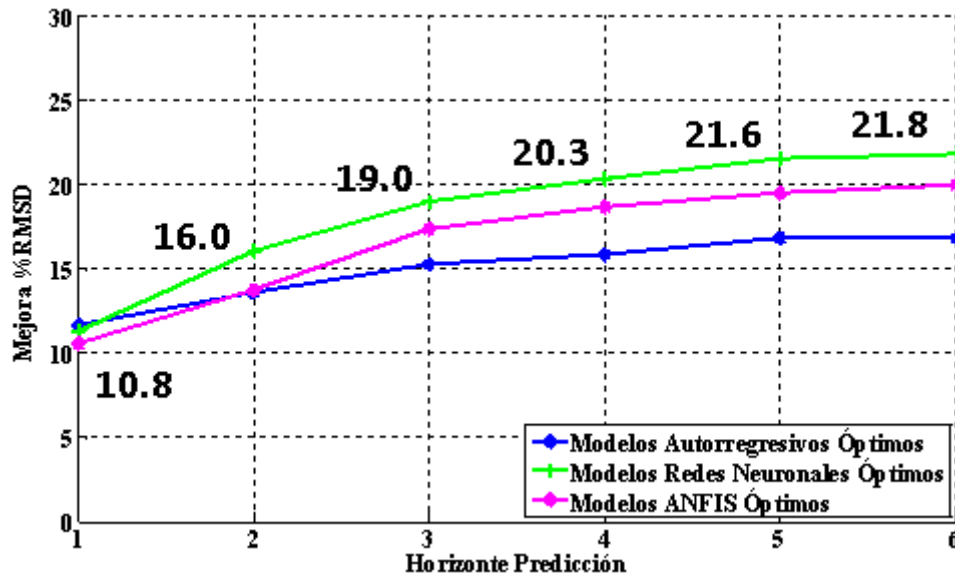


Figura 5.4 Madrid: mejora de los modelos ensayados frente al modelo de persistencia

En la estación de Madrid, para el primer horizonte temporal de predicción el modelo que presenta menor incertidumbre es el modelo autorregresivo, estimado a partir de la serie temporal LC. Esto se debe, a la autocorrelación lineal de dicha serie para el primer retardo temporal. Para el resto de horizontes temporales de predicción, los modelos lineales son los que presentan mayores errores. El modelo autorregresivo, que presenta mejores resultados para el primer horizonte de predicción, utiliza diez semidías para realizar la predicción, mientras que los modelos basados en redes neuronales utilizan nueve valores previos. La mejora de los modelos ensayados con respecto al modelo basado en la persistencia está por encima del 10%, en términos de RMSD relativo, para todos los horizontes de predicción, con valor medio del 18% para el conjunto del horizonte de predicción.

El predictor seleccionado en la estación de Lérida es el  $K_T$ . Para el primer horizonte temporal de predicción, el modelo con menor incertidumbre es el modelo autorregresivo  $K_T$  de orden cuatro, mientras que para el resto de horizontes el mejor modelo es el modelo de red neuronal con una única capa de orden nueve. Debido a que el índice de claridad presenta una dinámica lineal, la predicción de su evolución puede realizarse a partir de un modelo de red neuronal de una única capa, que realiza una regresión lineal entre la entrada y salida. Los modelos basados en redes neuronales son capaces de

extraer mayor información a partir de los patrones entre entrada y salida, y de esta forma, obtienen mejores resultados que los modelos autorregresivos. En esta estación, los errores cuadráticos medios, que presentan los diferentes tipos de modelos óptimos ensayados, son muy similares, variando entre ellos un 0.5% en términos de RMSD como valor máximo. La mejora de los modelos óptimos ensayados, con respecto al modelo basado en la persistencia, está por encima del 15%, en términos del RMSD, para todos los horizontes de predicción. El valor medio para el conjunto del horizonte de predicción es del 12%.

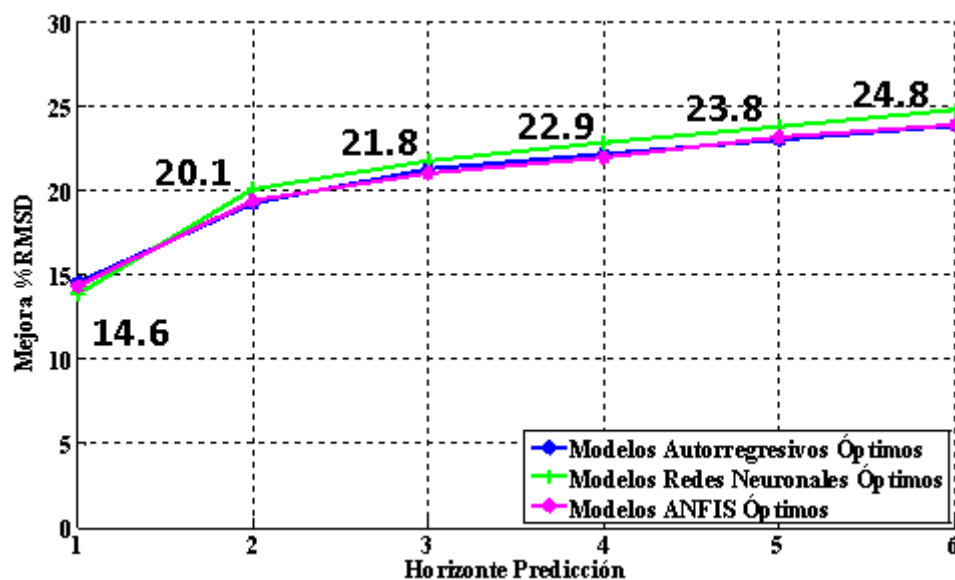


Figura 5.5 Lérica: mejora de los modelos ensayados frente al modelo de persistencia

Como resumen del análisis podemos señalar que la mejora porcentual en el error cuadrático medio (RMSD) para el conjunto de predicciones está en torno al 20%, lo que indica que la aplicación de sistemas de predicción como redes neuronales o modelos ANFIS supone una clara mejora en la utilización de modelos respecto a los métodos de persistencia.

## 5.5 Conclusiones

La bondad de los diferentes modelos predictivos ensayados depende del tipo de situaciones de cielo predominantes en el emplazamiento bajo estudio. Para un lugar

donde las condiciones de cielo predominantes son de cielo claro (función de distribución unimodal), los modelos presentan menores errores debido a que los parámetros de los modelos adquieren un mayor peso. Por otro lado, las estaciones radiométricas con una mayor heterogeneidad de tipos de estado de cielo (claro/nuboso) cuya función de distribución es bi-modal, presentan mayores errores debido a que los parámetros de la red neuronal no se especializan en un único tipo de situación de cielo para realizar las predicciones y a que los modelos tienen una menor probabilidad de acertar ya que pueden presentarse una mayor variabilidad atmosférica.

Las funciones de densidad de probabilidad de las series temporales  $K_T$  y LC (ver 4.6) muestran una alta predominancia de condiciones de cielo claro para las estaciones radiométricas analizadas. Así, la estación que presenta un mayor predominio de condiciones de cielo claro es la estación de Murcia, donde el error en la predicción para cada tipo de modelo ensayado es menor.

Por otro lado, a partir de las figuras de autocorrelación parcial de los diferentes emplazamientos analizados (ver 4.6), se ha podido comprobar que existe una alta autocorrelación parcial significativa para el primer retardo temporal semidiario. Los resultados de los diferentes modelos han permitido comprobar que la disminución de la incertidumbre de éstos, en base al aumento del orden o tamaño del vector de entrada, es muy limitada. Esto se debe a que la mayor parte de la información que relaciona un valor de la serie temporal con sus predecesores está contenida en el primer retardo temporal para ambas series temporales ( $K_T$  y LC).

El estudio de la linealidad de las series temporales  $K_T$  y LC nos ha permitido identificar que todas las series temporales tienen un comportamiento no-lineal a excepción de la serie temporal  $K_T$  en la estación de Lérida. Estos resultados nos han ayudado a conocer de antemano qué modelos en función de sus propiedades lineales ofrecen mejores resultados en la predicción de la irradiancia solar semidiaria en función de la linealidad de las series temporales  $K_T$  y LC.

Como era de esperar, a partir de los resultados presentados en el capítulo 4, los modelos lineales que utilizan como predictor los valores de  $K_T$  presentan menores errores frente a los modelo autorregresivos que utilizan LC para la estación de Lérida. En general, se puede afirmar que los modelos autorregresivos estimados a partir de la serie temporal  $K_T$  ofrecen mejores resultados; a partir de la afirmación anterior y los resultados de identificación de la linealidad presentados en el capítulo anterior, se demuestra que la serie temporal  $K_T$  tiene una menor no-linealidad que la serie temporal LC.

En el caso de los modelos basados en redes neuronales y modelos ANFIS, estos proporcionan en general mejores resultados al utilizar la serie temporal LC como vector de entrada; excepto en el caso de la estación de Lérida, donde los modelos que presentan menor incertidumbre son los que utilizan como predictor  $K_T$ , debido a que la serie temporal  $K_T$  presenta una mayor linealidad y una dinámica más sencilla que la serie temporal LC localmente.

El comportamiento general de la estación de Lérida difiere del resto de estaciones radiométricas. Una de las razones puede ser debido al tipo de clima dado que la región climática donde se encuentra esta estación presenta un comportamiento más homogéneo del  $K_T$ . Este hecho confirmaría al mismo tiempo la dependencia climática de los modelos sean regresivos lineales o basados en redes neuronales.

Una conclusión importante que se debe destacar es que no existe un modelo universal para todas y cada una de las estaciones analizadas. De esta forma, es preciso realizar un análisis local, siguiendo la metodología aquí presentada, para realizar la comparación y selección del mejor modelo local en cada emplazamiento.

La comparativa, realizada en este capítulo, sirve de base para seleccionar el modelo de predicción final. El objetivo que se persigue es identificar el modelo final de predicción óptimo, que se define como aquel modelo local cuyos parámetros han sido adaptados a un emplazamiento concreto y que presente la mayor sencillez posible (menor número de

parámetros) y que muestre menores errores en la predicción para cada uno de los horizontes temporales de predicción.



## 6 SELECCIÓN DEL MODELO FINAL DE PREDICCIÓN

Este capítulo presenta el modelo final de predicción óptimo elegido para cada estación radiométrica. El criterio de selección que se ha utilizado se basa en la elección de los modelos con menor RMSD relativo para los diferentes horizontes temporales de predicción. En el caso que varios modelos tengan RMSD relativos similares para un mismo horizonte temporal de predicción, se elige aquel modelo que tenga un *menor número de parámetros*. La utilización de este criterio no se fundamenta únicamente en elegir el modelo más sencillo, sino que además se intenta evitar problemas relativos a la redundancia de parámetros [Pankratz, 1983]. Aunque dos modelos (uno de menor y otro de mayor orden) se puedan utilizar indistintamente para realizar las predicciones, la estimación de los parámetros en los modelos de mayor orden puede llevar a inestabilidades, dando lugar a cambios significativos en el ajuste de los parámetros para diferentes muestras de datos [Farnum y Stanton, 1989]. Un último criterio que se utiliza para seleccionar el modelo final se realiza de acuerdo al mínimo número de modelos para los diferentes horizontes temporales de predicción y de esta forma simplificar el proceso de predicción.

### 6.1 Análisis de los modelos finales para cada estación radiométrica

Para cada estación radiométrica, se presentan los resultados del modelo final seleccionado en forma de tabla con los errores de predicción de la irradiancia solar semidiaria para cada horizonte temporal de predicción, en términos de MBD y RMSD relativo. Se presentan además los errores del modelo basado en la persistencia. En las dos últimas filas se muestra la mejora sobre el modelo basado en la persistencia para el modelo final seleccionado y los mejores modelos basados en redes neuronales. En color rojo y subrayado, se presenta el modelo que presenta menores errores para dicho horizonte temporal de predicción. En fondo verde, se señala el modelo final de predicción seleccionado en base a los criterios definidos anteriormente (Tabla 6.1, Tabla 6.2, Tabla 6.3 y Tabla 6.4).

Tabla 6.1 Murcia: errores del modelo final y del modelo basado en la persistencia

Modelos	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
LC NN2(10)	-0.55	<b>-0.85</b>	<b>-1.04</b>	<b>-1.11</b>	<b>-1.20</b>	<b>-1.21</b>	MBD%
K <sub>T</sub> NN3(8)	<b>-0.83</b>	2.43	3.32	3.80	4.13	4.21	
PERSISTENCIA (PER)	0.00	0.01	0.02	0.02	0.03	0.03	
LC NN2(10)	21.34	<b>23.38</b>	<b>24.16</b>	<b>24.55</b>	<b>24.76</b>	<b>24.90</b>	RMSD %
K <sub>T</sub> NN3(8)	<b>20.58</b>	23.62	24.66	25.30	25.54	25.70	
PERSISTENCIA (PER)	25.22	29.36	31.94	32.64	33.47	33.46	
RED NEURONAL	18.40	20.40	24.70	25.10	26.40	26.20	MSP%
MODELO FINAL	18.39	20.38	24.35	24.79	26.02	25.58	

A partir de la información anterior se puede señalar:

- En la estación de Murcia, los modelos con menor incertidumbre están basados en redes neuronales. Para el primer horizonte temporal de predicción, el modelo final seleccionado es el modelo de red neuronal basado en K<sub>T</sub>. Para el resto de horizontes temporales de predicción, el modelo elegido está basado en la serie temporal LC.
- El modelo final seleccionado corresponde con los modelos con menor incertidumbre para cada horizonte temporal de predicción. De esta forma, no se produce ninguna pérdida de precisión del modelo final. Para el primer horizonte temporal de predicción, se ha seleccionado el modelo K<sub>T</sub> NN(8), ya que tiene un error del 0.8% en términos de RMSD relativo menor que el modelo LC NN2(10), y a que el modelo K<sub>T</sub> NN(8) es más sencillo que el modelo LC NN2(10). Además, en caso que falten observaciones, el modelo K<sub>T</sub> NN(8) podrá utilizarse para realizar la predicción, para el resto de horizontes de predicción, al necesitar un vector de entrada de menor tamaño.
- Sobre el total de datos de validación, los resultados que se obtienen a partir del modelo final presentan una pequeña infraestimación de las predicciones (MBD entre -0.83% y -1.21%), que se ve reflejado en los gráficos de dispersión que se presentan en el APÉNDICE B. El modelo basado en la persistencia presenta un bias menor que el modelo final (entre 0% y 0.03%).

- El modelo final mejora considerablemente al modelo basado en la persistencia en términos de RMSD relativo. El error cuadrático medio relativo del modelo final varía entre 20.58%, para el primer horizonte temporal de predicción, y 24.90%, para el último horizonte temporal de predicción, frente al modelo basado en la persistencia que varía entre 25.22% y 33.46%.
- La mejora del modelo final con respecto al modelo basado en la persistencia, en términos de RMSD relativo, es de 18.39%, para el primer horizonte temporal de predicción, y del 25.58%, para el último horizonte temporal de predicción. Por lo tanto, se puede afirmar que a medida que aumenta el horizonte temporal de predicción, el modelo final tiene mayor precisión que el modelo basado en la persistencia.

Tabla 6.2 Albacete: errores del modelo final y del modelo basado en la persistencia

Modelos	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
LC NN1(6)	-0.44	-0.12	<b>0.07</b>	0.18	<b>0.25</b>	0.31	MBD%
LC NN3(2)	<b>0.20</b>	0.49	0.57	0.60	0.66	0.69	
LC ANFIS(1)	-0.29	<b>-0.47</b>	-0.59	<b>-0.58</b>	-0.63	<b>-0.63</b>	
PER	0.02	0.05	0.07	0.09	0.11	0.13	
LC NN1(6)	26.55	27.78	<b>30.24</b>	29.81	<b>31.26</b>	<b>30.39</b>	RMSD%
LC NN3(2)	<b>26.37</b>	28.35	30.61	30.26	31.46	30.63	
LC ANFIS(1)	28.17	<b>27.51</b>	31.80	<b>29.77</b>	32.37	30.42	
PER	27.90	32.06	35.91	36.30	38.39	38.19	
RED NEURONAL	5.50	14.20	15.80	18.00	18.60	20.40	MSP%
MODELO FINAL	4.84	14.19	15.78	17.88	18.58	20.43	

A partir de los resultados anteriores se puede destacar la siguiente información para la estación de Albacete:

- Los modelos que presentan menor incertidumbre son los modelos no-lineales (NN y ANFIS), a partir de la serie temporal LC.
- El modelo final seleccionado corresponde con los respectivos modelos con menor incertidumbre para cada horizonte temporal de predicción. A pesar de que el modelo LC NN1(6) ofrece los mejores resultados para un mayor número de horizontes temporales y la diferencia con respecto al resto de horizontes es

menor al 0.2%, en términos de RMSD, se ha optado por elegir también los otros dos modelos no-lineales (LC NN3(2) y ANFIS(1)), ya que necesitan un vector de entrada de tamaño mucho menor que el modelo (LC NN1(6)). De esta forma, en caso de que no se disponga de suficientes observaciones para formar el vector de entrada de tamaño seis (para el modelo LC NN1(6)) los modelos más sencillos (LC NN3(2) y ANFIS(1)) podrían ser utilizados.

- El bias relativo del modelo final es prácticamente cero.
- El modelo final mejora considerablemente al modelo basado en la persistencia, en términos de RMSD relativo. El RMSD relativo del modelo final varía entre 26.37%, para el primer horizonte temporal de predicción, y 30.39%, para el último horizonte temporal de predicción. El RMSD relativo del modelo basado en la persistencia varía entre 27.90% y 38.19%.
- La mejora del modelo final, con respecto al modelo basado en la persistencia, está entre el 4.84%, para el primer horizonte temporal de predicción, y el 20.43%, para el último horizonte temporal de predicción, en términos de RMSD relativo. Por lo tanto, conforme aumenta el horizonte temporal de predicción, el modelo óptimo aumenta el valor de la mejora con respecto al modelo basado en la persistencia.

Tabla 6.3 Madrid: errores del modelo final y del modelo basado en la persistencia

Modelos	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
LC AR(9)	<b><u>-0.20</u></b>	-0.21	-0.21	-0.22	-0.23	-0.24	MBD%
LC NN4(9)	0.08	<b><u>-0.07</u></b>	<b><u>-0.16</u></b>	<b><u>-0.23</u></b>	<b><u>-0.28</u></b>	<b><u>-0.32</u></b>	
PER	0.00	0.00	0.00	0.00	0.00	0.00	
LC AR(9)	<b><u>23.21</u></b>	26.81	28.29	29.21	29.73	30.12	RMSD %
LC NN4(9)	23.59	<b><u>26.07</u></b>	<b><u>27.08</u></b>	<b><u>27.68</u></b>	<b><u>28.07</u></b>	<b><u>28.31</u></b>	
Persistencia (PER)	26.17	31.05	33.43	34.75	35.79	36.20	
RED NEURONAL	10.80	16.00	19.00	20.30	21.60	21.80	MSP%
MODELO FINAL	9.86	16.05	19.01	20.36	21.57	21.80	

Para la estación de Madrid, y a partir de la información anterior se puede señalar:

- El modelo final seleccionado es el modelo LC NN4(9). Este modelo corresponde con el modelo óptimo para todos los horizontes temporales de predicción, excepto para el primer semidía de predicción, donde se produce una reducción de la precisión del 0.38%, en términos de RMSD, para el conjunto de datos de validación ensayados. La diferencia del error, en términos de RMSD relativo, del modelo LC NN4(9) con respecto al modelo LC AR(9), para el resto de horizontes en los que el modelo LC NN4(9) es mejor, varía entre 0.74% y 1.81%.
- Sobre el total de datos de validación, los resultados que se obtienen a partir del modelo final presentan una pequeña infraestimación de las predicciones (MBD entre -0.07% y -0.32%), para los horizontes temporales de predicción desde 2 hasta 6 semidías. Para el primer horizonte temporal de predicción, existe una sobreestimación de las predicciones.
- El modelo final mejora considerablemente al modelo basado en la persistencia en términos de RMSD relativo. El error cuadrático medio relativo del modelo final varía entre 23.59%, para el primer horizonte temporal de predicción, y 28.31%, para el último horizonte temporal de predicción; frente al modelo basado en la persistencia, que varía entre 26.17% y 36.20%.
- La mejora del modelo final con respecto al modelo basado en la persistencia está entre el 9.86%, para el primer horizonte temporal de predicción, y el 21.80%, para el último horizonte temporal de predicción. Por lo tanto, al igual que en las estaciones anteriores, conforme aumenta el horizonte temporal de predicción, el modelo final aumenta su mejora con respecto al modelo basado en la persistencia.

Tabla 6.4 Lérida: errores del modelo final y del modelo basado en la persistencia

Modelos	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
<b>K<sub>T</sub> NN1(9)</b>	<b>0.40</b>	1.08	<b>1.37</b>	<b>1.45</b>	<b>1.61</b>	1.73	<b>MBD%</b>
<b>K<sub>T</sub> NN2(9)</b>	0.62	<b>1.19</b>	1.47	1.56	-1.70	<b>1.80</b>	
<b>PER</b>	0.03	0.06	0.10	0.13	0.16	0.20	
<b>K<sub>T</sub> NN1(9)</b>	<b>23.65</b>	27.09	<b>28.54</b>	<b>28.88</b>	<b>29.13</b>	29.33	<b>RMSD %</b>
<b>K<sub>T</sub> NN2(9)</b>	23.94	<b>26.98</b>	28.58	28.91	29.15	<b>29.31</b>	
<b>PER</b>	27.45	33.78	36.50	37.45	38.24	38.99	
<b>RED NEURONAL</b>	14.60	20.10	21.80	22.90	23.80	24.80	<b>MSP%</b>
<b>MODELO FINAL</b>	13.84	20.12	21.80	22.88	23.83	24.82	

A partir de la información, que se presenta en las tablas y figuras anteriores, se puede señalar la siguiente información para la estación de Lérida:

- El modelo final seleccionado está basado en un modelo de red neuronal y el vector de entrada  $K_T$  de tamaño 9 ( $K_T$  NN1(9)). El modelo de red neuronal que ofrece mejores resultados es el modelo basado en una única neurona, que se comporta como un modelo lineal ponderando el valor de los parámetros de entrada.
- El modelo final seleccionado tiene una pérdida de precisión con respecto al modelo con menor incertidumbre ( $K_T$  NN2(9)) entre 0.02% y 0.11% en términos de RMSD relativo, para los horizontes temporales de predicción dos y seis. Debido a que la diferencia es pequeña y a que el modelo ( $K_T$  NN1(9)) presenta una menor complejidad computacional, este modelo ha sido seleccionado como modelo final de predicción para todos los horizontes temporales de predicción.
- Sobre el total de datos de validación, los resultados que se obtienen, a partir del modelo final, presentan una pequeña sobreestimación de las predicciones (MBD relativo entre 0.40% y 1.73%) que se ve reflejado en los gráficos de dispersión. El modelo basado en la persistencia presenta un bias entre 0.03% y 0.20%; por lo tanto, el bias del modelo final se sitúa ligeramente por encima del bias del modelo basado en la persistencia.
- El modelo final mejora considerablemente al modelo basado en la persistencia en términos de RMSD relativo. El error cuadrático medio del modelo final varía

entre 23.65% para el primer horizonte temporal de predicción y 29.33% para el último horizonte temporal de predicción. El RMSD relativo del modelo basado en la persistencia está entre 27.45% y 38.99%.

- La mejora en términos de RMSD relativo del modelo final con respecto al modelo basado en la persistencia está entre el 13.84%, para el primer horizonte temporal de predicción, y el 24.82%, para el último horizonte temporal de predicción. De esta forma, al igual que ocurre en el resto de estaciones, conforme aumenta el horizonte de predicción, el modelo final aumenta también su mejora con respecto al modelo basado en la persistencia.

### Diagrama Box-Whisker de los modelos finales

Los diagramas Box-Whisker nos permiten explorar la mediana, la dispersión y la asimetría de las series temporales de datos observados y predichos. Las siguientes figuras presentan el diagrama Box-Whisker (ver sección 4.3) de la serie temporal de datos observados y predichos para cada horizonte temporal de predicción y cada una de las estaciones radiométricas utilizadas en este trabajo. Cada caja del diagrama representa la serie de datos observados (Obs) y predichos (Pred) para cada horizonte temporal de predicción (1...6 semidías), de esta forma, Obs1 corresponde con la caja del diagrama Box-Whisker para la serie temporal de datos observados para el horizonte temporal de predicción de 1 semidía.

En el caso de la estación de Murcia se puede observar que los valores locales de la mediana de las series temporales de datos observados y predichos es similar, aunque para los últimos la mediana aumenta conforme se incrementa el horizonte temporal de predicción. Para ambas series temporales, la amplitud de los brazos en el diagrama Box-Whisker abarca prácticamente todo el dominio de la variable de irradiancia solar y se observa una ausencia de outliers. La variabilidad local de los datos predichos es menor que la de los datos observados ya que la amplitud de los brazos y la distancia intercuartílica es menor en los primeros para los sucesivos horizonte temporales de predicción.

Para la estación de Albacete, al igual que en la estación de Murcia, los valores locales de la mediana de las series temporales de datos observados y predichos es similar. Se puede observar que la variabilidad local de los datos predichos es menor que la de los datos observados y, además, existe una disminución de la amplitud de los brazos de los datos predichos conforme aumenta el horizonte temporal de predicción (disminución de la distancia intercuartílica y la amplitud de los brazos).

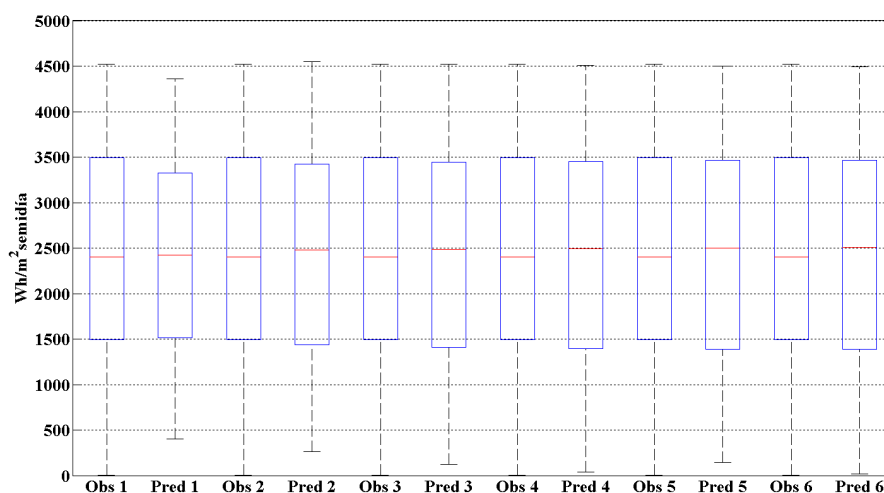


Figura 6.1 Murcia: Diagrama Box-Whisker datos observados y predichos

En el caso de la estación de Madrid los valores locales de la mediana de las series temporales de datos observados es menor que la de los datos predichos. Además, para estos últimos la mediana aumenta conforme se incrementa el horizonte temporal de predicción. La variabilidad local de los datos predichos es inferior a la de los datos observados (menor distancia intercuartílica y amplitud de los brazos). Se puede observar además que la diferencia entre los valores observados y predichos máximos es mayor que la diferencia entre los valores observados y predichos mínimos.

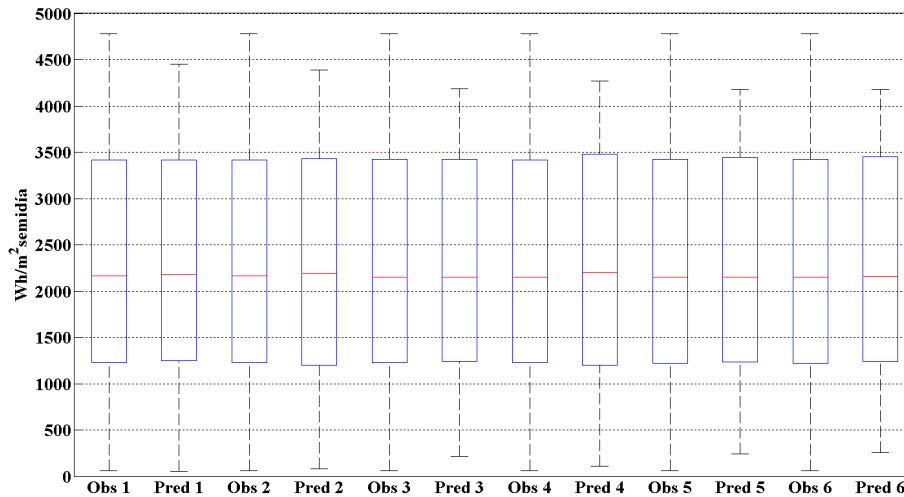


Figura 6.2 Albacete: Diagrama Box-Whisker datos observados y predichos

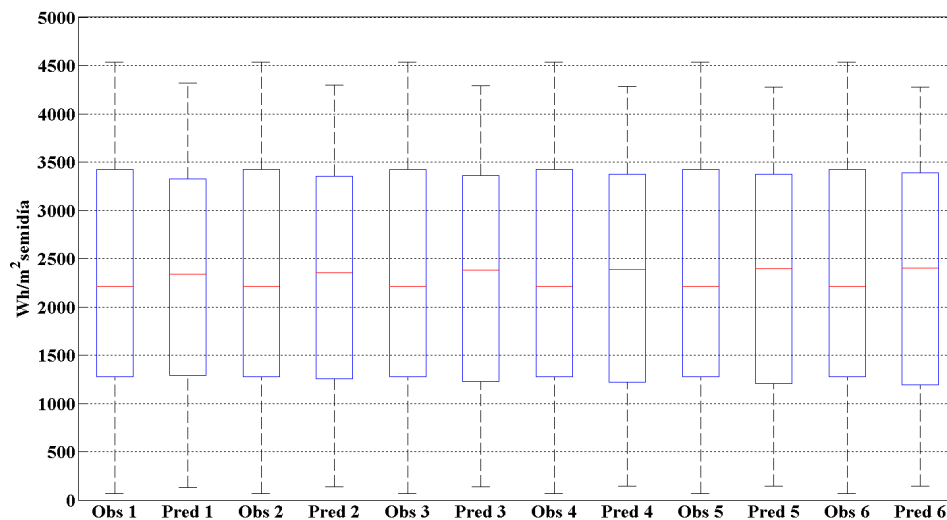


Figura 6.3 Madrid: Diagrama Box-Whisker de los datos observados y predichos

Los valores locales de la mediana para la estación de Lérida de las series temporales de datos observados y predichos son similares. La dispersión de los datos predichos es inferior que la de los datos observados, siendo dicha diferencia la mayor entre todas las estaciones radiométricas analizadas.

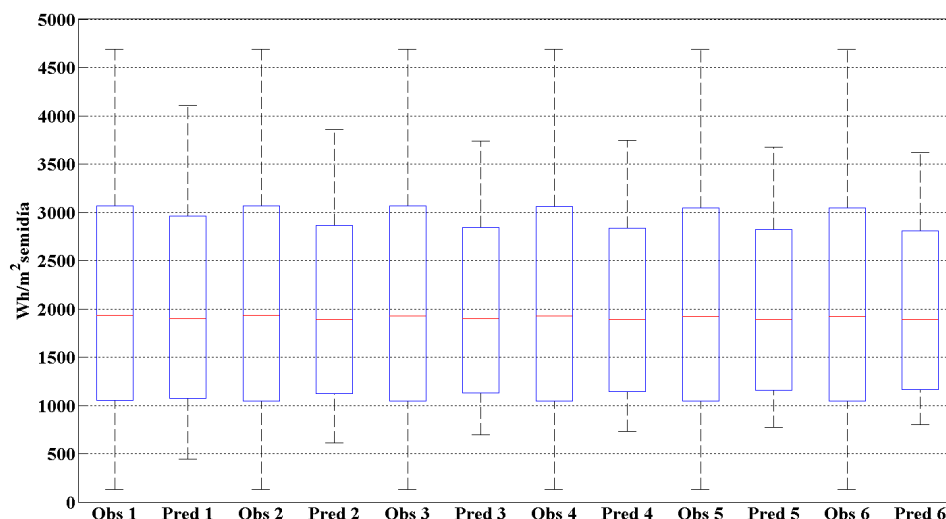


Figura 6.4 Lérida: Diagrama Box-Whisker datos observados y predichos

En la figura anterior se puede observar que el modelo para la estación de Lérida al consistir en una única neurona genera predicciones con una menor variabilidad que los modelos finales en el resto de estaciones.

#### Análisis de los errores sistemáticos de las predicciones

Además de los resultados exhibidos anteriormente, el APÉNDICE B presenta diversas gráficas que se han obtenido durante la fase de validación de los modelos finales para cada estación radiométrica. Al analizar dichas gráficas se puede observar que los modelos finales presentan una serie de errores sistemáticos al realizar las predicciones. A modo de ejemplo, se muestra a continuación este tipo de errores para el caso particular de los resultados del modelo final en la estación de Madrid y el horizonte temporal de predicción 5 semidías.

La Figura 6.5 presenta la serie temporal de datos observados y datos predichos donde se puede observar que los valores obtenidos a partir del modelo de predicción siguen la dinámica determinística de la irradiancia solar semidiaria, aunque existen valores extremos que no son simulados de forma adecuada por el modelo.

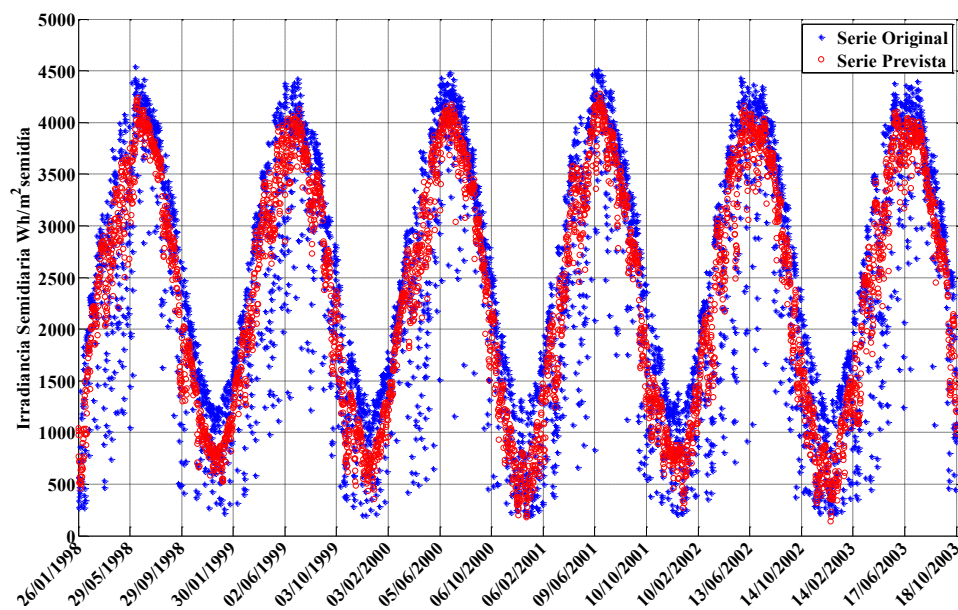


Figura 6.5 Serie temporal de datos observados y predichos utilizando el modelo final de la estación de Madrid y horizonte temporal de predicción 5 semidías

La Figura 6.6 presenta los errores absolutos en la predicción, donde se puede observar que dichos errores se producen para el mismo periodo en años sucesivos.

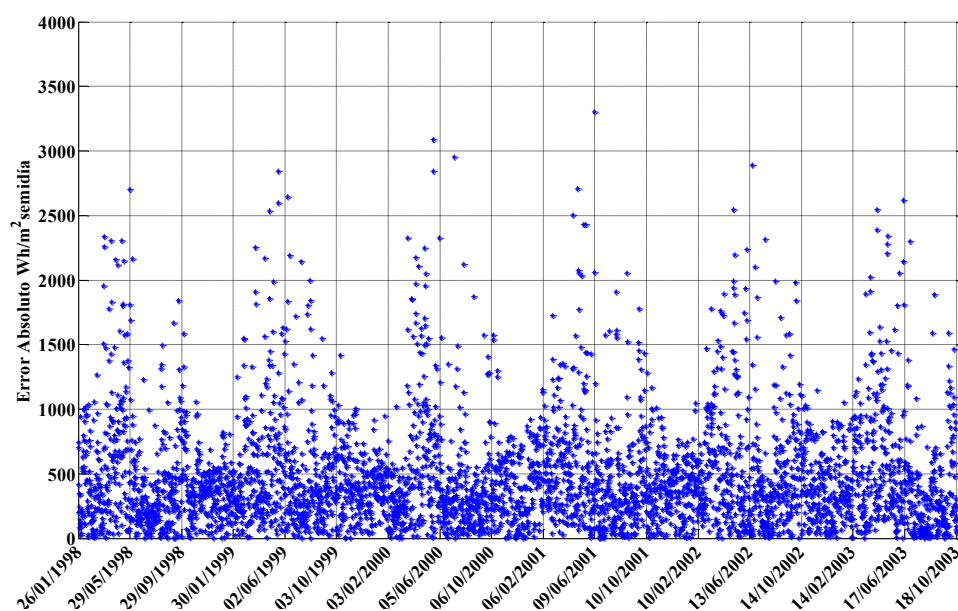


Figura 6.6 Errores absolutos para el modelo final en la estación de Madrid y horizonte temporal de predicción 5 semidías

La Figura 6.7 presenta los errores relativos de la predicción para cada día Juliano correspondiente a cada uno de los años de las series temporales de datos observados y predichos correspondientes al conjunto de validación. En dicha figura se puede observar que los menores errores se producen durante el verano, mientras que para el resto de estaciones los errores son significativamente mayores.

En la Figura 6.7 se puede observar que los mayores errores medios se producen para los meses de Mayo y Enero, existiendo una gran variabilidad de los mismos. Por el contrario, los menores errores se producen en los meses de Junio y Julio.

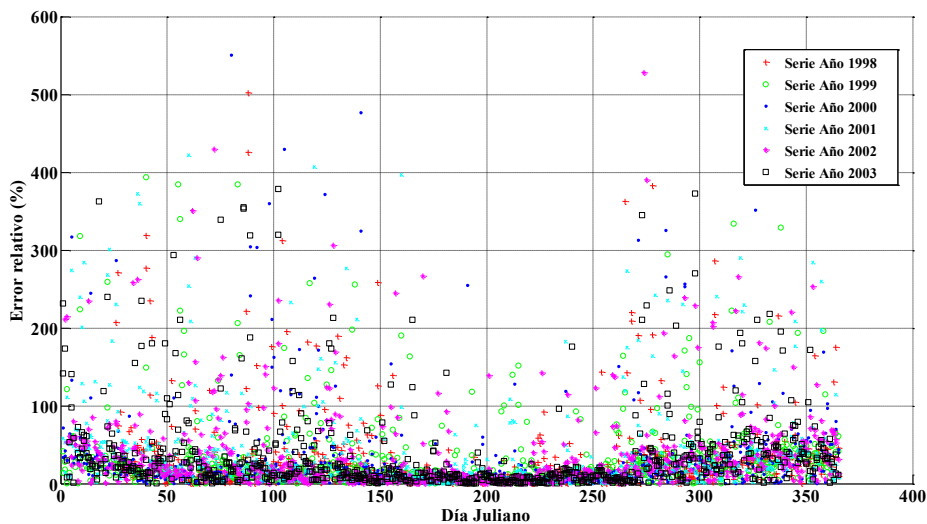


Figura 6.7 Errores relativos para cada día juliano en la estación de Madrid y horizonte temporal de predicción 5 semidías

La Figura 6.8 presenta la media mensual y la desviación de los errores relativos de la predicción para cada año de datos de validación. El error relativo se define mediante la siguiente expresión:

$$\text{Error relativo} = \frac{(\text{observaciones} - \text{predicciones})}{\text{observaciones}} \cdot 100 \quad (6.1)$$

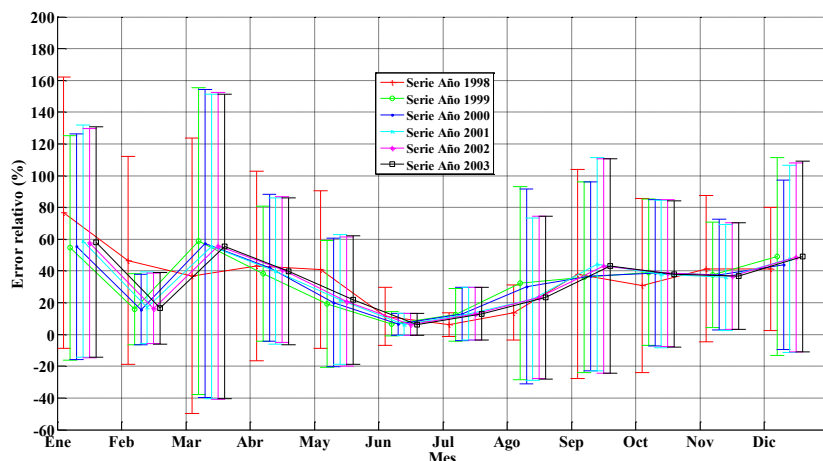
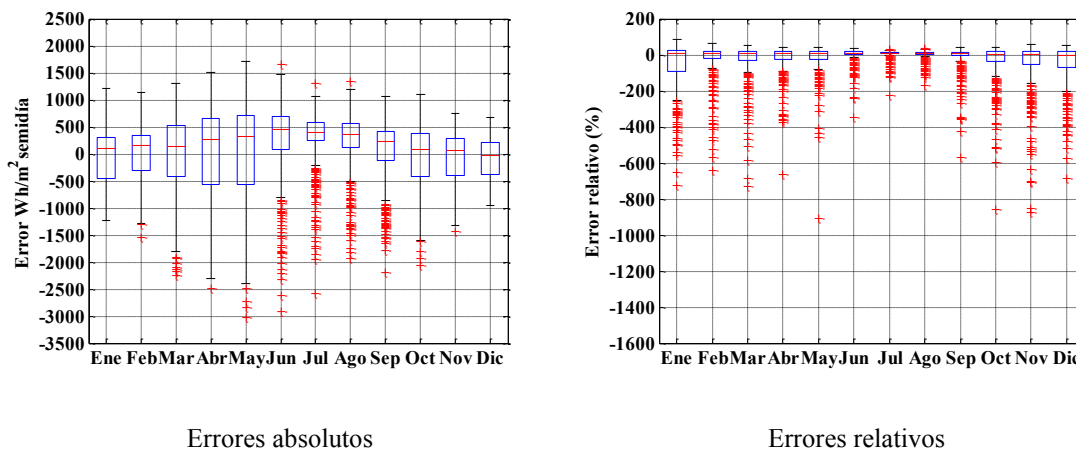


Figura 6.8 Media mensual y desviación estándar del error relativo porcentual de las predicciones de la radiación solar semidiaria para cada año

La Figura 6.9 presenta los errores absolutos y relativos de la predicción de la radiación solar semidiaria para el horizonte de predicción de 5 semidías en la estación de Madrid. Como se puede observar los mayores errores en valor relativo se producen para el periodo de invierno, otoño y primavera.



Errores absolutos

Errores relativos

Figura 6.9 Madrid: diagrama Box-Whisker de los errores absolutos y relativos de predicción para el horizonte de predicción 5 semidías

La siguiente figura presenta un zoom entre -100 y 100 de los errores relativos de la predicción de la radiación solar semidiaria para el horizonte de predicción de 5 semidías en la estación de Madrid.

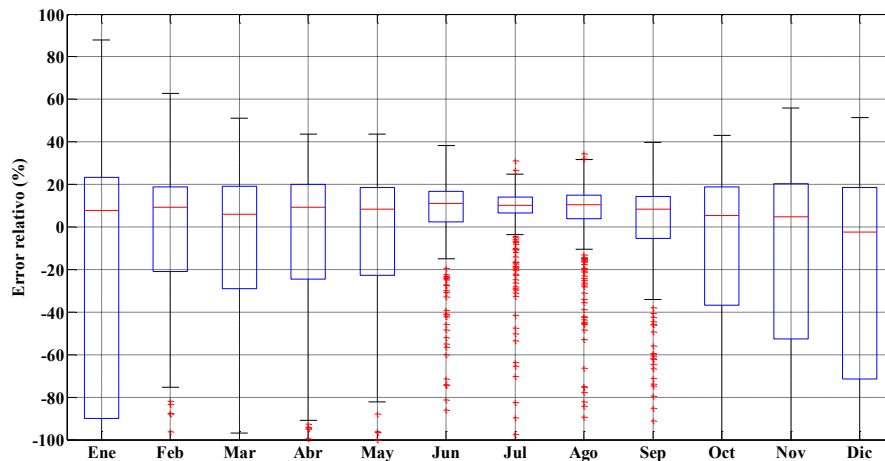


Figura 6.10 Madrid: Zoom del diagrama Box-Whisker de los errores relativos de predicción para el horizonte de predicción 5 semidías

En la Figura 6.11 y Figura 6.12 se presenta la distribución de los errores absolutos y relativos respectivamente para la estación de Madrid y horizonte temporal de predicción cinco semidías. Se puede observar que la mayoría de los errores se producen para valores observados bajos de irradiancia solar semidiaria, donde el modelo sobreestima las predicciones. Estos errores se producen en los periodos con alta variabilidad, que se pueden considerar como valores extremos al ser comparados con la dinámica determinística de la radiación solar semidiaria.

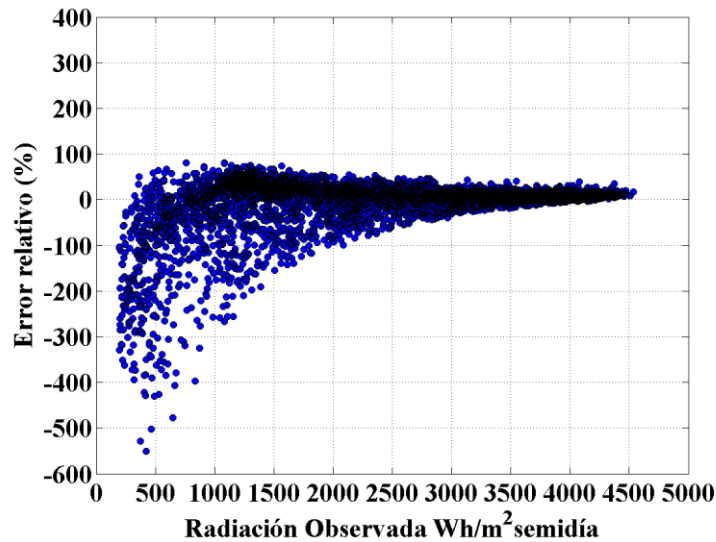


Figura 6.11 Pauta de errores detectados al comparar los errores absolutos de predicción frente a las observaciones

Las Figura 6.12 presenta el diagrama de dispersión para los valores observados y predichos de la radiación solar semidiaria junto con a la recta de regresión y los límites de predicción para un nivel de confianza del 95%. En dicha figura se pueden observar los errores sistemáticos que comete el modelo de predicción y que han sido destacados en las gráficas previas, aunque al estudiar el diagrama de contorno se puede apreciar que el número de dichos valores no es elevado.

Como se ha comentado anteriormente existen periodos durante los cuales hay grandes fluctuaciones de los valores de la irradiancia solar semidiaria que se recibe en la superficie, cuya dinámica no es simulada por los modelos finales ensayados. Así, se puede afirmar que los modelos fallan al realizar la predicción entre días sucesivos con variaciones significativas, por encima del 100% del nivel de radiación semidiaria observada en días consecutivos que provoca que los errores tengan una varianza heterogénea.

Se proponen dos alternativas para las predicciones de la irradiancia solar semidiaria para el tipo de error comentado anteriormente:

- Utilizar un esquema de predicción multietapa en el que se utilice un modelo diferente en función de las predicciones de nubosidad para el día siguiente.
- Utilizar diferentes modelos predictivos para cada periodo o para cada mes, de forma que cada modelo se especialice en la dinámica particular de la irradiancia solar semidiaria durante dicho periodo.

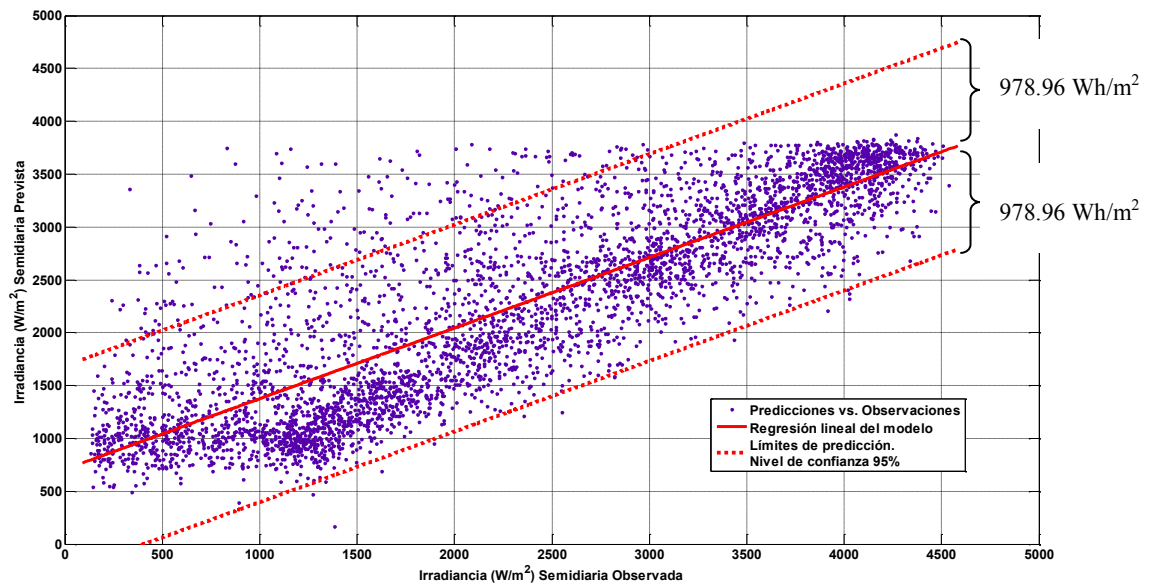


Figura 6.12 Pauta de errores detectados en el gráfico de dispersión de valores predichos frente a los valores observados ( $\text{Wh/m}^2$ )

Los errores predictivos del modelo basado en la persistencia y del modelo final se presentan en las Figura 6.13 y Figura 6.14 para cada estación radiométrica. La Figura 6.13 muestra el error RMSD relativo para el modelo final y el de persistencia en cada estación radiométrica para cada horizonte temporal de predicción.

En la Figura 6.14 se puede observar la mejora en términos de RMSD relativo del modelo final con respecto al modelo basado en la persistencia para cada estación radiométrica y cada horizonte temporal de predicción.

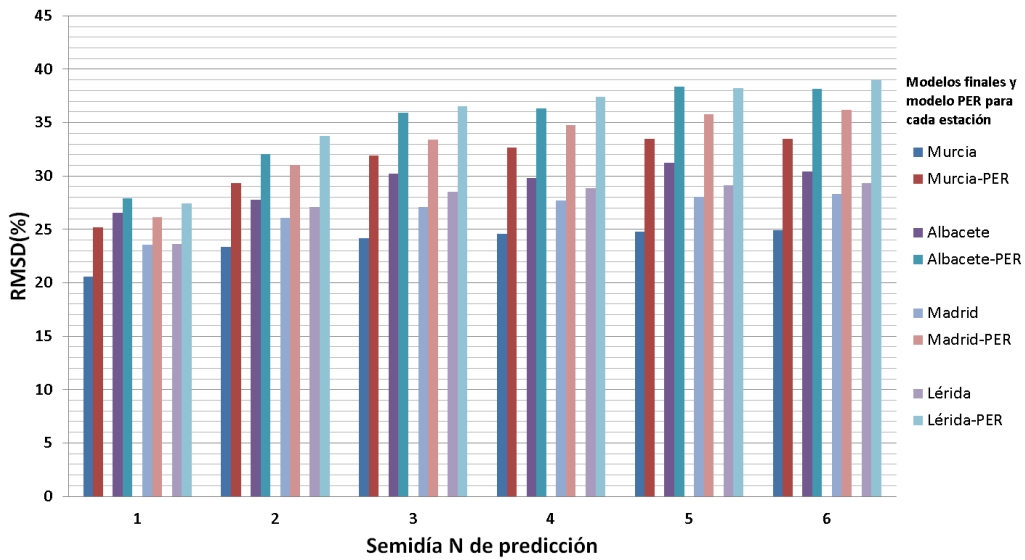


Figura 6.13 RMSD (%) de los modelos ensayados y del modelo basado en la persistencia

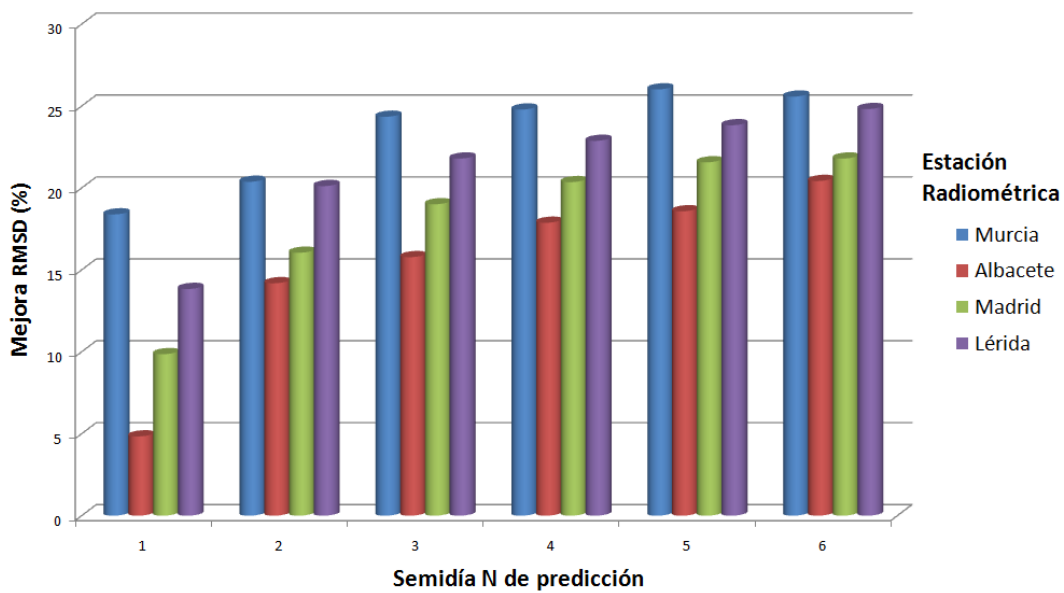


Figura 6.14 MSP (%) de los modelos finales

A partir de la información presentada en la Figura 6.13 y Figura 6.14, se puede señalar:

- Todos los modelos finales mejoran al modelo basado en la persistencia.
- Para el caso de la estación de Murcia, la mejora del modelo final con respecto al modelo basado en la persistencia es mayor que para el resto de estaciones, siendo la mejora mucho mayor el primer semidía de predicción.

- La diferencia entre los modelos finales y el modelo basado en la persistencia se va incrementando conforme aumenta el alcance de las predicciones de forma más notable para Albacete, Madrid y Lérida. Hay que exceptuar el caso de Murcia, donde, para el horizonte temporal de predicción de seis semidías, se produce una disminución de la mejora con respecto al horizonte temporal de predicción anterior.
- Los modelos basados en la persistencia que mayores errores presentan son los de las estaciones de Albacete y Lérida.

A partir de la información presentada en la Figura 6.14, se puede señalar:

- El modelo final, para la estación de Murcia, es el que presenta una mayor mejora porcentual con respecto al modelo basado en la persistencia. En el caso de la estación de Lérida, el modelo final seleccionado es el segundo mejor; sin embargo, en líneas generales se puede afirmar que su modelo basado en la persistencia presenta los peores resultados junto al modelo basado en la persistencia para Albacete.
- Por el contrario, el modelo final de la estación de Albacete es el que muestra una menor mejora con respecto al modelo basado en la persistencia.
- Para todas las estaciones, conforme aumenta el horizonte temporal de predicción, aumenta la mejora de los modelos finales con respecto al modelo basado en la persistencia para dicha estación.

#### Resultados de la validación de las predicciones semidiarias previas al cenit solar para cada mes

En este apartado se presentan los resultados mensuales obtenidos tras la validación de las predicciones previas al cenit solar para los modelos finales en los sucesivos horizontes temporales de predicción.

Para ilustrar lo que se va a hacer en este apartado, la Figura 6.15 presenta la irradiancia solar extraterrestre horaria de tres días sucesivos en un emplazamiento determinado. Para cada día, en color de relleno negro se puede observar la irradiancia solar semidiaria

previa al cenit solar para cada día en los que se realiza la predicción y sin relleno, la irradiancia solar semidiaria posterior al cenit solar.

De esta forma, se validarán de forma independiente los valores semidiarios previos al cenit solar, que corresponderán con los horizontes de predicción 1, 3 y 5. Por otro lado, se validarán los valores semidiarios predichos posteriores al cenit solar que corresponderán con los horizontes temporales de predicción 2, 4 y 6 semidías.

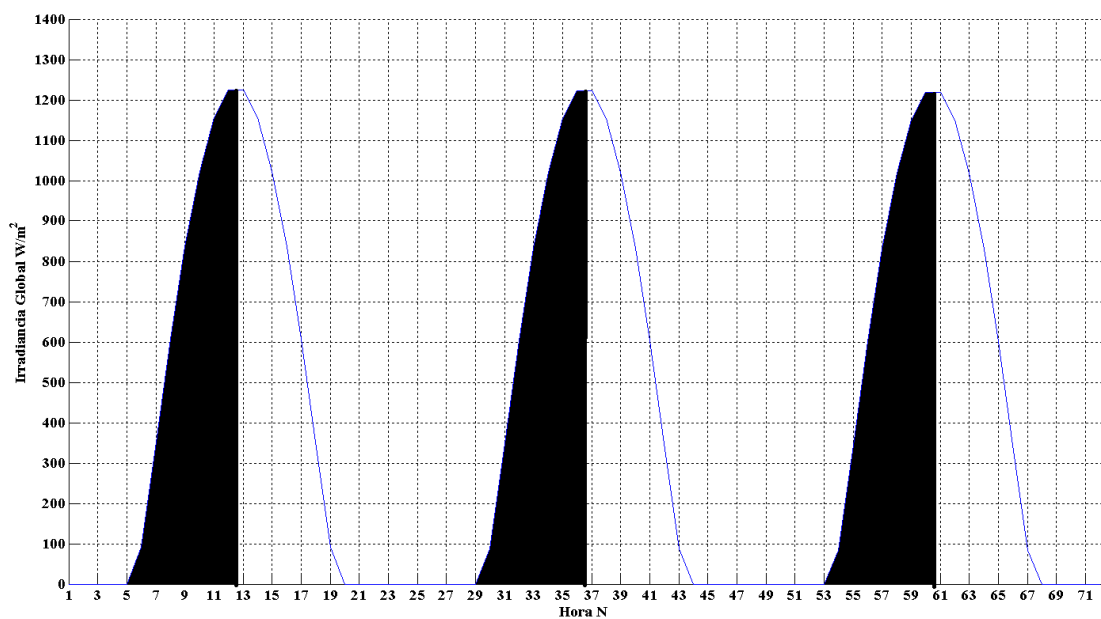


Figura 6.15 Irradiancia solar semidiaria previa al cenit solar para tres días sucesivos

Las Tablas 6.5, 6.6, 6.7 y 6.8 muestran los resultados mensuales de los estadísticos de error (MBD y RMSD porcentual) obtenidos para cada una de las estaciones radiométricas ensayadas y cada horizonte temporal de predicción de los modelos finales de predicción, únicamente para el primer semidía de predicción de cada día.

Tabla 6.5 Murcia: errores mensuales predicciones semidiarias previas al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
ENERO	-0.09	7.92	11.52	13.11	14.48	15.05	MBD(%)
FEBRERO	-0.22	0.66	2.13	2.48	2.84	2.70	
MARZO	-0.24	1.16	0.97	0.60	0.21	-0.07	
ABRIL	0.16	-1.64	-2.10	-2.29	-2.47	-2.50	
MAYO	1.20	-3.24	-3.98	-4.28	-4.48	-4.59	
JUNIO	3.26	0.28	-0.26	-0.54	-0.87	-0.99	
JULIO	3.58	1.27	0.83	0.69	0.64	0.67	
AGOSTO	3.14	1.52	1.09	0.95	0.79	0.74	
SEPTIEMBRE	1.94	1.65	1.46	1.41	1.38	1.39	
OCTUBRE	-0.52	6.20	6.76	7.28	7.78	8.05	
NOVIEMBRE	-0.03	8.10	11.09	12.46	13.80	14.44	
DICIEMBRE	-2.36	7.36	11.51	12.93	14.77	15.20	
ENERO	25.00	30.81	32.60	33.98	34.89	35.52	RMSD(%)
FEBRERO	22.54	27.34	28.34	28.31	28.11	27.88	
MARZO	25.51	29.66	30.26	30.68	30.86	31.26	
ABRIL	21.39	23.89	24.41	24.91	24.89	24.83	
MAYO	21.29	23.34	24.70	25.59	25.77	26.24	
JUNIO	12.19	13.32	13.45	13.65	14.07	13.76	
JULIO	10.55	10.56	10.86	10.98	11.02	10.96	
AGOSTO	14.08	15.56	15.75	15.79	15.83	15.74	
SEPTIEMBRE	16.31	18.55	19.52	20.01	19.64	19.57	
OCTUBRE	26.02	33.01	33.21	32.04	32.19	33.71	
NOVIEMBRE	27.14	34.10	35.37	36.03	37.41	38.92	
DICIEMBRE	27.03	33.24	35.73	36.88	38.56	38.48	

Tabla 6.6 Albacete: errores mensuales predicciones semidiarias previas al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
ENERO	8.61	20.41	19.98	26.89	24.26	25.32	MBD(%)
FEBRERO	-0.17	4.70	2.21	6.46	2.88	3.07	
MARZO	0.88	-0.34	0.83	-2.15	-0.91	-1.28	
ABRIL	-3.07	-5.49	-3.06	-6.54	-3.21	-3.22	
MAYO	-3.10	-5.94	-5.44	-7.45	-6.08	-6.39	
JUNIO	-0.12	-2.94	-2.17	-4.26	-3.40	-3.87	
JULIO	1.75	-0.19	0.03	-0.29	-0.22	-0.29	
AGOSTO	-5.33	-2.59	-1.01	-3.15	-1.42	-1.58	
SEPTIEMBRE	0.84	-0.53	-0.21	-0.59	-0.94	-1.36	
OCTUBRE	3.84	3.99	4.96	4.10	5.10	5.46	
NOVIEMBRE	8.62	18.03	18.67	25.17	21.65	21.35	
DICIEMBRE	10.95	23.91	20.22	29.22	25.25	27.54	
ENERO	40.51	44.97	39.30	47.25	42.92	43.95	RMSD(%)
FEBRERO	29.72	35.32	34.47	37.21	36.05	35.93	
MARZO	28.33	33.76	37.11	37.14	37.44	37.87	
ABRIL	23.03	27.56	27.72	29.27	28.14	27.22	
MAYO	22.83	25.26	27.94	28.33	29.92	29.84	
JUNIO	17.07	15.78	19.14	17.69	19.58	19.00	
JULIO	9.89	10.26	8.32	8.74	8.68	8.29	
AGOSTO	31.59	21.18	29.68	22.58	29.78	25.45	
SEPTIEMBRE	19.54	22.51	23.66	24.42	24.94	24.36	
OCTUBRE	29.60	33.45	35.59	36.16	37.99	37.57	
NOVIEMBRE	45.23	52.39	51.18	58.40	51.17	50.79	
DICIEMBRE	39.93	47.02	47.01	53.41	51.39	52.30	

Tabla 6.7 Madrid: errores mensuales predicciones semidiarias previas al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
ENERO	3.60	10.65	14.34	16.77	18.85	20.26	MBD(%)
FEBRERO	-1.53	1.60	3.18	4.07	4.99	5.47	
MARZO	-1.70	-2.13	-2.57	-2.99	-3.29	-3.56	
ABRIL	-2.41	-4.22	-5.36	-6.06	-6.62	-6.95	
MAYO	-3.64	-5.52	-6.66	-7.38	-7.87	-8.32	
JUNIO	0.89	-0.22	-0.78	-1.17	-1.45	-1.63	
JULIO	2.12	1.47	1.06	0.80	0.64	0.53	
AGOSTO	2.28	1.90	1.65	1.52	1.37	1.30	
SEPTIEMBRE	1.32	1.22	1.24	1.18	1.11	1.06	
OCTUBRE	3.71	4.80	5.63	6.30	6.63	6.94	
NOVIEMBRE	2.91	8.91	12.43	14.47	16.25	17.65	
DICIEMBRE	3.07	11.69	16.55	19.50	21.79	23.53	
ENERO	34.11	39.87	41.73	43.48	44.96	45.39	RMSD(%)
FEBRERO	27.32	30.84	31.58	32.16	32.56	33.26	
MARZO	27.32	31.77	32.95	33.66	34.42	34.63	
ABRIL	26.76	29.98	30.62	31.43	31.74	31.85	
MAYO	25.23	26.78	28.01	28.90	29.62	29.69	
JUNIO	16.77	17.89	18.56	18.51	18.69	18.66	
JULIO	12.67	13.28	13.40	13.54	13.40	13.30	
AGOSTO	14.59	15.63	15.94	16.14	16.52	16.31	
SEPTIEMBRE	18.93	21.02	21.67	22.09	22.17	22.87	
OCTUBRE	29.96	33.86	35.14	36.29	37.21	37.83	
NOVIEMBRE	33.49	39.13	40.62	41.69	41.76	42.65	
DICIEMBRE	35.75	41.55	44.00	45.49	46.49	47.37	

Tabla 6.8 Lérida: errores mensuales predicciones semidiarias previas al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
ENERO	-0.76	-7.31	-10.48	-11.92	-13.68	-15.01	MBD(%)
FEBRERO	-3.65	-4.36	-4.46	-5.27	-5.66	-6.22	
MARZO	4.13	6.92	8.46	9.07	9.77	10.52	
ABRIL	-1.97	-1.36	-0.89	-0.02	0.56	0.33	
MAYO	5.89	8.70	9.84	10.14	10.35	10.57	
JUNIO	2.31	6.44	8.25	9.12	10.05	11.03	
JULIO	2.77	5.31	6.51	7.01	7.54	8.09	
AGOSTO	1.88	4.17	5.18	5.70	6.21	6.64	
SEPTIEMBRE	1.14	3.53	4.57	4.99	5.52	6.02	
OCTUBRE	-0.33	-2.23	-3.08	-3.46	-3.98	-4.37	
NOVIEMBRE	-1.60	-6.65	-8.89	-10.21	-11.34	-12.57	
DICIEMBRE	-7.04	-14.14	-16.89	-19.07	-19.37	-20.62	
ENERO	38.07	42.16	45.16	47.00	48.36	48.90	RMSD(%)
FEBRERO	35.16	41.16	43.75	45.72	45.95	45.69	
MARZO	16.63	17.93	18.75	18.73	19.02	19.84	
ABRIL	27.31	30.99	30.41	28.27	26.66	28.17	
MAYO	23.48	26.46	26.45	24.88	24.34	26.26	
JUNIO	12.58	15.33	16.98	17.44	17.82	18.41	
JULIO	19.61	21.85	22.84	23.11	22.93	22.34	
AGOSTO	15.99	18.88	20.36	20.80	20.99	20.98	
SEPTIEMBRE	18.05	20.46	21.02	20.97	21.44	21.73	
OCTUBRE	27.63	30.92	31.78	32.47	32.98	33.20	
NOVIEMBRE	35.93	39.99	42.26	43.67	43.32	42.12	
DICIEMBRE	43.84	47.96	51.08	52.29	52.48	52.91	

Resultados de la validación mensual predicciones semidiarias posteriores al cenit solar

Las Tablas 6.9, 6.10, 6.11 y 6.12 muestran los resultados mensuales de los estadísticos de error (MBD y RMSD porcentual) considerando únicamente las predicciones para los semidías posteriores al cenit solar.

Tabla 6.9 Murcia: errores mensuales predicciones semidiarias posteriores al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
<b>ENERO</b>	2.15	6.62	9.21	11.62	12.48	13.79	MBD(%)
<b>FEBRERO</b>	2.47	-1.31	-1.16	0.20	-0.37	0.54	
<b>MARZO</b>	2.81	-1.59	-2.23	-2.04	-2.89	-2.88	
<b>ABRIL</b>	1.87	-2.93	-3.74	-3.74	-2.89	-3.92	
<b>MAYO</b>	3.66	-5.94	-6.90	-6.98	-7.42	-7.23	
<b>JUNIO</b>	3.27	0.16	-0.24	-0.56	-0.77	-1.11	
<b>JULIO</b>	4.41	0.78	0.27	0.15	0.02	0.16	
<b>AGOSTO</b>	3.23	1.30	0.96	0.72	0.68	0.53	
<b>SEPTIEMBRE</b>	4.46	-0.49	-1.24	-0.81	-1.29	-0.82	
<b>OCTUBRE</b>	-0.39	5.75	6.56	6.80	7.63	7.56	
<b>NOVIEMBRE</b>	1.76	6.73	8.79	11.24	11.71	12.92	
<b>DICIEMBRE</b>	-2.32	6.26	10.49	12.08	13.52	14.44	
<b>ENERO</b>	27.57	31.54	32.01	34.65	35.02	36.45	RMSD(%)
<b>FEBRERO</b>	25.64	28.45	27.77	29.14	28.37	28.69	
<b>MARZO</b>	28.20	30.72	30.58	31.94	31.82	33.30	
<b>ABRIL</b>	23.67	23.69	23.66	24.96	24.96	24.69	
<b>MAYO</b>	23.14	24.47	25.12	26.69	26.75	27.35	
<b>JUNIO</b>	12.29	13.21	13.00	13.37	13.56	13.55	
<b>JULIO</b>	10.72	11.39	11.70	12.04	11.76	12.00	
<b>AGOSTO</b>	15.87	15.55	15.60	15.71	15.60	15.79	
<b>SEPTIEMBRE</b>	18.23	18.48	18.73	20.01	19.61	19.57	
<b>OCTUBRE</b>	31.26	32.30	31.90	30.66	29.70	33.35	
<b>NOVIEMBRE</b>	30.63	33.45	34.12	35.70	35.45	38.71	
<b>DICIEMBRE</b>	30.82	34.13	36.28	37.66	39.08	39.09	

Tabla 6.10 Albacete: errores mensuales predicciones semidiarias posteriores al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
<b>ENERO</b>	13.83	19.26	25.91	25.32	29.66	28.09	MBD(%)
<b>FEBRERO</b>	-5.78	6.15	-2.55	8.39	-1.62	-0.01	
<b>MARZO</b>	-0.37	-3.15	2.48	-5.29	1.12	1.21	
<b>ABRIL</b>	-8.92	-4.07	-8.66	-4.45	-8.76	-6.83	
<b>MAYO</b>	-9.77	-3.45	-12.07	-4.27	-12.63	-11.42	
<b>JUNIO</b>	-10.43	0.36	-12.01	-0.20	-13.03	-10.84	
<b>JULIO</b>	-1.67	0.91	-4.17	1.63	-4.39	-3.15	
<b>AGOSTO</b>	-27.90	3.13	-22.08	4.34	-22.07	-16.21	
<b>SEPTIEMBRE</b>	-2.82	-0.27	-3.99	-0.21	-4.82	-3.84	
<b>OCTUBRE</b>	-3.17	4.81	0.21	5.20	0.42	2.93	
<b>NOVIEMBRE</b>	3.20	17.20	15.24	24.20	19.62	20.23	
<b>DICIEMBRE</b>	5.51	24.89	17.18	29.77	20.85	24.40	
<b>ENERO</b>	27.49	53.48	35.83	55.24	42.02	40.65	RMSD(%)
<b>FEBRERO</b>	28.31	32.05	36.13	36.10	38.28	38.37	
<b>MARZO</b>	21.63	38.19	30.88	40.37	31.29	32.80	
<b>ABRIL</b>	22.01	27.12	29.19	29.51	29.67	28.75	
<b>MAYO</b>	23.80	20.71	29.93	25.07	32.49	32.43	
<b>JUNIO</b>	21.99	7.57	25.50	9.60	26.59	26.23	
<b>JULIO</b>	7.24	10.02	9.47	5.99	10.21	10.27	
<b>AGOSTO</b>	43.27	16.97	36.77	18.06	37.07	31.81	
<b>SEPTIEMBRE</b>	14.64	22.12	23.29	24.71	25.97	24.74	
<b>OCTUBRE</b>	24.84	31.69	37.02	34.59	39.24	39.77	
<b>NOVIEMBRE</b>	30.55	54.37	46.72	59.38	49.51	49.28	
<b>DICIEMBRE</b>	28.60	51.89	44.32	56.59	49.13	51.75	

Tabla 6.11 Madrid: errores mensuales predicciones semidiarias posteriores al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
<b>ENERO</b>	3.93	9.68	13.69	15.79	18.52	19.43	MBD(%)
<b>FEBRERO</b>	-1.54	1.85	3.06	4.30	5.11	5.76	
<b>MARZO</b>	-2.49	-2.31	-3.34	-3.13	-3.99	-3.73	
<b>ABRIL</b>	-2.75	-4.39	-5.69	-6.25	-6.99	-7.10	
<b>MAYO</b>	-0.62	-3.76	-3.73	-5.49	-5.07	-6.47	
<b>JUNIO</b>	2.72	1.09	1.10	0.14	0.43	-0.28	
<b>JULIO</b>	2.27	1.95	1.21	1.26	0.87	0.97	
<b>AGOSTO</b>	3.39	2.69	2.66	2.34	2.34	2.15	
<b>SEPTIEMBRE</b>	0.68	0.98	0.74	0.89	0.65	0.74	
<b>OCTUBRE</b>	2.67	4.70	4.73	6.41	5.50	6.83	
<b>NOVIEMBRE</b>	5.17	9.74	14.09	15.27	18.04	18.38	
<b>DICIEMBRE</b>	5.35	12.64	17.93	20.33	23.02	24.64	
<b>ENERO</b>	39.82	41.05	43.20	44.82	46.30	46.05	RMSD(%)
<b>FEBRERO</b>	30.75	31.80	32.02	32.92	33.24	33.92	
<b>MARZO</b>	32.07	32.44	34.29	34.93	35.68	35.53	
<b>ABRIL</b>	29.50	30.63	31.19	32.08	32.35	32.88	
<b>MAYO</b>	26.26	25.53	27.54	27.73	28.50	28.20	
<b>JUNIO</b>	16.52	16.46	17.14	16.39	17.23	16.49	
<b>JULIO</b>	13.05	13.03	12.90	12.85	12.57	12.50	
<b>AGOSTO</b>	15.57	15.06	15.72	15.61	16.07	15.67	
<b>SEPTIEMBRE</b>	20.71	20.88	21.87	22.01	22.39	22.94	
<b>OCTUBRE</b>	34.18	34.56	36.77	37.37	38.56	39.44	
<b>NOVIEMBRE</b>	36.46	38.37	40.22	41.23	41.80	42.11	
<b>DICIEMBRE</b>	39.60	38.37	44.54	44.98	46.75	47.02	

Tabla 6.12 Lérica: errores mensuales predicciones semidiarias posteriores al cenit solar

Mes	Horizonte Predicción (Semidías)						
	1	2	3	4	5	6	
<b>ENERO</b>	3.03	-4.35	-6.39	-8.94	-10.15	-11.72	MBD(%)
<b>FEBRERO</b>	3.11	-2.90	0.07	-4.00	-1.72	-4.34	
<b>MARZO</b>	7.80	8.81	11.46	10.69	12.78	12.42	
<b>ABRIL</b>	-1.97	-0.45	-0.27	1.53	0.95	1.03	
<b>MAYO</b>	5.10	7.38	8.32	7.79	8.66	8.34	
<b>JUNIO</b>	5.40	8.01	10.76	10.87	12.49	12.92	
<b>JULIO</b>	7.05	6.54	10.07	8.49	10.89	9.78	
<b>AGOSTO</b>	3.10	4.70	6.17	6.36	7.12	7.30	
<b>SEPTIEMBRE</b>	3.97	5.05	7.25	6.66	8.11	7.81	
<b>OCTUBRE</b>	-0.25	-2.22	-2.79	-3.45	-3.88	-4.29	
<b>NOVIEMBRE</b>	4.82	-4.59	-4.06	-7.97	-6.45	-9.81	
<b>DICIEMBRE</b>	3.57	-9.54	-8.36	-12.58	-10.94	-13.29	
<b>ENERO</b>	29.57	41.41	41.02	45.55	44.98	46.72	RMSD(%)
<b>FEBRERO</b>	23.31	40.72	40.05	45.45	44.64	44.97	
<b>MARZO</b>	16.24	19.35	20.93	19.68	21.40	21.12	
<b>ABRIL</b>	20.30	28.30	27.19	24.11	22.25	25.89	
<b>MAYO</b>	22.71	29.71	29.13	27.70	25.13	27.41	
<b>JUNIO</b>	9.82	14.12	15.96	16.04	17.05	17.30	
<b>JULIO</b>	15.84	19.66	22.12	21.26	22.46	21.18	
<b>AGOSTO</b>	15.79	18.80	20.76	21.18	21.74	21.29	
<b>SEPTIEMBRE</b>	14.44	19.66	19.92	20.17	20.50	21.29	
<b>OCTUBRE</b>	24.73	30.52	31.01	33.13	32.72	33.44	
<b>NOVIEMBRE</b>	29.38	39.78	39.67	43.31	42.12	41.16	
<b>DICIEMBRE</b>	30.99	42.95	42.18	44.84	43.49	45.56	

A partir de las tablas de resultados previos y posteriores al cenit solar se puede observar que en las estaciones de Albacete y Lérica los errores predictivos son menores para las predicciones previas al cenit solar mientras que en las estaciones de Murcia y Madrid los resultados de las predicciones posteriores al cenit solar tienen menor incertidumbre.

#### Análisis de series de distribución (Test de Kolmogorov-Smirnov)

En el campo de la estadística es muy común la evaluación de alguna hipótesis sobre la distribución de una población a través de un test. Cuando el objetivo de dicho test está relacionado con la valoración de la concordancia entre la distribución de una muestra y una distribución teórica, la evaluación estadística se conoce con el nombre de test de bondad de un ajuste. El Test de Kolmogorov-Smirnov es un claro ejemplo de este tipo de evaluaciones estadísticas sobre la bondad de un ajuste [Massey F.J.Jr., 1951].

El Test de Kolmogorov-Smirnov básicamente consiste en comparar la distribución de una población con una distribución teórica de referencia. A partir de la función de distribución acumulada de una población,  $F_o(y)$ , (función de distribución acumulada de

referencia) y de un estimador de la función de distribución acumulada de una muestra de la población,  $S_N(y)$ , en este caso el conjunto de datos estimados que constituyen la respuesta de un modelo, se define el *estadístico*  $D$  como el valor máximo de la diferencia absoluta de ambas funciones de distribución acumulada [Press et al., 1998].

$$D = \max |F_o(y) - S_N(y)| \quad (9.1)$$

El estadístico  $D$  es independiente de  $F_o(y)$  si  $F_o(y)$  es una función continua.

Se formula la hipótesis nula de que la distribución real de la muestra es  $F_o(y)$ . Si el valor de este estadístico es inferior a un *valor crítico* definido al 99.9% de confianza como:

$$V_C = 1.63\sqrt{N} \quad | \quad N \geq 35 \quad (9.2)$$

siendo  $N$  el tamaño de la muestra ( $N=3407$  para Murcia,  $N=1263$  para Albacete,  $N=4265$  para Madrid y  $N=1122$  para Lérida), no se puede rechazar la hipótesis nula, es decir, ambos conjuntos de datos provienen de la misma distribución.

En este trabajo se aplicará una variación del Test de Kolmogorov-Smirnov. Dado que nuestra función de distribución depende de un conjunto de variables, elegimos la variable independiente para poder aplicar el test. Así pues, para  $n$  intervalos de variación de la variable dependiente  $y$  (intervalos de  $25 \text{ Wm}^{-2}$ ), se define el conjunto de estadísticos  $D_n$  como:

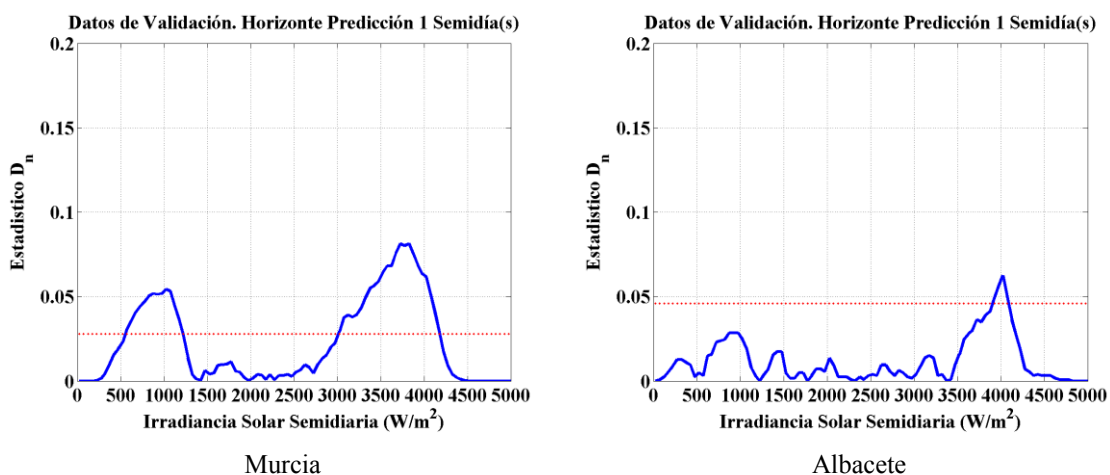
$$D_n = |F_o(y_n) - S_N(\hat{y}_n)| \quad (9.3)$$

Esta modificación del Test permite analizar la bondad del ajuste en función del dominio de la variable independiente. Además, habilita la posibilidad de analizar de forma gráfica las variaciones de comportamiento frente al Test según los intervalos definidos,

representando  $D_n$  frente a la variable independiente y cuyo valor se desea cuantificar, y una recta paralela al eje de abscisas definida por  $D_n=V_C$ .

En el APÉNDICE D se han recopilado los gráficos correspondientes a la aplicación del Test de Kolmogorov-Smirnov sobre los datos de validación de todas y cada una de las localidades analizadas. A modo de ejemplo se presentan los resultados de la aplicación del Test sobre los resultados de los modelos finales en cada estación radiométrica para el horizonte de predicción de 1 semidia.

En la Figura 6.16 se puede apreciar como prácticamente en todo el rango de variación de la irradiancia global semidiaria la mayor discrepancia entre los resultados de los modelos finales y los valores observados se producen para valores mayores a  $4000 \text{ Whm}^{-2}\text{semidia}^{-1}$ , excepto en el caso de la estación de Lérida donde el pico sucede para  $G=500\text{Whm}^{-2}$ . Ninguno de los modelos se aproxima al valor crítico definido por el Test (representado mediante una línea punteada paralela al eje de abscisas). Los modelos finales para las estaciones de Murcia, Madrid y Lérida presentan valores por encima del valor crítico para valores menores a  $1000 \text{ Whm}^{-2}\text{semidia}^{-1}$ . Prácticamente en todo el rango  $1000\text{-}3000 \text{ Whm}^{-2}\text{semidia}^{-1}$  los modelos finales se sitúan, en todos los casos, por debajo del valor crítico del Test.



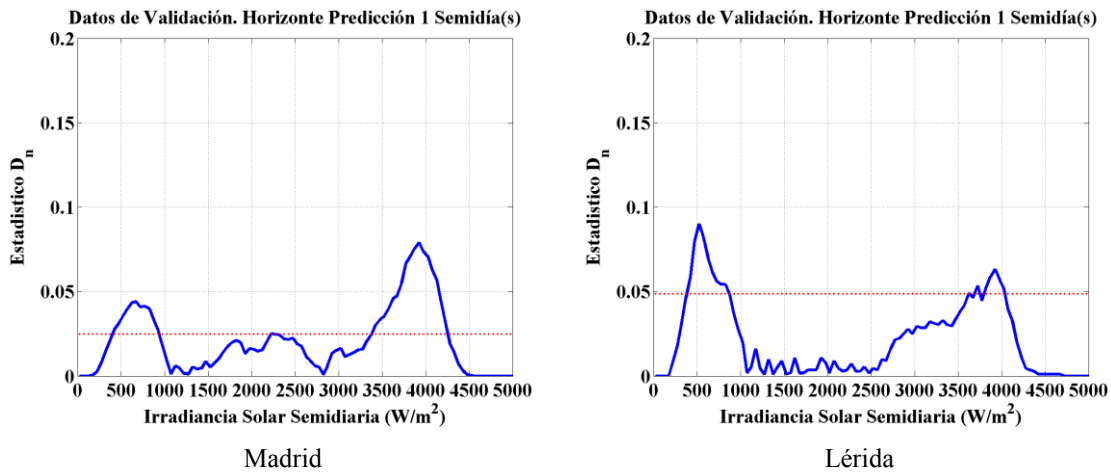


Figura 6.16 Test de de Kolmogorov-Smirnov de los modelos finales en cada estación para el horizonte temporal de predicción 1 semidía

## 6.2 Conclusiones

Tras la validación de los modelos finales para cada estación radiométrica se puede destacar la siguiente información:

- Los resultados de todos los modelos finales mejoran en todos los casos al modelo basado en la persistencia. Cuantitativamente todos los modelos propuestos mejoran en cuanto a MBD y RMSD. Cualitativamente el Test de Kolmogorov-Smirnov muestra un mejor comportamiento en el rango 1000-3000  $\text{Whm}^{-2}\text{semidía}^{-1}$ .
- Los valores locales de la mediana de las series temporales de datos observados es menor que la de los datos predichos. Además, la variabilidad local de los datos predichos es inferior a la de los datos observados (menor distancia intercuartílica y amplitud de los brazos).
- Existen ciertos periodos en los que la dinámica de la irradiancia solar semidiaria observada presenta una alta variabilidad y en los que los resultados de los modelos finales presentan errores elevados. La inclusión de la salida de los modelos dinámicos puede ser una vía para corregir estos errores sistemáticos de los modelos dinámicos.

- La predicción muestra mayores discrepancias en conjunto respecto al test de Kolmogorov-Smirnov a medida que aumenta el horizonte de la misma. Así, se puede observar que el valor integrado promedio para cada semidía va progresivamente disminuyendo.
- Se aprecia, por otra parte que en algunas estaciones, a medida que aumenta el horizonte de predicción el pico de error para bajas irradiancia tiende a desplazarse hacia irradiancias mayores, si bien el resultado no es homogéneo para todas las estaciones. Así, para Murcia el intervalo a partir del 2º semidía se acentúa en el rango 1000-1500 Whm<sup>-2</sup> mientras que para Madrid ocurre lo mismo y aparece un nuevo pico en torno a 2500 Whm<sup>-2</sup>. Para Albacete y Lérida este efecto no se aprecia, aunque en el caso de Lérida los picos tienden a elevarse y ensancharse.
- Los modelos presentan peores resultados para valores de irradiancia solar semidiaria medida bajos y altos, ya que estos son los semidías en los que hay una menor correlación entre valores sucesivos.
- Los modelos tienden a sobreestimar el valor de la irradiancia semidiaria acumulada para bajos valores de  $K_T$ , en tanto que para valores altos de dicho índice, el modelo subestima el valor de la irradiancia semidiaria.
- Por otro lado, se comprueba que a medida que aumenta el valor de  $K_T$ , el rango de incertidumbre disminuye de manera apreciable, lo que indica que el modelo ajusta mejor las predicciones cuanto mayor es el valor de dicho índice.
- La estación de Albacete muestra el mejor ajuste debido a la propia dinámica de su clima continental. La estación de Madrid presenta también unos buenos resultados.

## 7 RESUMEN

Este trabajo se engloba dentro de las técnicas de análisis de datos conocida como minería de datos (*data-mining*), cuyo fundamento es la utilización de técnicas estadístico-computacionales para la resolución de problemas como la toma de decisiones, la predicción y la clasificación.

El objetivo principal de este trabajo es la propuesta y validación de modelos para la mejora de las predicciones de los valores semidiarios de radiación solar. Los valores semidiarios de las predicciones de la energía solar se definen como los valores de la irradiancia solar integrados para los intervalos temporales entre la salida del Sol y el cenit Solar, y entre este último y la puesta del Sol. Con la intención de facilitar la revisión de los resultados a los que se hace mención se seguirá el mismo orden que en el planteamiento general del trabajo, distinguiéndose los siguientes cuatro bloques:

- Estado actual y antecedentes de la predicción de la irradiancia solar.
- Metodología experimental de predicción.
- Análisis y evaluación de resultados.
- Selección de un modelo óptimo final de predicción a partir de los resultados de predicción.

Se han expuesto los principales trabajos realizados, en el ámbito de la predicción del recurso solar, por los diferentes autores hasta la fecha para un horizonte de predicción del corto plazo. Los modelos desarrollados se han presentado en función del tipo de modelo empleado (dinámico o estadístico) y del origen de los datos con los que se realiza la predicción, obteniéndose la siguiente clasificación:

- Predicción de la irradiancia solar mediante modelos numéricos o *downscaling* dinámico.
- Predicción de la irradiancia solar mediante técnicas de *downscaling* estadístico.
- Predicción de series temporales de irradiancia solar basada en la aplicación de técnicas estadísticas lineales y no-lineales.

- Predicción del coeficiente de cobertura nubosa mediante el estudio de campos de vectores de movimiento a partir de imágenes de satélite.

Una de las dificultades en la predicción de la irradiancia solar está relacionada con la baja precisión de las predicciones que ofrecen los modelos numéricos, debido a la complejidad de la dinámica atmosférica. Los modelos de predicción numérica tienen una alta incertidumbre, al predecir las situaciones de cambios continuos de nubosidad para las situaciones de cielo cubierto e intervalos nubosos. Estos modelos predicen únicamente las situaciones sinópticas o de mesoescala y no la formación de nubes o su paso a escala local. Debido a esto, los modelos físicos presentan grandes errores en la predicción de la irradiancia solar para emplazamientos donde no existe un predominio de situaciones de cielo claro.

Las líneas actuales de investigación combinan la salida de los modelos globales o de mesoescala con un post-procesado o *downscaling* estadístico para mejorar la calidad de las predicciones. Los modelos estadísticos son una parte fundamental, que integran las predicciones de los modelos numéricos de predicción y que corrigen posibles desviaciones con respecto a las condiciones meteorológicas locales a partir de las observaciones en tierra o estimaciones de la irradiancia solar a partir de satélite. Sin embargo, los resultados que se obtienen para el paso temporal horario no tienen la precisión suficiente para las aplicaciones que hacen uso del recurso solar como principal fuente energética; sobre todo en aquellos sistemas donde no se realiza una gestión del recurso energético, como sucede en las plantas con almacenamiento (térmico o cinético) o sistemas híbridos basados en gas natural o biomasa.

Se han presentado algunas aplicaciones que hacen uso del recurso solar como principal fuente energética para la producción de electricidad y sus necesidades predictivas en cuanto a horizontes temporales de predicción y resolución espacial. Este trabajo pretende ser de utilidad en la operación y planificación de plantas termosolares y fotovoltaicas que precisan de una gestión que debe estar encaminada a la maximización del beneficio económico. A partir de las necesidades que se han presentado, se ha

optado por realizar las predicciones de la irradiancia solar con una resolución temporal semidiaria.

La elección del paso temporal semidiario para las predicciones se debe a que es el paso temporal que precisan las plantas termosolares que disponen de un sistema de almacenamiento y contiene mayor información de la dinámica de la irradiancia solar que los valores diarios. Se ha propuesto programar la entrega de la producción de las plantas termosolares en base a las predicciones de energía solar que se espera recibir para un intervalo semidiario y la gestión del almacenamiento térmico mediante estrategias óptimas de operación. Este esquema de operación es mucho más realista y permite optimizar la operación de la planta, obteniendo el máximo rendimiento de la misma; ya que se incurre en un menor número de paradas de los diferentes componentes, principalmente en la turbina que tiene un alto coste de operación en las arrancadas. El horizonte temporal de las predicciones que se han generado es de seis semidías.

Los modelos que se pueden utilizar para realizar las predicciones de la irradiancia solar global semidiaria se dividen en modelos físicos y modelos estadísticos. Se han presentado las ventajas y desventajas de ambos tipos de modelos. En este trabajo se ha optado por utilizar únicamente modelos estadísticos, conocidos también como modelos clásicos, basados en el análisis de series temporales para realizar las predicciones de la irradiancia solar con un horizonte temporal de seis semidías. De esta forma, no se utiliza ningún tipo de información procedente de los modelos numéricos de predicción. Los valores semidiarios de irradiancia solar se han obtenido a partir de la irradiancia solar horaria acumulada desde la salida del Sol hasta el cenit Solar, y entre éste y la puesta del Sol. Así, se disponen de dos valores semidiarios acumulados de la irradiancia solar por cada día, que permiten tener un mayor detalle sobre las características dinámicas de la radiación solar subdividida por el cenit Solar para cada día.

Las razones por las que se ha elegido realizar las predicciones de la irradiancia solar semidiaria mediante modelos estadísticos clásicos, utilizando técnicas de análisis de series temporales, son las siguientes:

- Simplicidad de los modelos estadísticos basados en el análisis de series temporales. Son una alternativa simple frente a los modelos numéricos, que tienen grandes requisitos computacionales.
- Disponibilidad de una extensa base de datos piranométricos en emplazamientos repartidos por toda la geografía de la Península Ibérica.
- Mayor experiencia previa en la aplicación de los modelos estadísticos en otros campos de la ciencia.
- Necesidad de un aumento en el conocimiento de las propiedades estadísticas de la irradiancia solar y la aplicación de los modelos estadísticos, a fin de que la obtención de predicciones tenga en cuenta los efectos locales.
- Menores costes desde el punto de vista del desarrollo de los modelos y operación en planta.
- Desde el punto de vista del desarrollo y la operación en plantas que hacen uso de la energía solar tienen menores costes, ofreciendo predicciones fiables en el corto plazo.

El criterio de selección de los modelos estadísticos de predicción que se ha elegido está basado en la linealidad, ya que la bondad de los modelos de predicción dependerá del determinismo (lineal o no-lineal) que presente la serie temporal de irradiancia solar. De esta forma, en caso de que la radiación solar muestre un comportamiento lineal, la utilización de los modelos lineales será más conveniente, ya que, previsiblemente, los resultados de predicción tendrán menores errores y los modelos serán más sencillos en cuanto a su uso y complejidad.

La metodología experimental de predicción que se ha propuesto se basa en el estudio y validación de modelos en función de su linealidad para la mejora de las predicciones de los valores semidiarios de la radiación solar. Las etapas en las que se divide la metodología experimental de predicción que se ha aplicado han sido:

1. Preprocesado de datos experimentales. Se han filtrado los datos de irradiancia solar horarios y se han construido las series temporales semidiarias. Además, se ha eliminado la dinámica determinística de la serie temporal de irradiancia solar semidiaria para obtener series temporales con un comportamiento estacionario. A partir de este último paso, se han obtenido dos series temporales intermedias estacionarias semidiarias, índice de claridad ( $K_T$ ) y componente solar atenuada (LC). Posteriormente, se han identificado las propiedades lineales que nos han permitido interpretar los resultados de los diferentes modelos en base a sus propiedades de linealidad
2. Ajuste de los modelos estadísticos de predicción y generación de las predicciones. En este paso, se han ajustado los parámetros de los diferentes modelos que se han planteado. Para realizar la predicción de la irradiancia solar semidiaria a partir de los valores futuros de  $K_T$  y LC, se ha utilizado un modelo lineal autorregresivo y dos modelos no-lineales basados en los modelos de redes neuronales artificiales y la lógica difusa. Las series temporales se han dividido en dos conjuntos independientes, un primer conjunto para realizar el ajuste de los modelos de predicción y un segundo conjunto para realizar la validación de los modelos de predicción. En un primer paso, se ha utilizado el conjunto de datos de entrenamiento para ajustar los parámetros de los diferentes modelos de predicción. Posteriormente, se han obtenido las predicciones de las series temporales intermedias a partir del conjunto de datos de validación. El paso final que se ha realizado en esta fase ha consistido en restaurar las tendencias que se eliminaron durante la fase de preprocesado para obtener las predicciones de la irradiancia solar semidiaria.
3. Validación de los modelos de predicción de la irradiancia solar semidiaria. Se ha realizado la validación de los diferentes modelos de predicción planteados para comprobar la adecuación y representación de la dinámica de la serie temporal de irradiancia solar semidiaria. Este paso se ha realizado analizando la calidad de las predicciones a partir de los estadísticos error medio relativo (MBD%), error cuadrático medio relativo (RMSD%) y mejora en términos de RMSD relativo de los modelos en comparación con el modelo basado en la persistencia.

Los pasos que se han realizado en la fase de preprocesado han sido los siguientes:

- Construcción de una base de datos piranométricos locales. Se han utilizado datos de irradiancia global horaria procedentes de cuatro estaciones radiométricas pertenecientes a la Red Radiométrica Nacional de AEMet, registrados durante un periodo máximo de 10 años.
- Filtrado de los datos piranométricos horarios a partir de las recomendaciones de la red radiométrica BSRN (*Baseline Radiation Network*) [McArthur, 1998].
- Estudio y selección de los periodos temporales de datos radiométricos disponibles. La intención de este estudio previo ha sido la de asegurar la continuidad de las series temporales que se han utilizado en el ajuste y validación de los modelos de predicción.
- Construcción de las series temporales semidiarias. A partir de los datos de irradiancia global horarios, se han construido los valores semidiarios formados por la suma de los valores horarios antes y después del cenit solar. De esta forma, para cada día se tienen dos valores de irradiancia solar semidiarios.
- Estudio de la estacionariedad de la serie temporal de irradiancia solar semidiaria. En este paso, se ha separado la parte determinística de la irradiancia solar, que corresponde con los ciclos/día y los cambios estacionales de la irradiancia solar, y la parte no-determinística influenciada por la nubosidad y el contenido de aerosoles en la atmósfera. A partir de la parte no-determinística se obtienen dos series temporales estacionarias intermedias correspondientes al índice de claridad y componente solar atenuada semidiaria que se han utilizado como predictores en los modelos estadísticos.
- Análisis de las propiedades estadísticas de las series temporales semidiarias intermedias. Se ha estudiado la función de densidad de probabilidad muestral y función de autocorrelación parcial de las series temporales  $K_T$  y LC.
- Test de linealidad. El test de linealidad que se ha realizado se basa en el método de datos subrogados. Éste se ha realizado sobre las series temporales semidiarias  $K_T$  y LC y ha permitido conocer de antemano qué modelos predictivos obtendrían mejores resultados a partir de ambas series temporales ( $K_T$  y LC), de acuerdo a sus propiedades lineales.

El análisis de las propiedades estadísticas de ambas series temporales ( $K_T$  y LC) para las estaciones radiométricas estudiadas ha permitido comprobar que sus funciones de densidad de probabilidad muestran una alta predominancia de días con condiciones de cielo claro. Estos resultados han concedido presuponer que los modelos de predicción se comportarán mejor bajo este tipo de condiciones.

Por otro lado, a partir de las figuras de autocorrelación parcial de los diferentes emplazamientos analizados, se ha podido comprobar que existe una alta autocorrelación parcial exclusivamente para el primer retardo temporal semidiario.

El estudio de la linealidad de las series temporales  $K_T$  y LC nos ha permitido identificar que todas las series temporales tienen un comportamiento no-lineal, a excepción de la serie temporal  $K_T$  en la estación de Lérída.

A pesar de que se ha realizado el estudio de linealidad de las series temporales  $K_T$  y LC semidiarias, se han ajustado y validado todos los modelos lineales y no-lineales que se han propuesto, para contrastar las hipótesis nula planteadas en el test de linealidad.

Los modelos que se han utilizado en este trabajo son los siguientes:

- Modelos autorregresivos. Tienen como principal característica la generación de predicciones que conservan las propiedades lineales de las series temporales.
- Modelos basados en redes neuronales. Están basados en los sistemas nerviosos biológicos y se caracterizan por realizar un procesamiento no-lineal.
- Modelos basados en redes neuronales y lógica difusa. El otro tipo de modelos no-lineales que se han ensayado se basan en el paradigma de las redes neuronales y la lógica difusa. Estos modelos tratan de mimetizar el conocimiento humano y los procesos de razonamiento.

Una vez que se han realizado las predicciones de  $K_T$  y LC, el siguiente paso (post-procesado) ha consistido en la obtención de los valores futuros de irradiancia solar semidiaria. Para ello, se han aplicado las transformaciones inversas a las aplicadas en la

fase de preprocesado, para obtener los valores de  $K_T$  y LC a partir de la irradiancia solar semidiaria. La validación de los diferentes modelos predictivos se ha realizado utilizando las predicciones de la irradiancia solar.

En base a los resultados que se han obtenido, a partir de la validación de los modelos, se ha seleccionado el modelo final de predicción para cada estación radiométrica. Los criterios de selección se basan en elegir los modelos con menor RMSD y menor complejidad para cada horizonte temporal de predicción.

## 8 CONCLUSIONES

Dentro del contexto de este trabajo, se han obtenido una serie de resultados destacables que, junto con las conclusiones derivadas de los mismos, se presentan aquí. Las conclusiones que se han obtenido tras realizar el ensayo de los diferentes modelos predictivos son las siguientes:

- La irradiancia solar semidiaria tiene un comportamiento no-estacionario. Las transformaciones realizadas han permitido obtener dos series temporales con comportamiento estacionario y que incluyen información relativa a otras variables como la latitud o el día juliano. De esta forma, al separar la parte determinística de la cuasi-aleatoria de la radiación solar, se ha conseguido simplificar la complejidad de los modelos de predicción y optimizar sus resultados.
- El primer retardo temporal de las series temporales  $K_T$  y LC contiene la mayor información para realizar las predicciones. De hecho, al aumentar el tamaño del vector de entrada de los modelos, la mejora en la predicción está limitada a un valor máximo del 2% en términos de RMSD.
- La serie temporal  $K_T$  presenta una mayor linealidad que la serie temporal LC, a partir de los resultados obtenidos en el test de linealidad. Así, se ha comprobado como la aplicación de los modelos autorregresivos lineales, a partir de la serie temporal  $K_T$ , presenta una menor incertidumbre que al utilizar la serie temporal LC como entrada de los mismos modelos.
- En general, los modelos de predicción no-lineales obtienen mejores resultados a partir de la serie temporal componente solar atenuada. La excepción a esta regla es la estación de Lérida, donde la serie temporal  $K_T$  presenta una menor dinámica no lineal. En dicha estación, los modelos basados en redes neuronales con una única neurona -que realizan regresiones simples entre entrada y salida-, y como predictor la serie temporal  $K_T$ , presentan una menor incertidumbre. Por lo tanto, se puede afirmar que los modelos de predicción no-lineales aplicados a la predicción de la radiación solar semidiaria, a partir de la serie temporal LC,

obtienen mejores resultados, excepto en el caso de que los test aseguren la linealidad (o menor no-linealidad) de la serie temporal  $K_T$ , como es el caso de la estación de Lérida.

- La precisión de los modelos depende del tipo de condición de cielo predominante en el emplazamiento bajo estudio. Para una estación radiométrica donde las condiciones de cielo predominantes son de cielo claro, los errores predictivos son menores. Esto se debe a que este tipo de días toma mayor peso en los parámetros de los modelos, que se especializan en ellos. En cambio, para las estaciones radiométricas con función de distribución bi-exponencial, es de esperar que los errores predictivos sean mayores, ya que la existencia de diferentes condiciones de tipos de cielo (cielos claros y nubosos) hace que los parámetros de los modelos no se especialicen en un único tipo de condición de cielo.
- La calidad de los datos radiométricos y la continuidad de la serie temporal es fundamental para obtener buenos resultados de predicción. Esta afirmación se basa en el hecho de que la estación de Albacete presenta una gran diferencia en términos de RMSD con respecto al resto de estaciones radiométricas, cuyas causas pueden ser las siguientes:
  - Discontinuidad temporal de las medidas. Este es el principal motivo de que dicha estación presente mayores errores predictivos.
  - Un comportamiento diferente en la dinámica de la irradiancia solar semidiaria.
  - Problemas en los instrumentos de medida que han podido perturbar el comportamiento óptimo de los modelos predictivos ensayados.
  - Escasa cantidad de muestras disponibles para el entrenamiento y validación de los modelos. Esta causa se descarta, ya que, en la estación de Lérida, el conjunto de datos disponibles también es reducido y los errores son similares al del resto de estaciones radiométricas. Por lo tanto, en principio no se puede afirmar que el número de datos disponibles para la estimación de los parámetros de los modelos influya en los errores.

- El modelo final mejora considerablemente al modelo basado en la persistencia en términos de RMSD relativo. El error cuadrático medio relativo del modelo final para todas las estaciones radiométricas varía entre: 20.58% y 30.39%, para el primer y último horizonte temporal de predicción, respectivamente. En el caso del modelo basado en la persistencia los errores, oscilan entre 25.22% y 38.99%. La mejora del modelo final con respecto al modelo basado en la persistencia está entre 4.84%, para el primer horizonte de predicción, y 25.58%, para el último horizonte temporal de predicción, en términos de RMSD. Por lo tanto, conforme aumenta el horizonte temporal de predicción, el modelo óptimo aumenta su mejora con respecto al modelo basado en la persistencia.
- Los errores de predicción presentan una varianza mayor para ciertos periodos y su función de distribución no es gaussiana. Para ciertos periodos, los errores de predicción de los modelos son mayores que para el resto de periodos. Dichos errores se producen en los periodos de verano, primavera y otoño; aunque fundamentalmente en estas dos últimas. Durante estos periodos existen grandes fluctuaciones en los valores de la radiación solar semidiaria que se recibe, de manera que los modelos de predicción no son capaces de seguir la dinámica de la irradiancia solar. De esta forma, los modelos fallan al realizar la predicción entre días sucesivos que presentan grandes cambios en el nivel de radiación. Este hecho provoca que los errores tengan una varianza heterogénea. Así, se proponen dos alternativas para mejorar las predicciones de radiación solar semidiaria durante estos periodos:
  - Utilizar un esquema de predicción multietapa, en el que se utilice un modelo diferente en función de las predicciones de nubosidad para el próximo día.
  - Emplear diferentes modelos predictivos para cada periodo o para cada mes. Dichos modelos se especializarían en la dinámica particular de la radiación solar semidiaria característica de cada periodo.
- La utilización de otros modelos de predicción más complejos, variantes de los planteados, no consiguen mejorar los resultados de predicción. Estos otros modelos que se han ensayado son los siguientes:

- Modelo basado en redes neuronales *backpropagation* independientes para cada horizonte temporal de predicción.
- Modelo de redes neuronales *backpropagation* basado en entrada multivariable (día juliano, índice de turbidez de Linke y declinación solar).
- Redes neuronales autorregresivas no-lineales (NARX).
- Red neuronal dinámica con retardo temporal a la entrada.
- Redes neuronales *wavenet* con función de activación basada en la transformada Wavelet.

Los trabajos que se tiene previsto desarrollar en el futuro son los siguientes:

- Desarrollo de un modelo operacional, combinando las predicciones de los modelos estadísticos y las salidas de los modelos meteorológicos globales (GFS y ECMWF).
- Cuantificación de la calidad de las predicciones combinando modelos estadísticos y dinámicos para diferentes escalas temporales.
- Asociación, a la predicción de la irradiancia solar, de intervalos de confianza.
- Desarrollo de modelos probabilísticos; como, por ejemplo, redes bayesianas y modelos basados en predicción por conjuntos (*ensemble predicitions*).
- Desarrollo de nuevas técnicas de predicción para el horizonte de predicción *nowcasting* mediante cámaras de cielo (Total Sky Imager TSI y Whole Sky Cameras WSC) y campos de vectores de movimiento (MVF).
- Desarrollo de técnicas de identificación del campo de visión de las cámaras de cielo en los píxeles del satélite, para la integración espacial y temporal de las medidas de ambos instrumentos. Este tipo de medidas es de gran importancia para el control de plantas termosolares, ya que el uso de cámaras de cielo puede permitir modelar el movimiento de las nubes y detectar ocultamientos del disco solar y en consecuencia variaciones de la irradiancia solar directa incidente de forma mucho más precisa que utilizando exclusivamente radiómetros o imágenes de satélite.
- Desarrollo de nuevos métodos para el modelado y predicción del coeficiente de

cobertura nubosa obtenido a partir de imágenes de satélite y cámaras de cielo.

- Desarrollo e implementación de técnicas para la asimilación de datos estimados a partir de imágenes de satélite, para su procesado en modelos de predicción numérica.
- Utilización de imágenes de satélite como base para realizar la predicción a partir de los modelos estadísticos en emplazamientos donde los datos piranométricos terrestres no estén disponibles.
- Desarrollo de métodos de predicción a largo plazo de la irradiancia solar y su relación con los principales patrones de oscilación atmosférica.
- Evaluación del valor económico de las predicciones.
- Integración de las predicciones de la irradiancia solar en herramientas de simulación de plantas termosolares y fotovoltaicas para conocer la producción.



## REFERENCIAS BIBLIOGRÁFICAS

- Aguiar, R. y Collares-Pereira, M., 1992a. Statistical properties of hourly global radiation. *Solar Energy* 48, 157-167.
- Aguiar, R. y Collares-Pereira, M., 1992b. TAG: A time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation. *Solar Energy* 49, 167-174.
- Aguiar, R., Collares-Pereira, M. y Conde, J. P., 1988. Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices. *Solar Energy* 40, 269-279.
- Amato, U., Andretta, A., Bartoli, B., Coluzzi, B. y otros, 1986. Markov processes and Fourier analysis as a tool to describe and simulate daily solar irradiance. *Solar Energy* 37, 179-194.
- Amato, U., Cuomo, V., Fontana, F. y Serio, C., 1989. Statistical Predictability and Parametric Models of Daily Ambient Temperature and Solar Irradiance: An Analysis in the Italian Climate. *Journal of Applied Meteorology* 28, 711-721.
- Araujo da Costa, A., 2005. Matemathical/Statistical and physical/meteorological models for short-term prediction of wind farms output. Tesis Doctoral, Departamento de Ingenieria térmica y fluidomecánica. E.T.S. Ingenieros Industriales. Politécnica de Madrid (España).
- Armenta-Deu, C., 1998. Improving photovoltaic system sizing by using electrolyte circulation in the lead-acid batteries. *Renewable Energy* 13, 215-225.
- Armenta-Deu, C., 2003. Prediction of battery behaviour in SAPV applications. *Renewable energy* 28, 1671-1684.
- Baker, D. G. y Casper, M. A., 1981. Subjective forecasting of received solar radiation. Proc. First Workshop on Terrestrial Solar Resource Forecasting and on the Use of Satellites for Terrestrial Solar Resource Assessment 8-11.
- Beaudin, M., Zareipour, H., Schellenberglobe, A. y Rosehart, W., 2010. Energy storagenext term for mitigating the variability of renewable electricity sources: An updated review. *Energy for Sustainable Development* 14, 302-314.
- Beckman, W. A., Bugler, J. W., Cooper, P. I., Duffie, J. A. y otros, 1978. Units and symbols in solar energy. *Solar Energy* 21, 65-68.
- Bendt, P., Collares-Pereira, M. y Rabl, A., 1981. The frequency distribution of daily insolation values. *Solar Energy* 27, 1-5.

BOE núm.126. REAL DECRETO 661/2007, de 25 de mayo, por el que se regula la actividad de producción de energía eléctrica en régimen especial, 2007. 22846-22886.

BOE núm.184. RESOLUCIÓN de 28 de julio de 2008, de la Secretaria General de Energía, por la que se aprueban los procedimientos de operación de liquidaciones de los servicios de ajuste del sistema 14.1, 14.3, 14.4, 14.6, 14.7 y 14.8 y se deroga el Procedimiento P.O. 14.5 "derechos de cobro y obligaciones de pago por garantía de potencia", 2008. 32925-32967.

Bofinger, S. y Heilscher, G., 2004. Solar radiation forecast based on ECMWF and model output statistics. Technical Report ESA/ENVISOLAR, AO/1-4364/03/I.IW, EOEP-EOMD.

Bofinger, S. y Heilscher, G., 2007. Solar electricity forecast - Approaches and first results.

Bosch, J. L., Zarzalejo, L. F., Batlles, F. J. y López, G., 2008. Downscaling Meteosat derived solar radiation maps over complex topography areas. Ponencias de: The 4th International Conference on Solar Radiation and Daylighting, Solaris 2008, Hong Kong (China).

Bouhadou, H., Hassani, M. M., Zeroual, A. y Wilkinson, A. J., 1997. Stochastic simulation of weather data using higher order statistics. Renewable energy 12, 21-37.

Bourges, B., 1979. Reconstitution des courbes de fréquence cumulees de l'irradiation solaire globale horaire recue par une surface plane. Centre d'Energetique de l'Ecole Nationale Superieure des Mines de Paris, Paris (Francia).

Box, G. E. P. y Jenkins, G. M., 1976. Time Series Analysis: Forecasting and Control. Holden-Day, California.

Breitkreuz, H., Schroedter-Homscheidt, M., Holzer-Popp, T. y Dech, S., 2007. Application of Aerosol Forecasts for Solar Energy Industries. Geophysical Research Abstracts 9.

Brinkworth, B. J., 1977. Autocorrelation and stochastic modelling of insolation sequences. Solar Energy 19, 343-347.

Cao, J. y Lin, X., 2008. Application of the diagonal recurrent wavelet neural network to solar irradiation forecast assisted with fuzzy technique. Engineering Applications of Artificial Intelligence In Press, Corrected Proof.

Cao, S.-H., Chen, J.-B., Weng, W.-B. y Cao, J.-C., 2009. Study of daily solar Irradiance forecast based on chaos optimization neural networks. Natural Science 1, 30-36.

Casa-Nova, J., Boaventura-Cunha, J. y Moura-Oliveira, P. B., 2005. Solar Irradiation Forecast Model Using Time Series Analysis and Sky Images. Ponencias de: EFITA/WCCA 2005, Vila Real, Portugal.

Demuth, H., Beale, M. y Hagan, M., 2010. Neural Network Toolbox. For use with MATLAB. User's Guide. Version 6. The MathWorks, Inc., Natick.

- Diks, C., van Houwelingen, J. C., Takens, F. y DeGoede, J., 1995. Reversibility as a criterion for discriminating time series. *Phys. Lett.* 201, 221-228.
- Dorvlo, A. S. S., Jervase, J. A. y Al-Lawati, A., 2002. Solar radiation estimation using artificial neural networks. *Applied Energy* 71, 307-319.
- Duffie, J. A. y Beckman, W. A., 1991. *Solar engineering of thermal process*. Wiley-Interscience, New York (USA).
- EFE. El Marquesado acogerá el segundo mayor parque eólico de Europa, que abastecerá a 300.000 personas, 2008. *Granada Hoy*. Hueneja (Granada), Joly Digital.
- Elizondo, D., Hoogenboom, G. y McClendon, R. W., 1994. Development of a neural network model to predict daily solar radiation. *Agricultural and Forest Meteorology* 71, 115-132.
- Erbs D.G., Klein S.A. y Duffie J.A., 1982. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. *Solar Energy* 28, 293-302.
- Espinar, B., Ramírez, L., Drews, A., Beyer, H. G. y otros, 2009. Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations. *Solar Energy* 83, 118-125.
- ESRA, 2000. *The European solar radiation atlas. Vol. 2: Database and exploitation software*. Editado por: Scharmer, K. y Reif, J. Les Presses de l'Ecole des Mines, Paris (France).
- Farnum, N. y Stanton, L., 1989. *Quantitative Forecasting Methods*. PWS-Kent Publishing Company, Boston.
- Fröhlich, C. y Brusa, R. W., 1981. Solar radiation and its variation in time. *Sol. Phys.* 74, 209-215.
- Gautama, T., Mandic, D. P. y Van Hulle, M. M., 2003. Signal nonlinearity in fMRI: a comparison between BOLD and MION. *IEEE Transactions on Medical Imaging* 22, 636-644.
- Geijo, C. y Orfila, B., 2006. Verificación de predicciones a corto plazo de energía solar mediante datos de piranómetros y observaciones de satélite. Ponencias de: XXIX Jornadas Científicas de la AME (Asociación Meteorológica Española), Pamplona (España).
- Glahn, H. R. y Lowry, D. A., 1972. The use of Model Output Statistics (MOS) in Objective Weather Forecasting. *Journal of Applied Meteorology* 11, 1203-1211.
- Gutiérrez, J. M., Cano, R., Cofiño, A. S. y Sordo, C. M., 2006. *Redes probabilísticas y neuronales en las ciencias atmosféricas*. Universidad de Cantabria e INM, Madrid (España).
- Hamilton, J. D., 1994. *Time Series Analysis*. Princeton University Press, Princeton, NJ.

Hammer, A., Heinemann, D., Lorenz, E. y Luckehe, B., 1999. Short-term forecasting of solar radiation: a statistical approach using satellite data. Ponencias de: The 1999 Meteorological Satellite Data Users' Conference, Copenhagen (Denmark).

Harvey, A. y Phipps, G. D. A., 1979. Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika* 66, 49-58.

Haykin, S., 2005. *Neural Networks: A comprehensive foundation*. Pearson - Prentice Hall.

Heinemann, D., Lorenz, E y Girodo, M., 2006. Forecasting of solar radiation. Oldenburg University, Oldenburg (Alemania).

Herrmann, Ulf, 20-2-2009. The PCTrough Performance Model. Workshop on Thermal Storage for Trough Power Systems.

Holmlund, K., de Smet, D., Gustafsson, J., Elliott, S. y Dew, G., 2008. Status report on the operational derivation of atmospheric motion vectors at EUMETSAT. Ponencias de: 9th International Winds Workshop, Annapolis (USA).

Hontoria, L., Aguilera, J. y Zufiria, P., 2005. An application of the multilayer perceptron: Solar radiation maps in Spain. *Solar Energy* 79, 523-530.

Huang, X.-Y., 1995. Initialization of Cloud Water Content in a Data Assimilation System. *Monthly weather review* 124, 478-486.

Ibañez, M., Beckman, W. A. y Klein, S. A., 2002. Frequency Distributions for Hourly and Daily Clearness Indices. *J. Sol. Energy Eng.* 124, 28-33.

Iqbal, M., 1983. *An introduction to solar radiation*. Academic Press Canada, Toronto (Canada).

Jang, J. S., 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on semiconductors* 83, 378-406.

Jensenius, J. S., 1981. Insolation Forecasting, pp. 335-350. Contenido en: *Solar Resources*. Editado por: Roland L.Hulstrom. The MIT Press, Cambridge, Massachusetts, London (England).

Jensenius, J. S., 1983. Automated forecasts of daily global solar energy. *Proc. Annu. Meet. - Am. Sect. Int. Sol. Energy Soc.* 6, 859-864.

Kandel, A., 1991. *Fuzzy expert systems*. Editado por: Kandel, A. CRC Press, Florida (USA).

Kasten, F., 1996. The linke turbidity factor based on improved values of the integral Rayleigh optical thickness. *Solar Energy* 56, 239-244.

Kasten, F. y Young, A. T., 1989. Revised optical air mass tables and approximation formula. *Applied Optics* 28, 4735-4738.

---

Kazempour, S. J., Moghaddam, M. P., Haghifam, M. R. y Yousefi, G. R., 2009. Electric energy storage next term systems in a market-based economy: Comparison of emerging and traditional technologies. *Renewable energy* 34, 2630-2639.

Kemmoku, Orita, Nakagawa y Sakakibara, 1999. Daily insolation forecasting using a multi-stage neural network. *Solar Energy* 66, 193-199.

Kipp and Zonen. CM 21 Pyranometer - Manual, 2009. Delft (Holland), Kipp & Zonen B.V.

Kistner, R., Grethe, K., Geyer, M. y Nebrera, J. A., 2004. The Progress of the AndaSol projects in Spain. Ponencias de: Proceedings of 12<sup>th</sup> IEA SolarPACES Symposium, Oaxaca (Mexico). pp. S10-107.

Kondratyev, K. Y., 1969. Radiation in the atmosphere. Academic Press, New York (USA).

Kreider, J. F. y Kreith, F., 1978. Solar engineering. London: Hemisphere Publishing Corporation, Washington (USA).

Linke, F., 1922. Transmissions-Koeffizient und Trübungsfaktor. *Beitr. Phys. fr. Atmos.* 10, 91-103.

Lorenz, E. Irradiance forecasting. Comunicación personal recibida: 17-3-2009.

Lorenz, E., 2004. Short term forecasting of solar radiation based on satellite data. Ponencias de: The 2004 Eurosun Congress, Freiburg (Germany). pp. 841-848.

Lorenz, E. N., 1963. Deterministic Nonperiodic Flow. *Journal of Atmospheric Sciences* 130-141.

Louche A., Notton G., Poggi P. y Simonnot G., 1991. Correlations for direct normal and global horizontal irradiation on French Mediterranean site. *Solar Energy* 46, 261-266.

Mandic, D. P., Chen, M., Gautama, T., Van Hulle, M. M. y otros, 2008. On the characterization of the deterministic/stochastic and linear/nonlinear nature of time series. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 464, 1141-1160.

Martín, L., Zarzalejo L.F., Polo J., Navarro, A. y otros, 2010. Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning. *Solar Energy* 84, 1772-1781.

Martín, L., Zarzalejo, L. F., Polo, J., Espinar, B. y Ramirez, L., 2006. Predicción de la radiación solar mediante técnicas estadísticas. Ponencias de: Jornada de Becarios 2006 (CIEMAT), Madrid (España).

Martín, L., Zarzalejo, L. F., Polo, J. y Navarro, A. Global solar radiation forecasting with non-linear statistical techniques and qualitative predictions from Spanish National Weather Service. Comunicación personal recibida: 2008a.

Martín, L., Zarzalejo, L. F., Polo, J., Navarro, A. y Gueijo-Guerrero, C. Solar irradiance prediction with HIRLAM model from AEMet: preliminary results, 2009a. 8th Experts Meeting of the IEA TASK 36: Solar Resource Knowledge Management, Berlín (Germany).

Martín, L., Zarzalejo, L. F., Polo, J., Navarro, A. y Marchante, R., 2008b. Comparación de técnicas predictivas basadas en series temporales aplicadas al índice de claridad semidiario. Ponencias de: XIV Congreso Ibérico e IX Iberoamericano de Energía Solar, Vigo (España). pp. 1173-1178.

Martín, L., Zarzalejo, L. F., Polo, J., Navarro, A., Ramirez, L. y Espinar, B. Daily radiation forecasting by statistical methods: preliminary results. Comunicación personal recibida: 12-3-2007.

Martín, L., Zarzalejo, L. F., Polo, J. y Ramírez, L., 2008c. Solar radiation forecasting with non-linear statistical techniques and quantitative predictions from Spanish National Weather Service. Ponencias de: Eurosun 2008, Lisboa (Portugal).

Martín, L., Zarzalejo, L. F., Polo, J., Navarro, A. y Marchante, R., 2009b. Comparison of statistical predictive techniques applied on half daily time series values of global solar irradiance. Ponencias de: SolarPaces 2009, Berlin (Germany).

Marzban, C., Sandgathe, S. y Kalnay, E., 2006. MOS, Perfect Prog, and Reanalysis. Monthly weather review 134, 657-663.

Massey F.J.Jr., 1951. The Kolmogorov-Smirnov test for goodness of fit. Journal of the American Statistical Association 56, 68-78.

Maxwell A.L., 1987. A quasi-physical model for converting hourly global horizontal to direct normal insolation. Report SERI/TR. Solar Energy Research Institute, Goleen, CO. pp. 215-387.

McArthur, L. J. B., 1998. Baseline Surface Radiation Network. Operations Manual.

Mellit, A., Benghanem, M. y Kalogirou, S. A., 2006. An adaptive wavelet-network model for forecasting daily total solar-radiation. Applied. Energy 83, 705-722.

Mora-Lopez, Ll. y Sidrach-de-Cardona, M., 1997. Characterization and simulation of hourly exposure series of global radiation. Solar Energy 60, 257-270.

Mora-Lopez, Ll. y Sidrach-de-Cardona, M., 2003. Using probabilistic finite automata to simulate hourly series of global radiation. Solar Energy 74, 235-244.

Mueller, R. W., Dagestad, K. F., Ineichen, P., Schroedter-Homscheidt, M. y otros, 2004. Rethinking satellite-based solar irradiance modelling: The SOLIS clear-sky module. Remote Sensing of Environment 91, 160-174.

Orgill J.F. y Hollands K.G.T., 1977. Correlation equation for hourly diffuse radiation on a horizontal surface. Solar Energy 19, 357-359.

- 
- Orlanski, I., 1975. Subdivisiones racionales de escalas para procesos atmosféricos. *Bulletin of the American Meteorological Society* 56(5), 527-530.
- Palit, A. K. y Popovic, D., 2005. *Computational intelligence in time series forecasting: theory and engineering applications*. Editado por: Birkhäuser. Springer-Verlag, London (UK).
- Pankratz, A., 1983. *Forecasting with Univariate Box-Jenkins Methods: Concepts and Cases*. Wiley Publishing Company, New York (USA).
- Pedrycz, W., 1989. *Fuzzy control and fuzzy systems*. Editado por: Wiley. New York.
- Pelland, S. Comparison of environment Canada's irradiance forecasts against ground station data. Comunicación personal recibida: 17-3-2009.
- Perez, R., Moore, K., Wilcox, S., Renne, D. y otros, 2007. Forecasting solar radiation - Preliminary evaluation of an approach based upon the national forecast database. *Solar Energy* 81, 809-812.
- Pielke, R. A., 2002. *Mesoscale Meteorological Modeling*. Vol. 78. Serie: International Geophysics. Academic Press, London (UK).
- Polo, J., Zarzalejo, L. F., Martín, L. y Navarro, A. A., 2008. Corrección del coeficiente de turbidez de Linke a partir de imágenes de satélite Meteosat y su aplicación en la estimación de la irradiancia directa para condiciones de cielo despejado. Ponencias de: XIV Congreso Ibérico e IX Iberoamericano de Energía Solar, Vigo (España). pp. 1137-1142.
- Polo, J., Zarzalejo, L. F., Martín, L., Navarro, A. A. y otros, 2009. Estimation of daily Linke turbidity factor by using global irradiance measurements at solar noon. *Solar Energy* 83, 1177-1185.
- Pozo, D., Lara, V., Ruiz-Arias, J. A. y Santos, F. Preliminary benchmarking results of solar GHI and DNI forecast in southern Spain based on the WRF model. Comunicación personal recibida: 17-3-2009.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. y Vetterling, W. T., 1998. *Numerical recipes in C. The Art of scientific computing*. Cambridge University Press, Cambridge (UK).
- Ramirez Santigosa, L., Polo Martinez, J., Mora Lopez, L., Sidrach de Cardona Ortin, M. y otros, 2003. Fuzzy inference systems applied to the daily ultraviolet radiation evaluation (295-385 nm) from daily global radiation. *Solar Energy* 75, 447-454.
- Reindl D.T., Beckman W.A. y Duffie J.A., 1990. Diffuse fraction correlations. *Solar Energy* 45, 1-7.
- Renne, D., 2009. Semi-Annual Status Report: November 2009. Task 36: Solar Resource Knowledge Management. IEA SHC Task 36 Solar Resource Knowledge Managment. IEA.
-

Rigollier, C., Bauer, O. y Wald, L., 2000. On the clear sky model of the ESRA -- European Solar Radiation Atlas -- with respect to the heliosat method. *Solar Energy* 68, 33-48.

Ruiz, V., 2006. El reto energético. Almuzara, Cordoba.

Safi S. y Zeroual A., 1999. Identification of Nonminimum Phase Systems Using Third and Fourth order cumulants, Application: Prediction of Daily solar Radiation. Ponencias de: 1st rencontre sur les Systèmes Electroniques et Informatiques, (S.E.I) 13-14, Fez (Marocco). pp. 147-161.

Safi, S., Zeroual, A. y Hassani, M., 2002. Prediction of global daily solar radiation using higher order statistics. *Renewable energy* 27, 647-666.

Schreiber, T., 1999. Interdisciplinary application of nonlinear time series methods - the generalized dimensions. *Physics Reports* 308, 1-64.

Schreiber, T. y Schmitz, A., 1996. Improved surrogate data for nonlinearity tests. *Phys. Rev. Lett.* 77, 635-638.

Sugeno, M., 1985. Industrial applications of fuzzy control. Elsevier Science Publication Co., New York (USA).

Swami, A., Mendel, J. M. y Nikias, C. L. M., 1998. Higher-Order Spectral Analysis Toolbox User's Guide: for use with MATLAB. The MathWorks, Natick (USA).

Takagi, T. y Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on systems, Man and Cybernetics* 15, 116-132.

Takens, F., 1981. Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence, Lecture Notes in Mathematics* 898, 366-381.

The MathWorks, I., 2010. Fuzzy Logic Toolbox. User's Guide. The MathWorks, Inc, Natick.

Theiler, J. T., Eubank, S., Longtin, A., Galdrikian, B. y otros, 1992. Testing for nonlinearity in time series: The method of surrogate data. *Physica D* 58, 77-94.

Theiler, J. T. y Prichard, D., 1996. Constrained-realization Monte-Carlo method for hypothesis testing. *Physica D* 94, 221-235.

Walpole, R. E., Myers, R. H., Myers, S. L., Ye, K. y Yee, K., 2002. Probability and Statistics for Engineers and Scientists. Prentice Hall, New Jersey.

Weigend, A. S. y Gershenfeld, N. A., 1993. Time Series Prediction: Forecasting the Future and Understanding the Past. Vol. XV. Editado por: Jack Repcheck. Serie: Proceedings Volume, Santa Fe Institute Studies in the Sciences of Complexity. Addison-Wesley y Westview Press.

Wilks, D. S., 2005. Statistical methods in the atmospheric sciences. Vol. 91. Serie: International Geophysics. Academic Press, London (UK).

Wittmann, M., Breitzkreuz, H., Schroedter-Homscheidt, M. y Eck, M., 2008. Case Studies on the Use of Solar Irradiance Forecast for Optimized Operation Strategies of Solar Thermal Power Plants. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE 1, 18-27.

WMO, 1981. Annex: World maps of relative global radiation. Technical note N° 172, pp. 25-27. Contenido en: Meteorological aspects of the utilization of solar radiation as an energy source. Secretariat of the World Meteorological Organization, Geneva (Switzerland).

Zarzalejo, L. F., 2005. Estimación de la irradiancia global horaria a partir de imágenes de satélite. Desarrollo de modelos empíricos. Tesis Doctoral, Departamento de Física Atómica, Molecular y Nuclear. Facultad de Ciencias Físicas. Universidad Complutense de Madrid (España).

Zarzalejo, L. F., Polo, J., Martín, L., Ramírez, L. y Espinar, B. K<sub>c</sub>-n fit for deriving global solar radiation from satellite images in Spain. Comunicación personal recibida: 2007.

Zarzalejo, L. F., Polo, J., Martín, L., Ramírez, L. y otros, 2009. A new statistical approach for deriving global solar radiation from satellite images. Solar Energy 83, 480-484.

Zhang, Q. y Benveniste, A., 1992. Wavelet Networks. IEE Transactions on Neural Networks 3, 889-898.



## APÉNDICE A – FUNDAMENTOS DE IRRADIANCIA SOLAR

La trayectoria que describe el Sol en la bóveda celeste varía para cada día del año. De esta forma, la modelización, así como las medidas de irradiancia, precisan saber con exactitud la posición de este astro en cada instante del día y cada día del año.

El Sol se puede considerar como una esfera de  $1.39 \cdot 10^6$  km de diámetro y  $1.99 \cdot 10^{32}$  kg de masa en la cual se producen reacciones termonucleares que transforman núcleos de hidrógeno en núcleos de helio. Debido a estas reacciones nucleares, la temperatura de su superficie varía entre los 4700 ~ 7500 K, y sus emisiones corresponden al del espectro de radiación de un cuerpo negro a 5760K aproximadamente. La distancia media entre la Tierra y el Sol es de  $1.49 \cdot 10^8$  km y se denomina unidad astronómica (*UA*). El Sol describe un movimiento de rotación sobre su propio eje con un periodo de 4 semanas aproximadamente. Sin embargo, ya que no es un sólido rígido, el periodo de rotación puede variar entre 27 y 30 días (regiones ecuatoriales y polares respectivamente).

Suponiendo que la temperatura y el espectro de radiación del Sol es constante, la cantidad de energía que llega a la superficie atmosférica terrestre puede determinarse analíticamente a partir de las posiciones relativas entre el Sol y la Tierra. Debido a que la Tierra también tiene un movimiento de rotación sobre su propio eje y otro de translación alrededor del Sol, es necesario definir previamente un sistema de referencia espacio temporal que ubique el Sol y la Tierra, ya que son dos cuerpos en movimiento.

### A.1 Sistema temporal de referencia

El *día solar* se puede definir como el intervalo de tiempo en el que el Sol cruza dos veces un mismo meridiano local. La longitud del día solar no es constante, ya que varía a lo largo del año por las siguientes causas:

- La distancia Sol – Tierra, debido a la órbita elíptica que registra la Tierra alrededor del Sol, y que varía a lo largo del año.

- La inclinación del eje de rotación de la Tierra ( $\sim 23.5^\circ$ ) respecto al plano de traslación alrededor del Sol conocido también como plano de la eclíptica.

Al observar el Sol desde la Tierra se puede apreciar que éste describe un desplazamiento sobre la bóveda celeste en función del tiempo (a lo largo del día y día tras día a lo largo del año). El *mediodía solar* se define como el instante de tiempo en que el Sol intercepta el plano meridiano del lugar, en ese momento el Sol alcanza su elevación máxima sobre el horizonte. El amanecer y ocaso se distribuyen de forma simétrica respecto al mediodía solar.

El *tiempo solar verdadero* o *tiempo local aparente (TSV)* se define como el tiempo contado a partir del mediodía solar. Por otro lado, la *hora oficial* o *tiempo local medio (TLM)* se corresponde con la hora que utilizamos habitualmente en nuestros relojes y se establece en función de la conveniencia de los países, que supone la utilización de un mismo sistema de referencia temporal para determinadas áreas geográficas. La conversión de la hora oficial en hora solar local necesita de dos correcciones [Iqbal, 1983; Duffie y Beckman, 1991]:

- La diferencia en longitud geográfica entre el meridiano local del observador y el meridiano de referencia respecto del cual se define la hora oficial (4 minutos por grado geográfico).
- Los efectos que vienen expresados en la denominada ecuación del tiempo (*ET*) y son debidos a la excentricidad de la órbita terrestre, la constancia en la velocidad areolar (2ª Ley de Kepler) y los movimientos de precisión y nutación del eje de rotación terrestre.

El cálculo de la ecuación del tiempo en minutos puede realizarse mediante la siguiente expresión [Spencer, 1971]:

$$ET = 229.18(0.000075 + 0.001868 \cos \Gamma - 0.032077 \sin \Gamma - 0.014615 \cos 2\Gamma - 0.040849 \sin 2\Gamma) \quad (\text{A.1})$$

donde  $\Gamma$  representa el denominado *ángulo diario*; es función del día juliano ( $J_d$ ) y puede calcularse, en radianes, según:

$$\Gamma = 2\pi(J_d - 1)/365.24 \quad (\text{A.2})$$

En ciertas ocasiones, puede ser necesaria una tercera corrección temporal ( $C_h$ ) debido a los posibles cambios horarios por consideraciones de ahorro energético (en España peninsular  $C_h=1$  en invierno y en verano  $C_h=2$ ). De esta forma, para transformar el tiempo solar verdadero en hora decimal, se puede utilizar la siguiente expresión [ESRA, 2000] :

$$TSV = TLM + (\Lambda_{of} - \Lambda_{loc})/15 + ET/60 + C_h \quad (\text{A.3})$$

siendo:

$\Lambda_{of}$  = Longitud del meridiano de referencia horaria oficial.

$\Lambda_{loc}$  = Longitud del meridiano local.

Los valores que proporciona la ecuación del tiempo oscilan  $\pm 17$  minutos a lo largo del año, tal y como puede apreciarse en la Figura A.1.

El tiempo solar verdadero es el sistema de referencia temporal habitualmente utilizado en todas las aplicaciones solares y es el adoptado en esta memoria.

## **A.2 Sistema espacial de referencia. Geometría solar**

La dinámica de la Tierra alrededor del Sol necesita evaluar una serie de parámetros geométricos que describan la posición del Sol con respecto a una localización concreta de la Tierra para poder realizar cualquier cálculo para ese punto concreto.

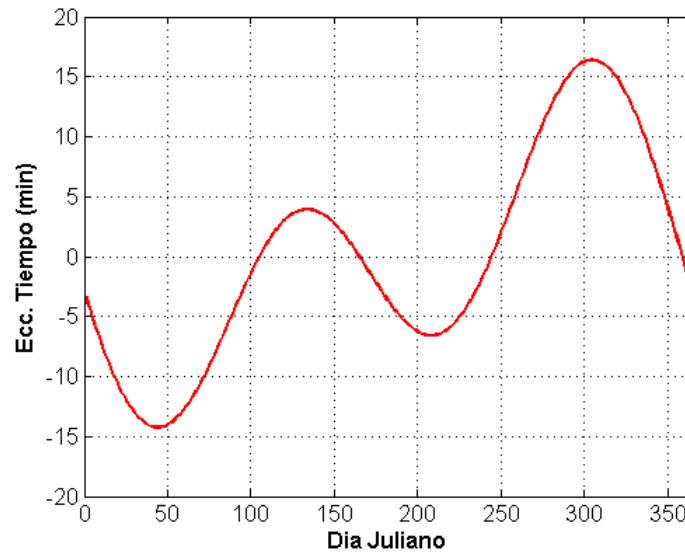


Figura A.1 Ecuación del tiempo en minutos

La Tierra describe un movimiento de traslación alrededor del Sol, con una órbita elíptica de baja excentricidad (conocida como eclíptica), y con un periodo de 365.25 días aproximadamente. La distancia entre la Tierra y el Sol no es constante, como se afirmó anteriormente, sino que varía a lo largo del año. Debido a que la órbita que describe nuestro planeta es excéntrica, existe una pequeña variación, 1.7% aproximadamente, entre el mínimo (perihelio,  $1.46 \cdot 10^8$  km.) y el máximo (afelio,  $1.52 \cdot 10^8$  km.). Además del movimiento de traslación, la Tierra gira alrededor de un eje que forma un ángulo de  $23^{\circ} 27'$  con el plano de la eclíptica.

Tras la definición del sistema de referencia temporal, a continuación se referencian las posiciones relativas del Sol y la Tierra. Por conveniencia y simplicidad se ha elegido el sistema planetario de Ptolomeo, en el que se supone que el Sol describe un movimiento de rotación en torno a la Tierra (sistema de coordenadas eclípticas geocéntrico).

Los parámetros fundamentales para determinar las relaciones geométricas el Sol respecto a un observador terrestre sobre una superficie horizontal a la superficie terrestre se presentan a continuación (ver Figura A.2):

- Latitud ( $\phi$ ). Es la posición angular, norte o sur, respecto al ecuador, positivo en el norte;  $-90^{\circ} \leq \phi \leq 90^{\circ}$ .

- Día Juliano ( $J_d$ ). Se define como el día del año ( $J_d=1\sim 366$ ). Siendo  $J_d=1$  para el 1 de Enero.
- Declinación ( $\delta$ ). Se define como el ángulo formado por el plano ecuatorial terrestre y el plano de translación de la Tierra, en el norte positivo;  $-23.45^\circ \leq \delta \leq 23.45^\circ$ . La declinación ( $\delta$ ) se puede obtener a partir de la siguiente aproximación (en radianes):

$$\delta = \frac{\pi}{180} 23.45 \sin \left[ 360(284 + J_d) / 365 \right] \quad (\text{A.4})$$

- Ángulo horario ( $\omega$ ). Se define como el desplazamiento angular del Sol, este u oeste, respecto al meridiano local. Es debido a la rotación de la Tierra sobre su propio eje. Varía a razón de  $15^\circ$  por hora. Por la mañana es positivo, y después de que el Sol pase por el meridiano local, negativo.
- Azimut solar ( $\psi$ ). Es el ángulo en el cenit solar entre el plano meridiano del observador y el plano de un gran círculo que pasa a través del cenit y el Sol. Se mide positivo al Este, negativo al Oeste (cero al Sur) y de esta forma varía entre  $0^\circ$  y  $\pm 180^\circ$ .
- Ángulo cenital solar ( $\theta_z$ ). Se puede definir como el ángulo de incidencia de los rayos solares directos sobre la superficie terrestre horizontal donde se sitúa el observador. El ángulo cenital solar se puede expresar mediante la siguiente ecuación:

$$\cos \theta_z = \sin \delta \sin \phi + \cos \delta \cos \phi \cos \omega = \sin \alpha \quad (\text{A.5})$$

- Altura o elevación solar ( $\alpha$ ). Se define como el ángulo que subtiende el vector entre el punto del observador terrestre y la posición del Sol, con el plano horizontal tangente a la superficie terrestre. Este ángulo es complementario al ángulo cenital solar ( $\theta_z + \alpha = \pi/2$ ).

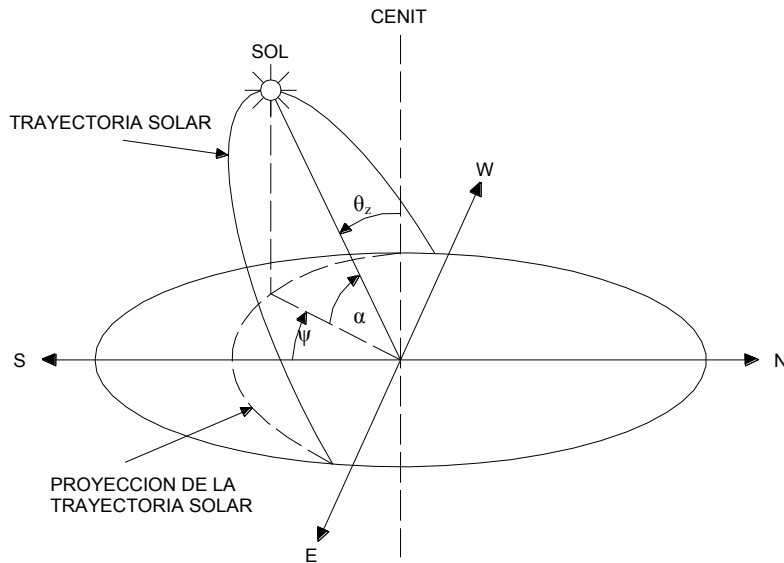


Figura A.2 Ángulos solares

Los ángulos cenital y azimutal pueden ser calculados mediante:

$$\theta_z = \cos^{-1}(\sin \delta \sin \phi + \cos \delta \cos \varpi \cos \phi) = (\pi/2) - \alpha \quad (\text{A.6})$$

$$\psi = \cos^{-1}[(\cos \theta_z \sin \phi - \sin \delta) / (\sin \theta_z \cos \phi)] = \sin^{-1}[\cos \delta \sin \omega / \sin \theta_z] \quad (\text{A.7})$$

### A.3 Variación de la irradiancia extraterrestre

La radiación emitida por el Sol y su relación espacial presenta una intensidad aproximadamente fija fuera de la atmósfera terrestre. Al valor de la irradiancia solar sobre superficie plana normal al vector posición del Sol, ubicada en el límite superior de la atmósfera terrestre se le conoce como *constante solar* ( $I_{SC}$ ). El valor admitido para esta constante ha variado en función de las distintas mediciones realizadas durante los últimos años. Sin embargo, el valor actualmente más aceptado, y utilizado en este trabajo, es  $1367 \text{Wm}^{-2}$ . Este valor es el adoptado por la Organización Meteorológica Mundial [WMO, 1981] y ha sido estimado por el Centro Radiológico Mundial (WRC), a partir de 25000 medidas realizadas con diferentes radiómetros de cavidad absoluta [Fröhlich y Brusa, 1981].

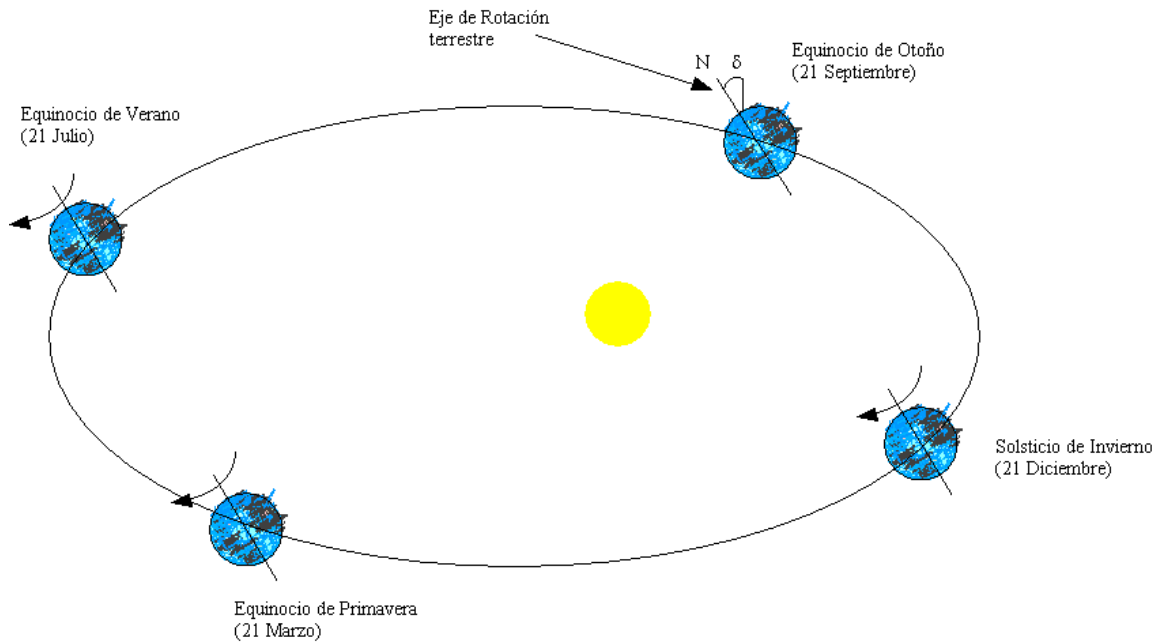


Figura A.3 Orientación del eje de la Tierra en su movimiento de translación

Debido a la *excentricidad* que describe la órbita terrestre en su movimiento de translación, la distancia Sol-Tierra varía un 1.7% aproximadamente cada año. De esta forma, la irradiancia solar que se recibe en el límite superior de la atmósfera terrestre no es constante, y se ve atenuada según la Ley del Cuadrado de la distancia, dando una variación estacional de  $\pm 3.3\%$ . La *excentricidad de la órbita terrestre* ( $\varepsilon$ ) se puede obtener mediante la siguiente expresión:

$$\varepsilon = 1.000110 + 0.034221 \cos \Gamma + 0.001280 \sin \Gamma + 0.000719 \cos 2\Gamma + 0.000077 \sin 2\Gamma \quad (\text{A.8})$$

donde  $\Gamma$  representa el denominado *ángulo diario* definido anteriormente.

La energía que se recibe en una superficie normal a la dirección del vector posición del Sol se puede obtener en función de la excentricidad. Este valor se conoce como *irradiancia solar extraterrestre* ( $I_{0n}$ ), y representa la variación estacional de la constante solar, debido a la variación de la distancia Sol-Tierra con respecto al valor medio ( $AU$ ):

$$I_{0n} = \varepsilon I_{CS} \tag{A.9}$$

Mediante la expresión anterior se puede obtener la irradiancia recibida por una superficie plana tangente a la superficie terrestre en el límite superior de la atmósfera ( $I_0$ ). El valor de  $I_0$  se puede obtener a partir de la siguiente expresión:

$$I_0 = I_{CS} \varepsilon \cos \theta_z \tag{A.10}$$

La relación anterior establece el límite superior de la irradiancia, que se puede recibir en un plano horizontal sobre la superficie terrestre.

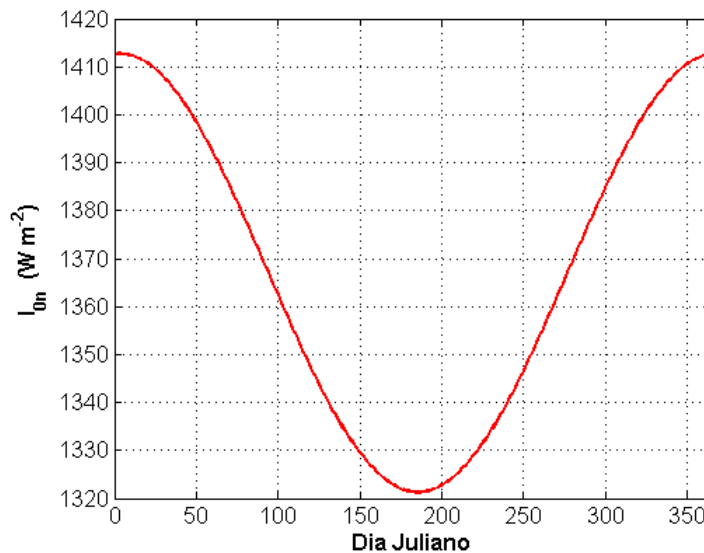


Figura A.4 Variación de la irradiancia solar extraterrestre a lo largo del año

#### A.4 Índice de claridad ( $k_t$ )

El índice de claridad ( $k_t$ ) define como el cociente entre los valores de irradiancia solar registrados en la superficie terrestre y los correspondientes valores de irradiancia solar extraterrestre, ambas para un mismo periodo temporal. Los valores de  $k_t$  se obtienen a partir de la irradiancia solar, aplicando la siguiente expresión:

$$k_t = \frac{G}{I_0} \quad (\text{A.11})$$

donde  $G$  es la irradiancia global que llega a la superficie terrestre sobre superficie horizontal e  $I_0$  es la irradiancia extraterrestre sobre superficie horizontal, ambas para el mismo periodo temporal.

#### **A.4 Componente solar atenuada ( $lc$ )**

La componente solar atenuada o *lost component* ( $lc$ ) se obtiene mediante la diferencia entre el valor de irradiancia solar extraterrestre y el valor correspondiente de la irradiancia global solar recibida en la superficie terrestre sobre superficie horizontal para un periodo determinado. Se obtiene utilizando la siguiente ecuación:

$$lc(t) = I_0(t) - G(t) \quad (\text{A.12})$$

donde la serie temporal  $lc$  representa la energía solar absorbida o reflejada a su paso por la atmósfera para el intervalo de tiempo  $t$ .

#### **A.6 Interacción radiación-atmósfera**

La radiación solar está compuesta por diferentes longitudes de onda. En este trabajo se considera aproximadamente igual a la emitida por un cuerpo negro, cuya temperatura es de 5760 K. El espectro de radiación que alcanza el límite de la atmósfera está compuesto por longitudes de onda ( $\lambda$ ) desde las 0.28 a las 5  $\mu\text{m}$  y suele subdividirse en tres regiones, ultravioleta ( $\lambda < 0.33 \mu\text{m}$ ), visible ( $0.33 < \lambda < 0.76 \mu\text{m}$ ) e infrarrojo ( $\lambda > 0.76 \mu\text{m}$ ).

En el espacio exterior no se produce pérdida de radiación por interacción con medio material alguno, tan solo atenuación debido a la Ley del cuadrado de la distancia. Sin embargo, a su paso por la atmósfera, la radiación solar sufre diversos procesos de

reflexión, atenuación y difusión como resultado de su interacción con los distintos componentes atmosféricos: aerosoles, nubes y moléculas de ozono, dióxido de carbono, oxígeno, vapor de agua, etc. Los principales efectos atmosféricos sobre la irradiancia solar son los siguientes:

- Disminución de la energía que se recibe a nivel del suelo debido a la interacción con los componentes atmosféricos.
- Modificación de las características espectrales de la irradiancia solar.
- Modificación de la distribución espacial de la irradiancia recibida en la superficie terrestre.

La *reflexión* de la irradiancia solar se debe principalmente a la interacción con las nubes y partículas en suspensión. La *absorción* de la irradiancia solar, debido los componentes atmosféricos, es responsable de la disminución de aproximadamente un 20% de la energía solar incidente. Los principales componentes que producen la atenuación son el ozono, el vapor de agua y el dióxido de carbono.

La *difusión* (o *scattering*) produce la atenuación de la radiación solar que llega al límite superior de la atmósfera, haciendo que ésta se distribuya en todas las direcciones. Los componentes atmosféricos que producen este efecto son el vapor de agua, los aerosoles y las componentes moleculares. El efecto de scattering está relacionado directamente con el tamaño del constituyente y su concentración. El tamaño medio de la partícula se define a partir del coeficiente adimensional  $\Theta$ :

$$\Theta = 2\pi q \lambda^{-1} \tag{A.13}$$

siendo  $q$  el tamaño de la partícula y  $\lambda$  la longitud de onda incidente. Se pueden distinguir tres tipos de difusión:

- *Difusión de Rayleigh*. Se origina cuando la longitud de onda de la irradiancia solar es mayor que la dimensión de las partículas responsables ( $\Theta \ll 1$ ). Este proceso se produce por moléculas de oxígeno y nitrógeno. El *scattering* de

Rayleigh es proporcional a  $\lambda^{-4}$ , por lo tanto, afecta a longitudes de onda cortas y es responsable del color azul del cielo. Este fenómeno se da principalmente en las capas altas de la atmósfera.

- Difusión de Mie. Se produce cuando la longitud de onda de la irradiancia solar tiene el mismo orden de magnitud que las moléculas causante del efecto ( $\Theta < 50$ ). Su causa fundamental se debe al vapor de agua, polvo y aerosoles. Tiene efecto sobre todas las longitudes de onda del canal visible y se da en las capas bajas de la atmósfera.
- Difusión no selectiva. Se produce cuando la longitud de onda es menor que la dimensión de las partículas ( $\Theta \gg$ ). Este efecto es causado principalmente por las gotas de agua que forman las nubes o nieblas.

Integrando a lo largo de todo el espectro radiativo la intensidad o potencia de la irradiancia solar, la energía solar que llega a la superficie terrestre dependerá del espesor de la capa atmosférica que tienen que atravesar los rayos antes de alcanzar la superficie terrestre, y de la concentración de componentes moleculares y partículas en suspensión que encuentra en su camino. La descripción física de la interacción de la irradiancia solar con la atmósfera no resulta un problema trivial y aparece ampliamente tratada en trabajos clásicos [Kondratyev, 1969] y en otros textos más actuales [Iqbal, 1983]. Una aproximación, adecuada a la resolución de este problema, consiste en la parametrización de las principales características atmosféricas según una serie de magnitudes que se describen a continuación.

La irradiancia solar a su paso por la capa atmosférica atraviesa un espesor variable. La *masa óptica relativa del aire* ( $m$ ) cuantifica la longitud del camino óptico que recorre la radiación solar. Este valor se puede estimar ignorando la curvatura terrestre y suponiendo una atmósfera uniforme con un índice de refracción igual a la unidad, mediante la siguiente expresión [Kasten y Young, 1989]:

$$m = (p / p_0) \left[ \sin \alpha^* + 0.50572 (57.29578 \alpha^* + 6.07995)^{-1.6364} \right]^{-1} \quad (\text{A.14})$$

siendo  $\alpha^*$  la *altura solar verdadera*, es decir, la altura solar corregida por los efectos de la refracción atmosférica, obtenida a partir de la siguiente ecuación:

$$\alpha^* = \alpha + 0.061359 \frac{0.1594 + 1.1230\alpha + 0.065656\alpha^2}{1 + 28.9344\alpha + 277.3971\alpha^2} \quad (\text{A.15})$$

El coeficiente adimensional ( $p/p_0$ ) es una corrección de la presión atmosférica debido a la altitud sobre el nivel del mar ( $z$ ) del lugar bajo estudio. El valor de ( $p/p_0$ ) se calcula de forma aproximada utilizando la siguiente ecuación:

$$(p/p_0) = \exp(-z/8400) \quad (\text{A.16})$$

donde  $z$  es la coordenada vertical referida al nivel del mar.

La influencia de los diferentes constituyentes atmosféricos puede estimarse a partir de la comparación entre la profundidad óptica, registrada en un instante determinado, y la profundidad teórica, para un cielo totalmente claro y seco. Este valor se conoce como *profundidad óptica de Rayleigh* ( $\delta_r$ ). A partir de una masa óptica relativa del aire ( $m$ ), se puede obtener la profundidad óptica de Rayleigh [Louche A. et al., 1991; Kasten, 1996]:

$$\delta_r(m) = \begin{cases} (6.6296 + 1.7513m - 0.1202m^2 + 0.0065m^3 - 0.00013m^4)^{-1} & \text{para } m < 20 \\ (10.4 + 0.718m)^{-1} & \text{para } m \geq 20 \end{cases} \quad (\text{A.17})$$

### **Ley de Bourguer-Lambert-Beer**

La ecuación que rige la atenuación de la radiación electromagnética al atravesar un medio se denomina ley de Bourguer-Lambert-Beer. Esta, aplicada a la radiación solar espectral al atravesar la atmósfera se puede expresar de la siguiente forma

$$I_n(\lambda) = I_0(\lambda) \exp(-k(\lambda)m) = I_0(\lambda) \tau(\lambda) \quad (\text{A.18})$$

donde

$I_n(\lambda)$  : irradiancia normal en la superficie de la tierra

$I_0(\lambda)$ : irradiancia extraterrestre en el límite de la atmósfera

$k(\lambda)$ : espesor óptico total en la dirección cenital

m: masa óptica relativa de aire

$\tau(\lambda)$ : transmitancia espectral total

### **Absorción de radiación solar en la atmósfera**

La radiación solar al atravesar la atmósfera es en parte absorbida por los constituyentes atmosféricos, aumentando la energía interna de los mismos, y por consiguiente su temperatura. Los procesos de absorción tienen lugar a determinadas longitudes de onda.

En las regiones ultravioleta y visible del espectro, los principales constituyentes que contribuyen a la absorción son oxígeno molecular, nitrógeno molecular, oxígeno atómico, nitrógeno atómico y ozono.

El nitrógeno y oxígeno atómicos absorben rayos X y otras radiaciones de corta longitud de onda, de forma continua hasta los 85 nm. Estos dos gases, se encuentran en las capas más altas de la atmósfera, y por tanto impiden que la radiación de longitud de onda inferior a los 85 nm alcance las capas bajas de la atmósfera. El nitrógeno y oxígeno moleculares presentan, por debajo de los 200 nm, una serie de bandas de absorción que se solapan. Con ello, la acción conjunta de estos gases produce que no llegue a la superficie terrestre radiación del ultravioleta lejano (por debajo de los 200 nm), que es muy perjudicial para los seres vivos.

El espectro de absorción del oxígeno molecular entre 260 y 200 nm se denomina banda de Herzberg, que es muy débil y se solapa con una banda mucho más fuerte del ozono. Sin embargo, es importante por el papel que juega en la formación del ozono. Adyacente a la banda de Herzberg, se encuentra la banda de Schumann\_Runge, que abarca desde 200 nm a 125 nm. La región por debajo de los 100 nm se denomina banda

de Hopfield. En la zona visible-infrarrojo próximo el oxígeno presenta tres bandas centradas en 630nm, 690 nm (banda B) y 760 nm (banda A), que se solapan con una extensa banda del ozono (la banda de Chappuis), de manera que su influencia sobre la radiación solar es pequeña.

El espectro de absorción del nitrógeno molecular se inicia en 145 nm y también se extiende hacia longitudes de onda más cortas. Hay dos bandas de absorción importantes: la banda de Lyman-Birge-Hopfield (145-100 nm) y la banda de Tanaka-Worley (100-80 nm).

El ozono presenta tres bandas de absorción a tener en cuenta en el cálculo de la atenuación de la radiación solar: la de Hartley (entre 210 y 300 nm), la banda de Huggins (300 a 340 nm), y la de Chappuis (entre 440 y 640 nm). Asimismo presenta bandas de absorción en el infrarrojo en 4800, 9600 y 14200 nm, aunque la contribución de estas bandas a la absorción es prácticamente despreciable.

El ozono presenta una banda de absorción muy acentuada entre los 220 nm a 300 nm (banda de Hartley). A continuación se encuentra la banda de Huggins (300-345 nm). En el visible se encuentra la banda de Chappuis (440-750 nm).

La absorción del vapor de agua juega un papel muy importante en la atenuación de la radiación solar. Presenta tres fuertes bandas de absorción en el infrarrojo centradas en 1000, 1400 y 1800 nm. Por encima de los 2300 nm la energía contenida en el espectro solar es muy pequeña (menor del 5%), y la acción combinada de absorción por el vapor de agua y el CO<sub>2</sub> hace que la energía recibida en la superficie terrestre, para estas longitudes de onda, sea muy pequeña.

### **Aproximaciones paramétricas para la transmitancia de la radiación solar**

Es frecuente expresar la ley de Bourguer-Lambert-Beer en función de las transmitancias de cada uno de los componentes atmosféricos: dispersión Rayleigh, dispersión por

aerosoles, absorción por gases moleculares, absorción por vapor de agua y absorción por ozono,

$$I_n(\lambda) = I_0(\lambda) \tau_R(\lambda) \tau_a(\lambda) \tau_g(\lambda) \tau_w(\lambda) \tau_o(\lambda) \quad (\text{A.19})$$

Existen diversas formulaciones y parametrizaciones para estimar las transmisividades correspondientes a cada uno de los constituyentes atmosféricos.

La transmisividad debida a dispersión Rayleigh viene dada por la expresión de Leckner,

$$\tau_R(\lambda) = \exp(-0.008735 \lambda^{-4.08} m) \quad (\text{A.20})$$

La dispersión de radiación solar por aerosoles viene dada por la formula de Angstrom para el coeficiente de turbiedad atmosférica,

$$k_a(\lambda) = \beta \lambda^{-\alpha} \quad (\text{A.21})$$

donde  $\beta$  es el denominado coeficiente de turbiedad de Angstrom, que representa el contenido de aerosoles presentes en la atmósfera en la dirección cenital, y se define como el espesor óptico de aerosoles apara una longitud de onda de 1  $\mu\text{m}$ . En esta expresión los parámetros  $\alpha$  y  $\beta$  se determinan a partir de medidas espectrales a distintas longitudes de onda. La transmisividad debida a la dispersión de los aerosoles queda pues definida por,

$$\tau_a(\lambda) = \exp\left(-\frac{\beta}{\lambda^\alpha} m\right) \quad (\text{A.22})$$

Para la transmitancia espectral por absorción de gases moleculares (CO<sub>2</sub>, O<sub>2</sub>, etc) se puede emplear la siguiente expresión:

$$\tau_g(\lambda) = \exp\left(-1.14a_\lambda \frac{m}{(1+118.93a_\lambda m)^{0.45}}\right) \quad (\text{A.23})$$

donde  $a_\lambda$  son los coeficientes espectrales de absorción de gases moleculares, que están tabulados en diversas referencias.

Para la transmisividad debida a la absorción por vapor de agua se puede emplear la expresión de Leckner,

$$\tau_w(\lambda) = \exp\left(-0.2385C_{w\lambda} w \frac{m}{(1+20.07C_{w\lambda} w m)^{0.45}}\right) \quad (\text{A.24})$$

donde  $C_{w\lambda}$  es el coeficiente espectral de absorción de vapor de agua y  $w$  es la cantidad de agua precipitable en un columna vertical de área unidad.

Finalmente, la atenuación debida al ozono se puede estimar por,

$$\tau_o(\lambda) = \exp(-C_{o\lambda} O_z m) \quad (\text{A.25})$$

donde  $C_{o\lambda}$  es el coeficiente de absorción para el ozono y  $O_z$  el contenido de ozono presente en la atmósfera.

### **A.7 Modelo de cielo claro del European Solar Radiation Atlas (ESRA)**

Los modelos de cielo claro son de gran utilidad en muchas aplicaciones en energía solar. En particular, resultan muy útiles para la evaluación de la radiación solar a partir de imágenes de satélite. Un modelo de cielo claro es básicamente una parametrización para estimar la radiación solar integrada en todo el espectro en un día completamente

despejado. Para días cubiertos o parcialmente cubiertos la estimación de la radiación solar por métodos físicos es muy compleja, ya que se requiere un profundo conocimiento de la morfología de la cubierta nubosa. En el apartado anterior se han presentado expresiones simples para la estimación de la irradiancia directa espectral en días despejados, sin embargo, su integración en todo el ancho de banda requiere del conocimiento de diversos parámetros que dependen de la longitud de onda, tales como los coeficientes de absorción de diversos constituyentes. Existen modelos espectrales bastante recientes para la estimación de la radiación global espectral en días despejados, como es el caso de SMARTS, sin embargo, el uso de modelos de cielo claro integrados es más sencillo para multitud de aplicaciones que no requieren el conocimiento del espectro.

El modelo del ESRA (European Solar Radiation Atlas) es una parametrización que solo requiere como variable a determinar el factor de turbiedad de Linke, que debe ser medido de forma experimental. Existen tablas del índice de turbiedad de Linke para toda Europa, que permiten el uso de este modelo de cielo claro con cierta precisión.

### **El coeficiente de turbiedad de Linke**

El *coeficiente de turbiedad de Linke* ( $T_L$ ) es un simple parámetro para expresar la atenuación de la radiación solar debido a los aerosoles y el vapor de agua [Linke, 1922; Polo et al., 2008; Polo et al., 2009]. Representa el nivel de transparencia de la atmósfera y cuantifica de forma adecuada los efectos de absorción y esparcimiento producidos por la atmósfera sobre la radiación solar. Fue desarrollado por Linke en el año 1922. Propone expresar el espesor óptico de una atmósfera sin nubes, como el producto de dos términos, el espesor óptico de una atmósfera libre de vapor de agua y aerosoles ( $k_C$ ) y el coeficiente de turbiedad de Linke ( $T_L$ ), de manera que

$$I_n = I_0 \exp(-k_C T_L m) \quad (\text{A.26})$$

Linke definió  $k_C$  como el espesor óptico integrado para una atmósfera ideal libre de nubes, aerosoles y vapor de agua, y calculó su valor a partir de consideraciones teóricas,

---

validándolo para una atmósfera pura y seca. De esta forma,  $T_L$  representa el número de atmósferas limpias ideales que es necesario para producir la atenuación observada en la radiación, debido a los efectos conjuntos del vapor de agua y los aerosoles.

La Tabla A.1 presenta distintos valores habituales para este índice bajo diferentes condiciones atmosféricas.

Tabla A.1. Valores frecuentes del factor de turbidez de Linke [ESRA, 2000]

Tipo de atmósfera	$T_L$
Muy clara (despejado, nivel humedad bajo, pocas partículas en suspensión)	~ 2
Clara y cálida	~ 3
Húmeda y cálida	4 ~ 6
Con polución	> 6

Este valor se puede obtener a partir de medidas experimentales; aunque debido a la escasez de éstas, se consigue mediante ajustes empíricos. Aunque en la actualidad no se ha modificado la definición del coeficiente de turbiedad de Linke, sus valores se han visto modificados debido a la mejora en la instrumentación y en la precisión de las medidas.

**Radiación directa sobre superficie horizontal para cielo claro**

En el modelo de cielo claro del ESRA, la radiación global se divide en sus dos componentes, radiación directa de cielo claro,  $B_c$ , y radiación difusa de cielo claro,  $D_c$ . Cada componente se determina de forma separada.

La irradiancia directa para cielo claro, en unidades de  $W\ m^{-2}$ , viene dada por la siguiente expresión,

$$B_c = I_{CS} \varepsilon \sin(\alpha) \exp(-0.8662 T_L (AM2) m k_r(m)) \tag{A.27}$$

donde  $I_{CS}$  es la constante solar ( $1367\ W\ m^{-2}$ ),  $\varepsilon$  es la corrección de la distancia real tierra-sol respecto de su valor medio  $(r/R)^2$ ,  $\alpha$  es la altura solar,  $T_L(AM2)$  es el

coeficiente de turbiedad de Linke para una masa de aire igual a 2,  $m$  es la masa óptica relativa de aire y  $k_R(m)$  es el espesor óptico de Rayleigh.

**Radiación difusa sobre superficie horizontal para cielo claro**

La irradiancia difusa para cielo claro viene dada por:

$$D_C = I_{CS} \varepsilon T_{rd}(T_L(AM2)) F_d(\alpha, T_L(AM2)) \quad (A.28)$$

Esta ecuación representa la irradiancia difusa para una elevación solar dada como el producto de dos contribuciones, la función de transmisión difusa en el cenit ( $T_{rd}$ ) y la función angular difusa ( $F_d$ ).

$$T_{rd}(T_L(AM2)) = -1.5843 \cdot 10^{-2} + 3.0543 \cdot 10^{-2} T_L(AM2) + 3.797 \cdot 10^{-4} (T_L(AM2))^2 \quad (A.29)$$

$$F_d(\alpha, T_L(AM2)) = A_0 + A_1 \sin \alpha + A_2 (\sin \alpha)^2 \quad (A.30)$$

donde los coeficientes  $A_i$  solo dependen del índice de turbiedad de Linke

$$\left\{ \begin{array}{l} A_0 = .264 - .0616 T_L(AM2) + .00314 (T_L(AM2))^2 \\ A_1 = 2.04 + .0189 T_L(AM2) - .01116 (T_L(AM2))^2 \\ A_2 = -1.30 + 0.0392 T_L(AM2) + .00850 (T_L(AM2))^2 \end{array} \right\} \quad (A.31)$$

Con estas expresiones se puede estimar la radiación global instantánea para cielo claro. La determinación de valores horarios o diarios requiere de la integración numérica de estas expresiones para las distintas elevaciones solares.

### **A.8 Baseline Surface Radiation Network (BSRN)**

La determinación de una climatología global del balance radiativo sobre la superficie de la Tierra es fundamental para entender el sistema climático de la Tierra, la variabilidad climática y el cambio climático resultado de la influencia humana. Las estimaciones globales de balance radiativo en la superficie no pueden inferirse directamente a partir de las observaciones de satélite sin medidas sobre superficie de alta calidad en varios emplazamientos situados en regiones climáticas diversas para la calibración y validación. Las observaciones de largo plazo también son necesarias para evaluar las tendencias dentro de regiones climáticas. Tales medidas son esenciales para evaluar tratamientos teóricos de transferencias radiativa en la atmósfera, para verificar los cálculos de los modelos climáticos, y para el estudio de tendencias de la radiación en superficie en escalas menores a las normalmente asociadas a las regiones climáticas.

Para satisfacer estos requerimientos, el World Climate Research Programme (WCRP), conjuntamente con la World Meteorological Organization (WMO), el International Council of Scientific Unions (ICSU) y el Intergovernmental Oceanographic Commission (IOC) de la UNESCO iniciaron la Baseline Surface Radiation Network (BSRN). El objetivo de esta red es proveer medidas continuas, de largo plazo, con una alta frecuencia de muestreo y según el estado del arte de los estándares de los procedimientos de medida, calibración y precisión radiométrica.

En la actualidad, existe un rango amplio de estaciones dentro de la BSRN en zonas tan diversas como la zona ártica, bosques de latitudes medias y áreas llanas, regiones de alta montañas, bosques tropicales, desiertos, islas tropicales (ver Figura A.5 donde se muestra la distribución de estaciones existentes y planeadas).

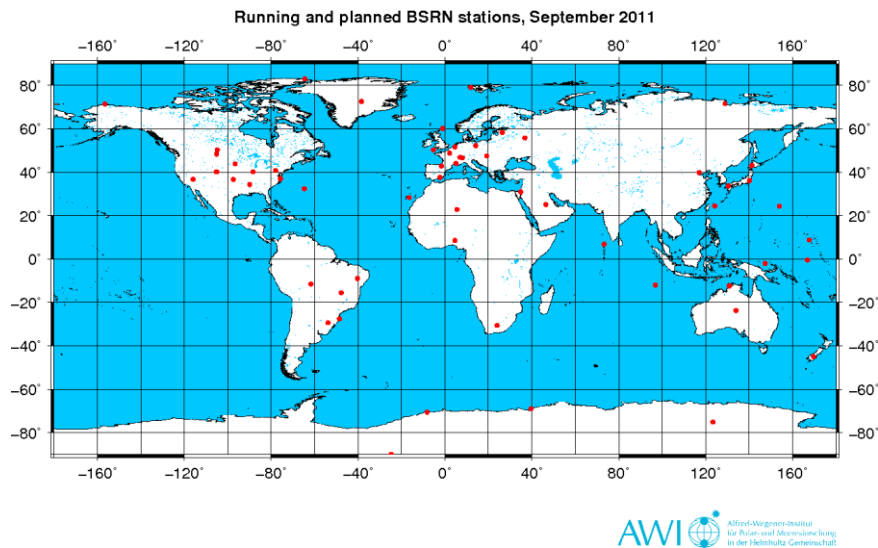


Figura A.5 Estaciones BSRN planificadas y en funcionamiento

Cada estación, o grupo de estaciones, dentro de una jurisdicción dada opera bajo la supervisión de un científico cualificado con experiencia en la medida de la radiación. Dentro de esta estructura, algunos emplazamientos se establecen específicamente con el propósito de realizar medidas de tendencias climáticas locales determinadas y proveer valores en superficie para las observaciones de satélite. Las estaciones BSRN forman parte de esta red no sólo por sus requerimientos de alta calidad, sino además a su importancia con respecto al clima o la geografía que representan.

La BSRN se desarrolló en un principio para satisfacer las necesidades de la comunidad que estudia el cambio climático y para realizar la validación de las estimaciones de radiación a partir de imágenes satélite. El plan inicial para establecer una red global de estaciones radiométricas fue desarrollado por el grupo de trabajo del WCRP sobre flujos radiativos (WGRF) en 1989, y fue redefinido en dos congresos posteriores sobre la implementación de la BSRN, el primero en Washington, DC, Estados Unidos en diciembre de 1990, y el segundo en Davos, Suiza en agosto de 1991. Los objetivos oficiales que se establecieron fueron los siguientes:

- Proporcionar datos para la calibración de las estimaciones a partir de satélite de los flujos radiativos en la superficie terrestre y la transferencia de radiación a través de la atmósfera.
- Monitorizar las tendencias regionales en los flujos de radiación en la superficie.

Además de la importante contribución a la investigaciones del clima global hecha por la BSRN, es de importancia el hecho que los países que asumen las responsabilidades de operar una estación BSRN se benefician significativamente de tener una estación de medida de la radiación en tierra de referencia, especialmente en el contexto de las naciones que dedican importantes esfuerzos a hacer uso del recurso de energía renovable y también, de algún modo, en la mejora de la producción agrícola. Las medidas de una estación BSRN son también fundamentales en la monitorización nacional y el estudio de las variaciones climáticas regionales, y en la evaluación de las implicaciones económicas asociadas. En países donde las redes radiométricas ya existen, la instrumentación y los procedimientos operacionales desarrollados para la BSRN pueden ser utilizados para mejorar las medidas actuales y mejorar la trazabilidad de la calibración al World Radiation Centre mediante el uso de las técnicas e instrumentos de la BSRN. En resumen, los conjuntos de datos BSRN tienen un amplio rango de aplicaciones más allá de la investigación climática.

Los objetivos específicos que la BSRN estableció en el informe de la reunión de Washington, D.C. son los siguientes:

- Medir las componentes de la radiación en superficie en emplazamientos estratégicos con una fiabilidad y precisión suficiente para revelar las tendencias de largo plazo.
- Obtener medidas simultáneas de los constituyentes atmosféricos tales como nubes, vapor de agua, ozono y aerosoles que afectan a la radiación solar en la superficie y en el borde superior de la atmósfera.
- Asegurar la uniforme adhesión a los mayores estándares alcanzables de procedimiento, precisión y calibración a través de la red BSRN.

Las actividades asociadas y los objetivos de investigación generales de la BSRN son los siguientes:

- Caracterización del emplazamiento: adquisición de información cuantitativa sobre propiedades tales como la naturaleza de la superficie, tipo y cubierta

nubosa media, aerosoles, etc, que caractericen el lugar para aplicaciones de satélite.

- Medidas de irradiancia infrarroja: adelanto en el estado del arte para obtener medidas precisas de irradiancia descendiente y medidas de irradiancia para cumplir los estándares en las medidas del balance de radiación en la superficie.
- Reflectancia en la superficie: desarrollo de métodos para medir la reflectancia medida en superficie en un área amplia (ejemplo 20 X 20 km) para recopilar información in situ que ayude en la validación de las medidas de detección remota.
- Inhomogeneidades atmosféricas: estudios enfocados en la mejora del conocimiento y la medida de las propiedades radiativas de nubes no homogéneas e intervalos nubosos.
- Medidas especiales: desarrollo de instrumentación económica y métodos para la medida espectral ultravioleta e infrarroja del balance radiativo que ayuden a la mejora del diseño y validación de los algoritmos de satélite para la determinación de los flujos radiativos.
- Mejora de la instrumentación: mejora del diseño y el rendimiento de la instrumentación “estándar” tal como foto sensores y piranómetros, y desarrollar instrumentación de detección remota sofisticada para perfeccionar las observaciones de nubes.

### **Procedimientos de control de datos radiométricos definidos por la BSRN**

Los procedimientos de control de calidad sobre los datos recomendados por la BSRN son los siguientes:

- Procedimiento 1: límites físicamente posibles.
- Procedimiento 2: límites extremadamente raros.
- Procedimiento 3: límites entre medidas.
- Procedimiento 4: comparación con un modelo.
- Procedimiento 5: control visual de las serie temporales de datos.

Cada uno de estos procedimientos consiste en un conjunto de subprocedimientos, que a su vez especifica un cierto límite para las diferentes componentes de la radiación solar. Los procedimientos 4 y 5 están en la actualidad bajo desarrollo y por lo tanto no implementados aún para su uso en la estimación de la calidad de los datos. A continuación se describen de forma detallada cada uno estos procedimientos para la estimación de la calidad de las medidas.

▪ **Procedimiento 1 (Límites físicamente posibles)**

Este es el procedimiento más sencillo y fundamental de todos por lo que puede ser aplicado, tanto a magnitudes de radiación independientemente como a variables meteorológicas.

La función de este procedimiento es la de marcar errores de medida extremadamente grandes causado por razones diversas, desconexión del instrumento, fallo del instrumento o del DAS,.... Además nos servirá para identificar errores aleatorios introducidos durante el post procesado y manejo de los datos. Los valores entre los límites superior e inferior que se muestran en la tabla siguiente se definen como límites físicamente posibles para la magnitud medida.

Tabla A.2. Filtros físicamente posibles

Límite inferior	Flujo Radiativo	Límite superior
0	Irradiancia global: medida con piranómetro... <i>DSGL2</i>	$I_{cs}$
0	Irradiancia difusa..... <i>DSDFS</i>	$I_{0+10}$
0	Irradiancia directa..... <i>DSDIR</i>	$I_{cs}$
0	Irradiancia reflejada (global: onda corta)..... <i>SW USR</i>	$I_0$
50 W/m <sup>2</sup>	Irradiancia descendente de onda larga..... <i>LW DL</i>	700 W/m <sup>2</sup>
50 W/m <sup>2</sup>	Irradiancia ascendente de onda larga..... <i>LW UL</i>	700 W/m <sup>2</sup>

Los valores límite, se pueden agrupar en dos tipos:

- Constantes: corresponden con los usados para las componentes de la radiación solar global, directa normal y para la radiación solar en la banda espectral del infrarrojo. Los valores límite para el IR, (50 W/m<sup>-2</sup> y superior 700 W/m<sup>-2</sup>) corresponden a las temperaturas de un cuerpo negro de -100 y 60° C

respectivamente. Estos valores de temperatura se definen a partir de las condiciones climáticas presentes actualmente en la tierra.

- Basados en funciones: se aplican al cálculo de la radiación difusa y a la radiación de onda corta (global) reflejada y es función de la irradiancia recibida por una superficie plana tangente a la superficie terrestre en el límite superior de la atmósfera ( $I_0$ ).

▪ **Procedimiento 2 (Valores extremadamente raros)**

Este test se aplica sobre los datos inmediatamente después del anterior y sus límites son mucho más estrechos que los definidos por el procedimiento 1, de forma que es posibles datos erróneos que hayan escapado al procedimiento 1. Los valores de radiación que estén fuera del rango definido por el procedimiento 2 pueden presentarse durante un corto periodo de tiempo y bajo condiciones extremadamente anormales. Como se sugiere en [Dutton E.G., 2002] los límites deben ser caracterizados para encajar mejor a los resultados extremadamente raros que se puedan presentar en una localización particular. En concreto, en la PSA después de ejecutar los filtros se ha observado que los valores registrados por los pirgeómetros ( $DL$  y  $UL$ ) no superaban los filtros por lo que se han tenido que modificar. A pesar de esto la evaluación del procedimiento 2 es aplicado por la BSRN a las estaciones en general y no de una en una de forma independiente. En la tabla siguiente se definen los diferentes límites para este procedimiento.

Tabla A.3. Filtros extremadamente raros

Flujo Radiativo	Límite superior
Irradiancia global: medida con piranómetro..... <i>DSGL2</i>	$I_0$ si $\theta_z < 80^\circ$ $I_0+0.56*(\theta_z-93.9)^2$ si $\geq 80^\circ$
Irradiancia global reflejada..... <i>SW USR</i>	$0.95 \cdot DSGL2$
Irradiancia difusa..... <i>DSDFS</i>	$700 \text{ W/m}^2$
Irradiancia directa..... <i>DSDIR</i>	$I_0 \cdot 0.9^m$
Radiación infrarroja descendente..... <i>LW DL</i>	$UL$
Radiación infrarroja ascendente..... <i>LW UL</i>	$DL+10$

Los límites superiores anteriores se definen bajo las siguientes suposiciones:

- En latitudes cercanas a los trópicos cuando las nubes dispersas pasan frente al sol la parte baja de la nube refleja parte de la radiación y se producen valores de radiación global superiores a la irradiancia  $I_{sc}$ . Este fenómeno se le conoce comúnmente como efecto lupa. Si la radiación global es mayor que el límite superior para un periodo considerable, los valores deben ser revisados.
- El albedo mayor que se puede registrar sobre la superficie terrestre corresponde a situaciones de nieve fresca, cielos totalmente cubiertos y nieblas. En todos estos casos el albedo puede llegar a superar valores de 0,93. Por lo tanto se establece como límite superior para la radiación reflejada de onda corta ( $USR$ ) 0,95 veces la radiación global que llega a la superficie terrestre. En caso de que se supere este límite puede deberse a que el instrumento este totalmente cubierto de nieve.
- Para la componente difusa de la radiación solar no se han registrado hasta la fecha valores mayores a  $700 \text{ W/m}^2$  en las diferentes estaciones de la BSRN. Sin embargo si se obtienen valores mayores puede ser un nuevo máximo o que el cristal de la cúpula del pirgeómetro esté cubierto con una fina capa de escarcha.
- La difusión dominante en las capas altas de la atmósfera bajo condiciones de cielo claro es la difusión de Rayleigh y representa un 9% de la extinción de la radiación directa al nivel del mar. Por lo tanto el coeficiente de transmisión de la radiación solar raramente registrará valores mayores a 0.9 sobre el nivel del mar. Aunque la altura de la PSA es de 500 metros sobre el nivel del mar es de considerar la gran cantidad de aerosoles, principalmente partículas de vapor de agua debido a su proximidad al mar, que atenúan la componente directa de la radiación solar por procesos de difusión de Mie [Luis Zarzalejo Tirado, 2005].
- Para los límites definidos por la BSRN sobre el balance energético de la radiación de onda larga se deduce que es difícil que se presente un balance neto positivo. El caso en el que se puede dar un balance positivo es en el que una capa más caliente que ha ascendido, como puede ser una nube muy caliente, se mueve sobre un superficie que se mantiene fría, como nieve, hielo o un glaciar. El balance neto raramente supera el límite de  $30 \text{ W/m}^2$ . Si se supera dicho límite es necesario revisar las cúpulas de los instrumentos que miran hacia arriba.

Para que se pueda aplicar este segundo procedimiento, es necesario tener medidas de irradiancia solar reflejada (de onda corta y onda larga) y de irradiancia solar descendente de onda larga.

▪ **Procedimiento 3 (Comparación entre medidas)**

Este procedimiento define límites aún más estrechos para identificar valores que hayan podido escapar a los procedimientos 1 y 2. El principio de este test es el de comparar instrumentos que miden la misma magnitud.

Los subprocesos están basados en relaciones empíricas generales de diferentes medidas y por lo tanto es posible adaptarlas a las condiciones locales que existan en el emplazamiento de la estación, al igual que sucedía con el procedimiento anterior. En la tabla siguiente se detallan los límites para el procedimiento 3.

Tabla A.4. Filtros para la comparación entre medidas

Límite inferior	Flujo Radiativo	Límite superior
$0.7 \sigma T^4$	Radiación descendente LW <i>DL</i>	$\sigma T^4$
$\sigma(T-10)^4$	Radiación ascendente LW <i>UL</i>	$\sigma(T+10)^4$
(DSGL2-DSDFS)-50 W/m <sup>2</sup>	DSDIR · cos $\theta_z$	(DSGL2-DSDFS)+50 W/m <sup>2</sup>
DSDIR · cos $\theta_z$ -50 W/m <sup>2</sup>	DSGL2-DSDFS	DSDIR · cos $\theta_z$ +50 W/m <sup>2</sup>

Donde *T* es la temperatura del aire en Kelvin a la misma altura que el instrumento de irradiancia solar descendente de onda larga.

▪ **Procedimiento 4 (Comparación con un modelo de cielo claro)**

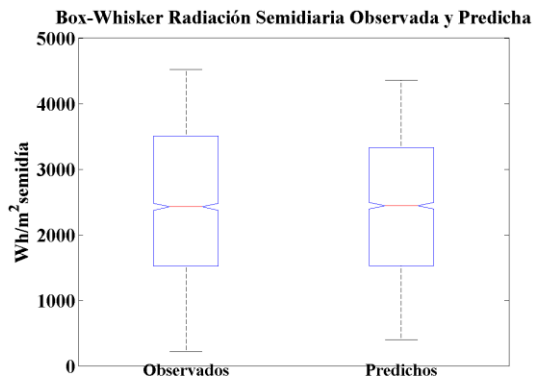
Este procedimiento se puede utilizar para comparar las medidas de días de cielos despejados con modelos que simulan las diferentes componentes de la radiación solar bajo condiciones de cielo claro, también conocidos como modelos de cielo claro [Iqbal, 1983] o modelos basados en el factor de turbidez de Linke [Linke, 1922]. Normalmente, las diferencias entre ambos valores no deben ser superiores a 10W/m<sup>2</sup>. Un valor mayor puede ser debido a que existen observaciones insuficientes de los componentes atmosféricos para caracterizar el cielo de la zona o que el día no sea totalmente despejado. Para aplicar este procedimiento se debe estar seguro de que los modelos de

cielo claro funcionan correctamente en la localización concreta de la estación, ya que de hecho, pueden encontrarse importantes discrepancias entre éstos y las medidas.

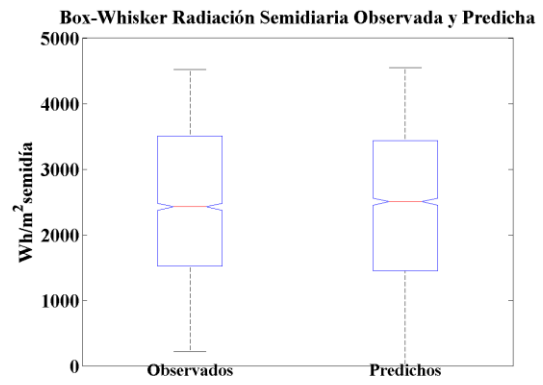
## APÉNDICE B – ANÁLISIS GRÁFICO DE LOS MODELOS FINALES

En este apéndice se presentan los resultados, en forma gráfica, obtenidos a partir de la aplicación de los modelos finales de predicción para cada estación radiométrica. La información que a continuación se presenta para cada horizonte temporal de predicción es la siguiente:

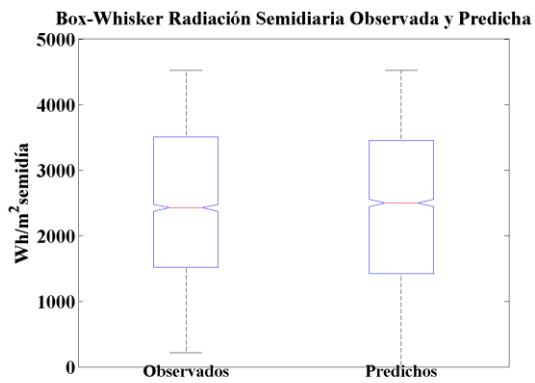
- Diagrama Box-Whisker de la serie temporal de datos radiométricos semidiarios observados y datos predichos.
- Figura con la serie temporal de datos radiométricos semidiarios observados y datos predichos.
- Figura con la serie temporal de errores absolutos de predicción.
- Diagrama de dispersión de serie la temporal observada frente a la serie temporal prevista.
- Figura con la radiación solar semidiaria observada frente al error relativo de predicción.
- Histograma de los errores de predicción y la función de densidad muestral de los errores. La integral de la función de densidad de probabilidad muestral sobre la línea real es igual a uno.



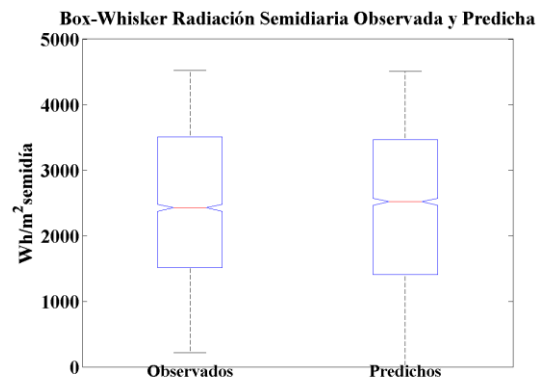
Horizonte de predicción 1 semidia



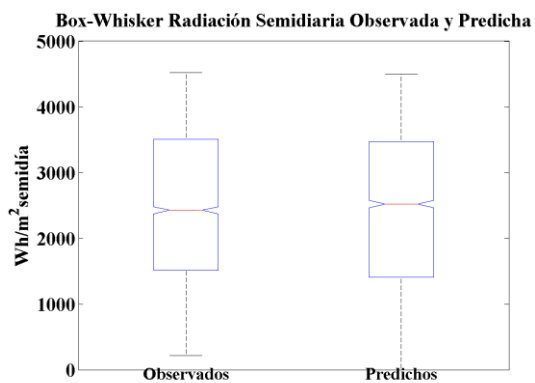
Horizonte de predicción 2 semidías



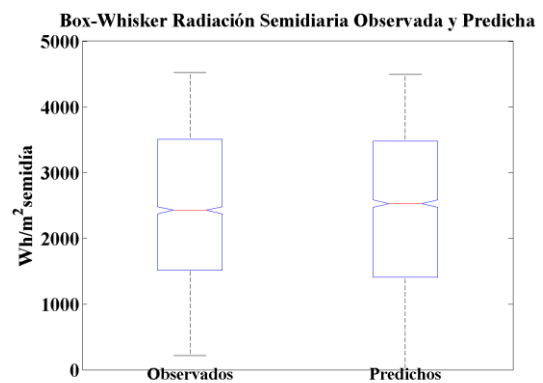
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

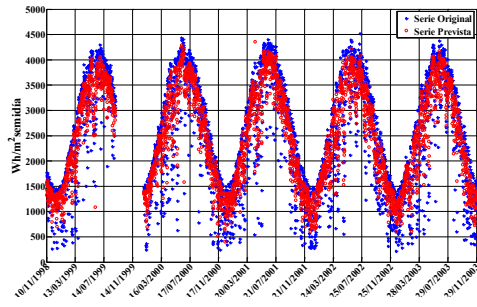


Horizonte de predicción 5 semidías

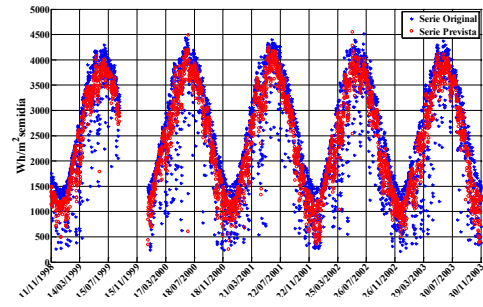


Horizonte de predicción 6 semidías

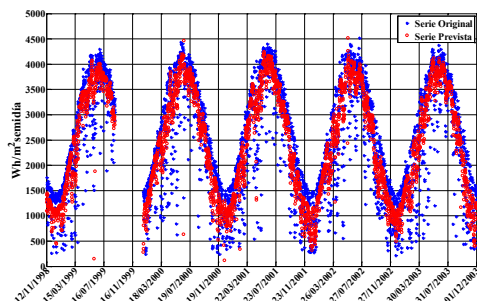
Figura B.1 Murcia: diagrama Box-Whisker de la serie temporal de datos radiométricos semidiarios observados y datos predichos



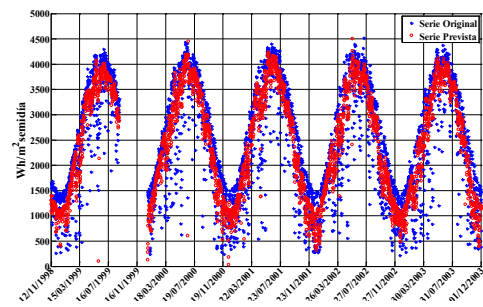
Horizonte de predicción 1 semidia



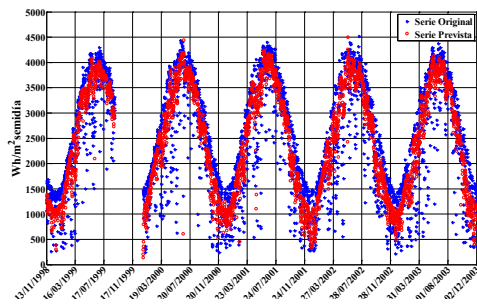
Horizonte de predicción 2 semidías



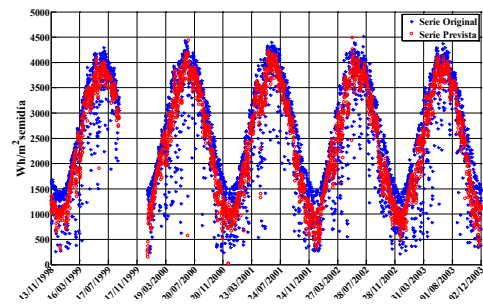
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

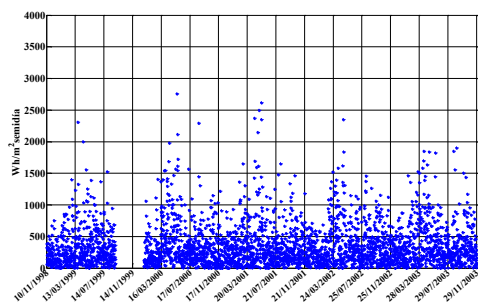


Horizonte de predicción 5 semidías

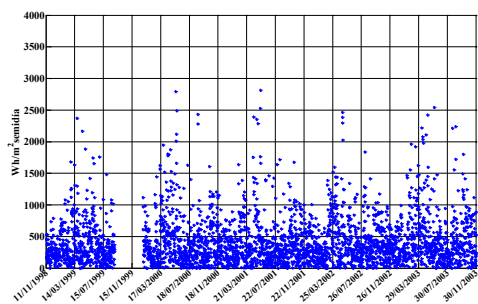


Horizonte de predicción 6 semidías

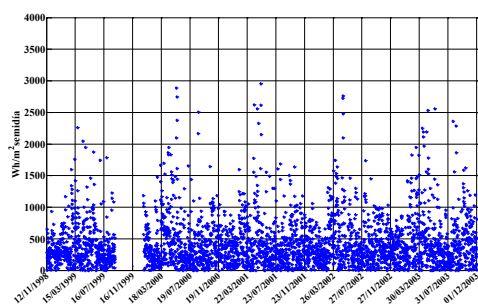
Figura B.2 Murcia: serie temporal de datos radiométricos semidiarios observados y datos predichos



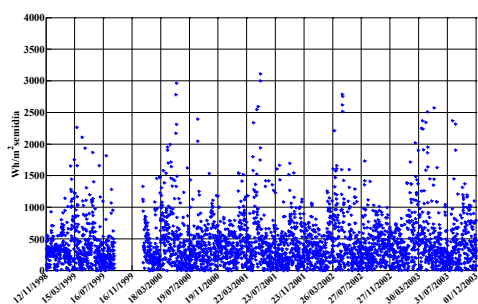
Horizonte de predicción 1 semidía



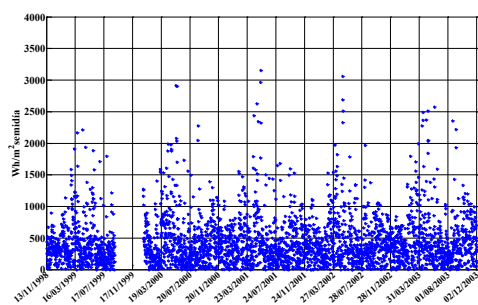
Horizonte de predicción 2 semidías



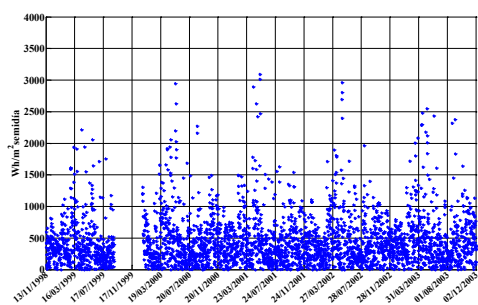
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

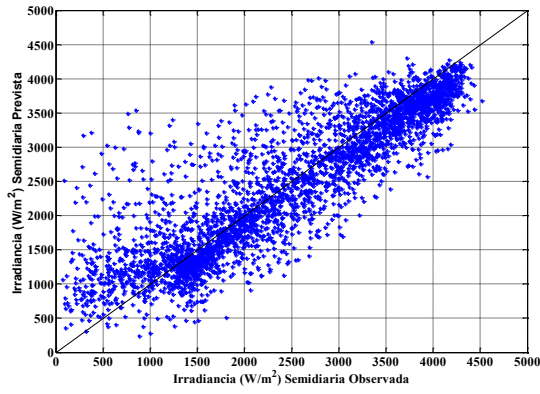


Horizonte de predicción 5 semidías

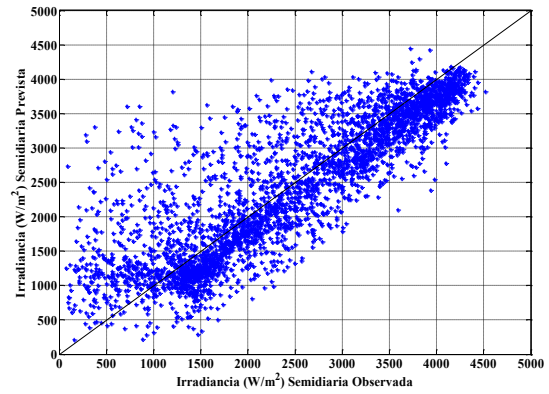


Horizonte de predicción 6 semidías

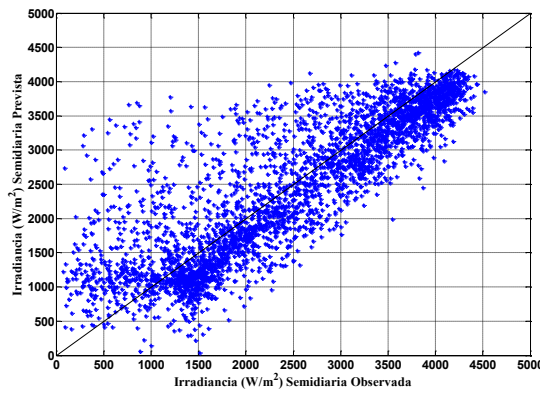
Figura B.3 Murcia: serie temporal de errores absolutos en la predicción



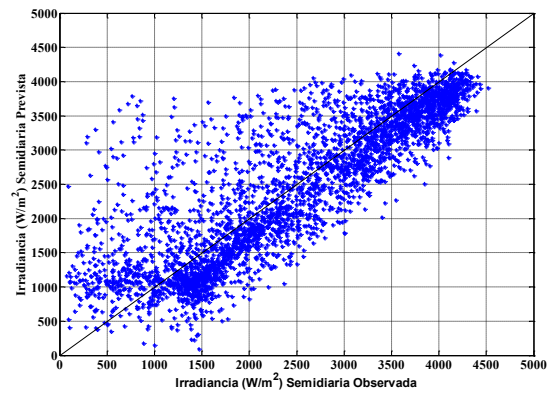
Horizonte de predicción 1 semidía



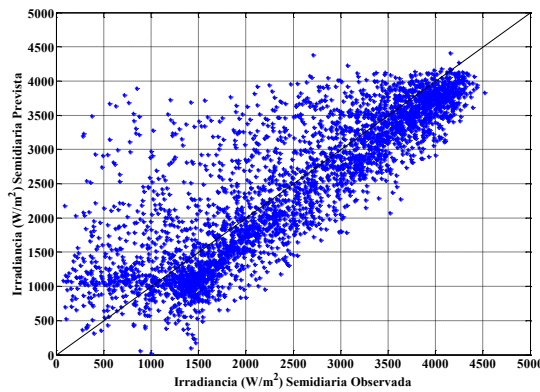
Horizonte de predicción 2 semidías



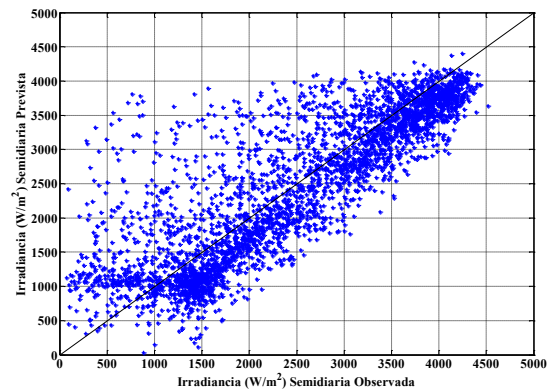
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

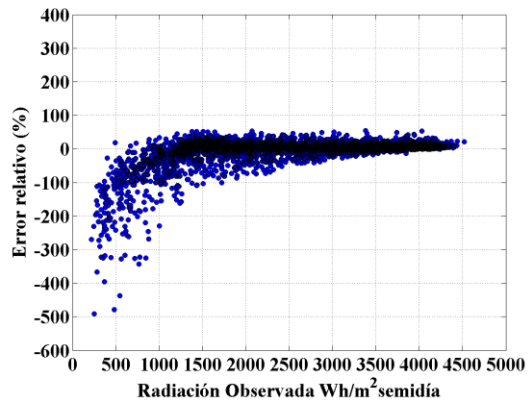


Horizonte de predicción 5 semidías

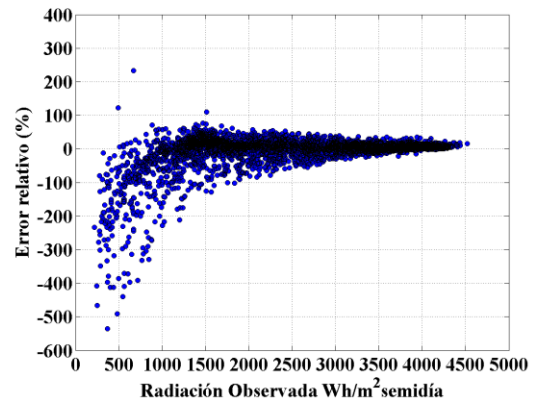


Horizonte de predicción 6 semidías

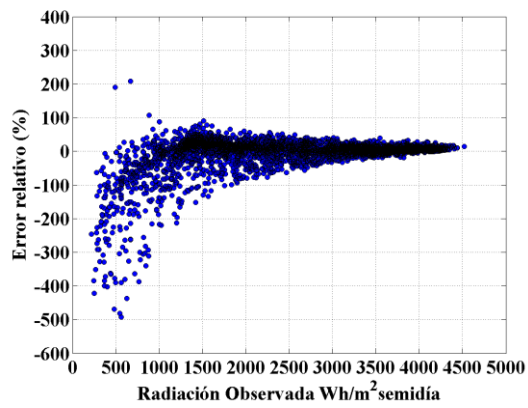
Figura B.4 Murcia: diagrama de dispersión serie temporal original / prevista



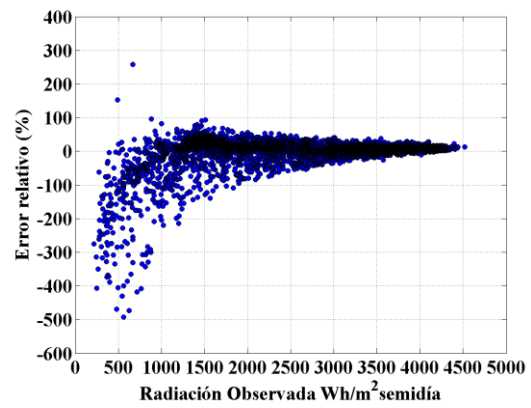
Horizonte de predicción 1 semidía



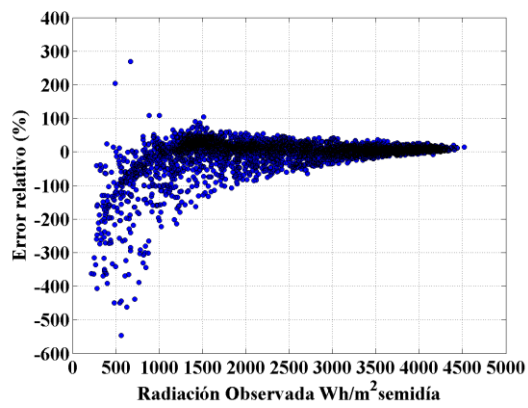
Horizonte de predicción 2 semidías



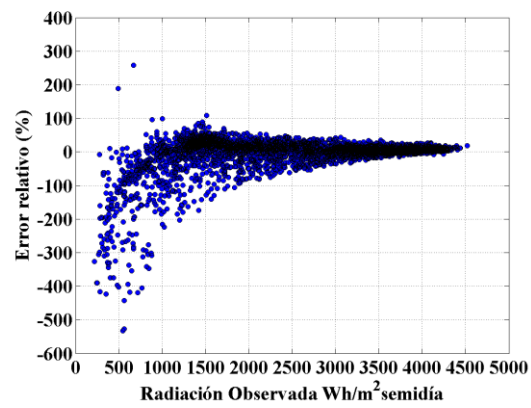
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

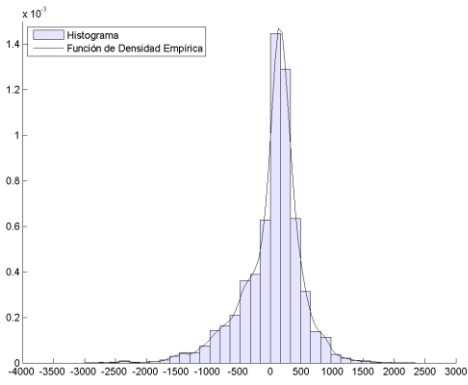


Horizonte de predicción 5 semidías

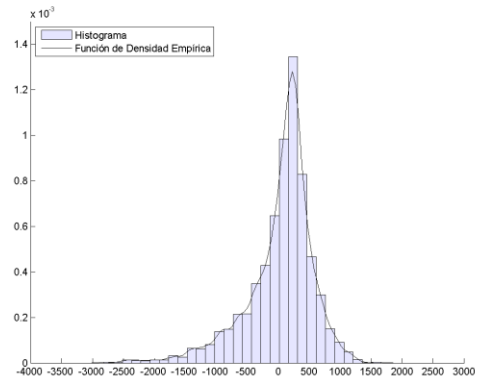


Horizonte de predicción 6 semidías

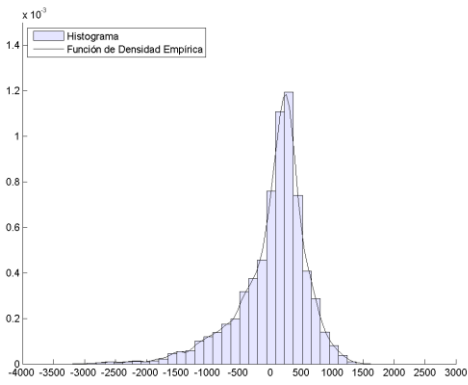
Figura B.5 Murcia: radiación solar semidiaria observada frente al error relativo de predicción



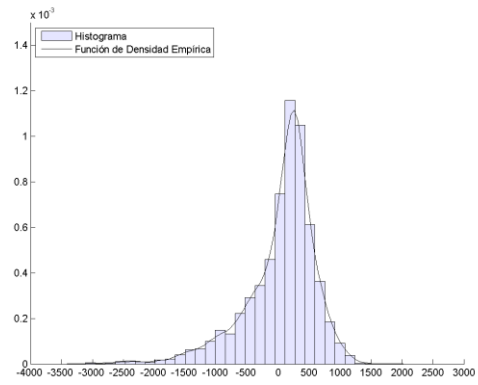
Horizonte de predicción 1 semidía



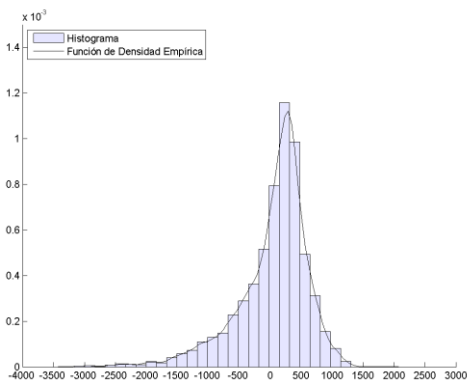
Horizonte de predicción 2 semidías



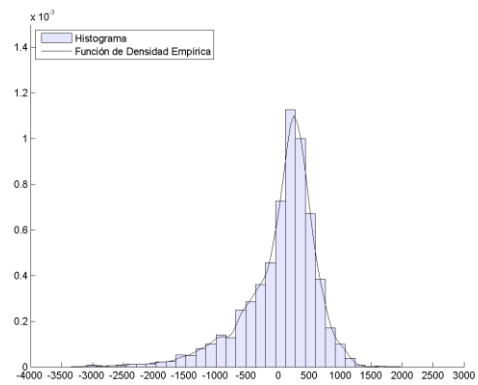
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

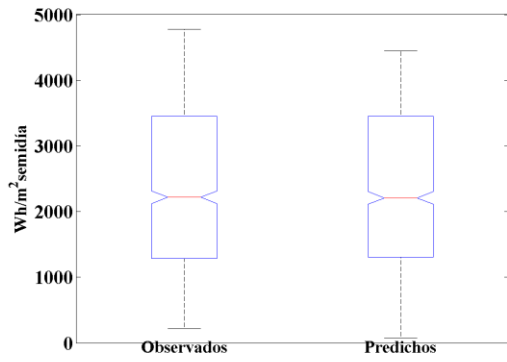


Horizonte de predicción 5 semidías

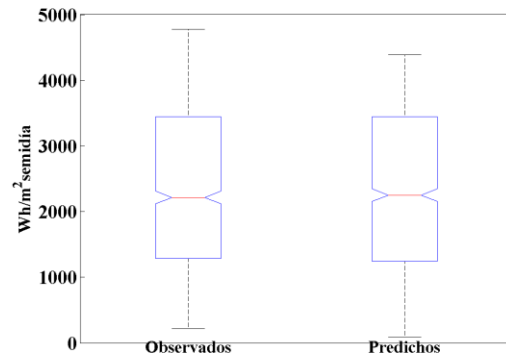


Horizonte de predicción 6 semidías

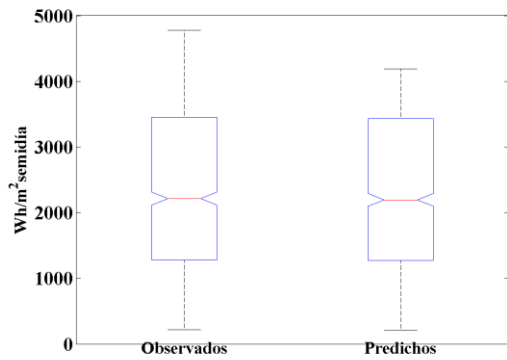
Figura B.6 Murcia: histograma de los errores de predicción y función de densidad muestral de los errores



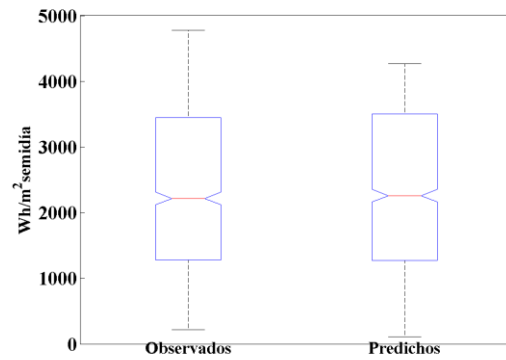
Horizonte de predicción 1 semidia



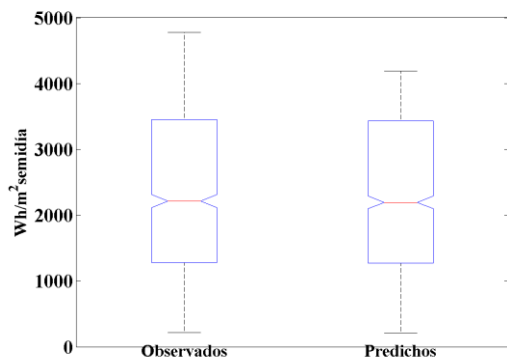
Horizonte de predicción 2 semidías



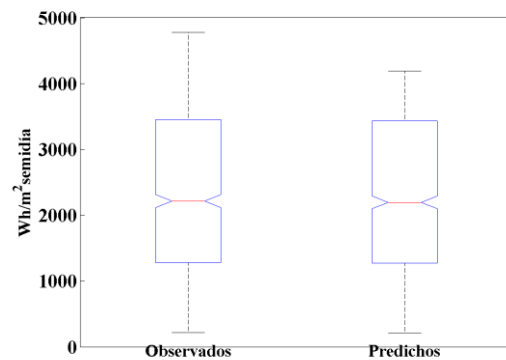
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

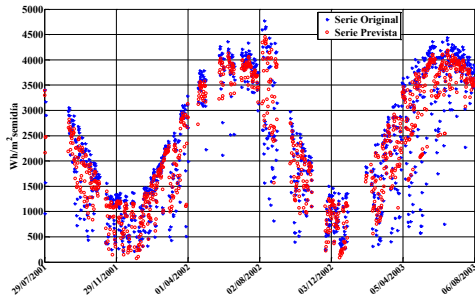


Horizonte de predicción 5 semidías

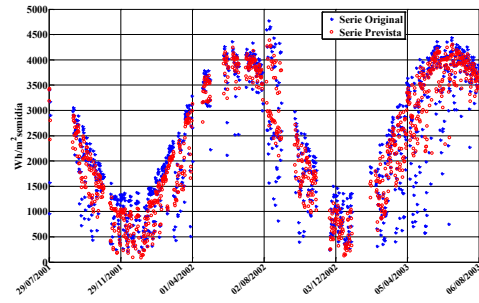


Horizonte de predicción 6 semidías

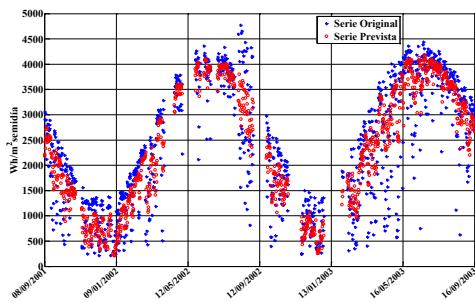
Figura B.7 Albacete: diagrama de Box-Whisker de la serie temporal de datos radiométricos semidiarios observados y datos predichos



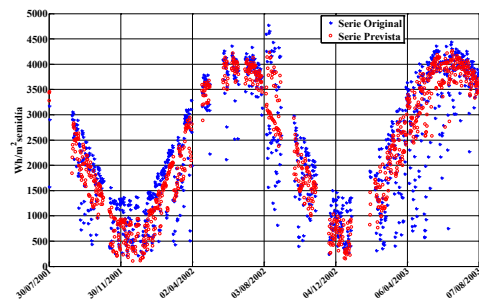
Horizonte de predicción 1 semidiá



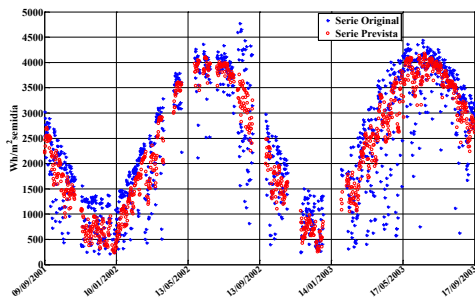
Horizonte de predicción 2 semidiás



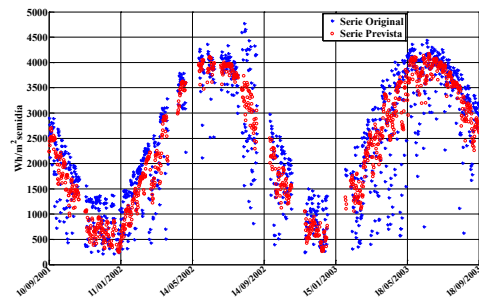
Horizonte de predicción 3 semidiás



Horizonte de predicción 4 semidiás

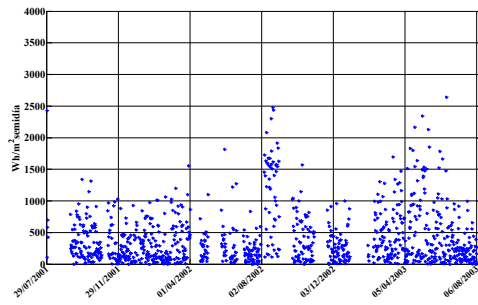


Horizonte de predicción 5 semidiás

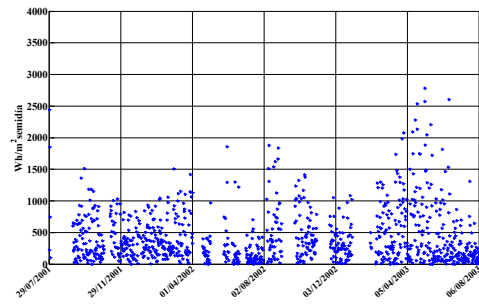


Horizonte de predicción 6 semidiás

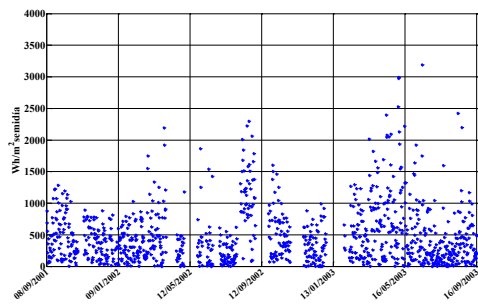
Figura B.8 Albacete: serie temporal de datos radiométricos semidiarios observados y datos predichos



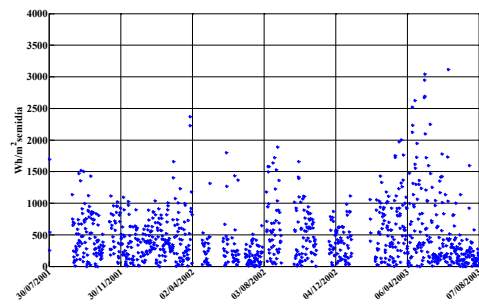
Horizonte de predicción 1 semidía



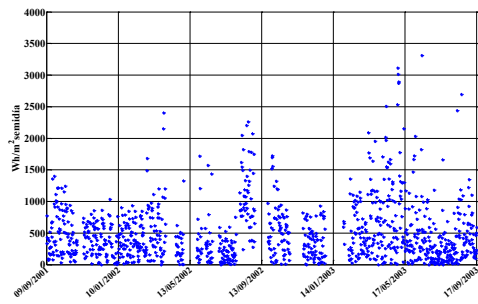
Horizonte de predicción 2 semidías



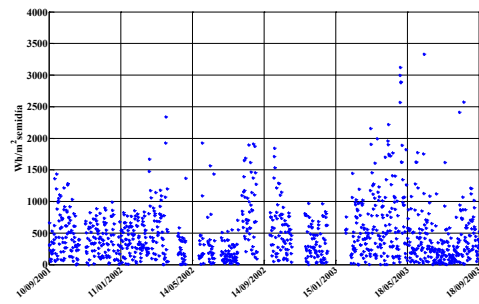
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

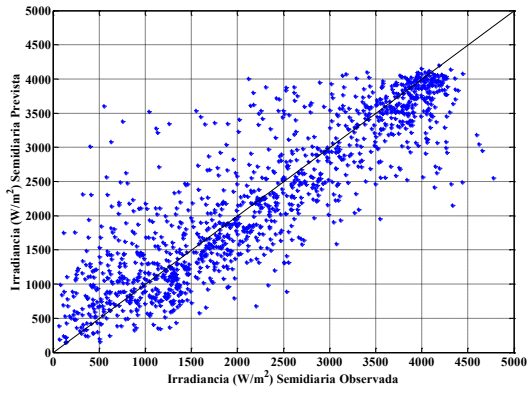


Horizonte de predicción 5 semidías

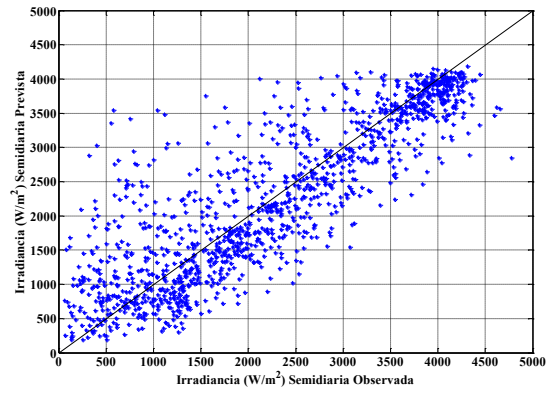


Horizonte de predicción 6 semidías

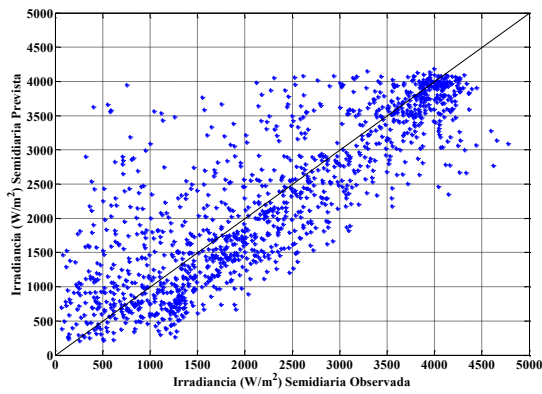
Figura B.9 Albacete: serie temporal de errores absolutos en la predicción



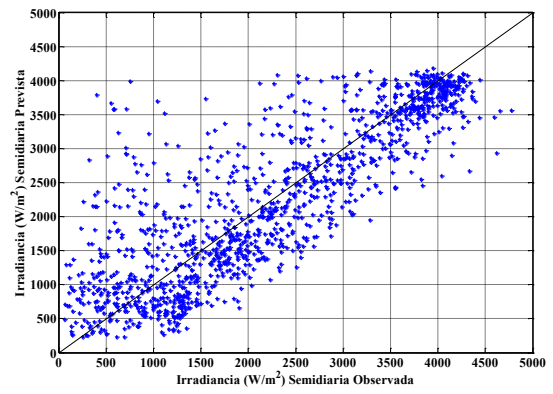
Horizonte de predicción 1 semidía



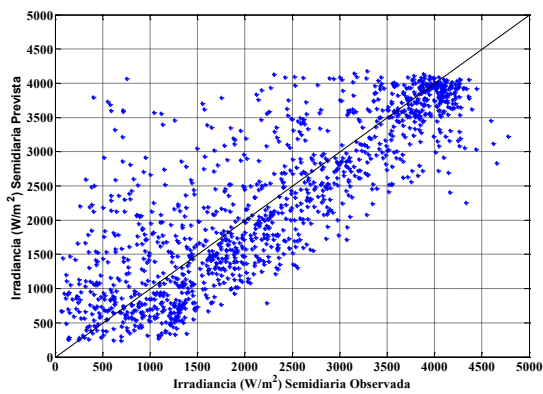
Horizonte de predicción 2 semidías



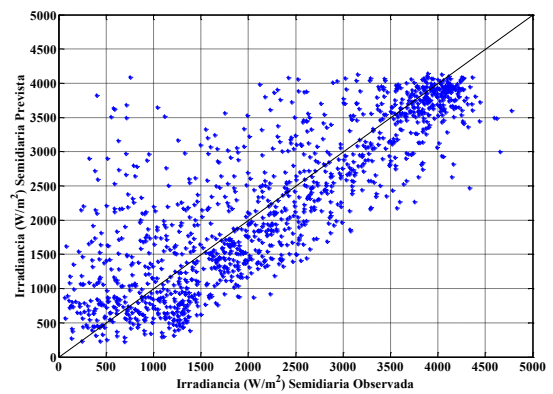
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

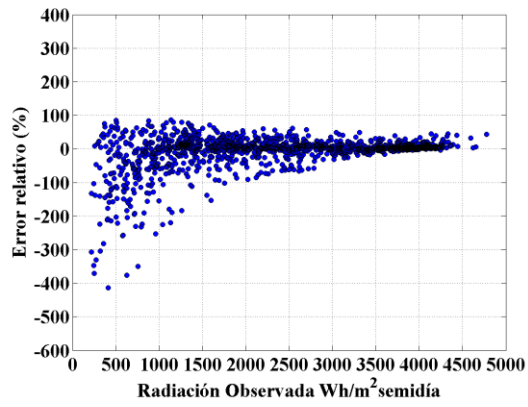


Horizonte de predicción 5 semidías

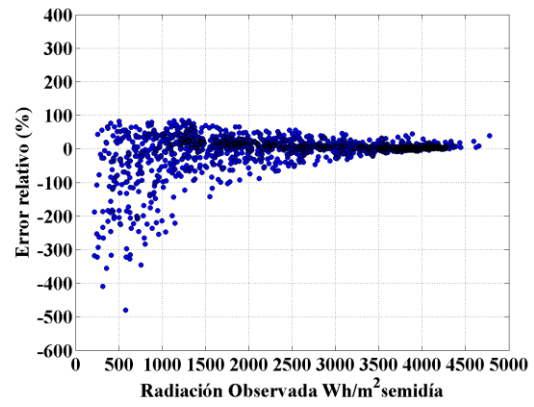


Horizonte de predicción 6 semidías

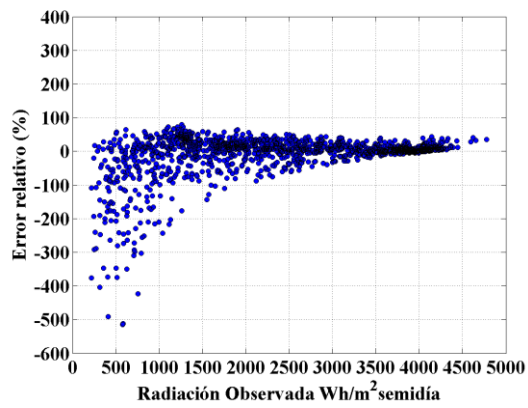
Figura B.10 Albacete: diagrama de dispersión serie temporal original / prevista



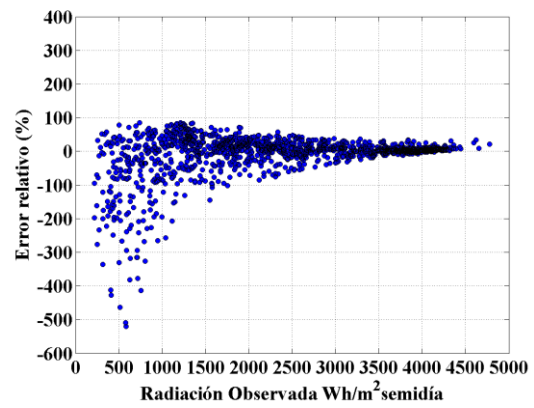
Horizonte de predicción 1 semidía



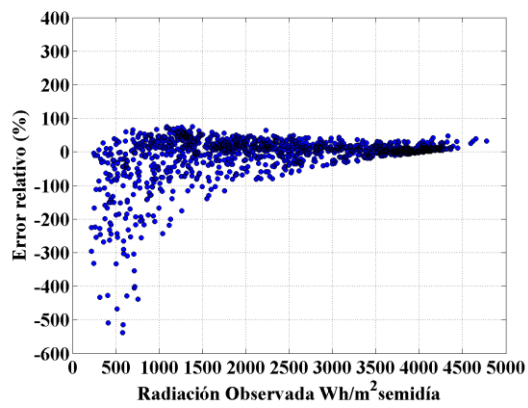
Horizonte de predicción 2 semidías



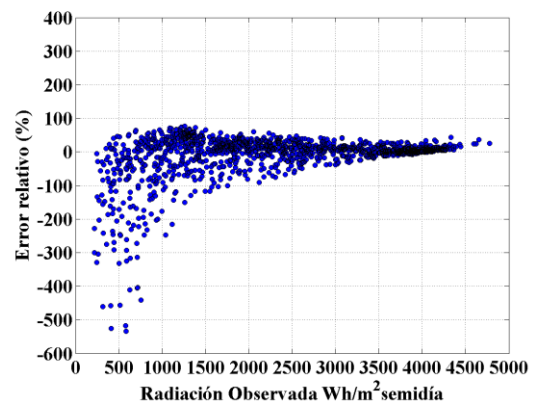
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

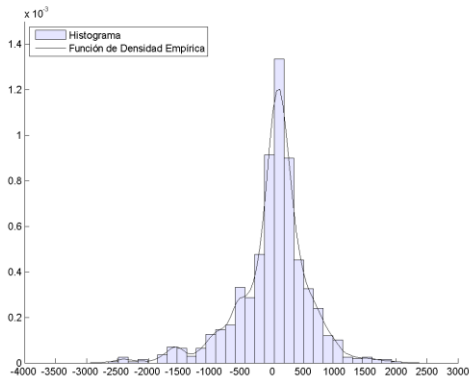


Horizonte de predicción 5 semidías

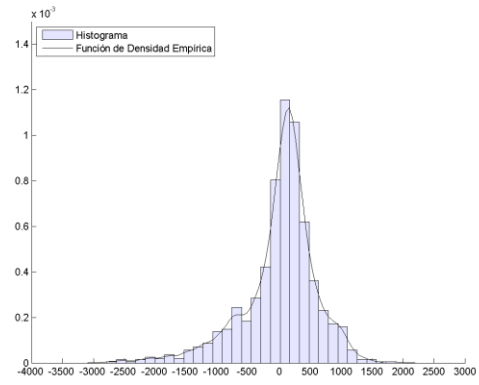


Horizonte de predicción 6 semidías

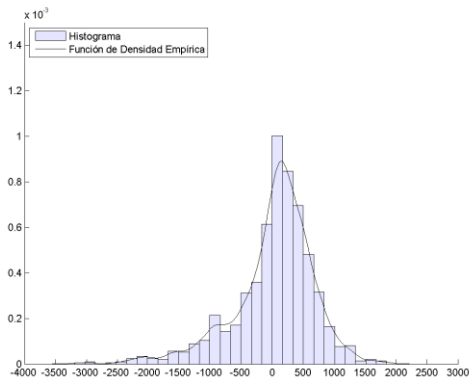
Figura B.11 Albacete: radiación solar semidiaría observada frente al error absoluto de predicción



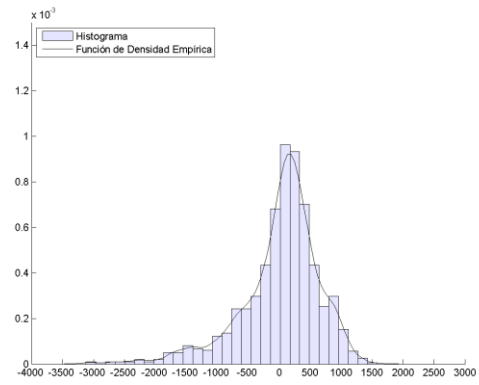
Horizonte de predicción 1 semidía



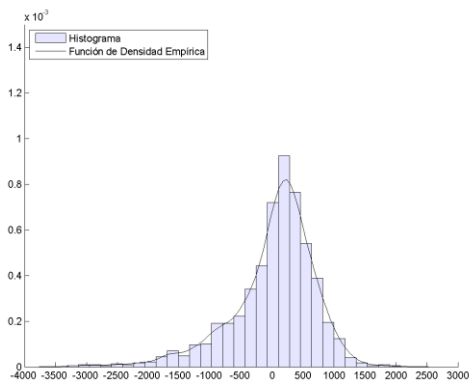
Horizonte de predicción 2 semidías



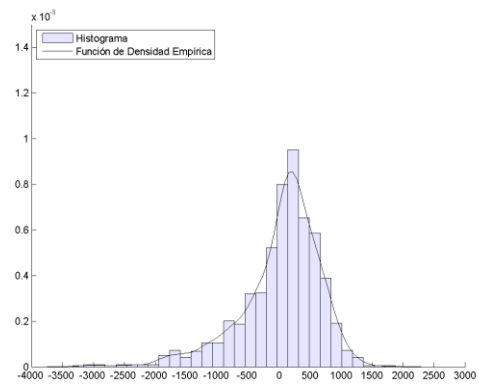
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

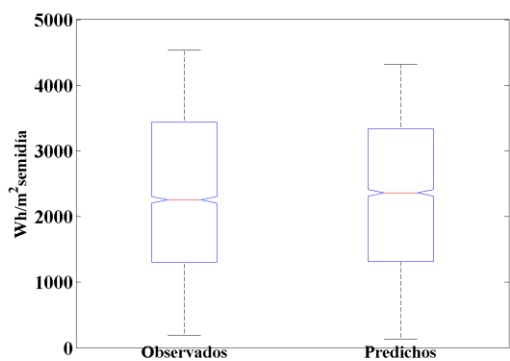


Horizonte de predicción 5 semidías

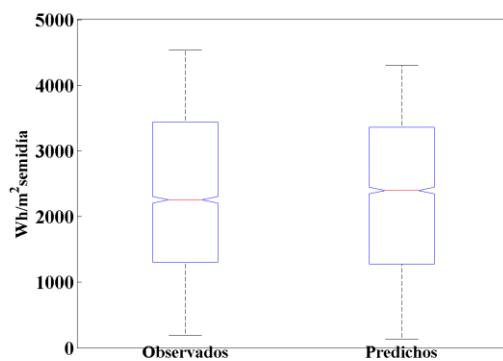


Horizonte de predicción 6 semidías

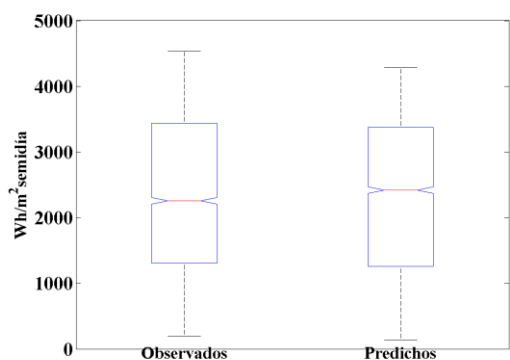
Figura B.12 Albacete: histograma de los errores de predicción y función de densidad muestral de los errores



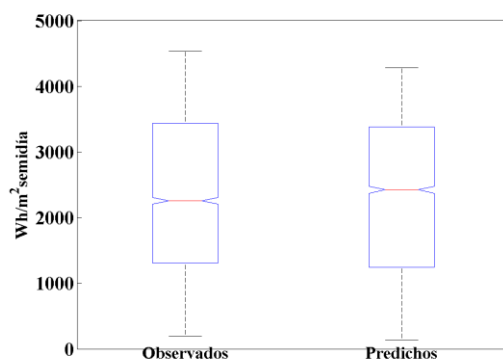
Horizonte de predicción 1 semidia



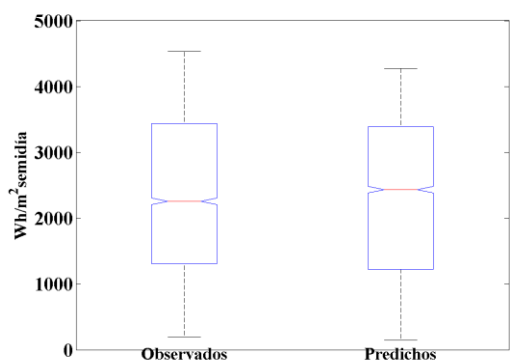
Horizonte de predicción 2 semidías



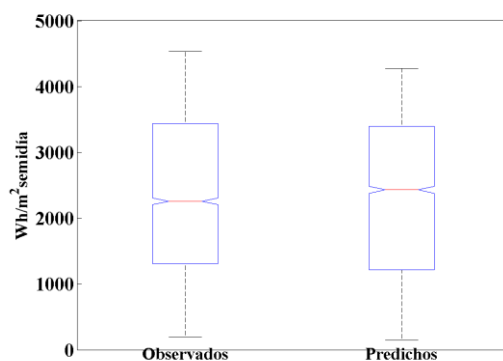
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

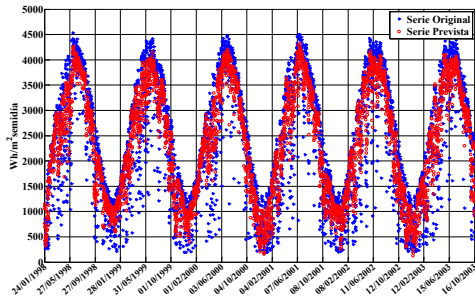


Horizonte de predicción 5 semidías

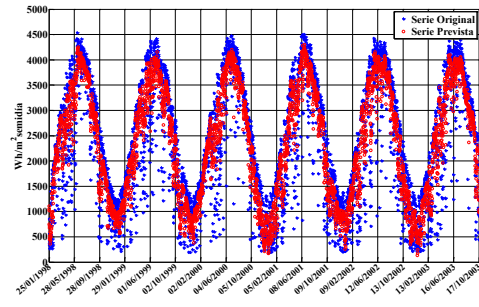


Horizonte de predicción 6 semidías

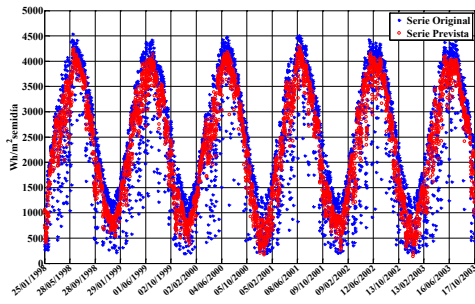
Figura B.13 Madrid: diagrama de Box-Whisker de la serie temporal de datos radiométricos semidiarios observados y datos predichos



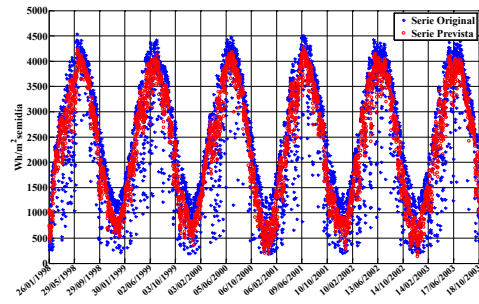
Horizonte de predicción 1 semidiá



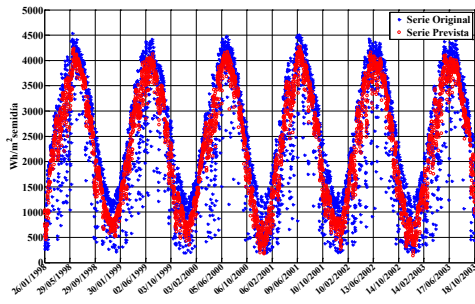
Horizonte de predicción 2 semidiá



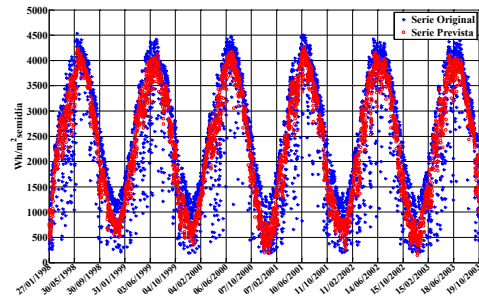
Horizonte de predicción 3 semidiá



Horizonte de predicción 4 semidiá

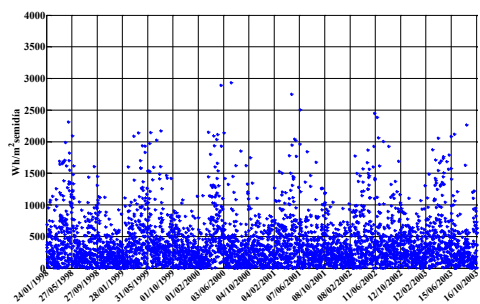


Horizonte de predicción 5 semidiá

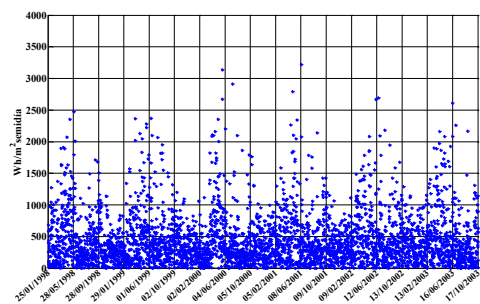


Horizonte de predicción 6 semidiá

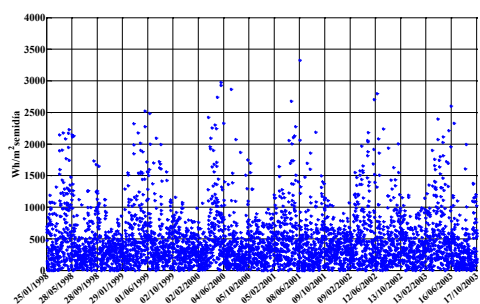
Figura B.14 Madrid: serie temporal de datos radiométricos semidiarios observados y datos predichos



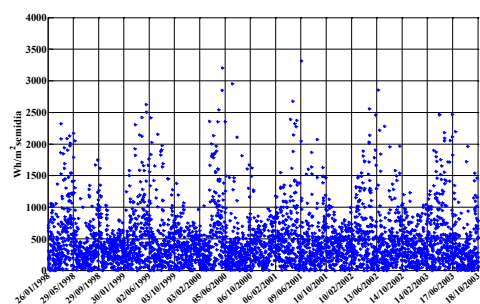
Horizonte de predicción 1 semidía



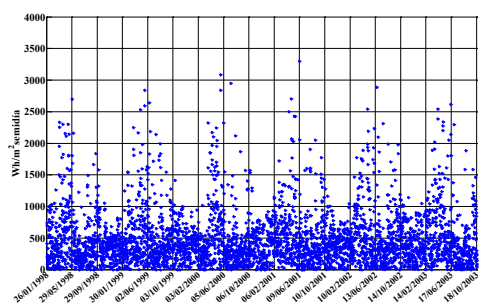
Horizonte de predicción 2 semidías



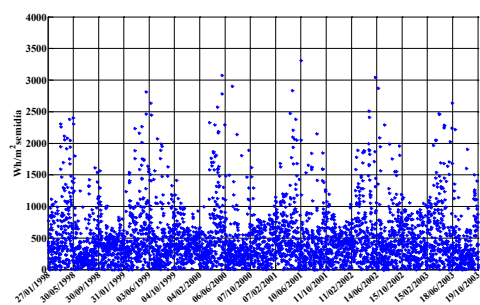
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

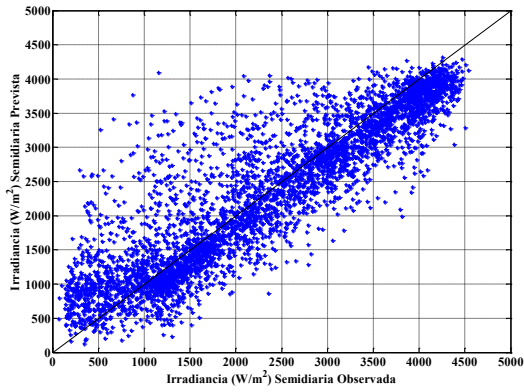


Horizonte de predicción 5 semidías

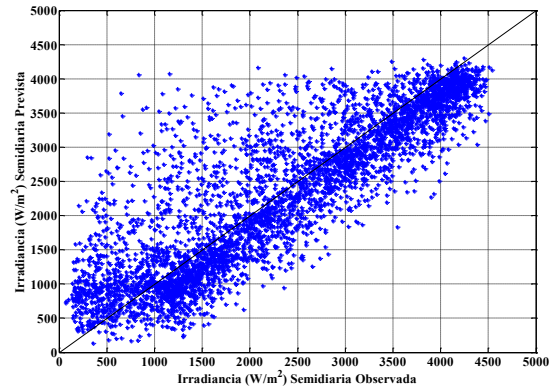


Horizonte de predicción 6 semidías

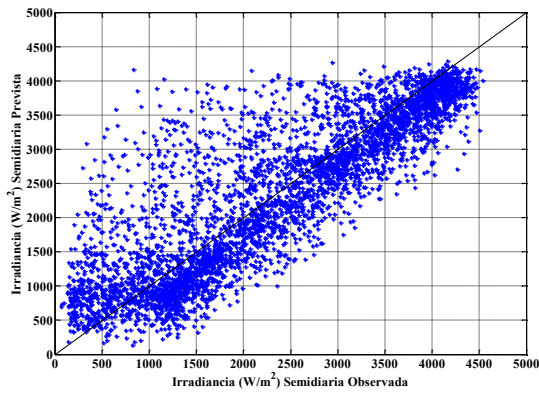
Figura B.15 Madrid: serie temporal de errores absolutos en la predicción



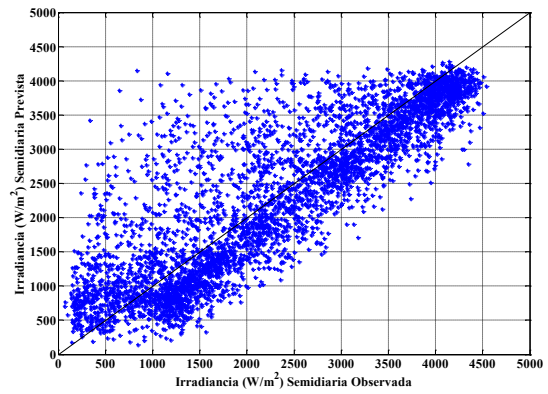
Horizonte de predicción 1 semidía



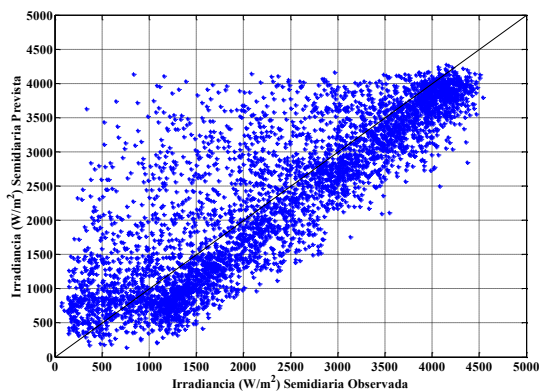
Horizonte de predicción 2 semidías



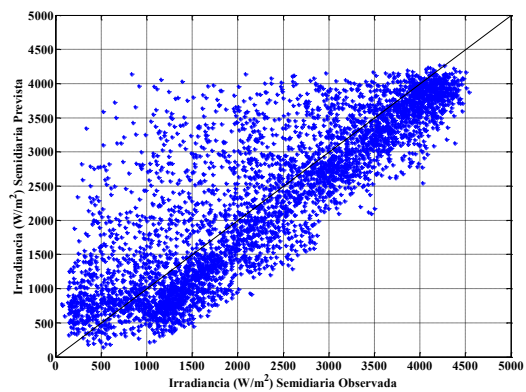
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

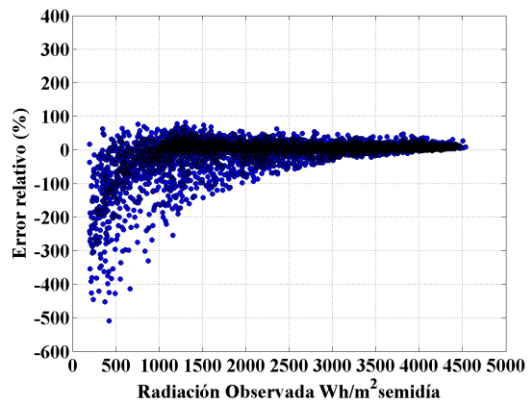


Horizonte de predicción 5 semidías

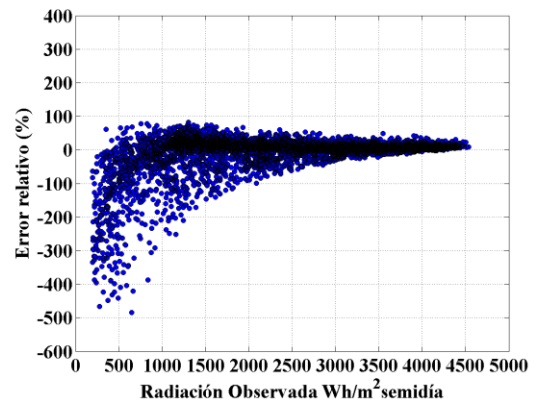


Horizonte de predicción 3 semidías

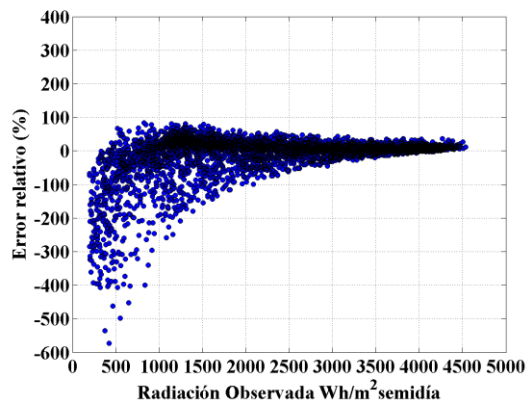
Figura B.16 Madrid: diagrama de dispersión serie temporal original / prevista



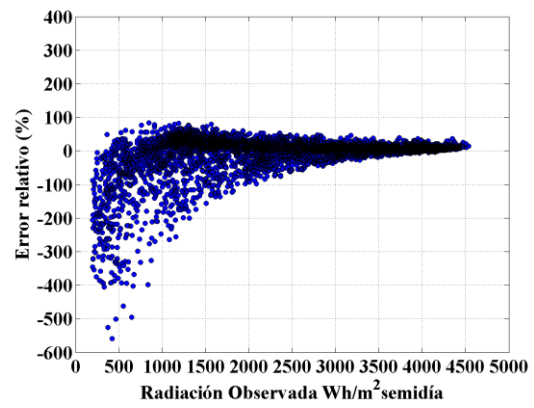
Horizonte de predicción 1 semidía



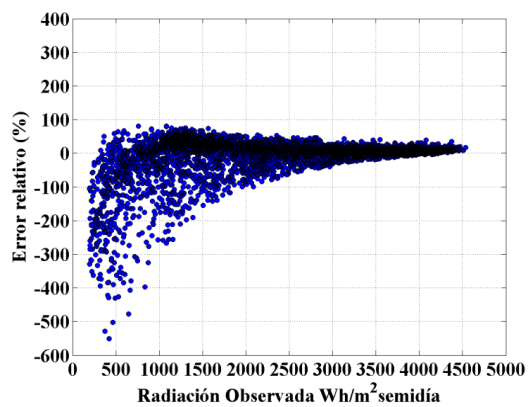
Horizonte de predicción 2 semidías



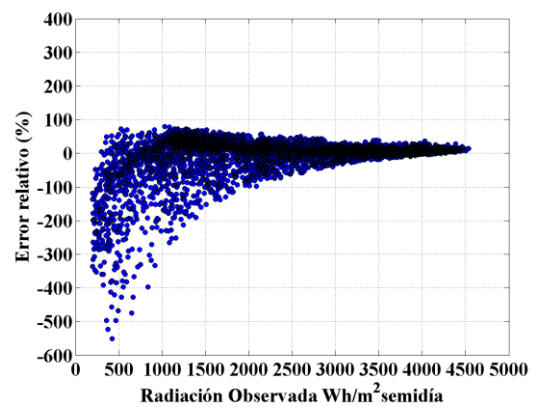
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

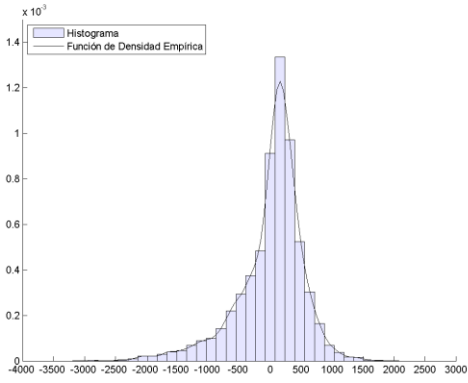


Horizonte de predicción 5 semidías

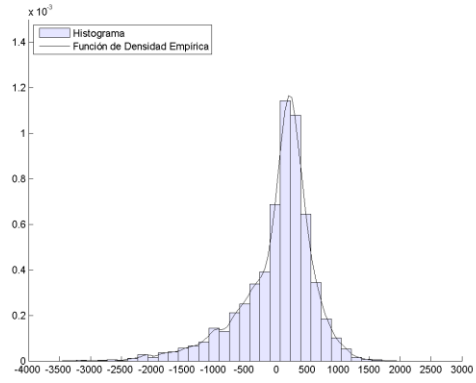


Horizonte de predicción 6 semidías

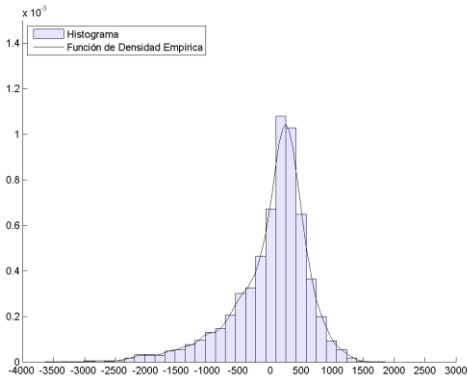
Figura B.17 Madrid: radiación solar semidiaria observada frente al error relativo de predicción



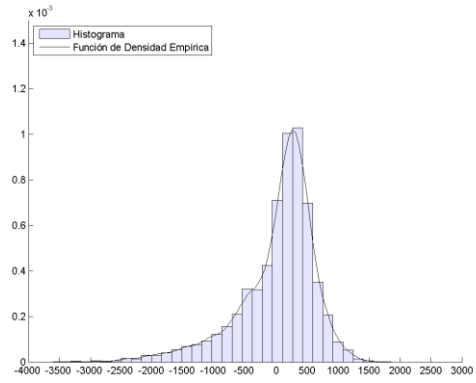
Horizonte de predicción 1 semidía



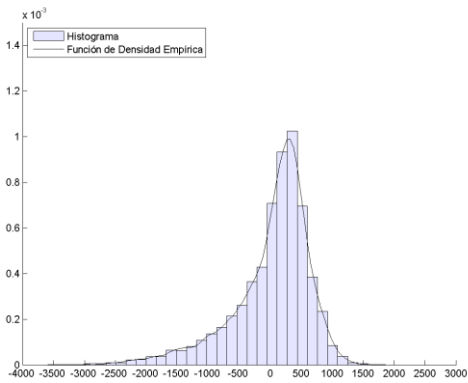
Horizonte de predicción 2 semidías



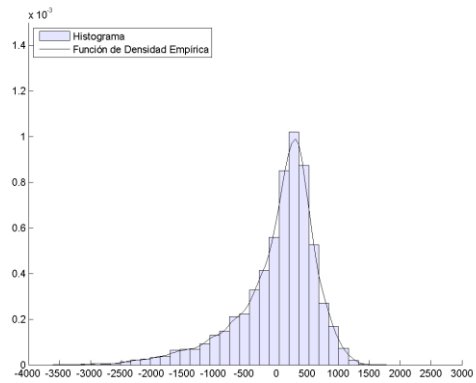
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

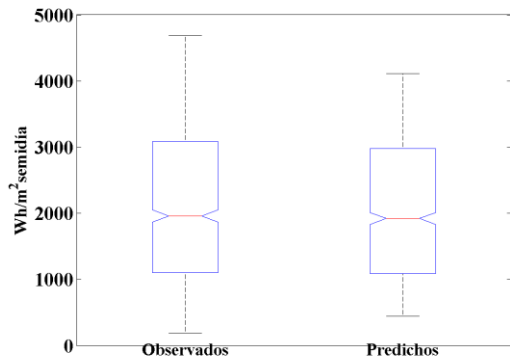


Horizonte de predicción 5 semidías

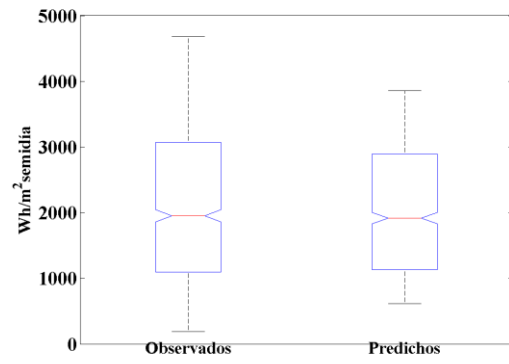


Horizonte de predicción 6 semidías

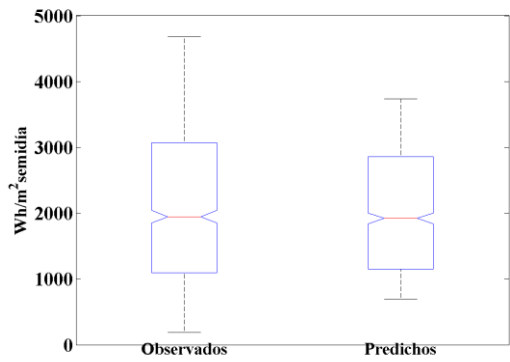
Figura B.18 Madrid: histograma de los errores de predicción y función de densidad muestral de los errores



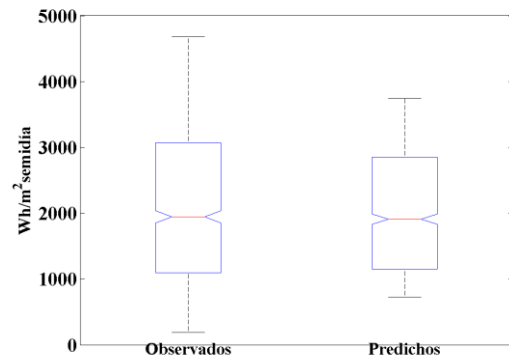
Horizonte de predicción 1 semidía



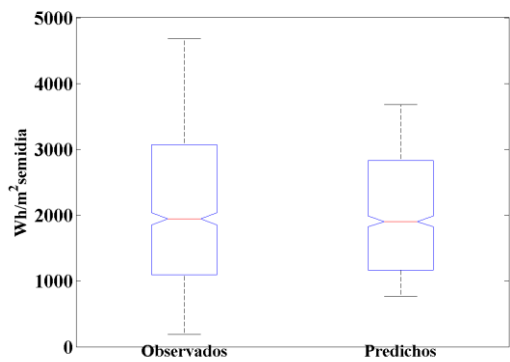
Horizonte de predicción 2 semidías



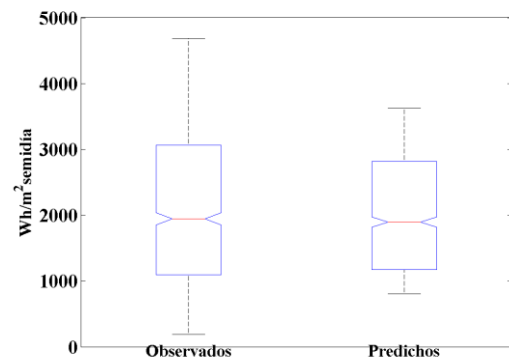
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

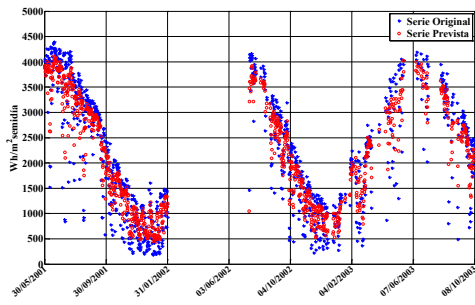


Horizonte de predicción 5 semidías

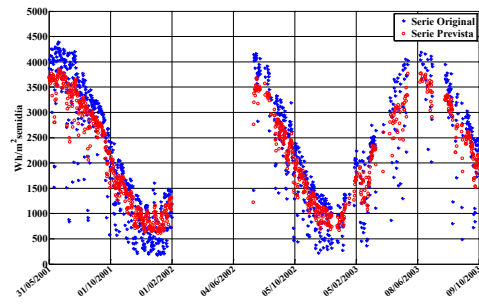


Horizonte de predicción 6 semidías

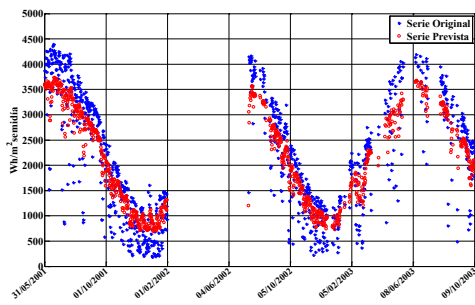
Figura B.19 Lérica: diagrama de Box-Whisker de la serie temporal de datos radiométricos semidiarios observados y datos predichos



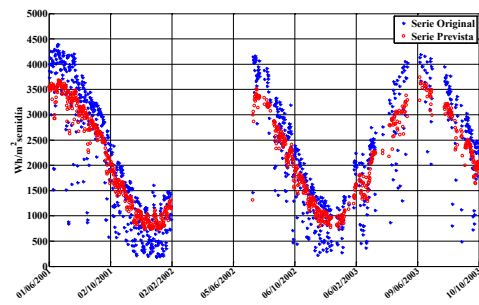
Horizonte de predicción 1 semidiá



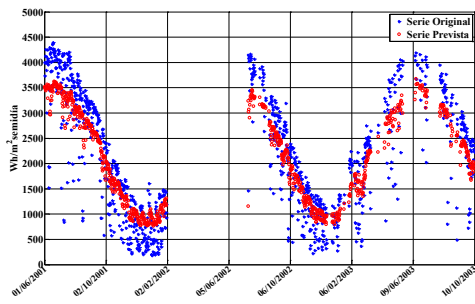
Horizonte de predicción 2 semidiás



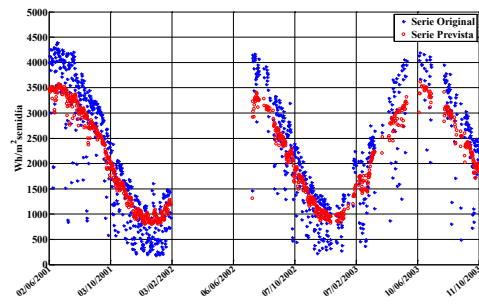
Horizonte de predicción 3 semidiás



Horizonte de predicción 4 semidiás

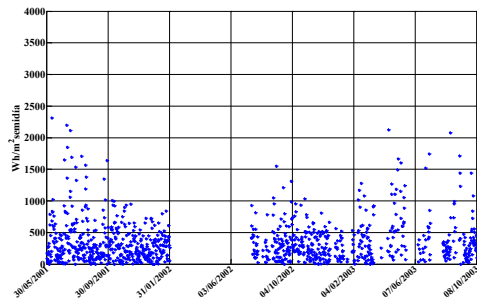


Horizonte de predicción 5 semidiás

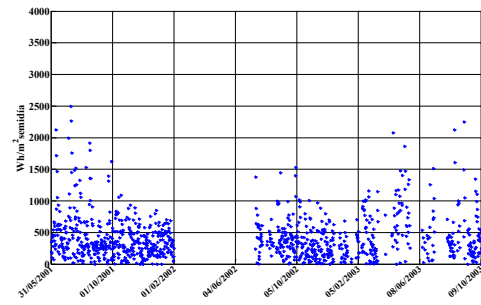


Horizonte de predicción 6 semidiás

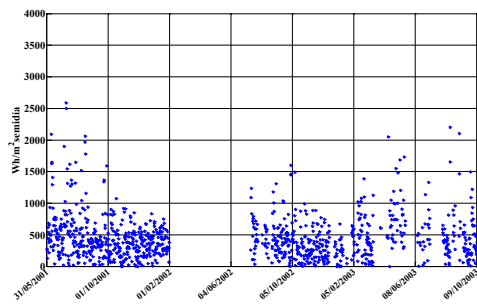
Figura B.20 Lérica: serie temporal de datos radiométricos semidiarios observados y datos predichos



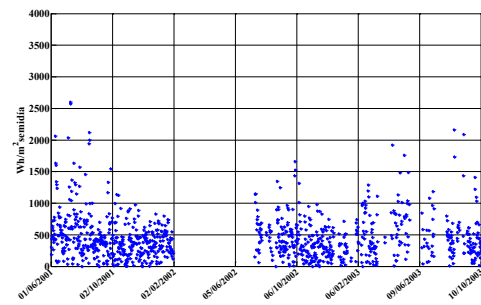
Horizonte de predicción 1 semidía



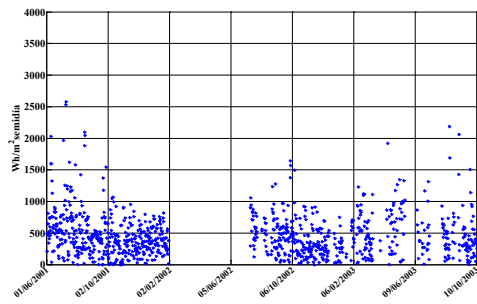
Horizonte de predicción 2 semidías



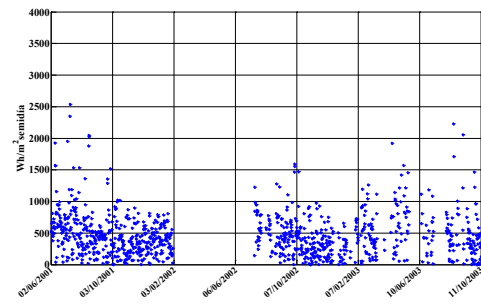
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

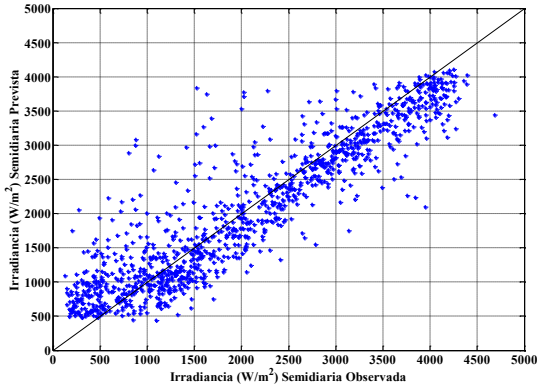


Horizonte de predicción 5 semidías

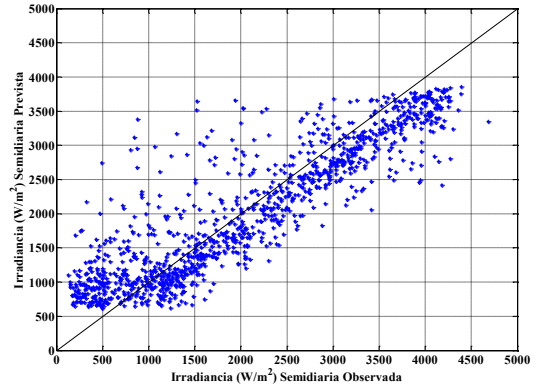


Horizonte de predicción 6 semidías

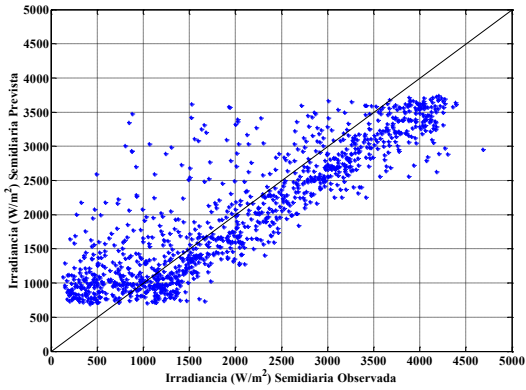
Figura B.21 Lérída: serie temporal de errores absolutos en la predicción



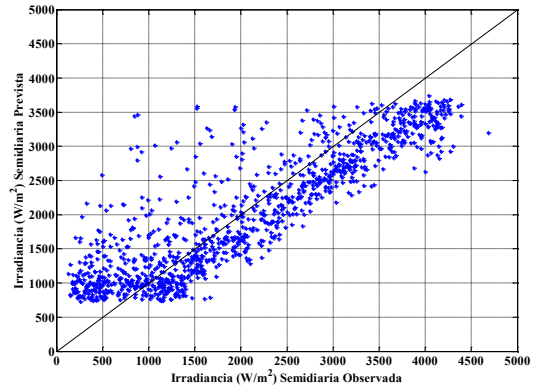
Horizonte de predicción 1 semidía



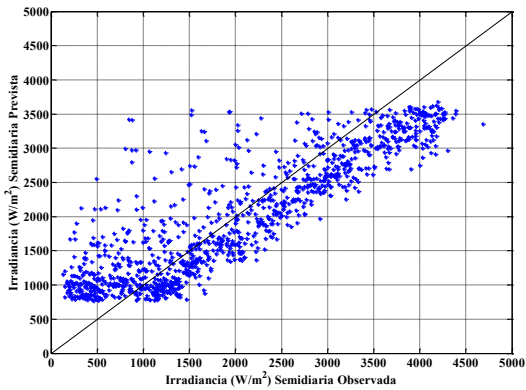
Horizonte de predicción 2 semidías



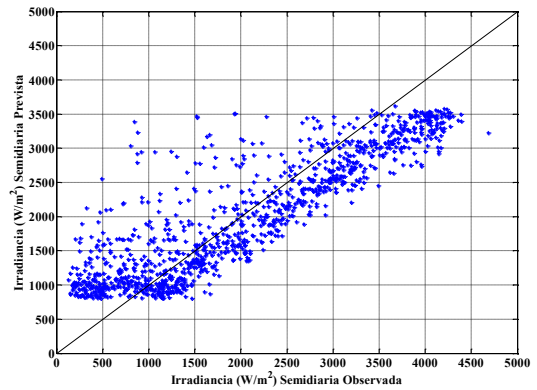
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

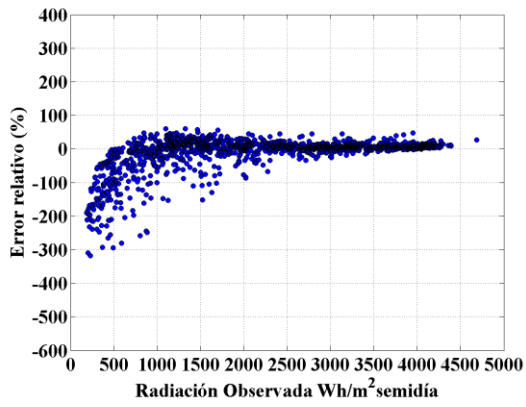


Horizonte de predicción 5 semidías

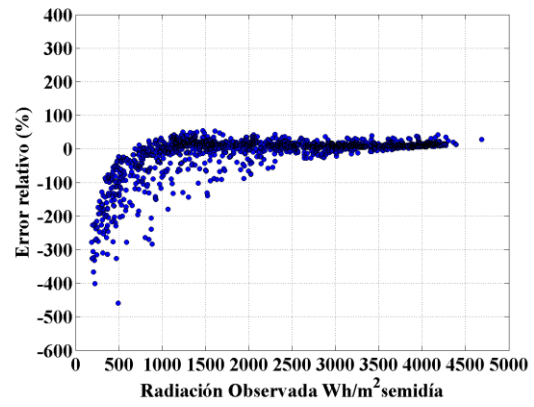


Horizonte de predicción 6 semidías

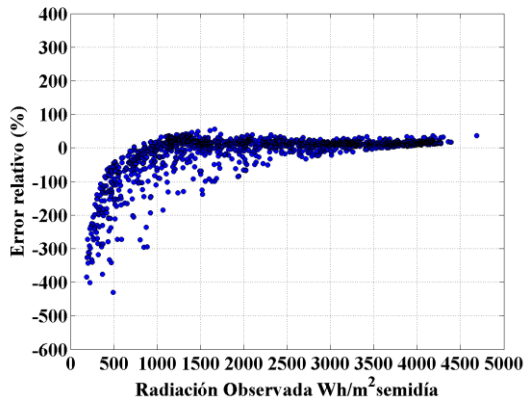
Figura B.22 Lérída: diagrama de dispersión serie temporal original / prevista



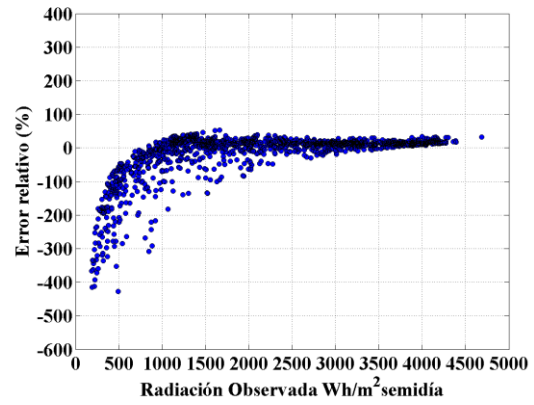
Horizonte de predicción 1 semidía



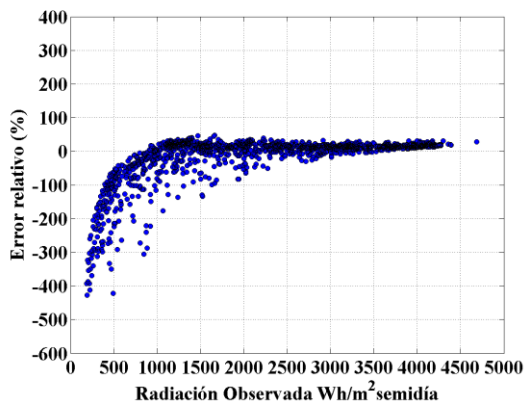
Horizonte de predicción 2 semidías



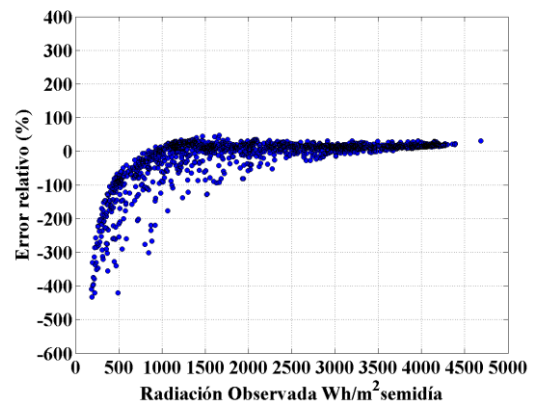
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

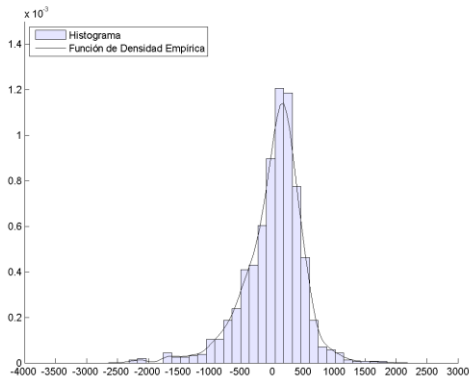


Horizonte de predicción 5 semidías

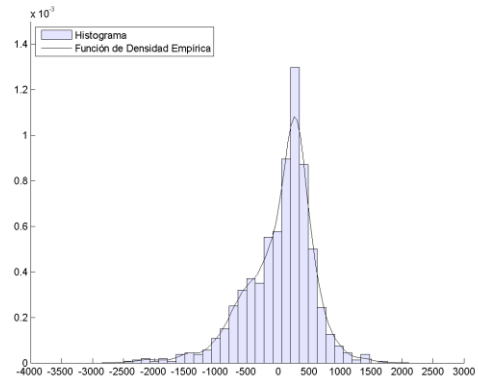


Horizonte de predicción 6 semidías

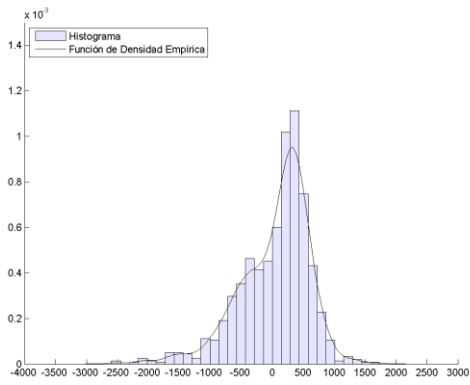
Figura B.23 Lérica: radiación solar semidiaria observada frente al error relativo de predicción



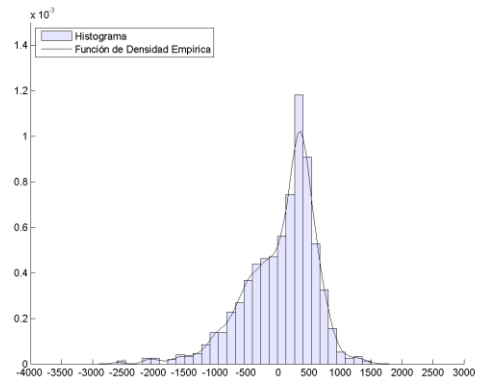
Horizonte de predicción 1 semidía



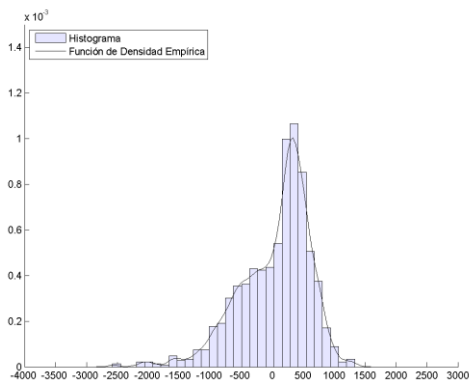
Horizonte de predicción 2 semidías



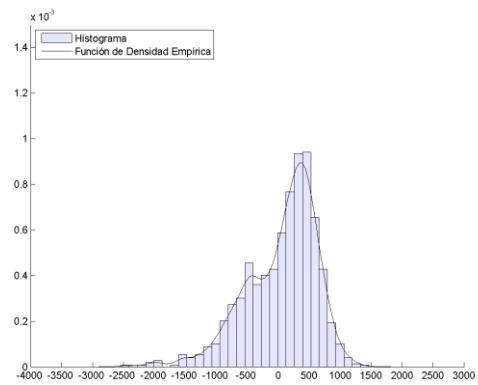
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías



Horizonte de predicción 6 semidías

Figura B.24 Lérída: histograma de los errores de predicción y función de densidad muestral de los errores



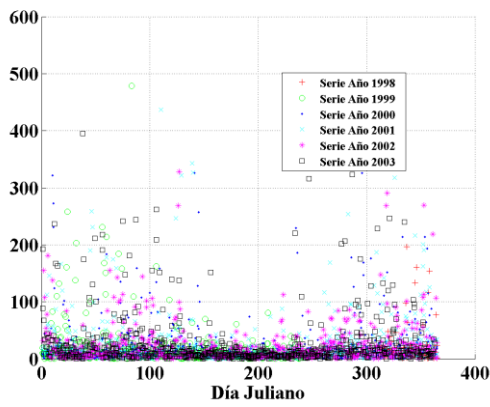
## APÉNDICE C – ANÁLISIS MENSUAL DE LOS RESULTADOS

En este apéndice se presentan los resultados a partir de los estadísticos de error mensuales obtenidos a partir de la aplicación de los modelos finales de predicción. La información que a continuación se presenta para cada estación radiométrica es la siguiente:

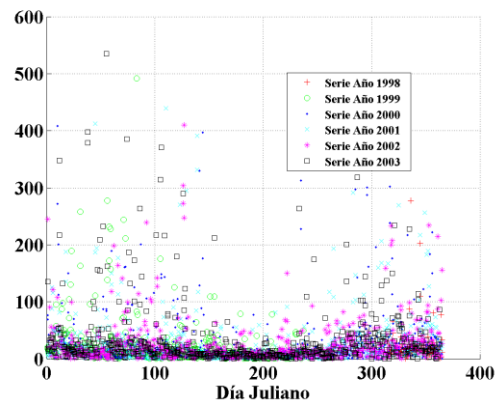
- Tabla con los estadísticos (mediana, media y desviación) de error relativo mensual para cada horizonte temporal de predicción.
- Figura con los errores relativos de predicción en términos absolutos para cada día juliano y cada año.
- Figura con el diagrama Box-Whisker de los errores de predicción para cada mes del año.
- Figura con el diagrama Box-Whisker de los errores relativos de predicción para cada mes del año.
- Figura con la media y desviación mensual de los errores de predicción para cada mes y cada año.
- Figura con la media y desviación mensual de los errores relativos de predicción para cada mes y cada año.

Tabla C.1 Murcia: estadísticos del error relativo (%) mensual

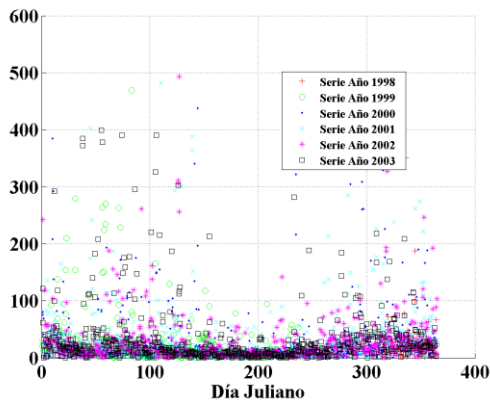
Mes	Horizonte Predicción (Semidías)						ESTADÍSTICO
	1	2	3	4	5	6	
<b>ENERO</b>	4.95	13.01	16.40	17.11	19.35	19.29	MEDIANA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	3.42	5.55	7.80	8.64	9.26	9.49	
<b>MARZO</b>	2.81	6.83	5.84	5.22	5.38	6.43	
<b>ABRIL</b>	4.22	2.48	2.10	2.48	2.47	2.92	
<b>MAYO</b>	4.36	4.39	4.17	4.53	4.34	4.14	
<b>JUNIO</b>	4.82	3.55	3.46	3.15	2.95	2.97	
<b>JULIO</b>	4.32	3.31	2.89	2.77	2.79	2.72	
<b>AGOSTO</b>	5.03	5.37	4.91	4.76	4.93	4.42	
<b>SEPTIEMBRE</b>	4.36	3.61	3.90	3.68	3.85	4.33	
<b>OCTUBRE</b>	4.08	10.48	11.13	12.17	13.56	14.30	
<b>NOVIEMBRE</b>	4.05	11.65	15.38	16.59	19.76	19.27	
<b>DICIEMBRE</b>	4.69	14.09	16.96	18.80	19.95	19.47	
<b>ENERO</b>	-13.59	-6.33	-2.64	-1.32	-0.26	-0.27	MEDIA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	-11.90	-13.48	-12.72	-12.22	-11.75	-11.84	
<b>MARZO</b>	-12.85	-13.62	-14.28	-14.90	-15.49	-16.11	
<b>ABRIL</b>	-8.99	-12.96	-13.87	-14.42	-14.67	-14.81	
<b>MAYO</b>	-9.87	-17.66	-19.47	-20.49	-20.85	-21.14	
<b>JUNIO</b>	1.55	-1.64	-2.28	-2.61	-2.99	-3.01	
<b>JULIO</b>	2.48	-0.09	-0.61	-0.76	-0.83	-0.78	
<b>AGOSTO</b>	-0.34	-2.76	-3.29	-3.42	-3.57	-3.60	
<b>SEPTIEMBRE</b>	-2.89	-3.99	-4.67	-4.99	-4.82	-4.79	
<b>OCTUBRE</b>	-13.92	-9.60	-9.38	-7.61	-7.26	-8.48	
<b>NOVIEMBRE</b>	-15.23	-8.28	-5.01	-3.75	-3.11	-3.78	
<b>DICIEMBRE</b>	-16.42	-7.22	-3.22	-2.24	-0.13	0.33	
<b>ENERO</b>	53.45	60.82	59.73	59.81	61.41	64.84	DESVIACIÓN TÍPICA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	54.06	72.14	72.65	72.44	71.89	70.51	
<b>MARZO</b>	54.32	62.57	63.29	63.79	63.54	64.90	
<b>ABRIL</b>	46.71	55.87	58.40	59.61	60.38	62.00	
<b>MAYO</b>	55.91	67.33	72.46	75.92	76.60	76.44	
<b>JUNIO</b>	16.58	21.41	21.30	21.60	22.26	21.70	
<b>JULIO</b>	12.74	14.04	14.60	14.77	14.85	14.69	
<b>AGOSTO</b>	27.39	33.01	33.57	33.09	32.78	32.75	
<b>SEPTIEMBRE</b>	31.68	37.21	39.36	40.87	39.32	39.30	
<b>OCTUBRE</b>	54.63	66.38	67.05	60.92	61.65	68.76	
<b>NOVIEMBRE</b>	64.23	68.55	66.42	67.28	68.78	73.94	
<b>DICIEMBRE</b>	52.81	58.19	60.45	60.43	60.63	59.46	



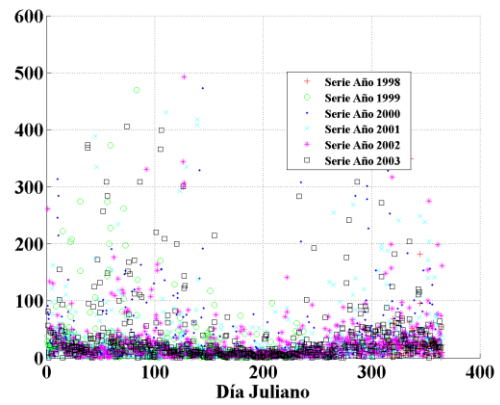
Horizonte de predicción 1 semidiá



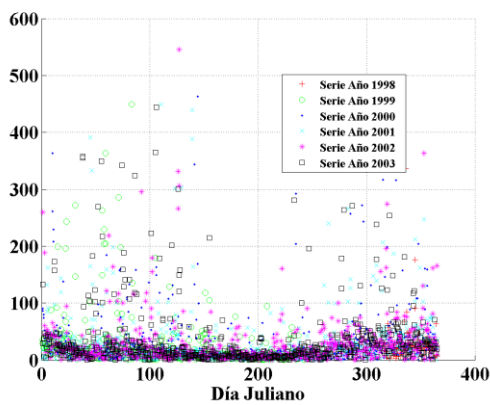
Horizonte de predicción 2 semidiás



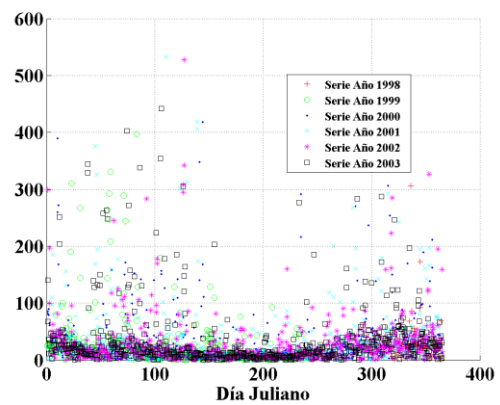
Horizonte de predicción 3 semidiás



Horizonte de predicción 4 semidiás

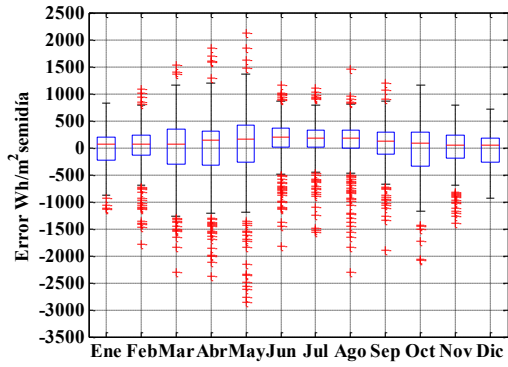


Horizonte de predicción 5 semidiás

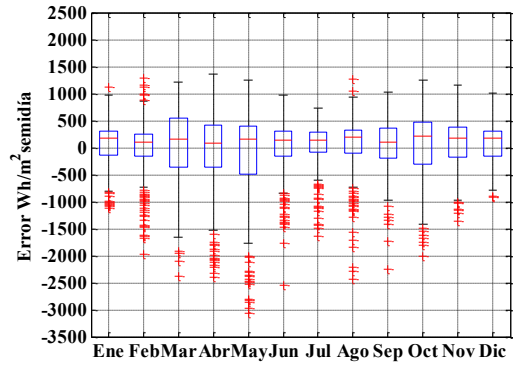


Horizonte de predicción 6 semidiás

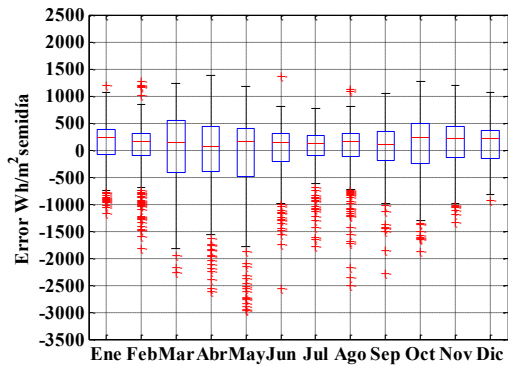
Figura C.1 Murcia: errores relativo de predicción en términos absolutos para cada día juliano y cada año



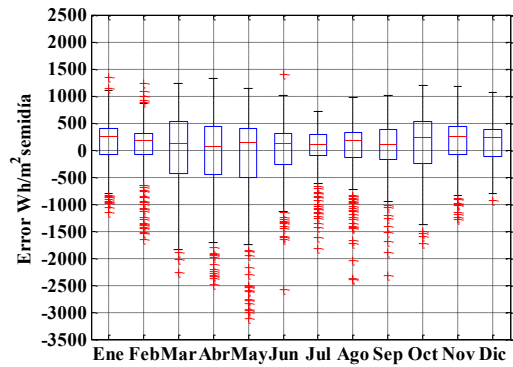
Horizonte de predicción 1 semidia



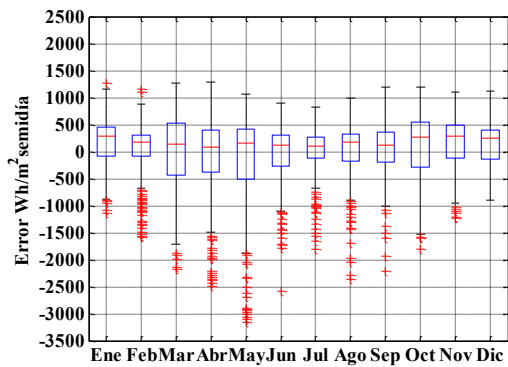
Horizonte de predicción 2 semidías



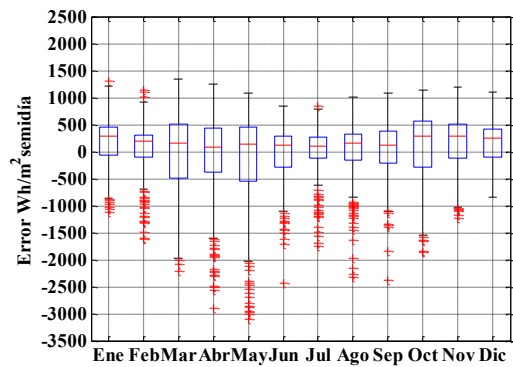
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

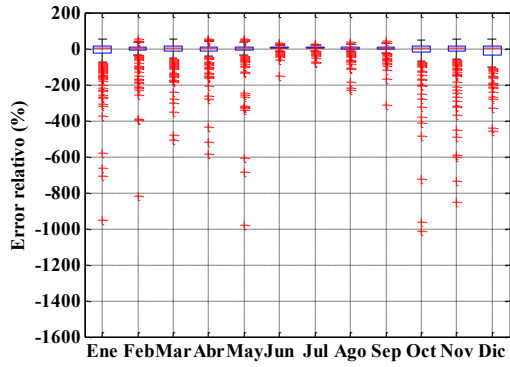


Horizonte de predicción 5 semidías

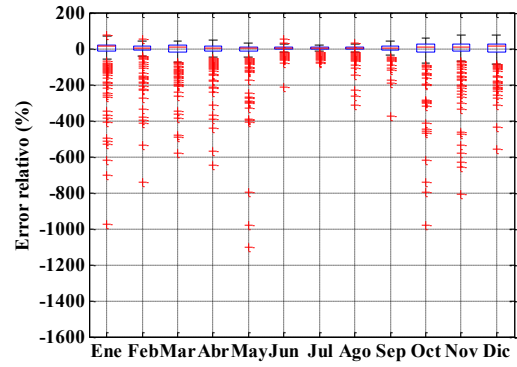


Horizonte de predicción 6 semidías

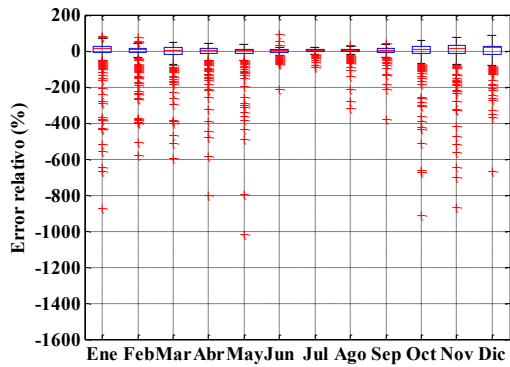
Figura C.2 Murcia: diagrama Box-Whisker de los errores de predicción para cada mes



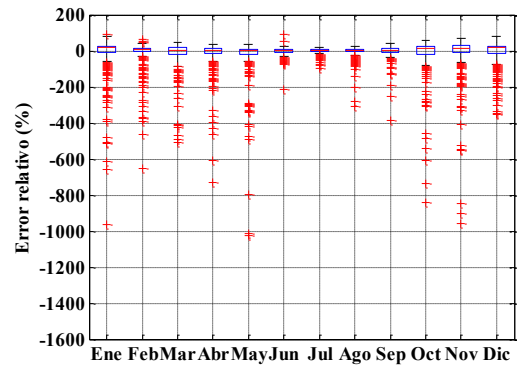
Horizonte de predicción 1 semidía



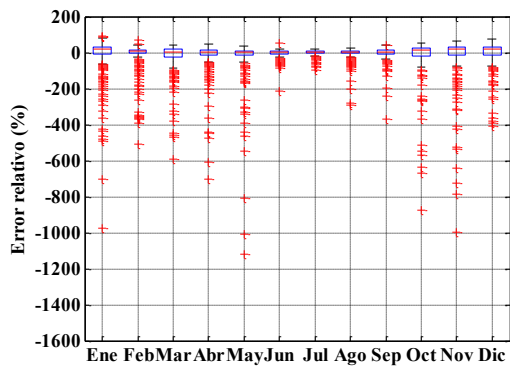
Horizonte de predicción 2 semidías



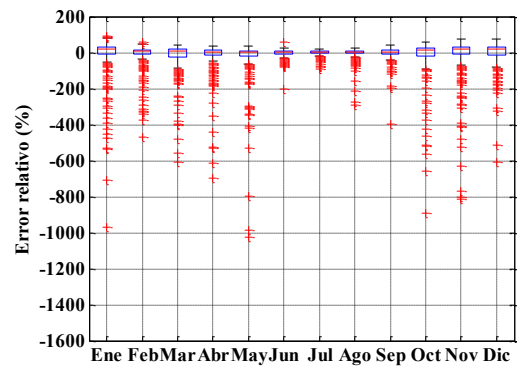
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

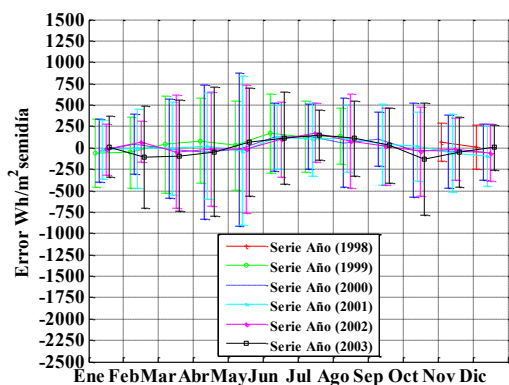


Horizonte de predicción 5 semidías

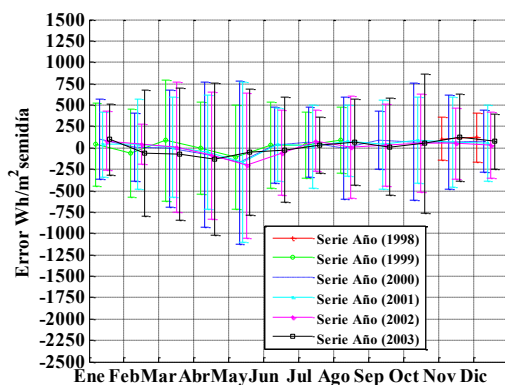


Horizonte de predicción 6 semidías

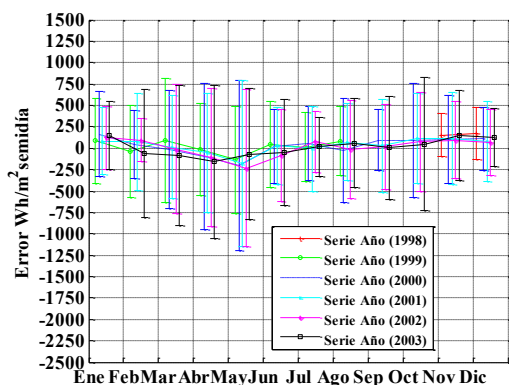
Figura C.3 Murcia: diagrama Box-Whisker de los errores relativos de predicción para cada mes



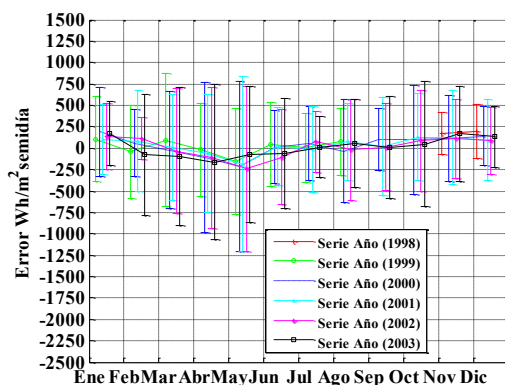
Horizonte de predicción 1 semidía



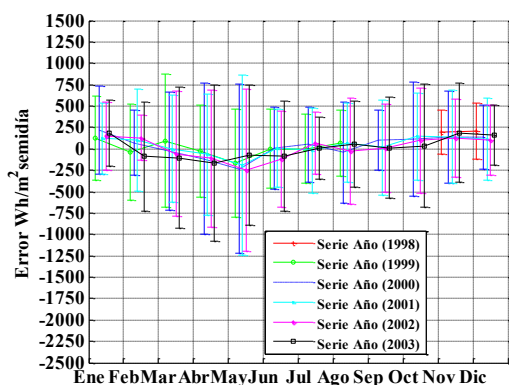
Horizonte de predicción 2 semidías



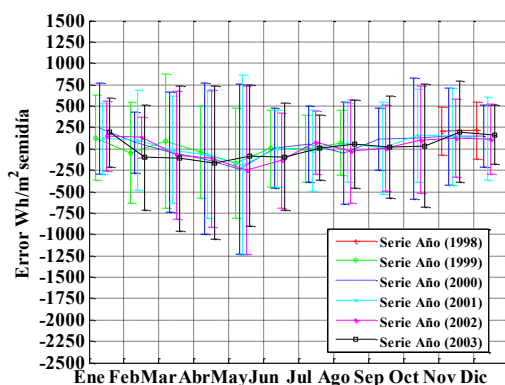
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

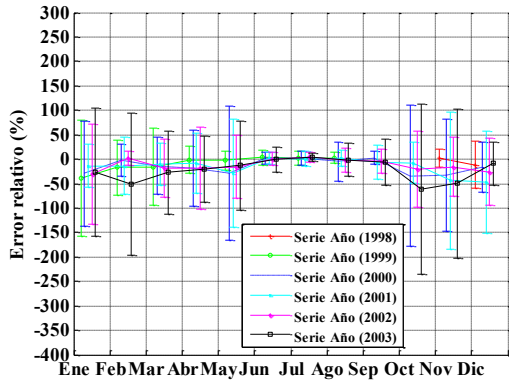


Horizonte de predicción 5 semidías

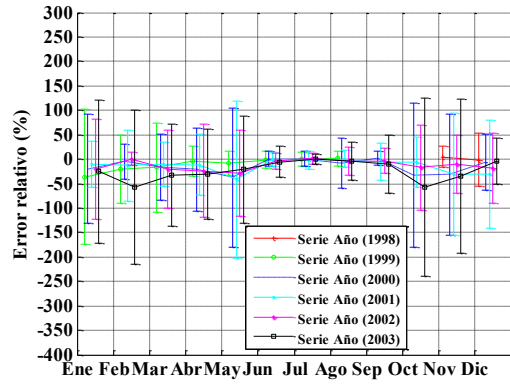


Horizonte de predicción 6 semidías

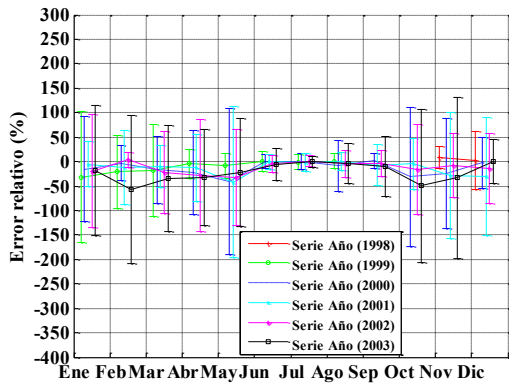
Figura C.4 Murcia: media y desviación mensual de los errores de predicción para cada mes y cada año



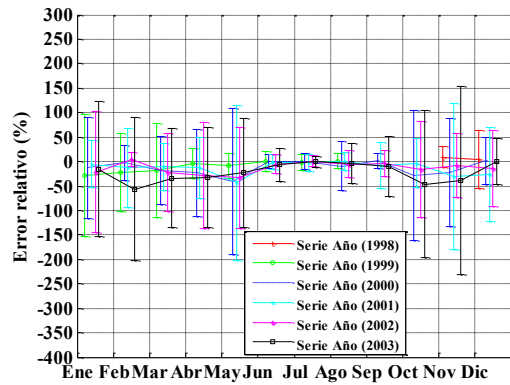
Horizonte de predicción 1 semidía



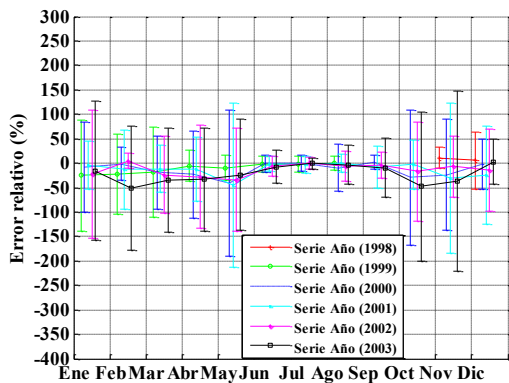
Horizonte de predicción 2 semidías



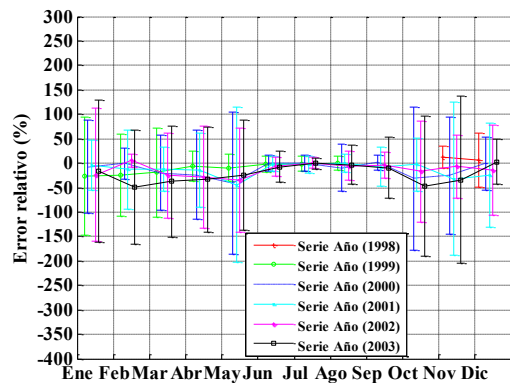
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías

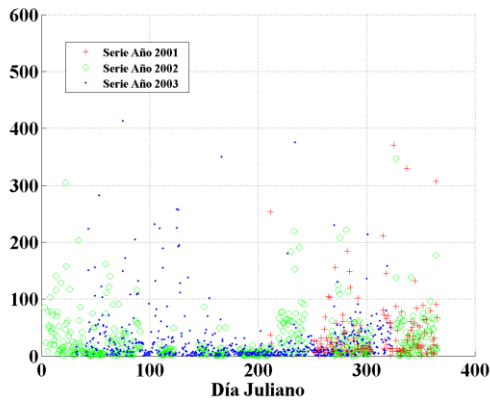


Horizonte de predicción 6 semidías

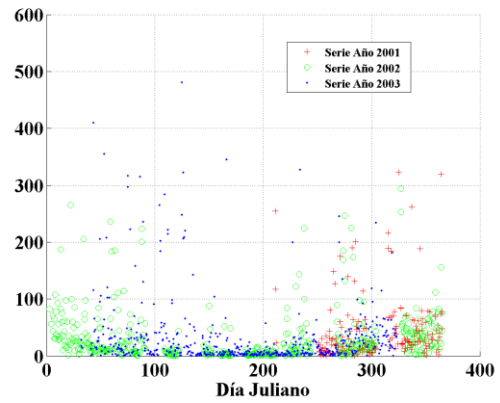
Figura C.5 Murcia: media y desviación mensual de los errores relativos de predicción para cada mes y cada año

Tabla C.2 Albacete: estadísticos del error relativo mensual

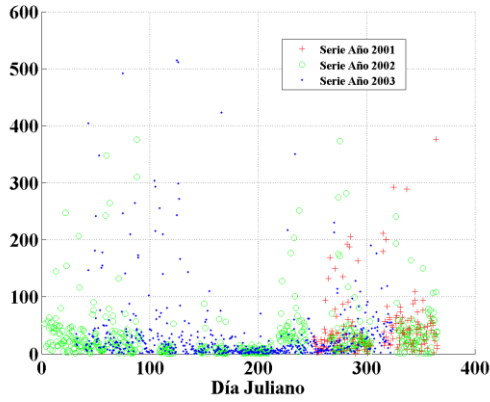
Mes	Horizonte Predicción (Semidías)						ESTADÍSTICO
	1	2	3	4	5	6	
ENERO	9.02	25.58	18.93	30.96	22.97	30.47	MEDIANA DEL ERROR RELATIVO (%)
FEBRERO	3.51	10.89	7.98	12.46	10.39	10.97	
MARZO	2.82	5.24	6.15	4.21	5.26	6.19	
ABRIL	1.46	1.80	3.23	1.68	2.91	3.17	
MAYO	1.22	1.29	1.65	1.99	2.69	3.07	
JUNIO	2.09	-0.28	1.39	0.62	0.66	0.82	
JULIO	2.61	1.65	0.99	1.05	0.88	1.47	
AGOSTO	1.46	2.02	2.52	1.83	3.72	3.48	
SEPTIEMBRE	3.32	4.43	5.10	4.09	5.22	4.04	
OCTUBRE	4.55	7.25	9.35	9.75	10.16	11.01	
NOVIEMBRE	8.59	21.09	22.25	22.80	28.39	25.79	
DICIEMBRE	9.43	24.23	19.18	29.94	27.85	25.41	
ENERO	-5.38	7.52	8.93	16.42	12.39	12.94	MEDIA DEL ERROR RELATIVO (%)
FEBRERO	-15.11	-15.74	-20.10	-15.11	-20.74	-20.73	
MARZO	-14.59	-21.36	-24.16	-28.15	-27.09	-28.08	
ABRIL	-14.86	-21.09	-19.37	-24.24	-19.46	-18.93	
MAYO	-19.64	-26.51	-29.95	-33.23	-33.69	-34.20	
JUNIO	-5.92	-8.54	-9.23	-10.89	-10.71	-11.04	
JULIO	-0.44	-2.72	-0.71	-1.49	-1.02	-1.06	
AGOSTO	-18.75	-12.74	-14.70	-14.45	-15.53	-14.12	
SEPTIEMBRE	-6.32	-9.80	-10.05	-11.14	-11.51	-11.81	
OCTUBRE	-7.61	-10.29	-11.96	-12.76	-14.34	-13.54	
NOVIEMBRE	-8.50	-3.39	-4.93	3.68	-0.96	-1.02	
DICIEMBRE	1.43	11.48	2.58	15.03	6.60	8.70	
ENERO	70.01	66.21	55.78	55.10	53.73	52.70	DESVIACIÓN TÍPICA DEL ERROR RELATIVO (%)
FEBRERO	58.95	79.24	80.65	76.00	81.61	83.00	
MARZO	61.34	75.92	92.02	94.58	97.03	97.89	
ABRIL	50.33	63.77	67.14	71.73	66.11	63.58	
MAYO	62.64	82.63	99.33	103.68	110.61	110.93	
JUNIO	42.88	41.52	50.62	48.61	52.27	51.92	
JULIO	24.43	26.36	10.37	13.79	10.85	10.47	
AGOSTO	59.76	49.49	59.16	53.75	62.39	57.67	
SEPTIEMBRE	37.15	45.89	45.61	49.26	46.99	46.15	
OCTUBRE	49.06	56.98	65.07	65.84	73.93	73.68	
NOVIEMBRE	75.45	80.34	75.65	71.86	74.58	74.07	
DICIEMBRE	61.69	61.09	65.13	55.44	62.82	63.16	



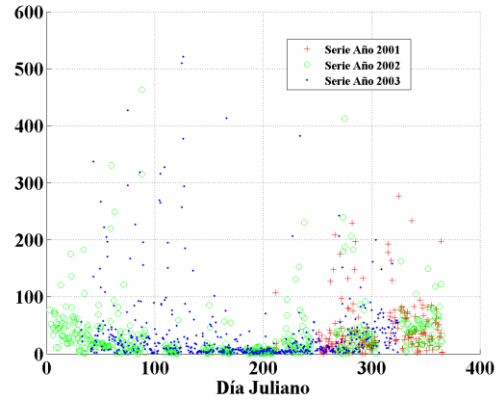
Horizonte de predicción 1 semidiá



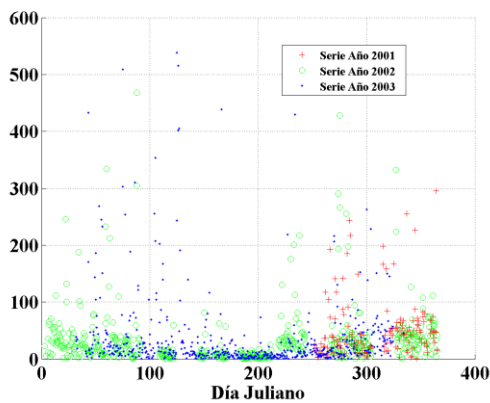
Horizonte de predicción 2 semidiás



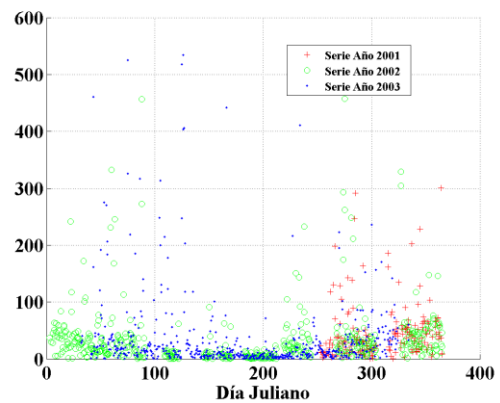
Horizonte de predicción 3 semidiás



Horizonte de predicción 4 semidiás

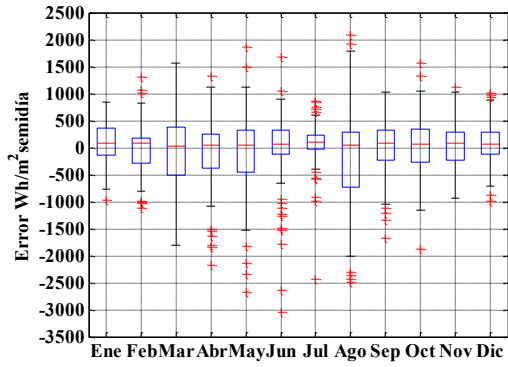


Horizonte de predicción 5 semidiás

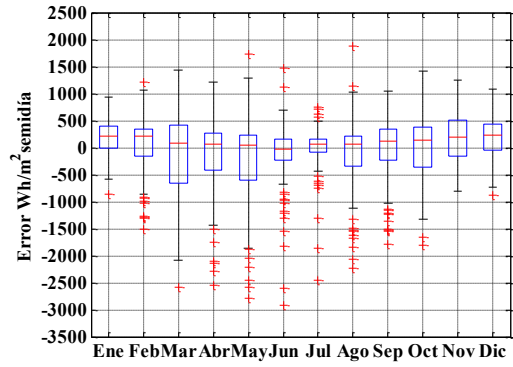


Horizonte de predicción 6 semidiás

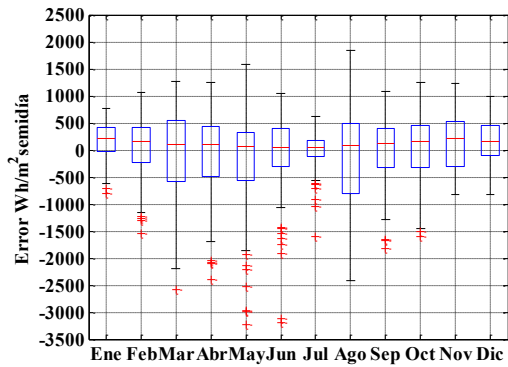
Figura C.6 Albacete: errores relativo de predicción en términos absolutos para cada día juliano y cada año



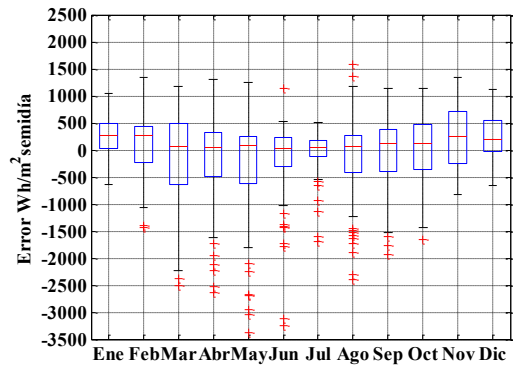
Horizonte de predicción 1 semidia



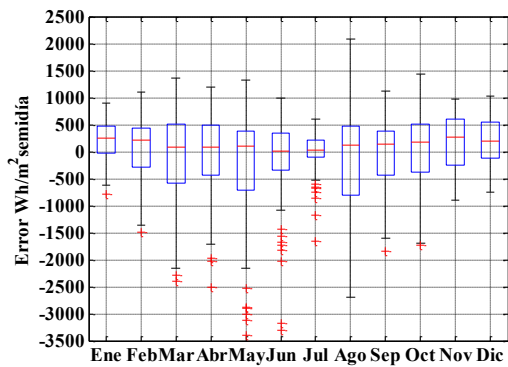
Horizonte de predicción 2 semidías



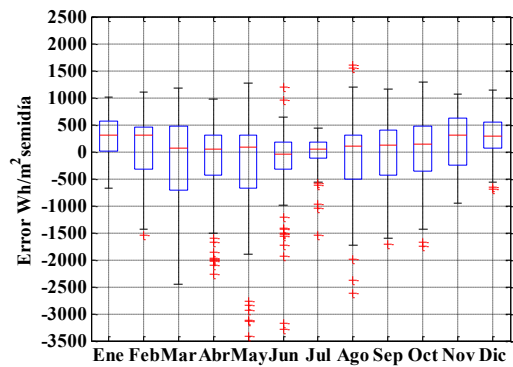
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

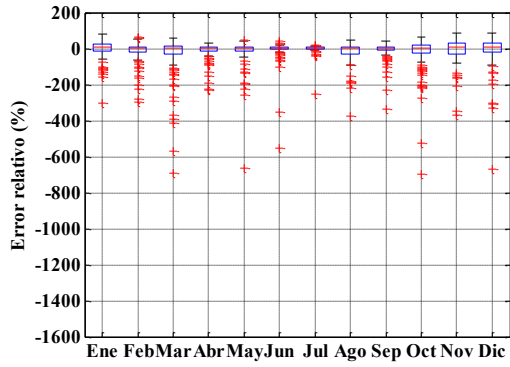


Horizonte de predicción 5 semidías

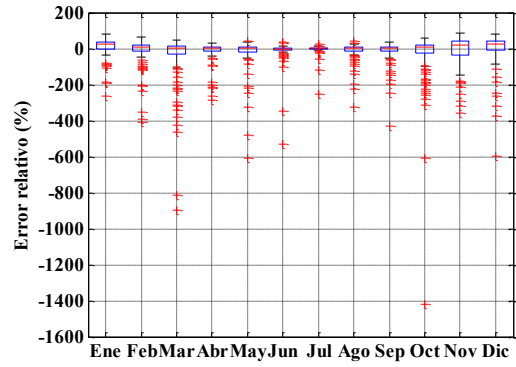


Horizonte de predicción 6 semidías

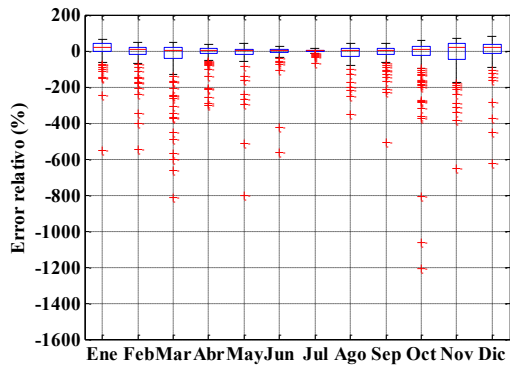
Figura C.7 Albacete: diagrama Box-Whisker de los errores de predicción para cada mes



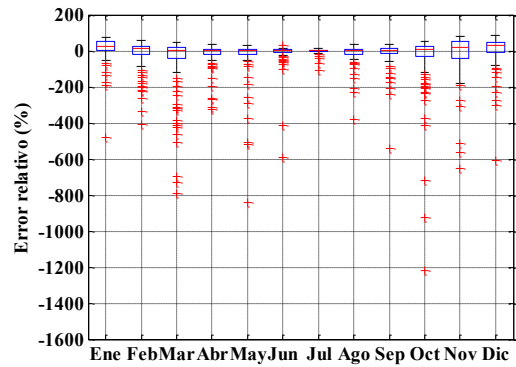
Horizonte de predicción 1 semidía



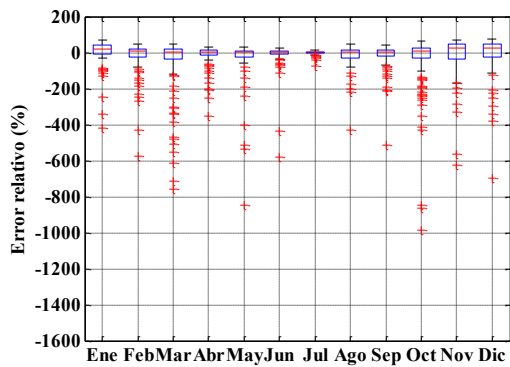
Horizonte de predicción 2 semidías



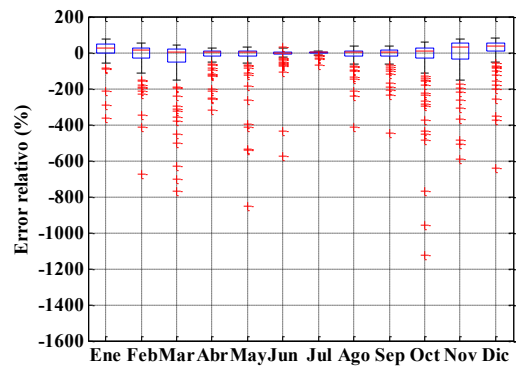
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

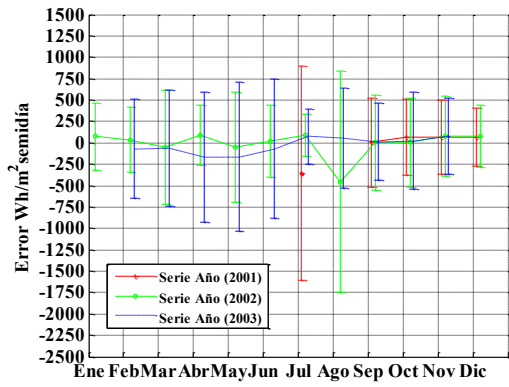


Horizonte de predicción 5 semidías

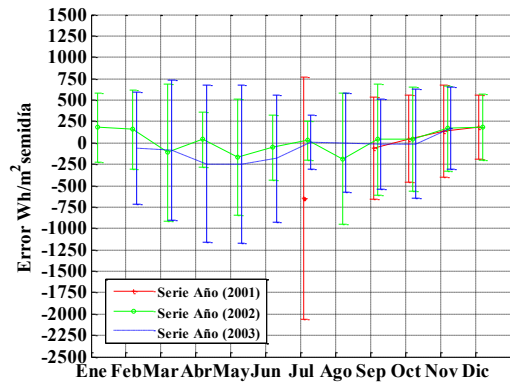


Horizonte de predicción 6 semidías

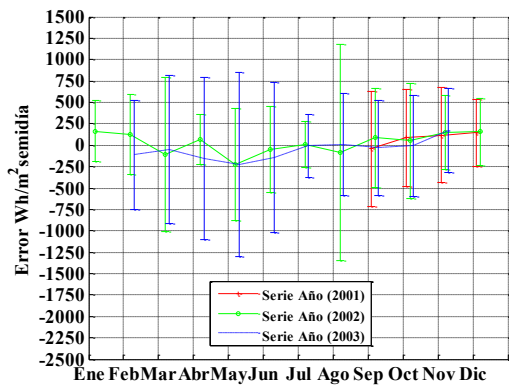
Figura C.8 Albacete: diagrama Box-Whisker de los errores relativos de predicción para cada mes



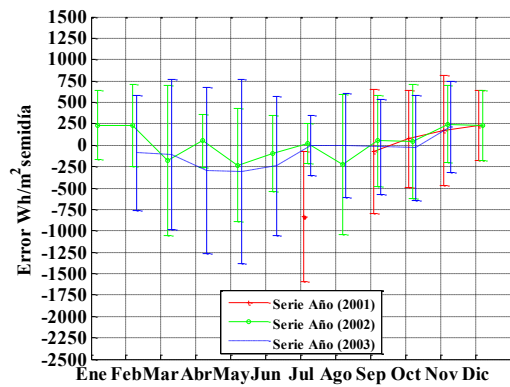
Horizonte de predicción 1 semidia



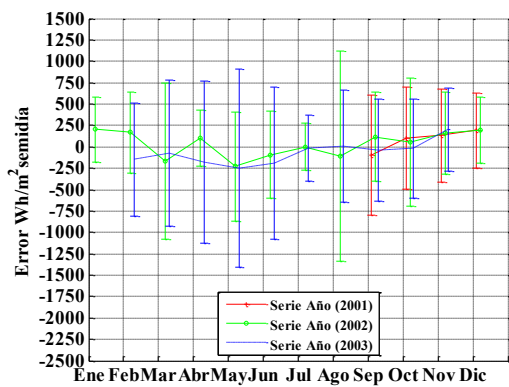
Horizonte de predicción 2 semidías



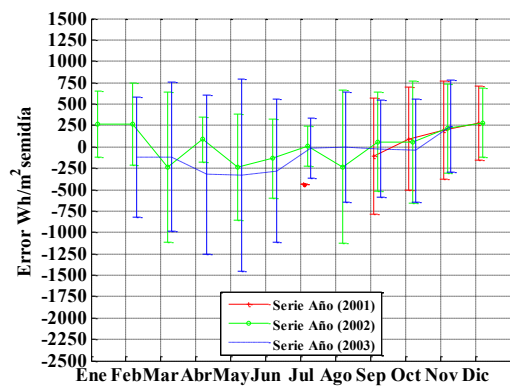
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

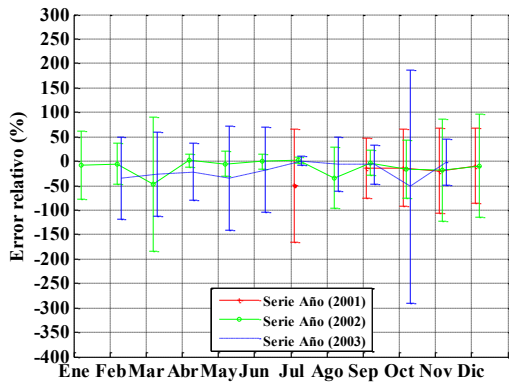


Horizonte de predicción 5 semidías

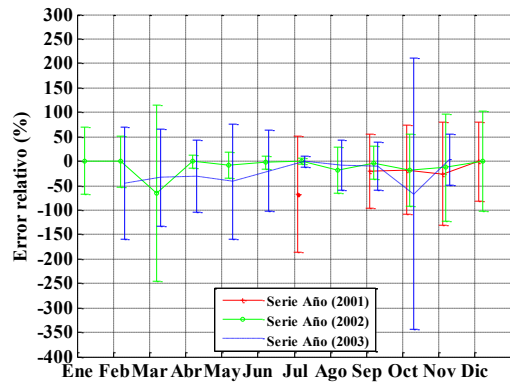


Horizonte de predicción 6 semidías

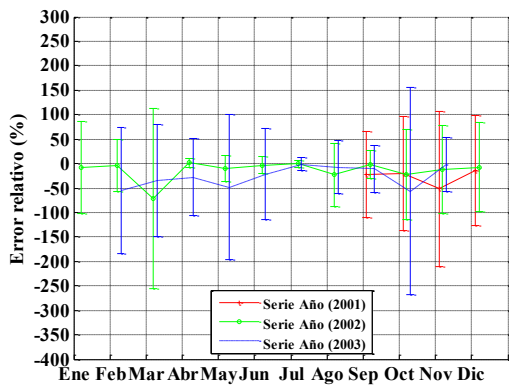
Figura C.9 Albacete: media y desviación mensual de los errores de predicción para cada mes y cada año



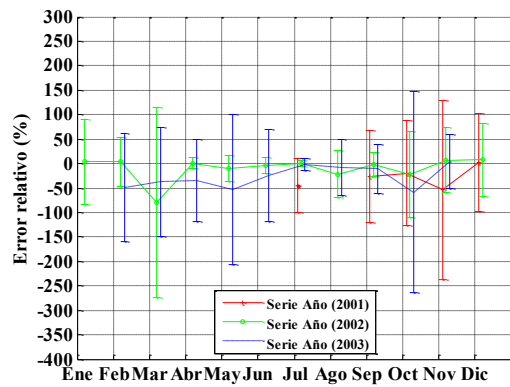
Horizonte de predicción 1 semidía



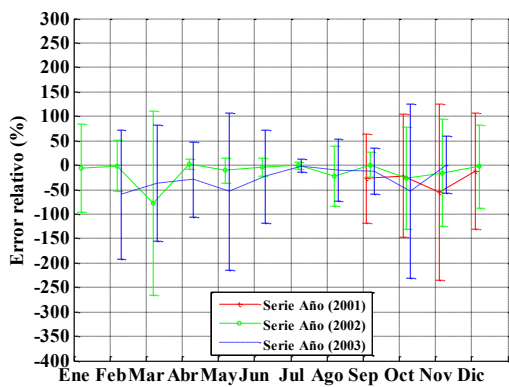
Horizonte de predicción 2 semidías



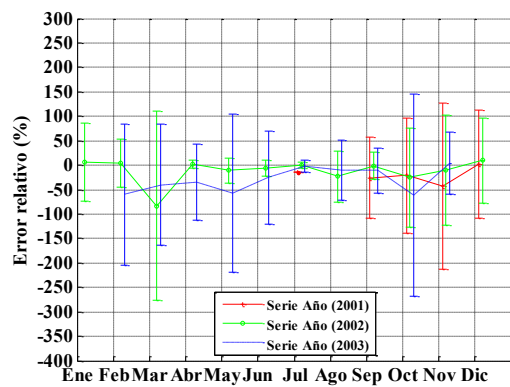
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías

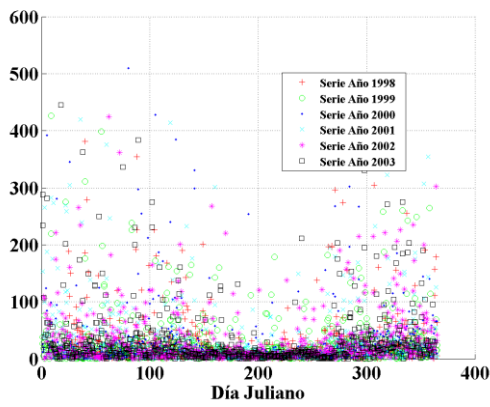


Horizonte de predicción 6 semidías

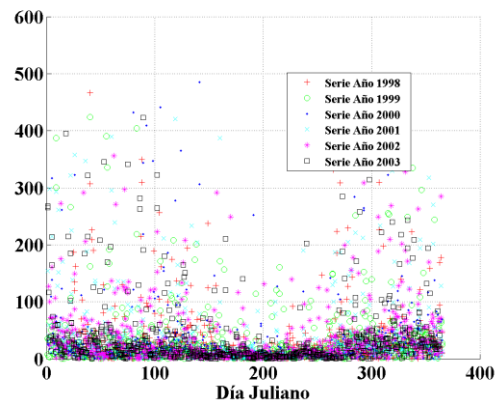
Figura C.10 Albacete: media y desviación mensual de los errores relativos de predicción para cada mes y cada año

Tabla C.3 Madrid: estadísticos del error relativo mensual

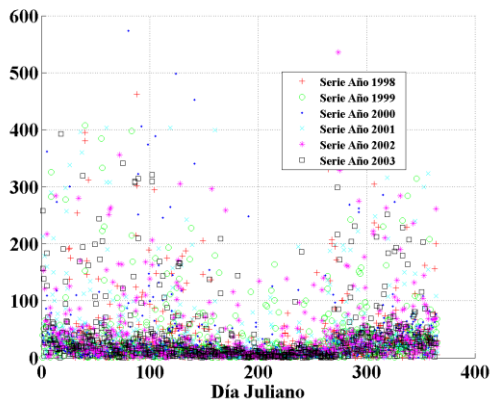
Mes	Horizonte Predicción (Semidías)						ESTADÍSTICO
	1	2	3	4	5	6	
<b>ENERO</b>	7.76	15.66	18.76	21.88	23.67	25.34	MEDIANA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	2.59	6.67	8.66	9.93	11.41	12.16	
<b>MARZO</b>	1.40	3.18	3.17	3.00	2.79	2.04	
<b>ABRIL</b>	3.56	3.01	3.04	2.86	3.03	2.20	
<b>MAYO</b>	2.89	2.06	0.20	0.52	0.50	0.46	
<b>JUNIO</b>	4.43	4.30	3.61	3.25	3.52	3.19	
<b>JULIO</b>	4.29	4.23	3.88	3.91	3.61	3.49	
<b>AGOSTO</b>	4.74	5.05	4.78	5.27	5.09	4.58	
<b>SEPTIEMBRE</b>	3.84	5.44	5.66	6.75	6.36	5.96	
<b>OCTUBRE</b>	6.00	8.95	11.39	13.12	13.68	11.85	
<b>NOVIEMBRE</b>	6.88	14.19	17.71	19.41	21.41	22.78	
<b>DICIEMBRE</b>	7.84	14.55	19.77	22.72	26.27	29.74	
<b>ENERO</b>	-21.04	-16.58	-12.64	-9.73	-8.18	-5.99	MEDIA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	-17.72	-17.33	-16.75	-15.68	-14.62	-14.55	
<b>MARZO</b>	-18.87	-22.55	-24.43	-25.69	-26.61	-27.71	
<b>ABRIL</b>	-16.65	-20.95	-22.65	-23.80	-24.62	-25.08	
<b>MAYO</b>	-16.04	-19.15	-21.35	-22.58	-23.48	-24.11	
<b>JUNIO</b>	-4.73	-6.53	-7.34	-7.71	-8.07	-8.37	
<b>JULIO</b>	-0.44	-1.29	-1.77	-2.13	-2.28	-2.38	
<b>AGOSTO</b>	-1.10	-1.83	-2.19	-2.38	-2.63	-2.66	
<b>SEPTIEMBRE</b>	-7.43	-9.29	-9.70	-10.16	-10.37	-10.90	
<b>OCTUBRE</b>	-11.14	-13.21	-13.57	-13.76	-13.96	-13.78	
<b>NOVIEMBRE</b>	-16.58	-13.22	-9.16	-7.01	-4.50	-3.09	
<b>DICIEMBRE</b>	-15.03	-8.10	-3.33	-0.42	2.22	3.78	
<b>ENERO</b>	82.98	86.75	84.22	80.89	83.16	81.26	DESVIACIÓN TÍPICA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	67.90	76.60	79.32	76.91	75.69	76.56	
<b>MARZO</b>	70.83	78.18	83.27	85.33	87.07	90.91	
<b>ABRIL</b>	59.31	68.09	69.00	69.77	70.67	71.32	
<b>MAYO</b>	52.02	57.20	61.80	62.94	63.25	63.79	
<b>JUNIO</b>	36.02	40.75	41.81	41.30	41.52	42.93	
<b>JULIO</b>	22.35	23.44	23.67	24.43	24.39	24.22	
<b>AGOSTO</b>	25.77	27.50	27.62	27.77	28.06	27.91	
<b>SEPTIEMBRE</b>	43.84	51.87	52.00	53.50	53.38	55.44	
<b>OCTUBRE</b>	60.57	70.68	75.27	77.49	78.58	77.99	
<b>NOVIEMBRE</b>	70.12	76.27	74.64	73.45	71.61	70.91	
<b>DICIEMBRE</b>	65.30	67.69	66.37	66.80	64.97	65.59	



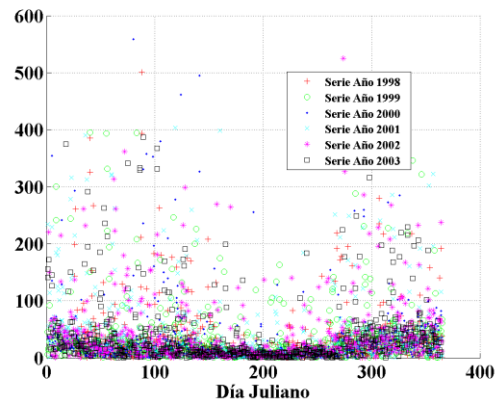
Horizonte de predicción 1 semidiá



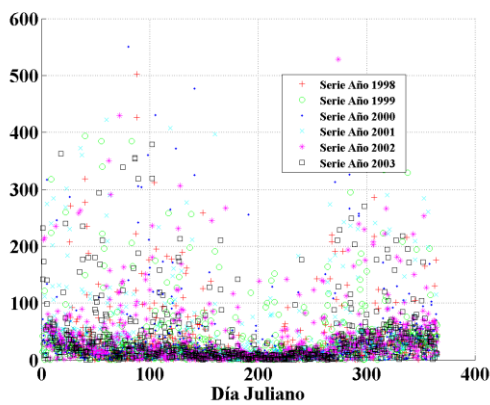
Horizonte de predicción 2 semidiás



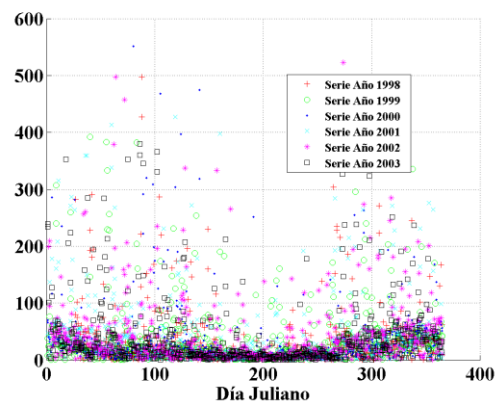
Horizonte de predicción 3 semidiás



Horizonte de predicción 4 semidiás

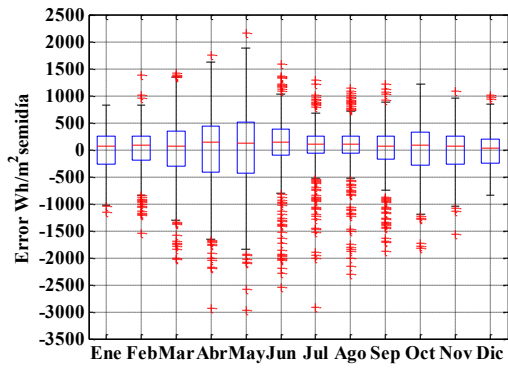


Horizonte de predicción 5 semidiás

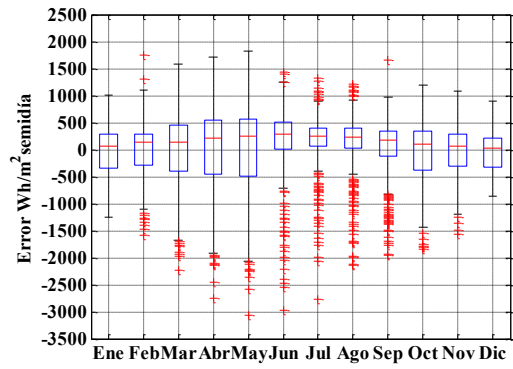


Horizonte de predicción 6 semidiás

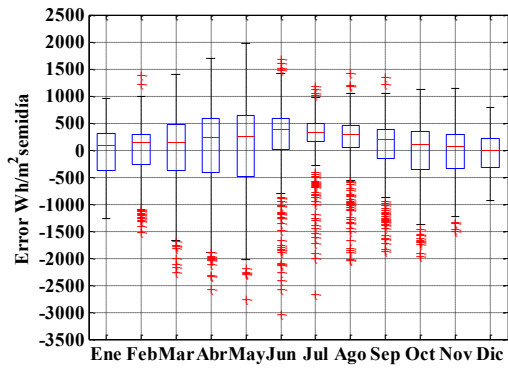
Figura C.11 Madrid: errores relativo de predicción en términos absolutos para cada día juliano y cada año



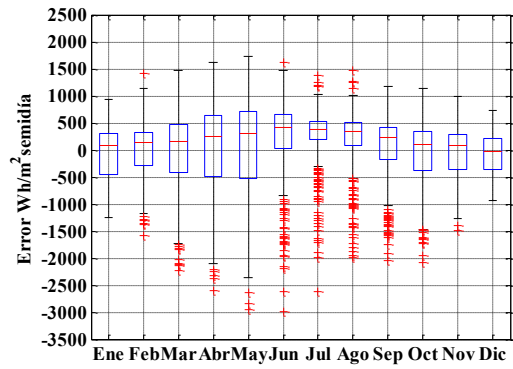
Horizonte de predicción 1 semidia



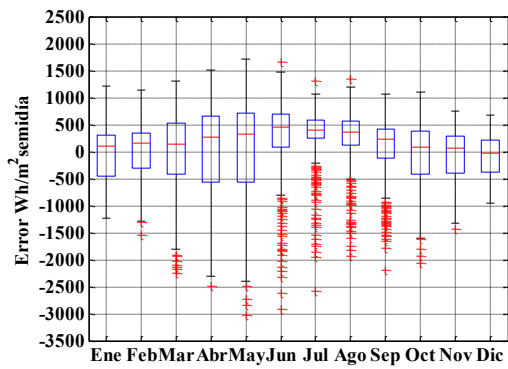
Horizonte de predicción 2 semidías



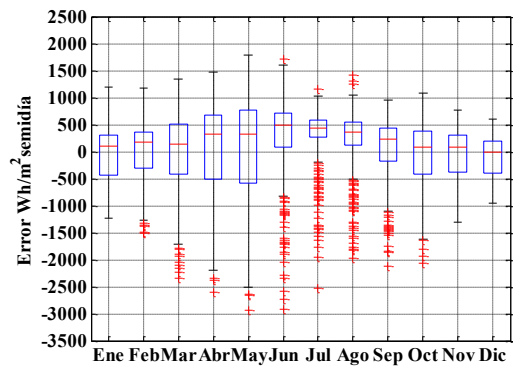
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

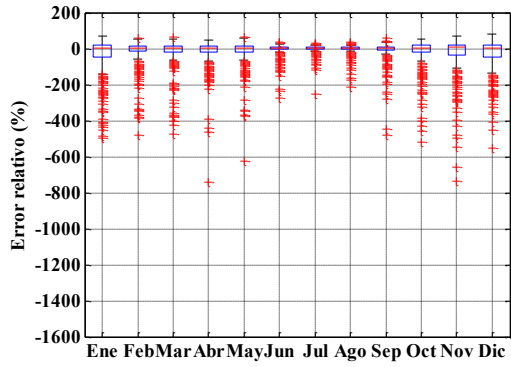


Horizonte de predicción 5 semidías

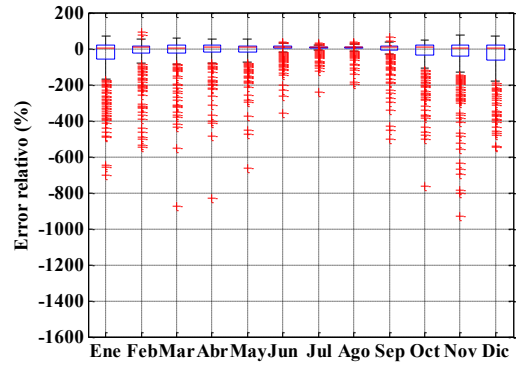


Horizonte de predicción 6 semidías

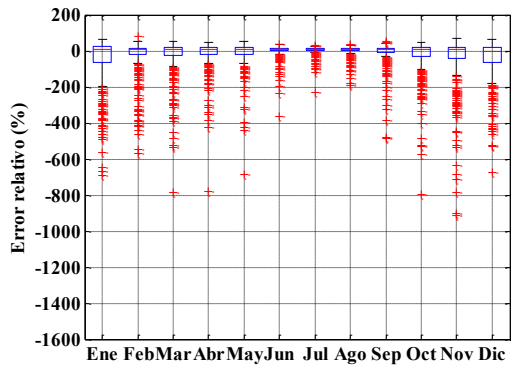
Figura C.12 Madrid: diagrama Box-Whisker de los errores de predicción para cada mes



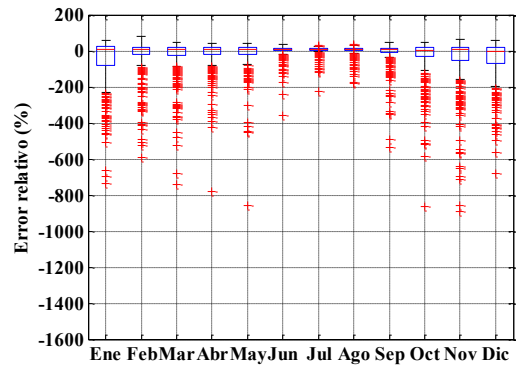
Horizonte de predicción 1 semidía



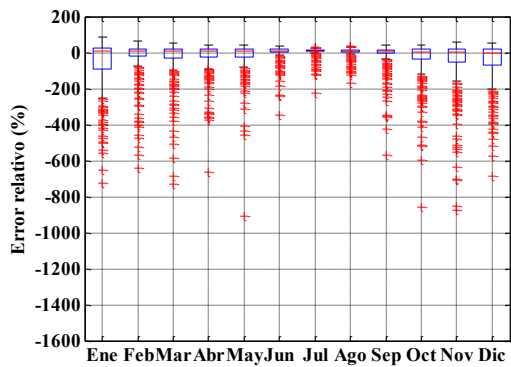
Horizonte de predicción 2 semidías



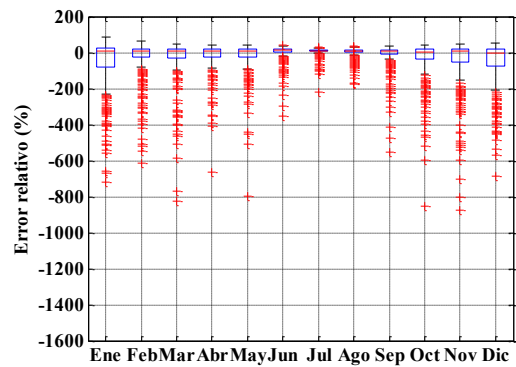
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

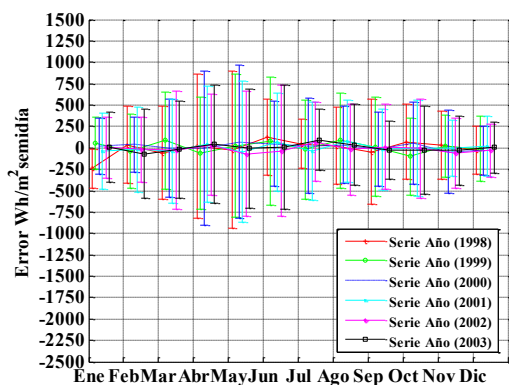


Horizonte de predicción 5 semidías

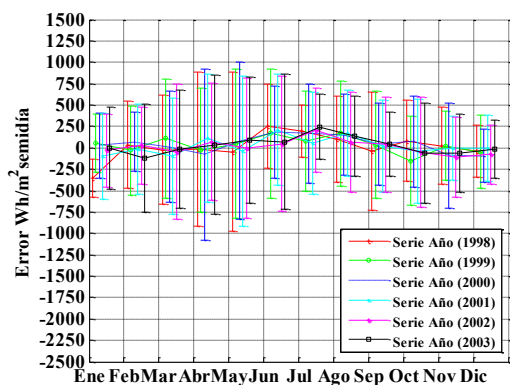


Horizonte de predicción 6 semidías

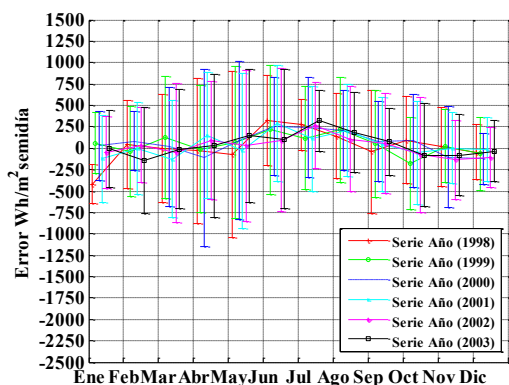
Figura C.13 Madrid: diagrama Box-Whisker de los errores relativos de predicción para cada mes



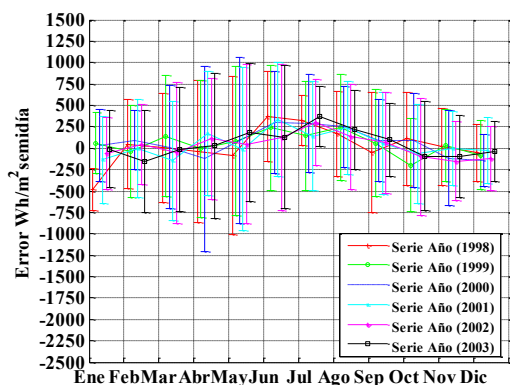
Horizonte de predicción 1 semidia



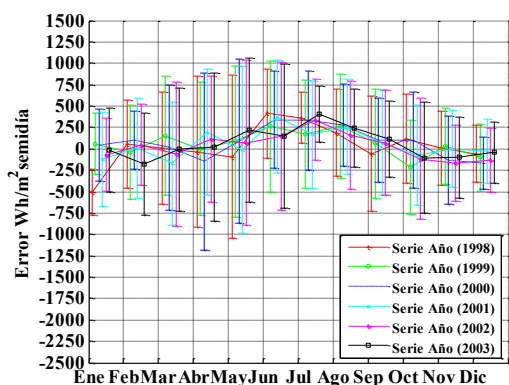
Horizonte de predicción 2 semidías



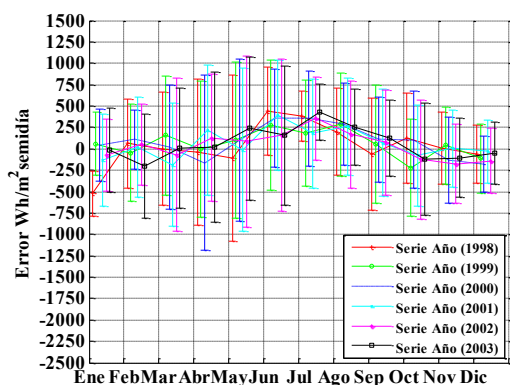
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

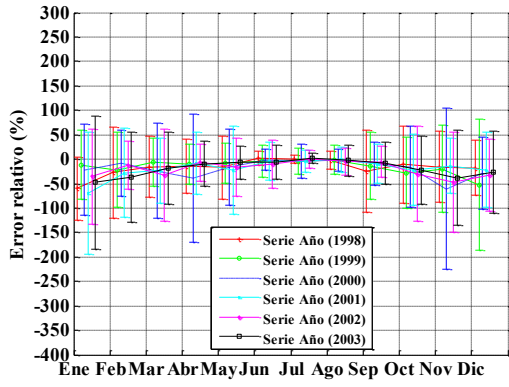


Horizonte de predicción 5 semidías

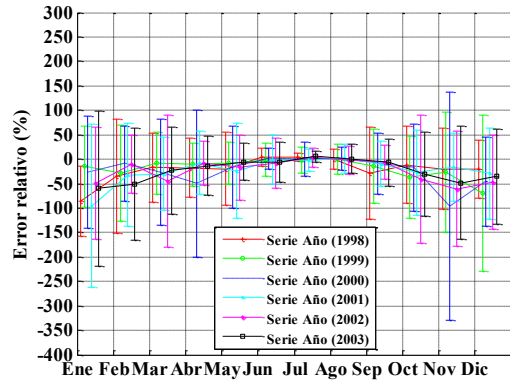


Horizonte de predicción 6 semidías

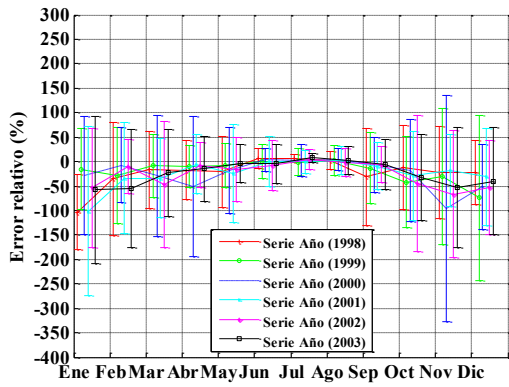
Figura C.14 Madrid: media y desviación mensual de los errores de predicción para cada mes y cada año



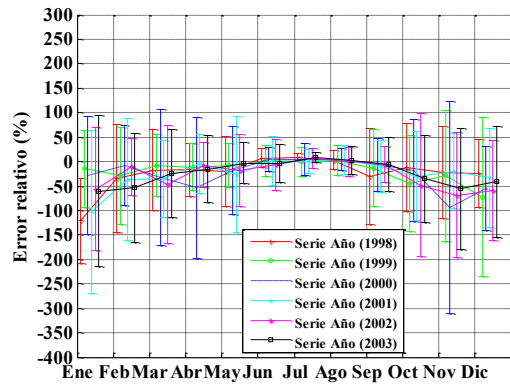
Horizonte de predicción 1 semidía



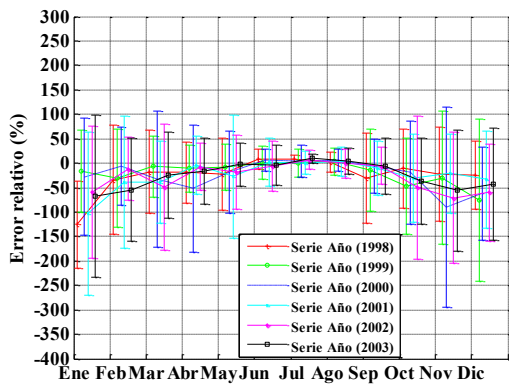
Horizonte de predicción 2 semidías



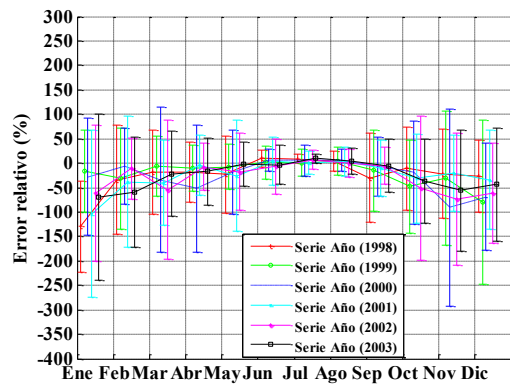
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías

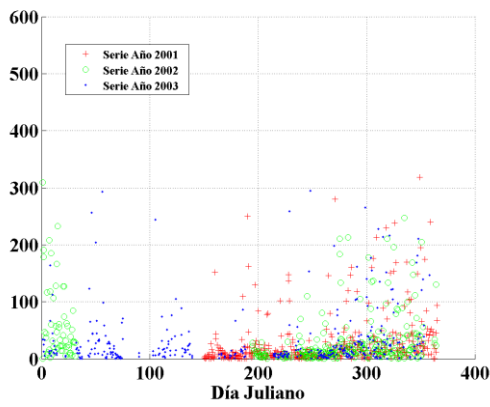


Horizonte de predicción 6 semidías

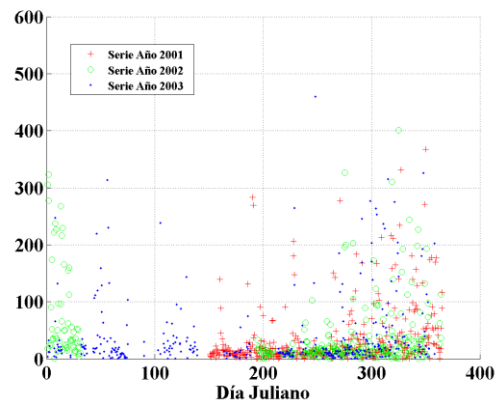
Figura C.15 Madrid: media y desviación mensual de los errores relativos de predicción para cada mes y cada año

Tabla C.4 Lérida: estadísticos del error relativo mensual

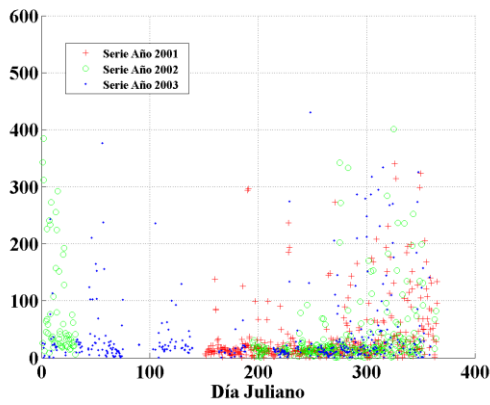
Mes	Horizonte Predicción (Semidías)						ESTADÍSTICO
	1	2	3	4	5	6	
<b>ENERO</b>	-2.53	-11.02	-16.03	-14.77	-18.60	-13.62	MEDIANA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	0.52	1.15	-4.12	-3.85	-1.37	-3.27	
<b>MARZO</b>	3.41	8.53	10.70	12.35	12.87	13.99	
<b>ABRIL</b>	6.29	7.93	13.39	8.46	5.74	6.23	
<b>MAYO</b>	6.86	14.59	16.79	16.09	16.99	15.17	
<b>JUNIO</b>	3.58	9.19	10.79	11.87	12.89	14.44	
<b>JULIO</b>	5.50	8.27	10.91	11.34	12.65	12.65	
<b>AGOSTO</b>	4.63	8.60	9.99	10.94	11.38	11.00	
<b>SEPTIEMBRE</b>	4.05	8.78	10.48	10.92	11.72	12.38	
<b>OCTUBRE</b>	3.38	4.15	4.75	5.45	3.35	4.94	
<b>NOVIEMBRE</b>	3.42	3.52	0.67	-2.02	-3.75	-6.16	
<b>DICIEMBRE</b>	-8.66	-17.33	-20.21	-26.79	-25.47	-23.63	
<b>ENERO</b>	-29.17	-45.59	-53.15	-57.36	-61.26	-63.82	MEDIA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	-27.02	-34.53	-37.55	-40.00	-40.07	-40.70	
<b>MARZO</b>	1.52	3.78	5.14	5.88	6.66	7.24	
<b>ABRIL</b>	-14.41	-15.32	-14.83	-12.90	-11.62	-12.45	
<b>MAYO</b>	-1.30	-0.42	0.72	1.42	2.10	1.42	
<b>JUNIO</b>	-0.35	3.37	4.91	5.79	6.71	7.69	
<b>JULIO</b>	-4.20	-3.02	-2.35	-1.99	-1.29	-0.43	
<b>AGOSTO</b>	-3.82	-3.04	-2.76	-2.46	-1.98	-1.55	
<b>SEPTIEMBRE</b>	-6.53	-5.90	-5.01	-4.48	-4.07	-3.67	
<b>OCTUBRE</b>	-15.15	-21.08	-23.59	-24.36	-25.61	-26.50	
<b>NOVIEMBRE</b>	-23.91	-36.56	-41.62	-45.16	-46.08	-46.45	
<b>DICIEMBRE</b>	-33.83	-46.69	-52.07	-55.77	-57.27	-59.15	
<b>ENERO</b>	74.68	94.05	103.50	109.61	114.49	116.87	DESVIACIÓN TÍPICA DEL ERROR RELATIVO (%)
<b>FEBRERO</b>	75.61	85.01	93.10	94.75	92.24	92.60	
<b>MARZO</b>	21.57	24.44	24.56	24.32	24.36	24.86	
<b>ABRIL</b>	56.93	58.71	58.49	54.84	52.68	53.83	
<b>MAYO</b>	35.65	43.16	41.71	39.86	36.18	40.31	
<b>JUNIO</b>	22.37	23.83	25.23	25.01	25.02	24.69	
<b>JULIO</b>	40.60	47.71	50.57	51.22	50.62	48.59	
<b>AGOSTO</b>	35.26	42.04	45.76	46.87	46.83	46.80	
<b>SEPTIEMBRE</b>	44.12	53.81	52.85	51.72	51.74	51.84	
<b>OCTUBRE</b>	54.85	66.67	73.43	73.51	75.67	77.51	
<b>NOVIEMBRE</b>	67.90	86.75	92.99	100.71	98.42	95.87	
<b>DICIEMBRE</b>	73.40	82.91	86.15	89.42	91.97	93.80	



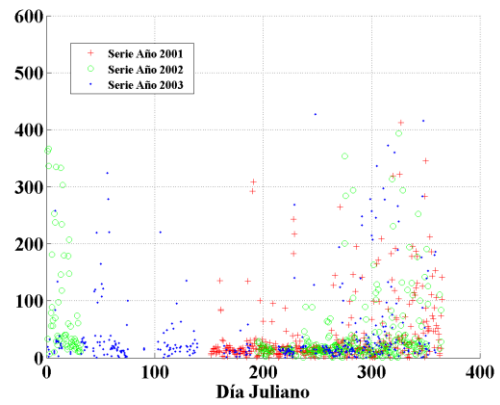
Horizonte de predicción 1 semidía



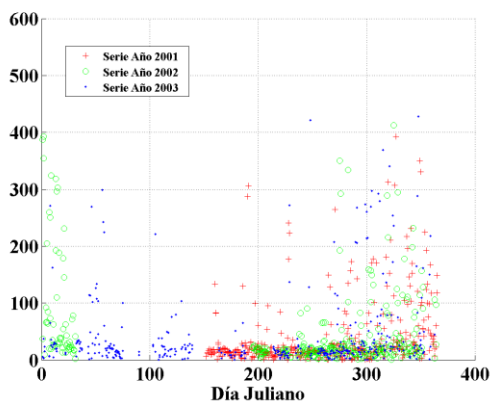
Horizonte de predicción 2 semidías



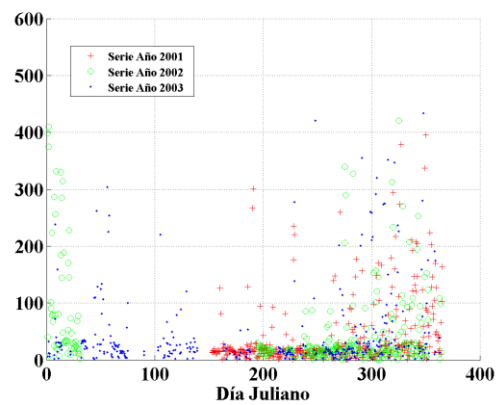
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

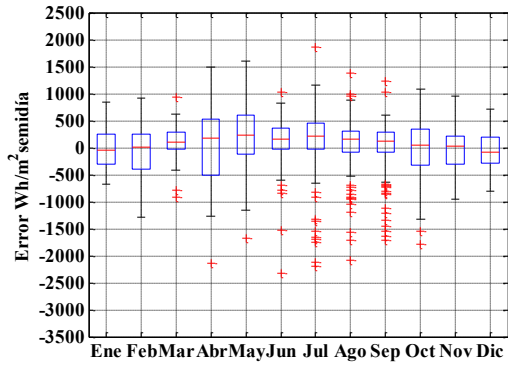


Horizonte de predicción 5 semidías

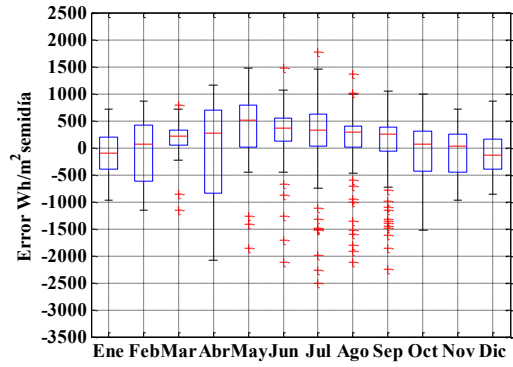


Horizonte de predicción 6 semidías

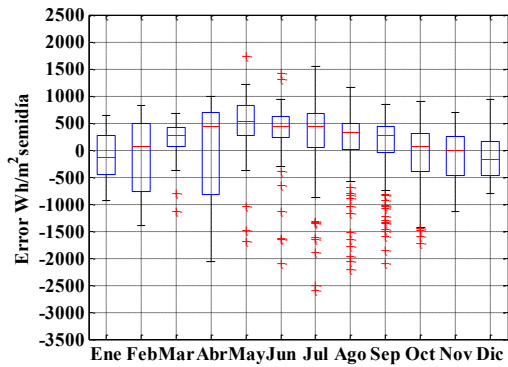
Figura C.16 Lérída: errores relativo de predicción en términos absolutos para cada día juliano y cada año



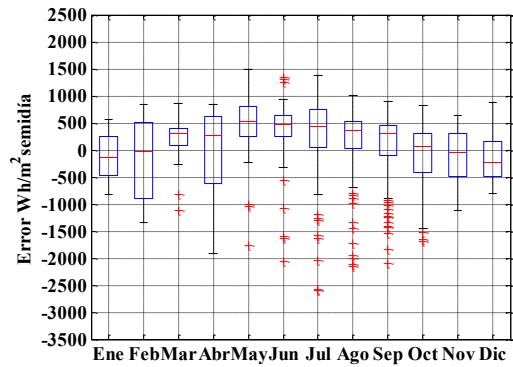
Horizonte de predicción 1 semidia



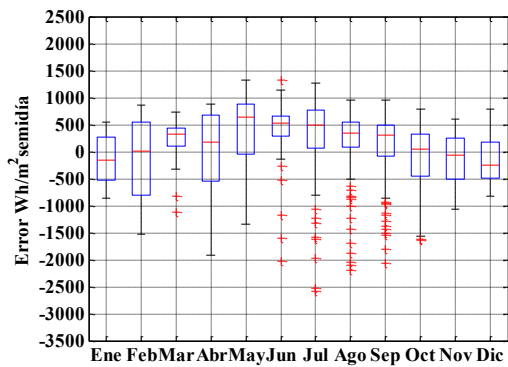
Horizonte de predicción 2 semidías



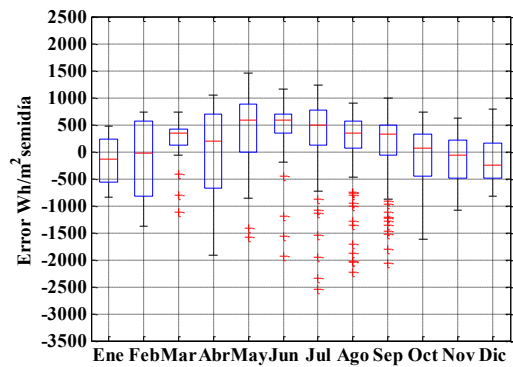
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

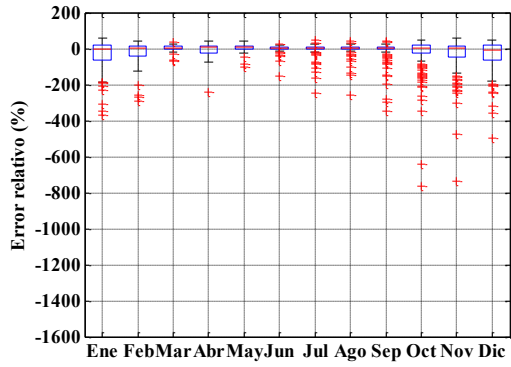


Horizonte de predicción 5 semidías

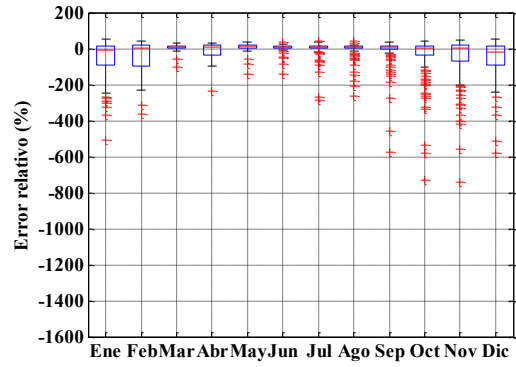


Horizonte de predicción 6 semidías

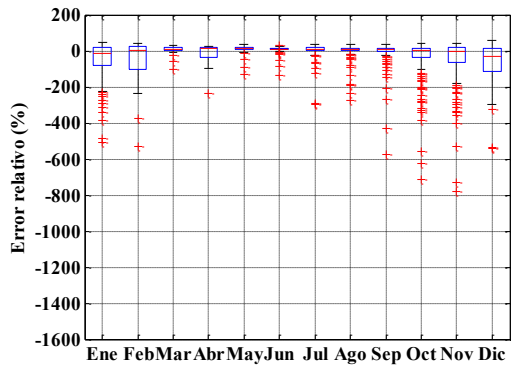
Figura C.17 Lérída: diagrama Box-Whisker de los errores de predicción para cada mes



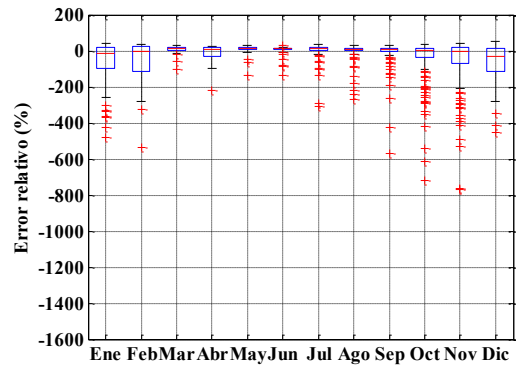
Horizonte de predicción 1 semidía



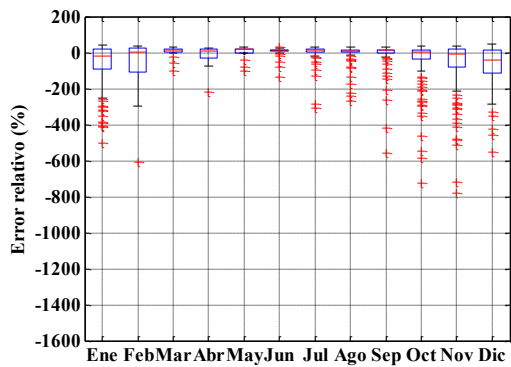
Horizonte de predicción 2 semidías



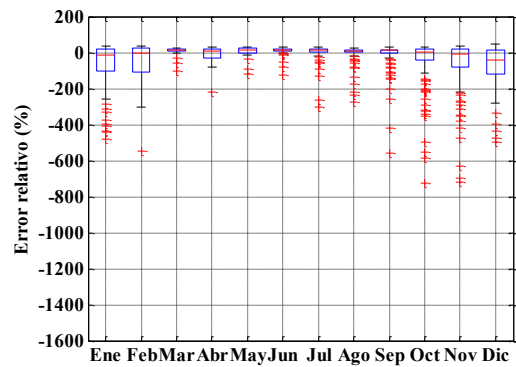
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

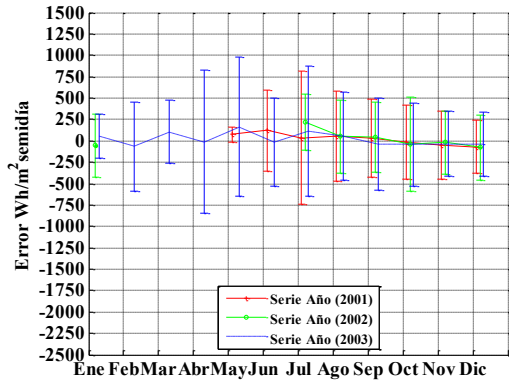


Horizonte de predicción 5 semidías

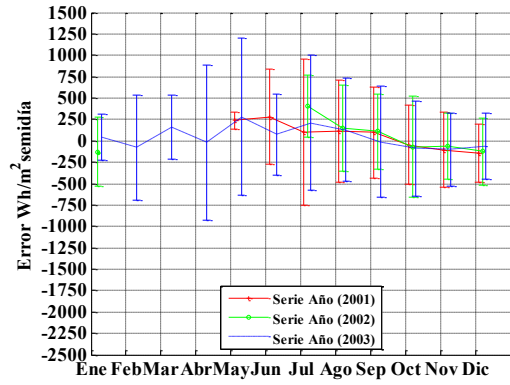


Horizonte de predicción 6 semidías

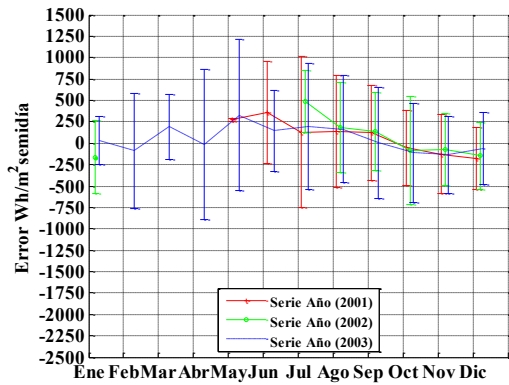
Figura C.18 Lérída: diagrama Box-Whisker de los errores relativos de predicción para cada mes



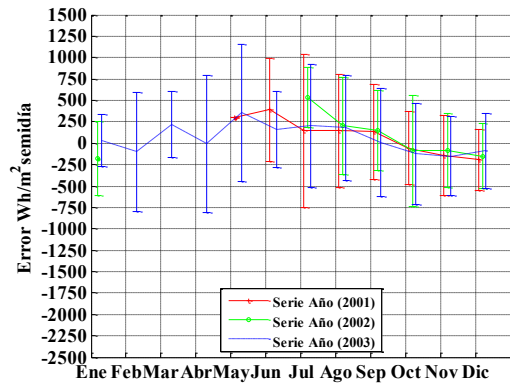
Horizonte de predicción 1 semidía



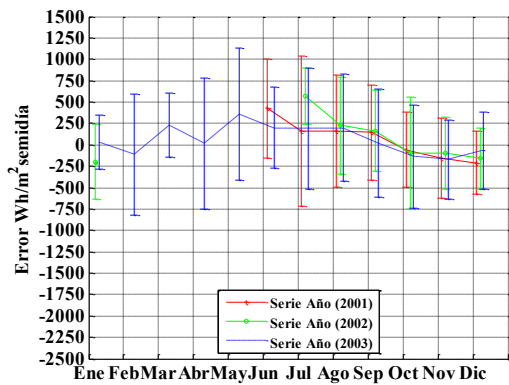
Horizonte de predicción 2 semidías



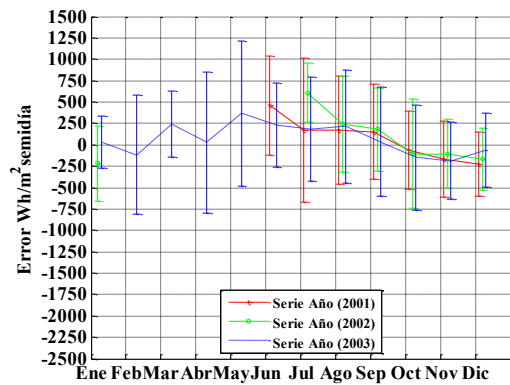
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías

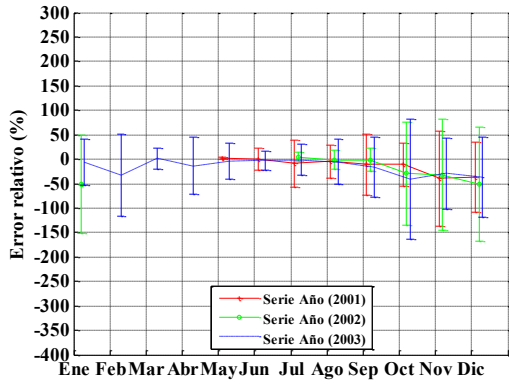


Horizonte de predicción 5 semidías

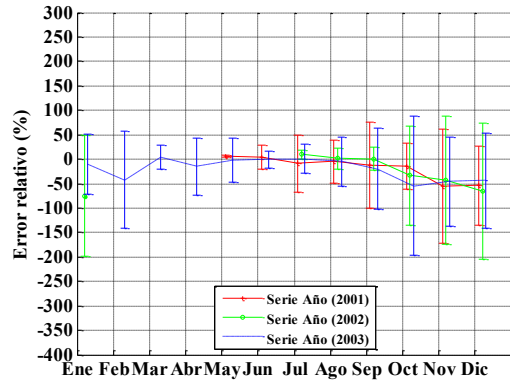


Horizonte de predicción 6 semidías

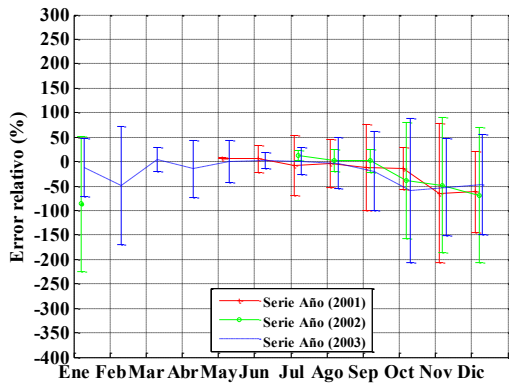
Figura C.19 Lérica: media y desviación mensual de los errores de predicción para cada mes y cada año



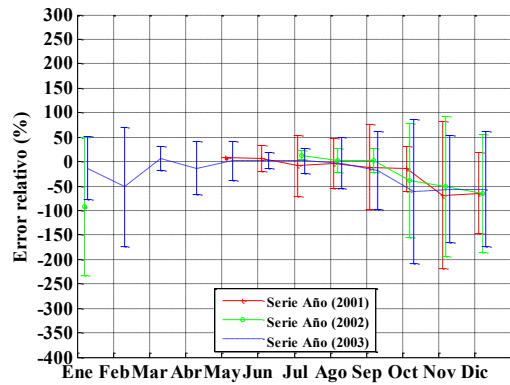
Horizonte de predicción 1 semidía



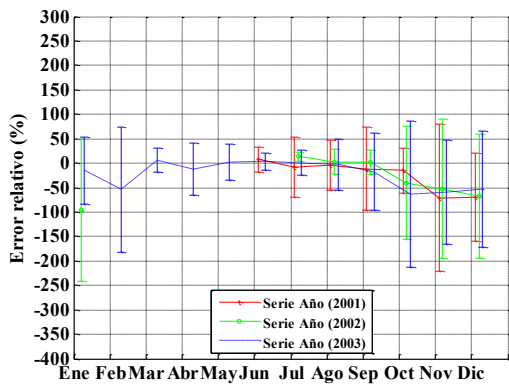
Horizonte de predicción 2 semidías



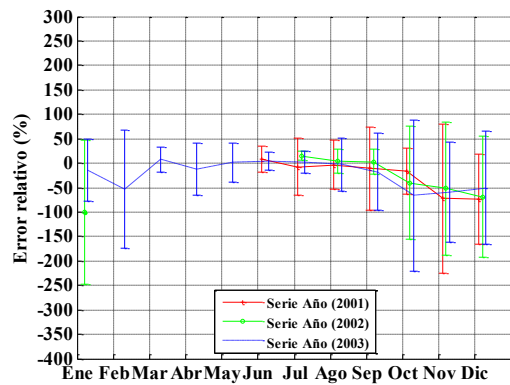
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías



Horizonte de predicción 6 semidías

Figura C.20 Llerida: media y desviación mensual de los errores relativos de predicción para cada mes y cada año



## **APÉNDICE D – ANÁLISIS DE LA FUNCIÓN DE DISTRIBUCIÓN ACUMULADA Y TEST DE KOLMOGOROV-SMIRNOV**

Se presenta en este apéndice los resultados de la aplicación del Test de Kolmogorov-Smirnov a los resultados de predicción de los modelos finales de predicción para cada una de las estaciones radiométricas utilizadas en este trabajo.

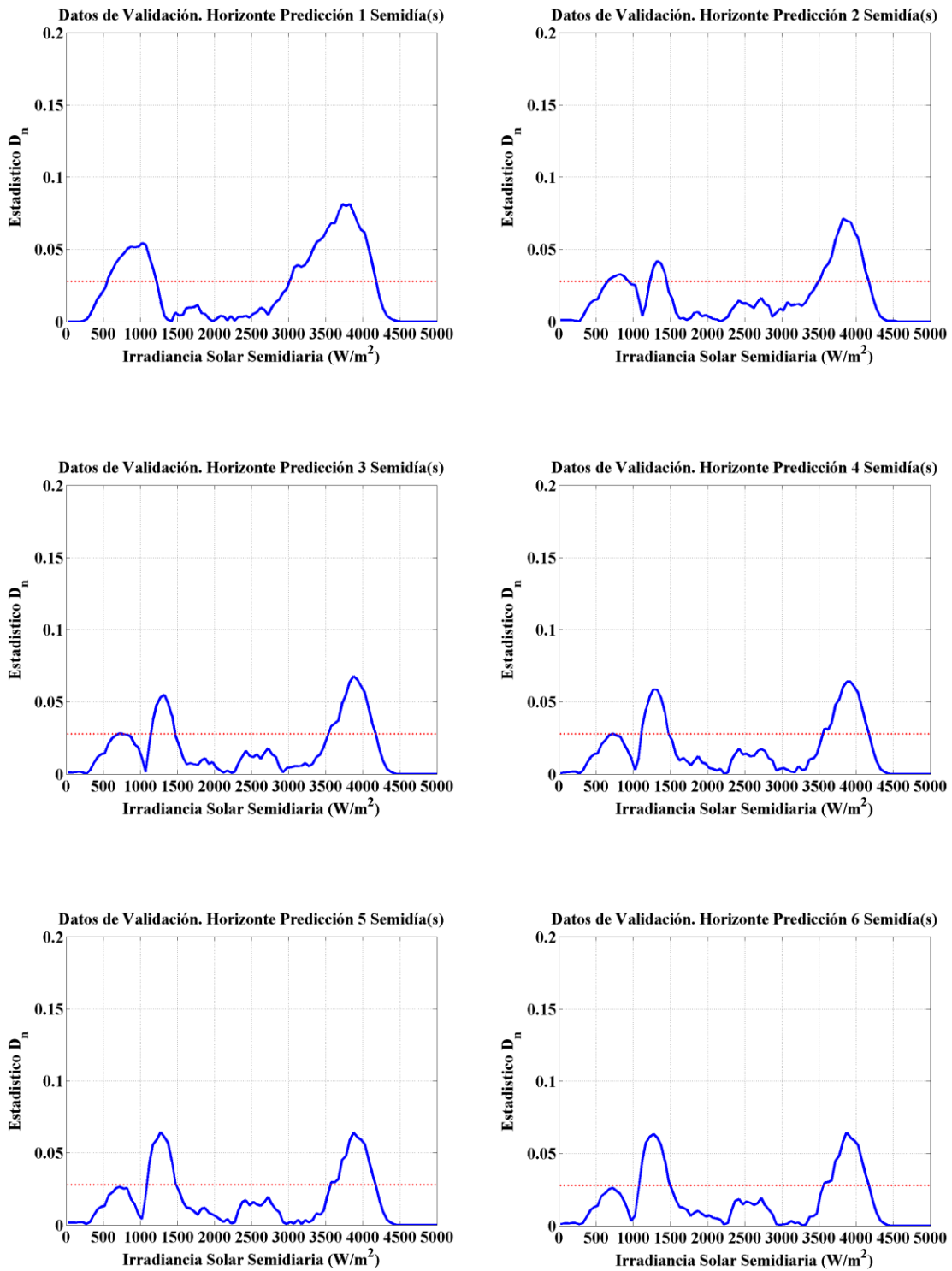
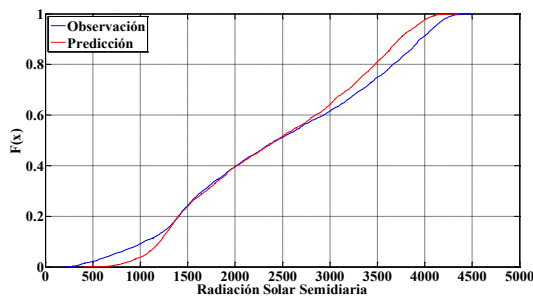
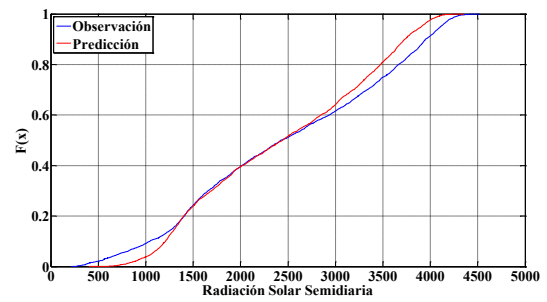


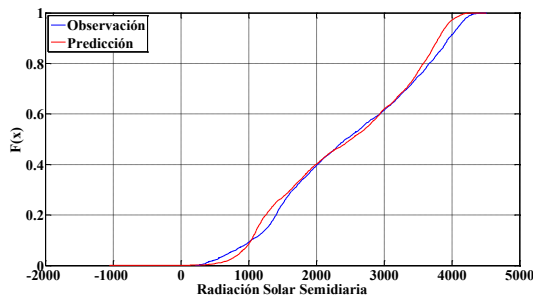
Figura D.1 Murcia: Test de de Kolmogorov-Smirnov irradiancia solar semidiaria



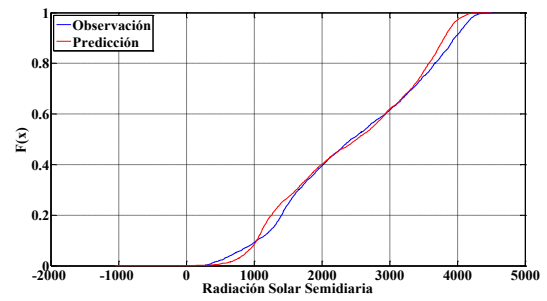
Horizonte de predicción 1 semidía



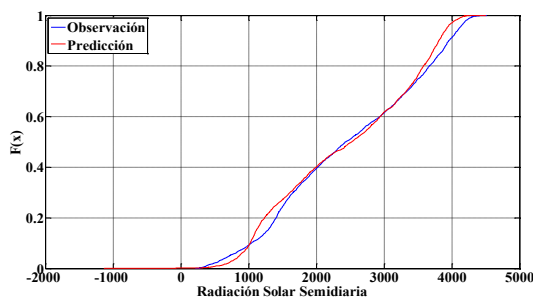
Horizonte de predicción 2 semidías



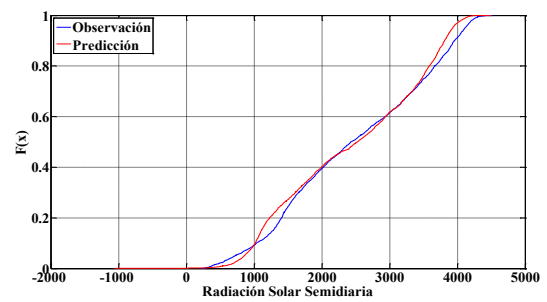
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías



Horizonte de predicción 6 semidías

Figura D.2 Murcia: CDF irradiancia solar semidiaria

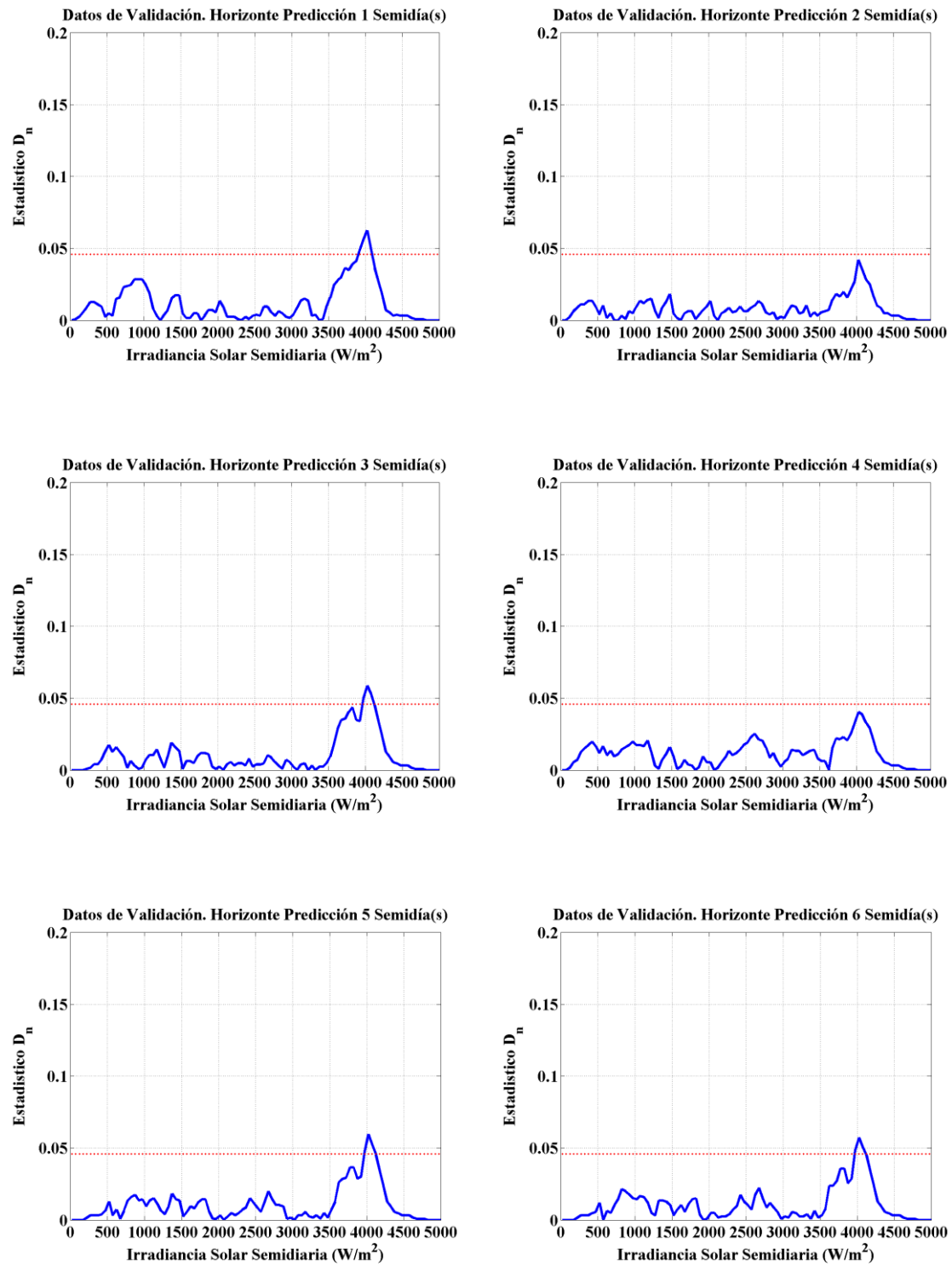
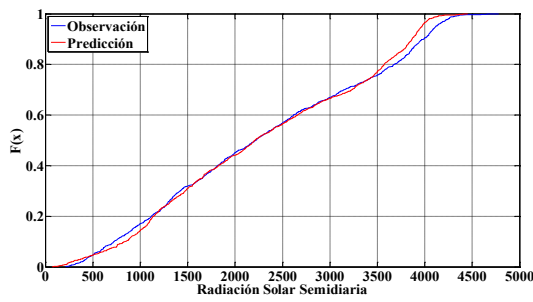
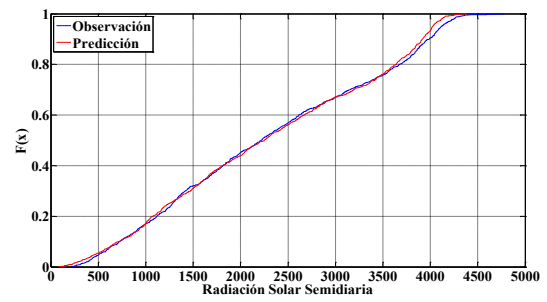


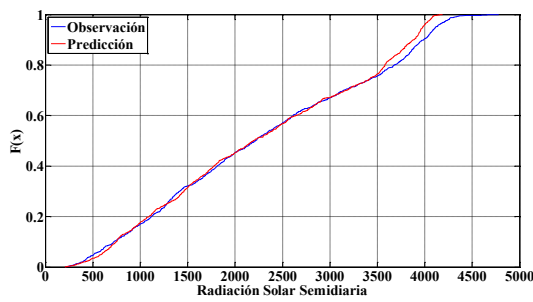
Figura D.3 Albacete: Test de de Kolmogorov-Smirnov irradiancia solar semidiaria



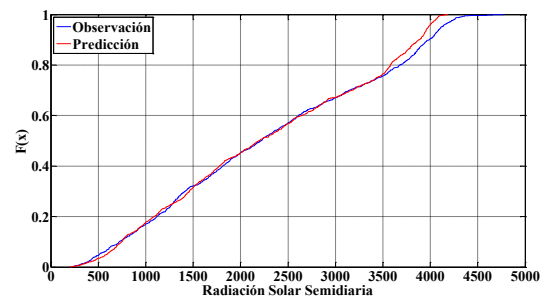
Horizonte de predicción 1 semidías



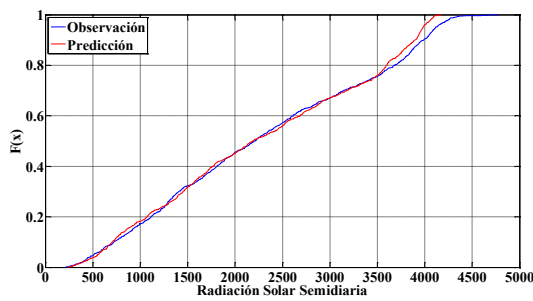
Horizonte de predicción 2 semidías



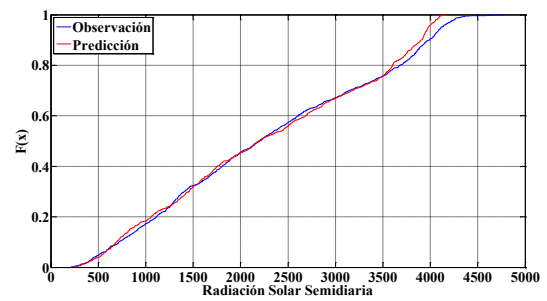
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías



Horizonte de predicción 6 semidías

Figura D.4 Albacete: CDF irradiancia solar semidía

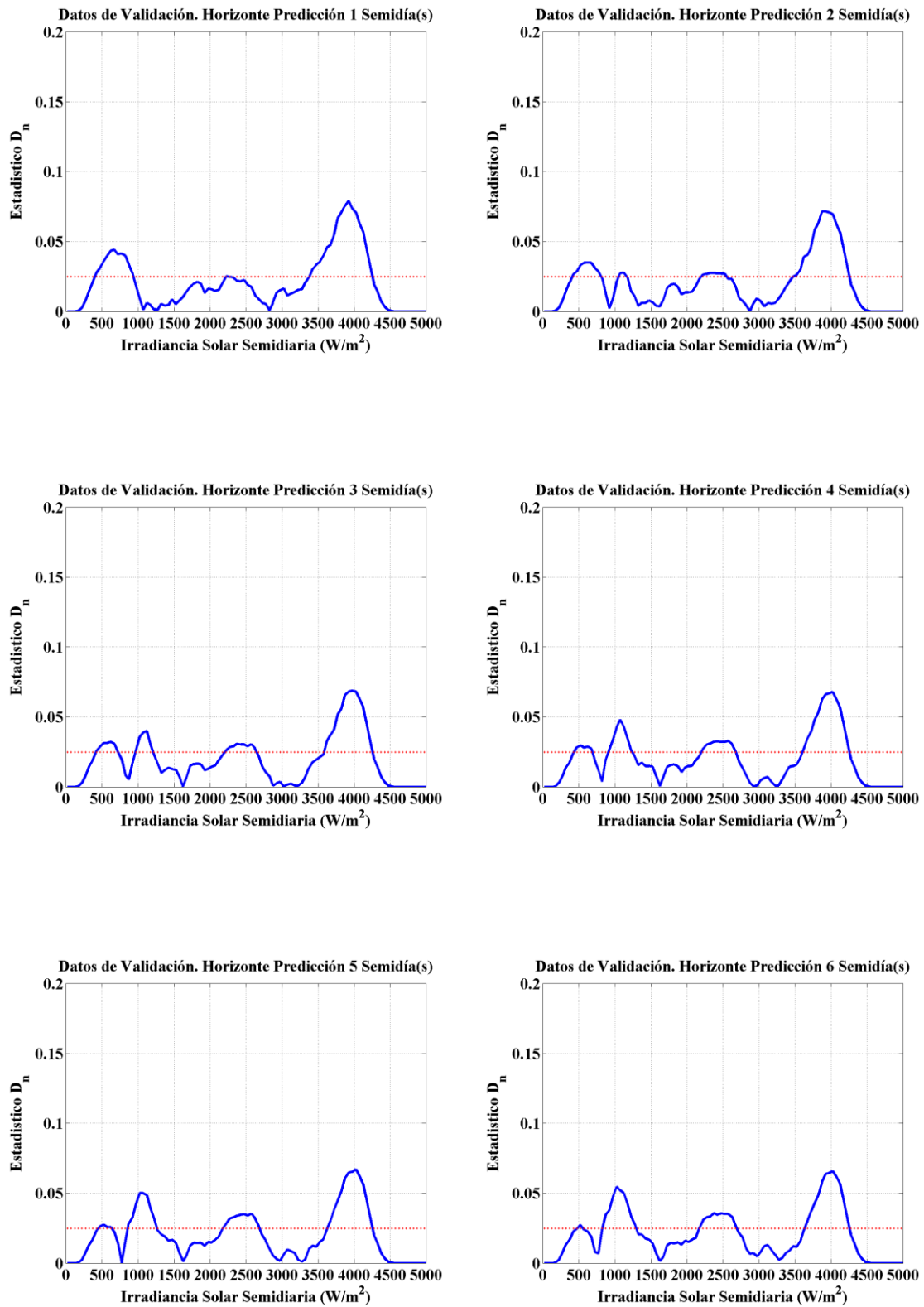
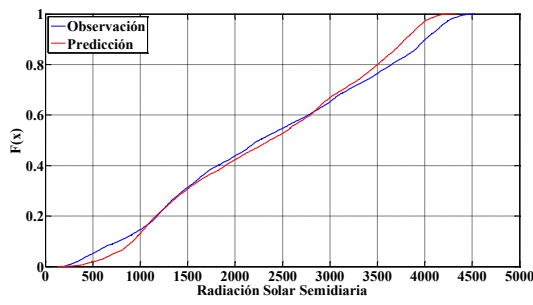
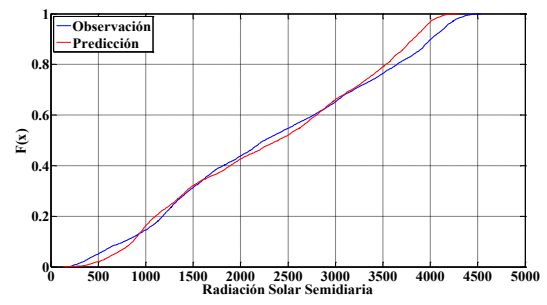


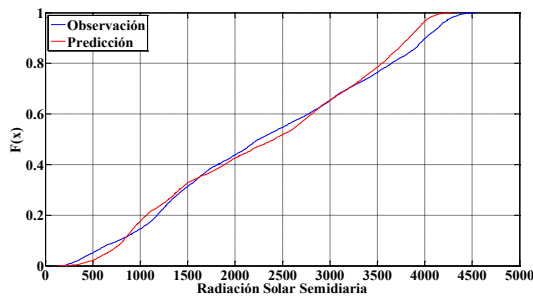
Figura D.5 Madrid: Test de de Kolmogorov-Smirnov irradiancia solar semidiaria



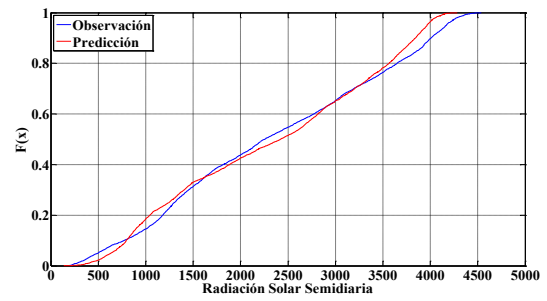
Horizonte de predicción 1 semidías



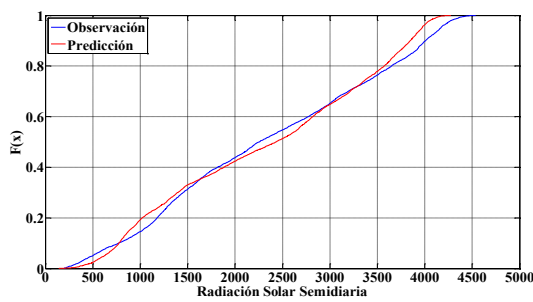
Horizonte de predicción 2 semidías



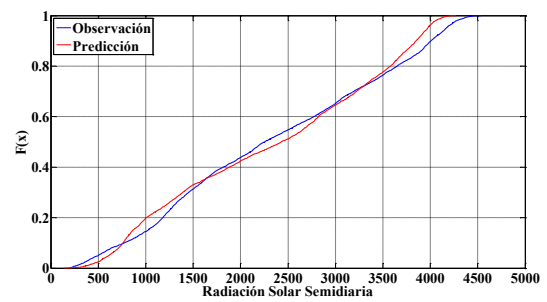
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías



Horizonte de predicción 6 semidías

Figura D.6 Madrid: CDF irradiancia solar semidiaria

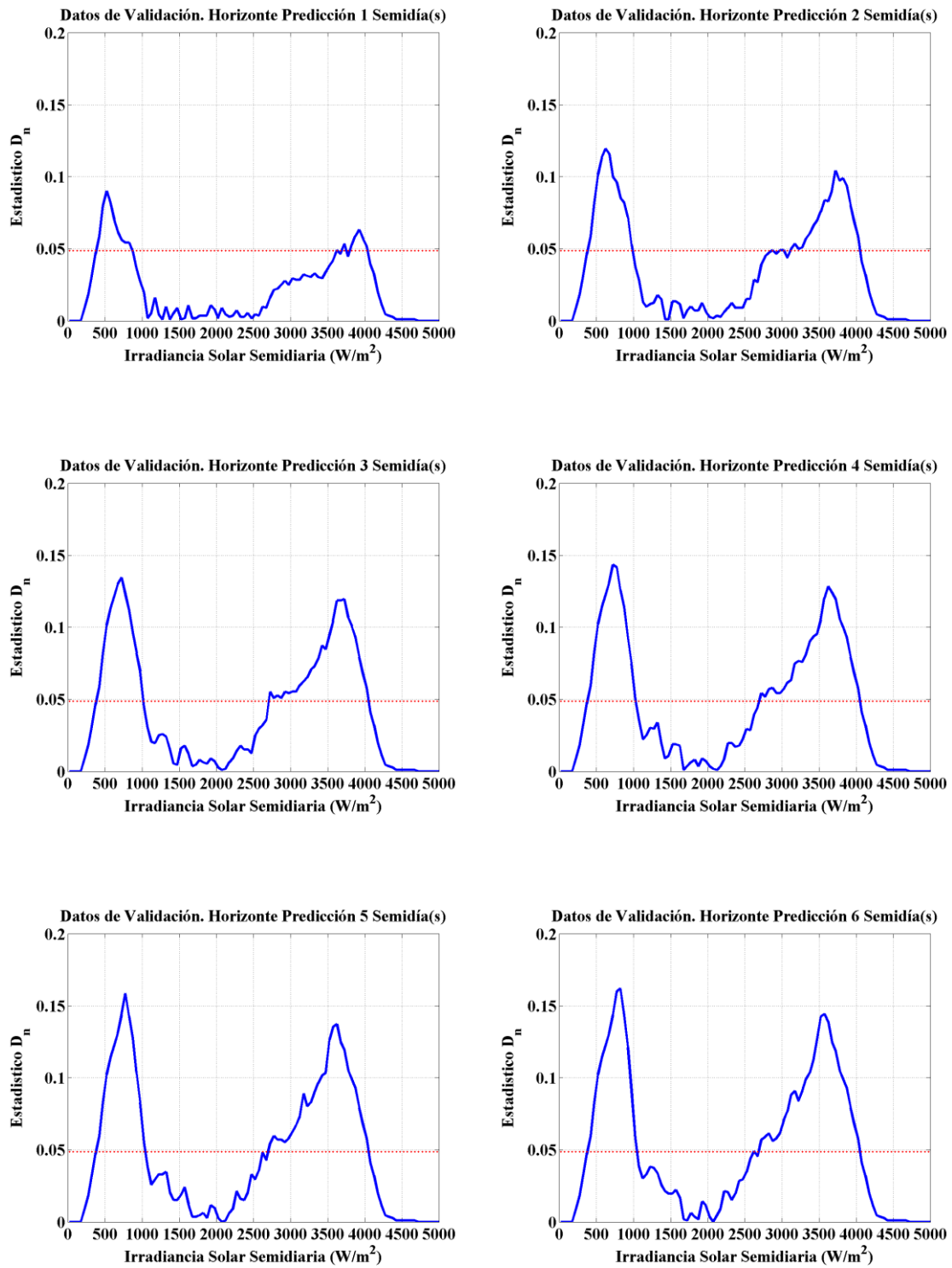
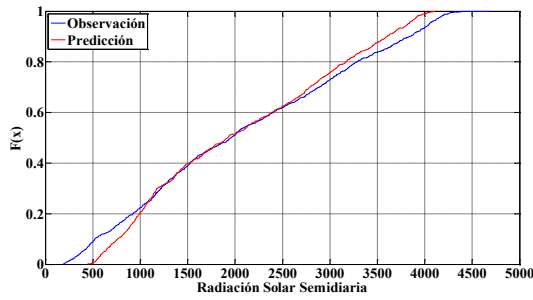
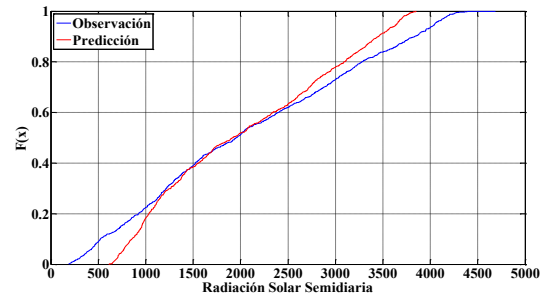


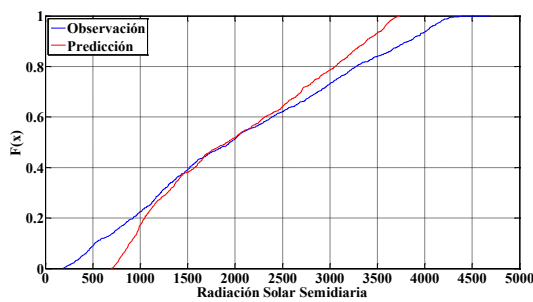
Figura D.7 Lérica: Test de de Kolmogorov-Smirnov irradiancia solar semidiaria



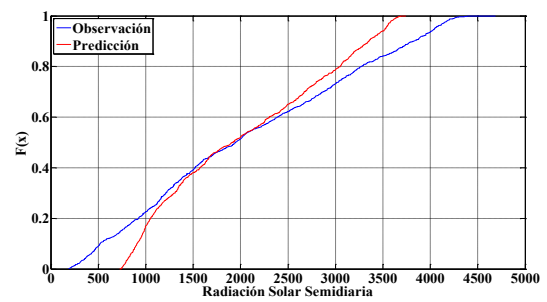
Horizonte de predicción 1 semidía



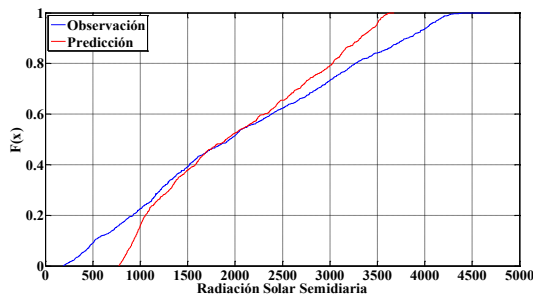
Horizonte de predicción 2 semidías



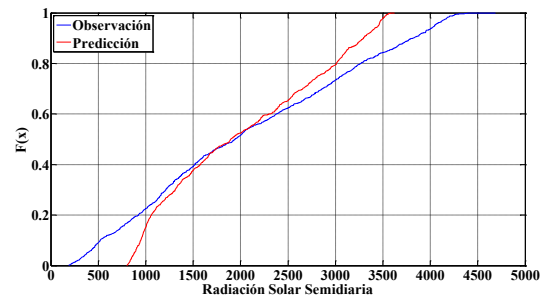
Horizonte de predicción 3 semidías



Horizonte de predicción 4 semidías



Horizonte de predicción 5 semidías



Horizonte de predicción 6 semidías

Figura D.8 Lérica: CDF irradiancia solar semidiaria