

MEMORIA PARA OPTAR AL GRADO DE DOCTOR EN CIENCIAS FÍSICAS  
QUE PRESENTA:

Llanos Mora López



\* 5 3 0 9 5 4 5 6 0 7 \*

UNIVERSIDAD COMPLUTENSE

**CARACTERIZACIÓN Y GENERACIÓN DE SERIES DE  
EXPOSICIÓN HORARIA DE RADIACIÓN GLOBAL**

TUTOR: Dr. JOSÉ DORIA RICO

DIRECTOR: Dr. MARIANO SIDRACH DE CARDONA

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE CIENCIAS FÍSICAS**  
**DEPARTAMENTO DE FÍSICA ATÓMICA, MOLECULAR Y NUCLEAR**

**A mis padres, Pedro y Llanos**

**A Mariano**

## **AGRADECIMIENTOS**

*Quiero agradecer a Mariano Sidrach de Cardona, director de este trabajo, sus valiosos comentarios y su constante apoyo; ha desarrollado un papel fundamental en la elaboración de esta tesis. Agradezco, también, a mi tutor, D. José Doria Rico, la ayuda y confianza que me ha prestado.*

*Quiero agradecer, de forma muy especial, la revisión del trabajo que ha hecho Elena Palomo. Sus numerosos comentarios han contribuido, de manera significativa, a la forma final que presenta el trabajo. Gracias a éstos, la redacción de muchos aspectos del estudio han ganado en claridad y precisión, de manera especial los capítulos 2 y 6.*

*Además quiero agradecer a Juan Mora, sus continuas "lecciones" en todos los conceptos estadísticos utilizados en el trabajo, así como su inmensa paciencia en la ardua tarea de corrección de las distintas versiones preliminares.*

*La mayor parte de los programas que se han elaborado para el desarrollo de este trabajo se han hecho utilizando la infraestructura de los Servicios Centrales de Informática de la Universidad de Málaga, a la que pertenezco desde el año 1992. Al personal de estos servicios quiero agradecer la ayuda que me han prestado en todo momento.*

*También quiero agradecer los valiosos comentarios de algunos de los asistentes a diversos seminarios. En concreto, resultados parciales de los capítulos 4, 5 y 6 se presentaron en el 12 European Photovoltaic Solar Energy Conference, celebrado en Amsterdam (1994) y en el VII Congreso Ibérico de Energía Solar, celebrado en Vigo (1994).*

*Cualquier error contenido en esta tesis es, por supuesto, responsabilidad exclusiva de su autora.*

# ÍNDICE

## 1. INTRODUCCIÓN

## 2. ANTECEDENTES

- 2.1 Caracterización de las tendencias
- 2.2 Propiedades estadísticas de las series residuales
- 2.3 Métodos de simulación utilizados
- 2.4 Justificación del trabajo

## 3. SERIES DE EXPOSICIÓN HORARIA DE RADIACIÓN GLOBAL

- 3.1 Introducción
- 3.2 Datos utilizados
- 3.3 Cálculo de la exposición horaria máxima
- 3.4 Análisis de las series de exposición horaria
  - 3.4.1 Descripción de las series del índice  $X_h$  de Badajoz
  - 3.4.2 Descripción de las series del índice  $X_h$  de Castellón
  - 3.4.3 Descripción de las series del índice  $X_h$  de Logroño
  - 3.4.4 Descripción de las series del índice  $X_h$  de Madrid
  - 3.4.5 Descripción de las series del índice  $X_h$  de Málaga
  - 3.4.6 Descripción de las series del índice  $X_h$  de Murcia
  - 3.4.7 Descripción de las series del índice  $X_h$  de Oviedo
  - 3.4.8 Descripción de las series del índice  $X_h$  de P. Mallorca
  - 3.4.9 Descripción de las series del índice  $X_h$  de Sevilla
  - 3.4.10 Descripción de las series del índice  $X_h$  de Tortosa
- 3.5 Conclusiones

#### **4. UTILIZACIÓN DE LA TEORÍA DE PROCESOS ESTOCÁSTICOS**

- 4.1 Descripción de las series utilizadas
- 4.2 Identificación del modelo
- 4.3 Estimación de los parámetros del modelo propuesto
- 4.4 Diagnóstico del modelo. Descripción de los contrastes
- 4.5 Resultados de la diagnosis
- 4.6 Análisis de los meses para los que no funciona el modelo propuesto
- 4.7 Interpretación de los resultados. Conclusiones

#### **5. MÉTODO DE GENERACIÓN DE SERIES DE EXPOSICIÓN HORARIA**

- 5.1 Generación de las series del índice  $D, X_h$
- 5.2 Integración
- 5.3 Cálculo de las series de exposición horaria de radiación global
- 5.4 Conclusiones

#### **6. PARÁMETROS QUE CARACTERIZAN LAS SERIES DE EXPOSICIÓN HORARIA DE RADIACIÓN GLOBAL**

- 6.1 Relación entre el índice de transparencia atmosférico diario y el índice propuesto,  $X_h$
- 6.2 Relación entre el índice  $X_{h,m}$  y el mes del año
- 6.3 Relación entre  $K_{d,m}$  y los parámetros del modelo ARMA
- 6.4 Tipos de meses en función del modelo ARMA asociado
- 6.5 Caracterización y generación de series anuales de exposición horaria de radiación global.
- 6.6 Conclusiones

#### **7. CONCLUSIONES**

## **ANEXO I**

Clasificación de días tipo en función de los valores de radiación global instantáneos recibidos. (Datos del Lier-CIEMAT)

## **ANEXO II**

Tablas con los parámetros estimados del modelo ARMA propuesto, para cada mes de las localidades de: Badajoz, Castellón, Logroño, Madrid, Málaga, Murcia, Oviedo, Palma de Mallorca, Sevilla y Tortosa.

## **ANEXO III**

- III.1 Contraste de igualdad de varianzas
- III.2 Contraste de igualdad de los parámetros estimados en un modelo ARMA multiplicativo  $(1,0) \times (0,1)$ ,
- III.3 Derivación de las covarianzas y correlaciones de un proceso ARMA $(1,0) \times$ ARMA $(0,1)$ , estacionario

## **NOMENCLATURA**

## **BIBLIOGRAFÍA**

# 1. INTRODUCCIÓN

El sol, como es bien sabido, es nuestra principal fuente de energía. Las dos características más importantes de esta fuente energética son: su baja densidad y su variabilidad con el tiempo. Estas dos características son de fundamental importancia cuando se intenta aprovechar la energía procedente del sol. Por esta razón, nos encontramos con numerosos trabajos cuyo objetivo es poder determinar, con más o menos precisión, qué cantidad de energía se recibe en un lugar concreto. Esta determinación es bastante difícil, por el hecho de que la radiación solar antes de llegar a la superficie de la tierra atraviesa la atmósfera.

La complejidad de los fenómenos que afectan a la radiación solar en su camino a través de la atmósfera es el principal problema que aparece a la hora de intentar cuantificar esta disponibilidad energética. Si bien podemos conocer con suficiente precisión la energía emitida por el sol en un momento determinado, no es tan sencillo llegar a estimar cuánta de esta energía alcanzará la superficie de la tierra. Desde que esta energía entra en la atmósfera dos tipos de factores influirán en su recorrido a través de la misma hasta que alcanza la superficie de

la tierra: unos que, por su naturaleza y efecto, se pueden cuantificar de forma más o menos precisa, y que, de forma intuitiva, podemos llamar deterministas; y otros, sin embargo, que por su naturaleza podemos considerar, en cierta manera, aleatorios.

Entre los factores "deterministas" se encuentran los factores astronómicos, que dependen de la geometría sol-tierra. Son función de la posición relativa de ambos, y del lugar de la tierra que consideremos. Estos factores condicionarán el recorrido de la radiación a través de la atmósfera y el ángulo de incidencia de la misma.

Los otros fenómenos que inciden en la cantidad de energía que se recibe en la superficie de la tierra son los factores climáticos. Estos serán los responsables de que se produzca una atenuación en la cantidad de energía que podría llegar a la superficie de la tierra. Estos factores, a diferencia de los factores astronómicos, no son tan fácilmente cuantificables. Los responsables de que se produzca esta atenuación son los componentes atmosféricos: vapor de agua, aerosoles, ozono, etc; y sobre todo las nubes.

El efecto de estos componentes en la radiación solar se manifiesta en tres fenómenos físicos: reflexión, absorción y difusión. De todos estos componentes el que mayor influencia tiene es la presencia de nubes. Esta presencia no es, sin embargo, constante, por lo que será muy importante determinar su influencia en la radiación solar.

El objetivo de la mayoría de los trabajos que existen sobre radiación solar es encontrar algún modelo, más o menos complejo, que explique razonablemente qué energía es la que se puede esperar recibir en un futuro.



Este problema se puede intentar resolver por dos caminos bien distintos, en función de la finalidad que se persiga:

- Utilizando modelos físicos, descriptivos, de la atmósfera y sus componentes, ver por ejemplo /32/, /36/ y /46/. Es decir, utilizando métodos que describan la influencia de cada componente atmosférico en la radiación solar, y caracterizar mediante las ecuaciones oportunas cada uno de ellos. El principal inconveniente que presentan es que, en general, se trata de modelos bastante complejos, que requieren el conocimiento de numerosos parámetros.
- Utilizando modelos matemáticos que describan el comportamiento de la radiación, pero sin dar cuenta de su naturaleza física, ver por ejemplo /2/, /3/, /4/.

Los primeros modelos se vienen utilizando cuando interesa conocer la composición espectral de la radiación y también cuando se quiere estudiar la variación de los distintos componentes atmosféricos.

La segunda vía es la que se utiliza, normalmente, cuando lo que se quiere es, únicamente, conocer el valor de radiación global o total que llega a la superficie de la tierra.

Nuestro trabajo se enmarca dentro de estos últimos, ya que nos acercaremos al problema desde el punto de vista de la utilización de modelos matemáticos con el objeto de caracterizar la radiación solar. Desde un punto de vista práctico, nuestro objetivo será encontrar un modelo estadístico con el que poder cuantificar la energía que se recibe del sol.

Normalmente no interesa conocer sólo valores puntuales de la radiación que alcanzará la superficie de la tierra, sino que es necesario conocer la evolución

de estos valores para periodos de tiempo suficientemente grandes. Para ello, será necesario utilizar series de valores de radiación integrados para periodos temporales que pueden variar desde un minuto, una hora, hasta un año.

En concreto, cuando se quiere abordar el estudio (dimensionado, diseño, evaluación, optimización) de sistemas de aprovechamiento de energía solar con memoria (capacidad de acumulación) de una a pocas horas nos encontramos, normalmente, con que los datos disponibles son insuficientes, ya que para este tipo de sistemas es necesario tener en cuenta las características estacionarias y secuenciales de la radiación solar.

Una forma de estudiar las propiedades secuenciales y estacionarias de la radiación solar es utilizar el análisis de series temporales. El análisis de series temporales es una disciplina matemática ampliamente utilizada en áreas como la economía, hidrología, meteorología, y su objetivo, básicamente es analizar los datos del pasado para determinar, a partir de esta información, las propiedades estadísticas que nos pueden servir para predecir el futuro.

Desde hace unos años esta metodología se ha empezado a utilizar para el estudio de series de radiación solar, ver por ejemplo /3/, /9/, /12/ y /16/. Una parte importante de este trabajo se centra en la aplicación del análisis de series temporales a datos de exposición horaria de radiación global para distintas localidades de España.

Partimos de la observación de que nuestra variable de estudio (exposición horaria) tiene cierta inercia en su evolución, y hacemos la hipótesis de que, aunque sus valores futuros son desconocidos, el estudio de su historia es informativo para prever su evolución futura. Por ello, el objetivo del trabajo es encontrar un modelo, utilizando la teoría de procesos estocásticos o series

---

temporales, que nos sirva para describir de forma adecuada las series de radiación. Además, utilizando los resultados obtenidos para cada localidad definiremos qué características son las mínimas necesarias para poder definir las series de exposición horaria de radiación de cualquier localidad.

A modo de introducción, en el capítulo segundo hacemos una revisión general de los trabajos que consideramos más significativos en los que se utiliza el análisis de series temporales para el estudio de series de radiación solar. En este capítulo señalamos los aspectos que quedan por resolver y justificamos, en función de éstos, los objetivos de este trabajo.

En el capítulo tercero analizamos las series de exposición horaria de radiación global. Para ello definimos un índice que es función de la radiación máxima esperable. Para las series de este índice estudiamos sus características más significativas, como son: valores medios, varianzas y funciones de distribución mensuales. Analizamos por separado las características de cada localidad. En este capítulo proponemos también una expresión general para los valores de radiación máxima esperable y definimos el índice con el que vamos a trabajar.

En el capítulo cuarto abordamos el estudio de las series del índice propuesto utilizando la teoría de procesos estocásticos o series temporales. Para las series de este índice proponemos un modelo ARMA con el que se pueden identificar y estimamos los parámetros del mismo. La aportación fundamental de este capítulo es que, en función del modelo propuesto, describimos dos características fundamentales de las series de exposición horaria: la relación entre el valor de una hora y el de la siguiente, y la relación entre la perturbación que se produce una hora y la que se produjo el día anterior para esa misma hora.

---

En el capítulo quinto desarrollamos un método que permite la generación de series sintéticas de valores de exposición horaria de radiación global. Es de destacar que en este capítulo desarrollamos un método para integrar series diferenciadas con el que se consigue que las series sintéticas tengan las mismas propiedades estadísticas que las series reales. Este método resuelve el problema clásico de la normalidad de las series generadas mediante modelos ARMA.

En el capítulo sexto estudiamos la relación que existe entre los parámetros que caracterizan las series de índices  $X_h$  (valor medio, coeficiente autorregresivo, coeficiente de medias móviles, varianza residual y mes o época del año) y los parámetros de las series de radiación normalmente disponibles para cada localidad (valores medios mensuales del índice de transparencia atmosférico diario). En este capítulo también analizamos cómo se pueden simplificar el número de modelos ARMA obtenidos, en función de la relación entre los parámetros estimados para los distintos meses. Como consecuencia de este estudio proponemos quince modelos ARMA con los que recogemos toda la información contenida en más del 90% de las series originales. Además, en función de las relaciones obtenidas entre los parámetros de cada modelo ARMA y las demás variables proponemos dos métodos con los que generar series de varios años; la utilización de uno u otro dependerá del tipo de clima (definido por otras variables) de la localidad para la que se vaya a utilizar. La clasificación que obtenemos de las distintas localidades consideradas coincide con la obtenida por otros autores, ver por ejemplo, M.Pardo /49/.

En el capítulo séptimo se resumen los resultados más importantes obtenidos y se resaltan las aportaciones originales más significativas de este trabajo.

Hemos incluido además tres anexos que consideramos de interés.

En el primero se muestra una clasificación de días en función de la forma que tiene la curva que representa los valores de radiación global registrados durante el día.

En el segundo se listan los estimadores de los parámetros del modelo propuesto para los distintos meses y localidades.

En el tercer anexo se describen los contrastes más significativos que hemos utilizado y se deriva una relación fundamental entre las covarianzas y correlaciones del tipo de modelo ARMA propuesto.

## **2. ANTECEDENTES**

En los artículos publicados hasta la fecha relacionados con el análisis de las series de radiación solar se estudian, entre otras, series de radiación global para intervalos horarios y diarios.

En ambos tipos de series encontramos la influencia de los dos factores descritos en el capítulo anterior: deterministas y aleatorios.

En las series de exposición diaria de radiación global se observa una tendencia que es debida a la influencia de la variación de la posición relativa sol-tierra para los distintos días del año; esta tendencia es de tipo estacional.

En las series de exposición horaria de radiación global aparece, además de esta tendencia estacional, una tendencia diaria debida a la sucesión día-noche. Así, en un mismo día los valores presentan una tendencia debida a la variación en la masa de aire (inversa del camino óptico) que atraviesan los rayos del sol en su paso a través de la atmósfera. En la figura 2.1 se muestra la tendencia diaria observada en estas series. En esta figura hemos representado los valores de exposición horaria de radiación global para tres días de verano y tres días de

invierno. En la figura 2.2 se puede observar la tendencia estacional que presentan las series de exposición horaria de radiación global. En esta figura se muestran los valores de exposición horaria registrados a las 12 horas durante el año 1977, en la localidad de Málaga.

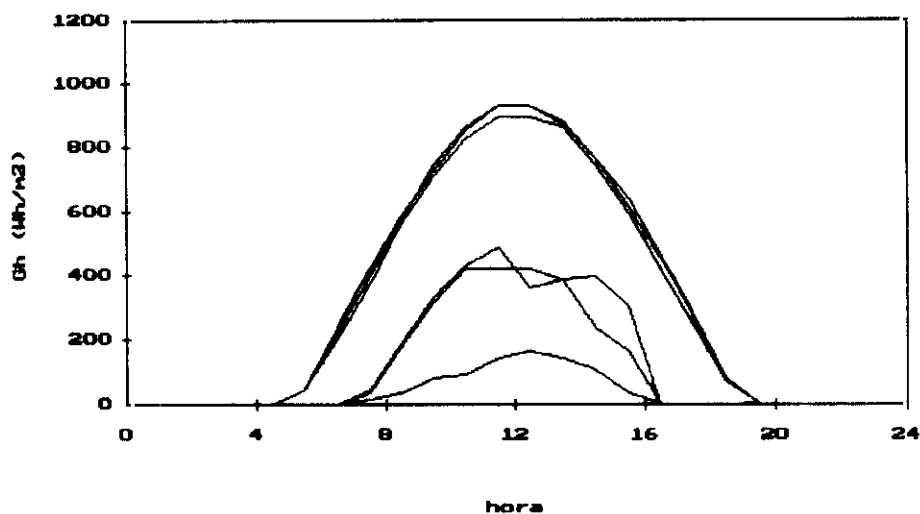


Figura 2.1. Exposición horaria de radiación global. Málaga, Enero-1977 (curvas inferiores) y Julio-1977 (curvas superiores).

Cuando se quiere utilizar el análisis de series temporales estacionarias, es necesario que las series que se van a estudiar sean estacionarias. Por esto y de acuerdo con las anteriores observaciones, el estudio de las series de exposición horaria, o diaria, de radiación global (con el objeto de caracterizarlas y reproducirlas, con la teoría de procesos estocásticos o series estacionarias) se puede realizar según el siguiente esquema:

- Caracterizar las tendencias observadas en las series y eliminar éstas de las series originales.
- Analizar las propiedades estadísticas de las series resultantes (series residuales): valores medios, varianzas, funciones de distribución de

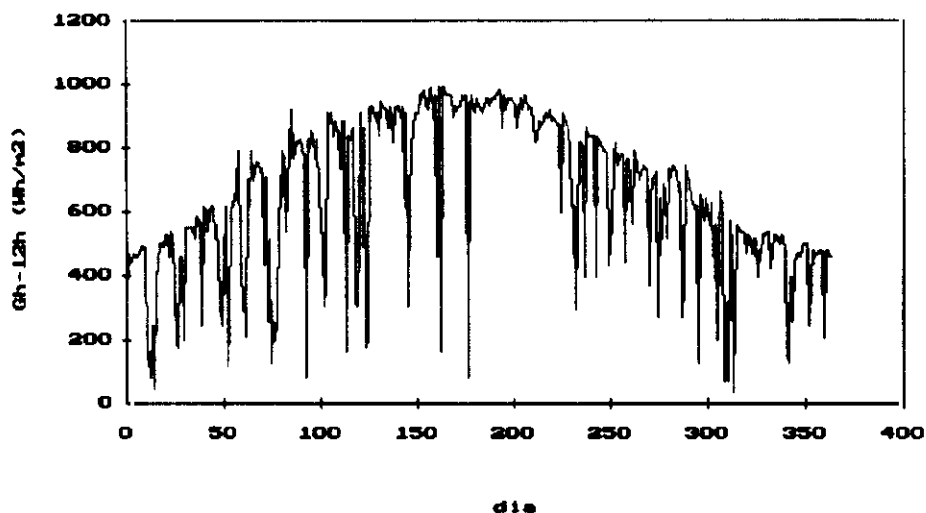


Figura 2.2. Exposición horaria de radiación global a las 12 h, Málaga (1977)

probabilidades de ocurrencia, propiedades secuenciales, etc.

- En función de las propiedades observadas proponer modelos con los que se puedan simular o reproducir las series originales.

Analizaremos en los siguientes apartados las técnicas y modelos propuestos que más se han utilizado, para cada uno de los pasos descritos en este esquema.

## 2.1 CARACTERIZACIÓN DE LAS TENDENCIAS

Vamos a hacer primero un estudio de las técnicas que se han utilizado para eliminar las tendencias estacional y diaria de las series de exposición horaria de radiación global.

Si analizamos los trabajos publicados vemos que los métodos que se han utilizado, más frecuentemente, para describir y poder eliminar, por tanto, las tendencias estacional y diaria de las series de exposición global horaria son:



- (a) Cálculo de la media (o media móvil) de la serie, utilizando funciones de Fourier.
- (b) Utilización de la radiación extraterrestre horaria.
- (c) Utilización de modelos físicos que permitan estimar los valores de radiación.
- (d) Utilización de la envolvente o valores máximos de exposición horaria de radiación global.

El primero de estos métodos lo utilizan, por ejemplo, Balouktsis et al., /10/. Estos autores calculan los coeficientes de Fourier para cada hora a lo largo del año (año medio horario) mediante una función del tipo:

$$G_{h,d}^F = A_{o,h} + A_{1,h} \cos \frac{2\pi(d-1)}{L} + B_{1,h} \operatorname{sen} \frac{2\pi(d-1)}{L} \quad (2.1)$$

donde:  $G_{h,d}^F$  es el valor de la serie de Fourier para la hora  $h$  del día  $d$ ;  $B_1$ ,  $A_1$  son los coeficientes de Fourier; y  $L$  el período de la hora considerada.

Utilizando estos valores de  $G_{h,d}^F$  calculan un índice  $F$  que viene dado por la expresión:

$$F_{h,d} = \frac{G_{h,d}}{G_{h,d}^F} \quad (2.2)$$

Esta nueva serie no es, sin embargo, normal y es necesario emplear técnicas gaussianas para conseguir la normalidad.

Caridad et al., /18/, determinan una tendencia media en los valores máximos de exposición global horaria, mediante la técnica de media móvil. Así consiguen obtener unos valores máximos de radiación no distorsionados por el efecto de coincidencias de días con el mismo clima durante los años considerados.

El segundo método de eliminación de tendencias es utilizar la radiación extraterrestre horaria. Las series que se obtienen al utilizar la radiación horaria extraterrestre son las del índice de transparencia horario,  $k_h$ , que vendrá dado por la expresión:

$$K_h = \frac{G_h}{G_{0,h}} \quad (2.3)$$

donde  $G_h$  es la exposición horaria de radiación global y  $G_{h,0}$  es la radiación extraterrestre horaria.

De los trabajos que utilizan este método hay que destacar el de Amato et al., /6/, el de Olseth et al., /42/, y el de Aguiar et al., /3/.

Karaeb et al., /34/, también trabajan con el índice de transparencia horario. En su estudio afirman que éste se puede interpretar como el producto de dos: un índice de transparencia medio horario y un coeficiente diario de transparencia máximo, que se calcula suponiendo un día claro (sin nubes):

$$K_h = T T_c \quad (2.4)$$

donde:  $K_h$  es el índice de transparencia horario,  $T$  es el coeficiente de transparencia y  $T_c$  el coeficiente diario medio de nubosidad.

Respecto al tercer método empleado, Engels et al., /20/, afirman que las series de radiación horaria se pueden caracterizar, de manera sencilla, utilizando modelos de cielo claro. El método que proponen consiste en construir una serie de valores horarios a partir de un modelo físico de cielo claro y modular esta serie con una atenuación aleatoria generada con modelos autorregresivos. El principal problema que presenta este método es el de conocer los parámetros físicos que definen el modelo de cielo claro.

---

El último de los métodos citados se basa en la utilización de los valores máximos esperables de exposición horaria o "envolventes". Esta técnica de eliminación de tendencias ha sido utilizada, por ejemplo, por D.Guinea et al., /31/, Ll.Mora, /39/, y E. Peleona, /46/.

Para las series de exposición diaria de radiación global se han utilizado métodos similares. Entre los trabajos que utilizan la radiación extraterrestre diaria podemos citar, por ejemplo, los de B.Bartoli, /11/, E.Boileau, /15/, Amato et al., /4/, y Aguiar et al., /1/. Por otra parte, nos encontramos también con trabajos en los que se utilizan los valores de exposición diaria máxima o "radiación recibida en días claros"; como ejemplo, se puede ver el estudio de Olseth, /42/.

En general, una vez eliminada la tendencia observada en las series originales el siguiente paso es analizar, desde el punto de vista estadístico, las series resultantes. Normalmente este análisis incluye también un estudio de las propiedades secuenciales de las series utilizando la teoría de procesos estocásticos estacionarios.

Se puede observar, en la mayoría de los trabajos citados, que las series residuales obtenidas no son, sin embargo, estacionarias. Será, por tanto, necesario emplear algún tipo de transformación para conseguir series estacionarias. Las transformaciones que más se han utilizado son:

- Restar a las series el valor medio mensual del parámetro analizado (índice de transparencia atmosférico horario o diario, etc.) Esta técnica la utiliza, por ejemplo, B. Bartoli, /11/, o E.Boileau, /15/. Definen un nuevo índice,  $R_d$ , que resulta de restar a los valores del índice de transparencia diarios,  $K_d$ , que se dan en un mes el valor medio de este índice a lo largo del mes,  $K_{d,m}$ . El objetivo de utilizar este nuevo índice es la eliminación de la

tendencia anual que presentaban las series del índice de transparencia diario. Este nuevo índice se puede expresar como:

$$R_d = K_d - K_{d,m} \quad (2.5)$$

- Utilizar la función error de las series. Esta técnica se ha utilizado para las series de índice de transparencia. La función error viene definida por la expresión:

$$erf(u) = F(K_d) \quad (2.6)$$

donde  $erf(u)$  es la función de distribución de una normal tipificada, es decir:

$$erf(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u \exp\left[-\frac{x^2}{2}\right] dx \quad (2.7)$$

- Utilizar variables normalizadas. Esta técnica la utilizan, por ejemplo, Amato et al., /5/, para series diarias, y Boch et al., /13/, para series horarias.

Amato et al. caracterizan la componente estacional de las series de exposición global diaria utilizando series de Fourier. A partir de las series así obtenidas, proponen obtener nuevas series definidas mediante la expresión:

$$X_{r,d} = \frac{G_d - G_{d,m}}{\sigma_{d,m}} \quad (2.8)$$

donde:  $X_{r,d}$  es la serie residual resultante,  $G_d$  es la serie de datos de exposición global diaria,  $G_{d,m}$  es la serie de valores medios de exposición global diaria y  $\sigma_{d,m}$  son las varianzas medias diarias. La nueva serie,  $X_{r,d}$ ,

tiene que ser todavía transformada utilizando su función error para que sea estacionaria.

Boch et al., /12/, emplean la técnica de modificar las series de índices de transparencia horarios utilizando para ello los valores medios horarios de estas series y su varianza. El índice que calculan viene dado por la expresión:

$$B_{h,d} = \frac{K_{h,d} - K_{h,m}}{\sigma_{h,m}} \quad (2.9)$$

La nueva serie de índices  $B_{h,d}$  sigue modelos autorregresivos.

- Diferenciar las series originales: esta técnica es utilizada, por ejemplo, por E.Boileau, /15/, para series de valores diarios. El operador diferencias utilizado es de orden uno. Este procedimiento no ha sido utilizado para series de valores horarios.

En el siguiente apartado describiremos las propiedades más significativas de las series resultantes.

## 2.2 PROPIEDADES ESTADÍSTICAS DE LAS SERIES RESIDUALES

Las propiedades que se utilizan para caracterizar y describir las series residuales resultantes son: valor medio, varianza, función de distribución de probabilidades de ocurrencia, estacionariedad, función de autocorrelación y función de autocorrelación parcial.

Normalmente este estudio se hace no sólo para conocer las características de estas series, sino también con el objetivo de determinar, en función de éstas,

---

modelos con los que se puedan identificar las mismas.

Puesto que se trata de series temporales (sucesiones de valores horarios o diarios) su estudio se hace, normalmente, utilizando la teoría de procesos estocásticos estacionarios. Por ello, y para que ésta sea aplicable, como ya apuntamos, las series utilizadas han de ser estacionarias, tanto en media como en varianza.

Vamos ahora a describir las propiedades de las series residuales descritas en el apartado anterior.

Las series obtenidas en los trabajos de B. Bartoli, /11/, o E. Boileau, /15/, (resultantes de restar a los valores del índice de transparencia diario el valor medio mensual de éstos) presentan funciones de distribución de probabilidades que no son gaussianas. Las series son estacionarias, según los autores, en media y varianza. Las propiedades secuenciales las estudian mediante la función de autocorrelación; encuentran que sólo el primer coeficiente de autocorrelación es significativamente distinto de cero. Esto quiere decir que las series residuales se pueden identificar con modelos autorregresivos de primer orden. Sin embargo, el hecho ya apuntado de que las funciones de distribución de estas series no son normales (condición necesaria para poder utilizar modelos de este tipo), hace necesario modificar estas series. Esta nueva modificación hace demasiado complejo el modelo propuesto.

Las series de índices de transparencia diaria son estacionarias en media y varianza, si se consideran períodos mensuales. Las funciones de distribución de estas series son, normalmente, bimodales. Esto hace que, si no se utiliza alguna otra de las modificaciones o transformaciones anteriormente descritas, no puedan ser identificadas con modelos autorregresivos. Sí pueden ser consideradas, sin embargo, como procesos de Markov de primer orden, ya que las características

---

que presentan sus funciones de autocorrelación simple y parcial presentan las restricciones que caracterizan estos procesos. Además estos modelos no imponen ninguna restricción a la forma de las funciones de distribución. Este modelo ha sido propuesto por diversos autores. Uno de los trabajos más aceptados y utilizados es el desarrollado por Aguiar et al., /1/. Sin embargo, según otros autores, como Biga et al., /13/, las series de índices de transparencia diarios no pueden describirse mediante modelos autorregresivos de primer orden. Esta hipótesis la encontramos también en los trabajos de Gordon et al., /28/, que afirman que los modelos autorregresivos son más flexibles que el método de las matrices de Markov, ya que con aquellos se pueden tener en cuenta la influencia de los valores de radiación global diaria para desfases mayores que uno.

Las series de exposición horaria obtenidas utilizando el método propuesto por Balouktsis et al., /10/, (valores de exposición horaria de radiación global entre valores de la función de Fourier propuesta para cada hora) presentan, también, funciones de distribución que no son normales. Los autores proponen la utilización de técnicas gaussianas para conseguir la normalidad.

Las series horarias obtenidas al utilizar la radiación máxima esperable o "envolvente", ver por ejemplo D. Guinea et al., /31/, Ll. Mora, /39/, E. Peleona, /46/, pueden considerarse estacionarias para períodos mensuales, pero presentan funciones de distribución que no son normales. En estos trabajos se afirma que se pueden identificar con modelos o procesos de Markov de primer orden. En un capítulo posterior de este trabajo analizaremos si estas series se pueden identificar con procesos de Markov de primer orden.

Otro enfoque que se puede hacer al estudiar las propiedades de las series de exposición horaria de radiación global es el que hacen Aguiar et al., /3/. Analizan primero las propiedades de las series de índices de transparencia diarios. Encuentran que, para períodos de un mes, estas series son estacionarias en media

---

y varianza, tienen funciones de distribución bimodales y, sus propiedades secuenciales, pueden ser identificadas como procesos de Markov de primer orden. Para cada día proponen que la estructura subyacente de la serie de valores horarios depende únicamente del valor del índice de transparencia diario. Ésta la caracterizan a partir del valor del coeficiente de autocorrelación. Proponen, para cada rango o intervalo de valores del índice de transparencia diario un valor único del coeficiente de autocorrelación, calculado como el valor medio de todos los estimados para ese rango.

Este trabajo está siendo muy utilizado por su sencillez; encontramos, sin embargo, varias cosas que quedan sin resolver:

- Se parte de un hecho que no se contrasta: los días con similar índice de transparencia tienen la misma distribución de valores del índice de transparencia horario.
- Se calculan estimadores de la función autocorrelación y autocorrelación parcial para series que tienen sólo de 8 a 16 valores. El problema de este planteamiento es que, desde el punto de vista estadístico, cualquier contraste que se haga no tiene potencia; esto es, aunque la hipótesis de que se parte fuera falsa, por los pocos datos utilizados la aceptaríamos como verdadera.

## 2.3 MÉTODOS DE SIMULACIÓN UTILIZADOS

Los dos tipos de modelos más utilizados para la generación de series de exposición horaria y diaria de radiación global vienen determinados por las características de las series residuales obtenidas, descritas en el apartado anterior.

Para todas las series que presentaban funciones de distribución normales



los modelos utilizados son, en general, modelos autorregresivos de primer orden, que responden a una función del tipo:

$$R_d = \rho_1 R_{d-1} + \varepsilon_d \quad (2.10)$$

donde  $\rho_1$  es el coeficiente de autocorrelación para desfase uno y  $\varepsilon_d$  es un ruido blanco.

Este tipo de modelos ha sido utilizado, por ejemplo, por B. Bartoli, /11/, E. Boileau, /15/, Amato et al., /4/; en todos los casos, tras modificar las series de índices de transparencia diarios por alguna de las técnicas descritas en el apartado anterior.

Otros modelos que también se han utilizado son los procesos de Markov de primer orden. La ventaja que presentan respecto a los modelos autorregresivos es que no imponen ninguna restricción a la forma de las funciones de distribución de probabilidades de ocurrencia. Este tipo de modelos ha sido utilizado, por ejemplo, por Aguiar et al., /1/, para series del índice de transparencia atmosférico diario, y por Palomo, /46/, y Mora, /39/, para series del índice que se obtiene al dividir las series de exposición horaria por los valores máximos esperables para cada hora.

El método se basa en la construcción de matrices de transición. Estas matrices parecen ser universales cuando se trabaja con valores diarios. Por el contrario, no está probada su universalidad cuando se utilizan valores horarios, ver Palomo, /46/.

Por último, hay que citar el trabajo de Aguiar et al., /3/, en el que se propone la generación de los valores horarios de un día utilizando un modelo autorregresivo de orden uno. El coeficiente de autocorrelación se calcula en función del índice de transparencia atmosférico de ese día. En un apartado

---

anterior ya comentamos las deficiencias que encontramos en el desarrollo de este método. Además, para poder determinar la serie de valores horarios es necesario conocer el valor del índice de transparencia diario, dato que, normalmente no está disponible, por lo que se hace necesaria la utilización de un modelo.

## 2.4 JUSTIFICACIÓN DEL TRABAJO

De acuerdo con lo expuesto anteriormente, el estado actual de los estudios realizados sobre caracterización y generación de series de exposición horaria de radiación global, se pueden observar las siguientes limitaciones o cuestiones que quedan por resolver:

- Si se utilizan las series del índice de transparencia horario, como series modificadas, podemos observar que estas series presentan una tendencia diaria. Las técnicas utilizadas para eliminar ésta, o bien hacen necesaria la utilización de parámetros normalmente no disponibles o bien complican demasiado los métodos de análisis y generación de nuevas series. No encontramos, entre los trabajos que siguen este esquema, ninguno que, partiendo sólo de parámetros disponibles, proponga un modelo universal para caracterizar y generar nuevas series.
- Si se analizan las series de índices que se obtienen al dividir los valores de exposición horaria de radiación global por los valores máximos esperables o envolventes, se observa que también estas series presentan una tendencia diaria. A pesar de ello, en los trabajos en los que se utiliza este índice se propone la utilización de procesos de Markov de primer orden. Respecto a las matrices calculadas no se puede afirmar que sean universales, ver, por ejemplo, Palomo, /46/.
- El tercer método es el propuesto por Aguiar et al., /3/,. Si bien es un

---

método que parece ser universal, carece de una adecuada contrastación, desde el punto de vista estadístico (al menos, en los puntos que comentamos anteriormente). Además para llegar a obtener las series de exposición horaria se utilizan varios modelos o aproximaciones: para generar series de valores diarios del índice de transparencia atmosférico (modelos de Markov de primer orden), para calcular el coeficiente autorregresivo del modelo que genera los valores horarios de índices de transparencias de un día (primero se calcula un valor medio de este coeficiente, para cada intervalo de valores de índice de transparencia diario, y después se ajustan estos valores a una expresión). Esto hace que en cada paso se vaya perdiendo precisión/información. Tampoco se contrasta, desde el punto de vista estadístico, la bondad de los distintos ajustes o modelos (la pérdida de información que supone la utilización de cada uno de ellos).

Teniendo en cuenta todas estas consideraciones, el objetivo que pretendemos en este trabajo es, primero, caracterizar las series de exposición horaria de radiación global utilizando sólo parámetros normalmente disponibles. Una vez hecho esto analizaremos la posibilidad de generar nuevas series de estos valores utilizando el tipo de modelo encontrado y la información normalmente disponible.

Para ello, vamos a estudiar los resultados que se obtienen cuando se utiliza como técnica de eliminación de tendencias, primero, dividir las series originales por los valores de exposición máxima esperables, y después diferenciar estas series.

Los resultados obtenidos los evaluaremos de acuerdo con estos tres criterios:

- Rigor estadístico del método utilizado. Calidad del modelo predictivo resultante.
- Complejidad: número de parámetros necesarios para su utilización.
- Universalidad del modelo propuesto.

En el capítulo de conclusiones se hace una síntesis de los resultados obtenidos utilizando estos criterios de evaluación.

## **3. SERIES DE EXPOSICIÓN HORARIA DE RADIACIÓN GLOBAL**

### **3.1 INTRODUCCIÓN**

Como se ha señalado en el capítulo anterior, todos los trabajos que intentan describir las series de exposición horaria de radiación global utilizando un análisis estocástico proponen algún método, más o menos complejo, de modificación de estas series. Esto es necesario porque estas series presentan dos tipos bien diferenciados de tendencia:

- una diaria, debida a la sucesión día-noche.
- una estacional, debida a la variación en la posición relativa sol-tierra.

Así pues, para poder utilizar el análisis de series temporales será necesario eliminar, de alguna manera, estas dos tendencias. El método que proponemos en

---

este trabajo es dividir las series de exposición horaria por los valores de exposición horaria máximos esperables para cada hora. Justificamos esta elección en el siguiente hecho: cualquier disminución en los valores de exposición global horaria respecto al valor máximo es debida, fundamentalmente, o bien a variaciones en la presencia de los distintos componentes atmosféricos (vapor de agua, aerosoles, ozono, etc.) o bien a la presencia de nubes.

Algunos autores estiman que las variaciones en el contenido de vapor de agua y aerosoles en la atmósfera sólo suponen, como máximo, una variación de un tres a un cinco por ciento en la cantidad de radiación que alcanza la superficie de la tierra, ver Paltridge et al., /47/. El efecto de las nubes es, por tanto, mucho más importante, y queda recogido en las series de índices que resultan al utilizar el método que proponemos.

La hipótesis que proponemos es la de que el efecto de las nubes en los valores de radiación global puede describirse de forma adecuada utilizando modelos autorregresivos. Es decir, que si utilizamos el índice que resulta al dividir los valores de exposición horaria por los valores máximos de radiación esperables para cada hora, las nuevas series resultantes son, tras una sencilla transformación, estacionarias y pueden analizarse utilizando la teoría de procesos estocásticos estacionarios.

Para poder utilizar la teoría de series temporales será necesario primero definir cuál es esa exposición horaria máxima esperable con la que calculamos el índice que vamos a utilizar. En principio nos parece razonable suponer que ésta será sólo función de factores de tipo determinista, como posición relativa sol-tierra, camino óptico y localidad considerada, y será la que se registrará en días en los que la atmósfera esté limpia o clara, es decir, sobre todo en días sin nubes. La presencia de nubes supondrá, pues, disminuciones respecto a ese nivel máximo, por el efecto de los fenómenos de difusión, reflexión y absorción que

las nubes ocasionarán en la radiación. Si esto es así, podemos afirmar que los valores máximos esperables de exposición horaria ( $G_{h,max}$ ) serán, sólo, función de la altura solar (ángulo que recoge la posición relativa sol-tierra) y de la latitud para un momento y localidad determinados. Es decir:

$$G_{h,max} = f(\gamma, \phi) \quad (3.1)$$

donde  $\gamma$  representa la altura solar y  $\phi$  representa la latitud.

A partir de la ecuación (3.1) podemos definir la serie de índices  $X_h$  según la expresión:

$$X_h = \frac{G_h}{G_{h,max}} \quad (3.2)$$

donde  $G_h$  es la exposición horaria de radiación global.

En este trabajo utilizaremos modelos estadísticos de series temporales para analizar el comportamiento de esta serie. En primer lugar estudiaremos qué transformación debe realizarse para conseguir estacionariedad. Posteriormente propondremos un modelo ARMA (autorregresivo y/o de medias móviles) para la serie estacionaria resultante utilizando estimadores de las funciones de autocorrelación simple y parcial. Después estimaremos los parámetros del modelo propuesto y contrastaremos su validez. Por último, con este modelo propuesto ya verificado analizaremos la generación de nuevas series de exposición global horaria.

### 3.2 DATOS UTILIZADOS

Los datos que vamos a utilizar en este trabajo son valores de exposición global horaria de la radiación solar registrados a lo largo de varios años. Estos datos pertenecen a la Base de Datos del Instituto de Energías Renovables del CIEMAT. En la Tabla 3.I se muestra el número total de años disponibles, el período a que corresponden y la latitud de cada localidad. Puesto que no todos los años están completos para todas las localidades en la última columna de esta tabla se recoge el número de meses disponibles para cada localidad.

Con el fin de poder apreciar las diferencias climatológicas, en lo que a niveles de radiación se refiere, entre las localidades que vamos estudiar, en la Tabla 3.II se muestran los valores medios mensuales de exposición global diaria, en kWh/m<sup>2</sup>.

Localidad	Período	Nºtot.años	Latitud	Tot.meses
Badajoz-BA	1976-1983	8	38.89	84
Castellón-CS	1979-1984	6	39.95	55
Logroño-LO	1981-1984	4	42.46	48
Madrid-M	1979-1986	8	40.45	96
Málaga-MA	1977-1984	8	36.66	94
Murcia-MU	1977-1984	8	38.00	96
Oviedo-OV	1977-1984	8	43.35	96
P.Mallorca-PM	1977-1984	8	39.33	94
Sevilla-SE	1977-1984	8	37.42	78
Tortosa-TR	1980-1984	5	40.81	52

Tabla 3.I



Loc\mes	01	02	03	04	05	06	07	08	09	10	11	12
BA	2.3	2.9	4.4	5.4	6.5	7.1	7.4	6.5	5.1	3.5	2.6	1.8
CS	2.1	2.8	4.3	5.1	6.2	6.7	6.7	5.7	4.7	3.6	2.4	2.1
LO	1.7	2.5	4.1	4.9	5.7	6.8	7.0	6.1	4.9	3.0	2.0	1.4
M	2.0	2.9	4.3	5.4	6.5	7.3	7.6	6.7	5.3	3.6	2.4	1.8
MA	2.4	3.3	4.7	5.3	6.7	7.2	7.2	6.3	5.1	3.7	2.8	2.1
MU	2.3	3.2	4.7	5.4	6.3	7.2	7.4	6.1	5.0	3.7	2.7	2.1
OV	1.4	2.0	3.1	4.0	4.1	4.7	4.6	4.1	3.6	2.4	1.6	1.2
PM	2.0	2.6	4.2	4.9	6.0	6.7	6.7	5.9	4.6	3.3	2.4	1.8
SE	2.5	3.1	4.7	5.4	6.9	7.6	7.5	6.6	5.3	3.9	2.9	2.1
TR	2.0	2.6	4.0	4.8	5.8	6.7	6.7	5.5	4.5	3.5	2.2	1.9

Tabla 3.II

Para cada día los valores registrados de exposición global horaria comienzan a la hora de salida del sol y terminan a la hora de puesta del sol. Como hemos comentado, la atmósfera atenúa la radiación solar extraterrestre. Para períodos cercanos a la salida y puesta del sol la atenuación es mayor que para las horas cercanas al mediodía, debido al efecto de la mayor masa de aire que atraviesan los rayos del sol. Este efecto se refleja en el hecho de que, si bien la radiación extraterrestre sigue una distribución tipo coseno, cuando atraviesa la atmósfera esta forma de curva se distorsiona en las colas por la variación de la masa de aire. Es decir, la transmitancia de la radiación solar es mucho mayor durante las horas centrales del día que en las horas cercanas a la salida y la puesta del sol.

Por este motivo, en nuestro trabajo no vamos a utilizar las primeras y últimas horas del día. De esta forma se eliminan los problemas que aparecen en las medidas que se registran en las mismas por efecto de la masa de aire que atraviesa la radiación y también por la fiabilidad de los equipos de medida en

estas horas. Además, la mayoría de los sistemas solares presentan un aprovechamiento muy pequeño o nulo de la energía recibida a esas horas (umbral de funcionamiento, pérdidas, etc;).

En la Tabla 3.III se recoge, para cada mes, el número de horas que se van a considerar, centradas respecto al mediodía solar. En esta tabla también aparece el porcentaje de energía que se recibe durante las horas no consideradas respecto al total diario recibido; para su cálculo se han utilizado todos los datos disponibles para cada localidad.

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
h	8	8	10	10	12	12	12	12	10	10	8	8
BA	1.0	2.6	1.8	4.9	1.8	2.9	2.8	1.3	2.5	0.5	1.6	0.6
CS	1.1	2.8	1.6	5.1	2.1	2.9	2.0	0.8	1.7	0.2	1.0	0.6
LO	1.6	4.2	2.7	4.2	3.8	4.8	4.9	2.5	4.9	0.9	2.1	0.7
M	1.0	1.7	3.5	2.4	3.0	2.5	3.2	1.3	3.0	0.4	1.3	0.5
MA	1.3	3.4	1.8	4.2	1.6	2.5	2.1	0.8	2.1	0.4	1.7	0.8
MU	1.5	3.6	2.0	3.9	1.8	3.0	2.5	1.1	2.7	0.7	1.9	1.1
OV	1.7	2.6	1.5	4.5	2.5	3.8	3.1	1.2	2.9	0.5	1.2	0.5
PM	1.1	2.9	1.8	4.9	2.1	3.3	2.8	1.2	0.7	0.6	1.6	0.6
SE	1.4	3.3	1.6	4.4	1.6	2.3	2.0	0.9	2.3	0.6	1.8	1.0
TR	1.4	4.0	2.2	3.8	2.5	4.4	3.8	1.3	3.3	0.5	1.4	0.7

TABLA 3.III

Como se puede observar, para todas las localidades el porcentaje medio de energía que consideramos en las series que vamos a utilizar es, aproximadamente, del 98% respecto al total. Si tenemos en cuenta además que para el diseño de la mayoría de las aplicaciones de energía solar se utilizan como parámetros de entrada datos de radiación sobre superficie inclinada, calculados a partir de los de

superficie horizontal, el porcentaje de energía que no se considerará será todavía menor y, prácticamente, no influirá en el comportamiento final del sistema.

### 3.3 CÁLCULO DE LA EXPOSICIÓN HORARIA MÁXIMA

La energía que llega a la superficie de la tierra procedente del sol dependerá, como ya se ha comentado en apartados anteriores, de factores que podemos considerar de tipo determinista y de otros factores más difícilmente predecibles. Si la composición de la atmósfera fuera constante para todos los días podríamos saber con bastante aproximación qué energía se recibiría en cualquier lugar de la tierra, ya que ésta sólo dependería de la energía que se recibe en la parte exterior de la atmósfera y del camino que recorre a través de la atmósfera.

De hecho, si medimos la energía que llega en "días claros o limpios", con ausencia total de nubes, contaminación, etc., podemos comprobar que la radiación que alcanza la superficie de la tierra es sólo función de los dos factores anteriormente descritos: radiación solar extraterrestre y altura solar. Partiendo de estas observaciones hay numerosos trabajos que intentan cuantificar esos niveles de radiación para días claros, ver por ejemplo, /20/, /25/ y /44/. En este trabajo vamos a definir la exposición horaria máxima esperable como la energía que se recibe durante una hora con el cielo limpio.

Así, si expresamos la radiación que alcanza la superficie de la tierra en un instante determinado como una función de la radiación extraterrestre y de unos factores que recojan la influencia de los componentes atmosféricos y el efecto de las nubes, mediante una expresión del tipo:

$$G = I_0 \cdot (fac.atenuac.comp.atmosf.) \cdot (fac.atenuac.nubes) \quad (3.3)$$

en el supuesto de ausencia de nubes, y con atmósfera limpia, la expresión (3.3)

nos dará esos valores máximos de radiación esperables. En este caso podemos poner:

$$I_{\max} = I_0 \cdot (\text{fac. aten. comp. atmof. dias claros}) \quad (3.4)$$

De acuerdo con este esquema, Perrin de Brichambaut et al., /52/, caracterizan los valores de  $G_{h,\max}$  según el tipo de clima. Para localidades con un clima continental proponen la siguiente expresión:

$$G_{h,\max} = 1120(\text{sen } \gamma)^{1.2} \quad (3.5)$$

donde  $\gamma$  es la altura solar correspondiente a una hora.

J.M.Caridad et al., /19/, proponen un modelo basado en calcular una envolvente para cada hora, a lo largo del año. Si representamos los valores de exposición global para una hora concreta durante varios años, observamos que estos valores presentan un límite superior o envolvente. Los autores demuestran que esta envolvente puede describirse mediante una función de Fourier de la forma:

$$G_{h,\max} = \frac{A_{0,h}}{2} + A_{1,h} \cos \frac{2\pi d}{L} + B_{1,h} \text{sen} \frac{2\pi d}{L} + A_{2,h} \cos \frac{4\pi d}{L} + B_{2,h} \text{sen} \frac{4\pi d}{L} \quad (3.6)$$

en la que el período,  $L$ , será el número de días en los que  $G_h$  es mayor que cero (durante la hora  $h$ ).

El método que proponemos en este trabajo se basa en el hecho de que al representar los datos de una serie de exposición horaria frente al seno de la altura solar, se observa una clara envolvente de los mismos, como se puede observar en la figura 3.1, en la que se hemos representado los datos correspondientes a la localidad de Madrid para el año 1980.

## MADRID-1980

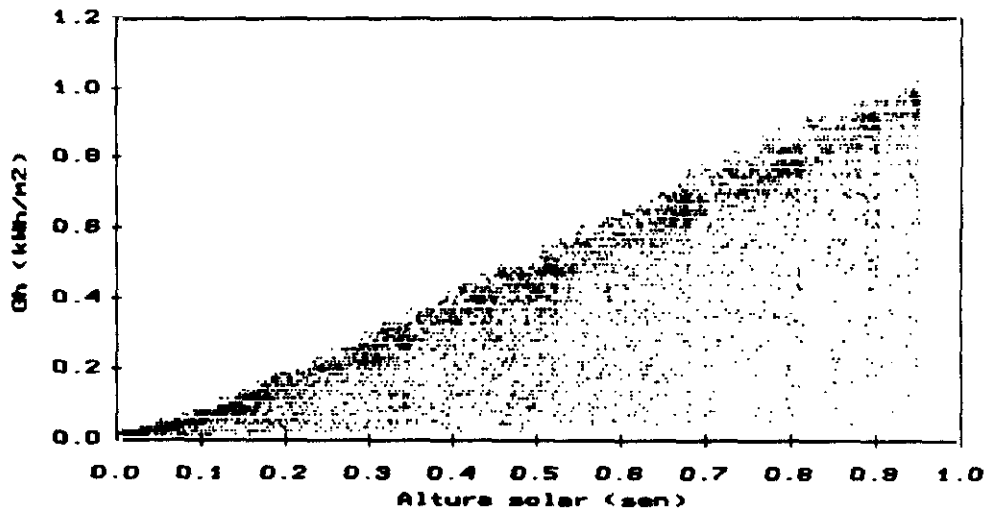


Figura 3.1. Valores de exposición horaria de radiación global para el año 1980. Madrid.

Esta función depende de la altura solar y es prácticamente independiente de la época del año, tal y como se describe en el trabajo de Ll. Mora et al, /39/. La expresión que define esta envolvente es del tipo:

$$G_{h,max} = A(\text{sen}\gamma)^B \exp(C/\text{sen}\gamma) \quad (3.7)$$

donde A, B y C son parámetros que se determinan en función de los datos de cada localidad.

Esta será la expresión que nosotros vamos a utilizar para calcular  $G_{h,max}$ . Con esta expresión propondremos unos valores para los parámetros A, B y C para cada localidad, calculados mediante el siguiente procedimiento:

- Dividir el rango de valores correspondiente a la altura solar en diez intervalos.
- Seleccionar para cada uno de estos intervalos el valor máximo de la serie de datos.

- Ajustar estos valores máximos medidos a una función del tipo anteriormente propuesto.

Ajustando por el método de mínimos cuadrados la ecuación (3.7), transformada logarímicamente, obtenemos estimadores de  $\ln A$ ,  $B$  y  $C$  y sus correspondientes desviaciones típicas. Estos valores, para cada localidad, se muestran en la Tabla 3.IV.

LOC	$\hat{\ln A}$	std $\hat{\ln A}$	$\hat{B}$	std $\hat{B}$	$\hat{C}$	std $\hat{C}$
BA	6.982	2.708	1.037	0.018	0.021	0.014
CS	6.965	3.091	0.999	0.034	0.040	0.020
LO	7.018	2.485	1.039	0.009	0.010	0.013
M	7.039	3.135	1.058	0.007	-0.043	0.022
MA	6.999	2.197	1.027	0.010	0.012	0.010
MU	7.013	2.639	1.05	0.008	0.017	0.011
OV	7.008	2.708	0.982	0.039	-0.017	0.012
PM	6.989	2.303	1.075	0.016	0.041	0.027
SE	7.031	2.944	1.03	0.013	0.012	0.01
TR	7.030	2.833	1.072	0.012	-0.004	0.01

Tabla 3.IV

Como se puede observar, para las distintas localidades el rango de variación de los parámetros  $A$ ,  $B$  y  $C$  calculados es pequeño. Esto lo podemos observar en la figura 3.2 en la que aparece la función  $\hat{G}_{h,\max}$  para cada localidad en función de la altura solar. En esta expresión  $\hat{G}_{h,\max}$  es:

$$\hat{G}_{h,\max} = e^{\hat{\ln A}} (\text{sen } \gamma)^{\hat{B}} \exp\left(\frac{\hat{C}}{\text{sen } \gamma}\right) \quad (3.8)$$

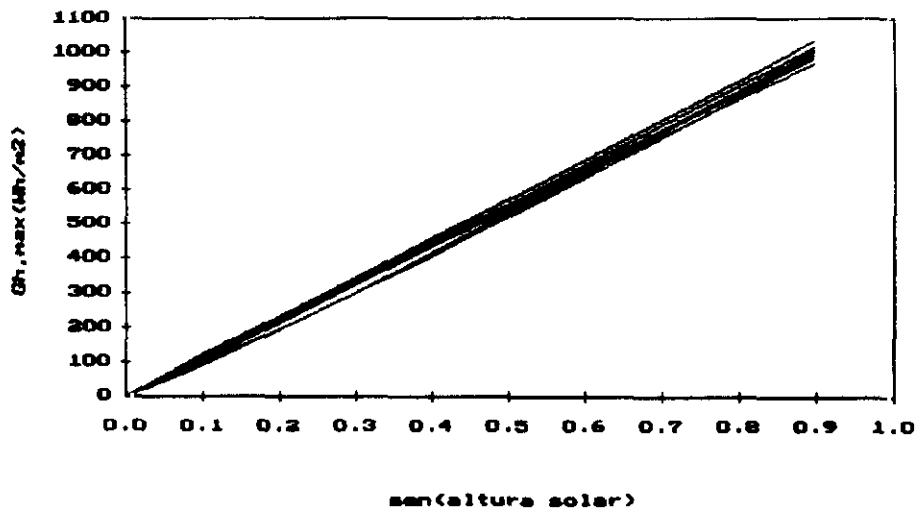


Figura 3.2.  $\hat{G}_{h,max}$  en función de la altura solar, para las distintas localidades.

Basándonos en esta observación y con el objeto de proponer una envolvente única para todas las localidades que estamos estudiando, hemos calculado unos coeficientes A, B y C generales que son el resultado de hacer un análisis de sensibilidad de las curvas de  $G_{h,max}$  de cada localidad en función de los valores de A, B y C. Estos valores son:

$$A=1100 \quad B=1.05 \quad C=0 \quad (3.9)$$

Para ver si podemos admitir que los parámetros toman estos valores en todas las localidades realizaremos un contraste estadístico. El estadístico del contraste que utilizamos es:

$$\frac{\hat{\theta}_c - \theta_p}{\hat{\sigma}_c} \approx N(0,1) \quad (3.10)$$

donde  $\theta_p$  es el valor general propuesto para cada parámetro (lna, B ó C),  $\hat{\theta}_c$  es el estimador del parámetro para cada localidad y  $\hat{\sigma}_c$  es la desviación típica de este estimador. Hay que hacer observar que para A, lo que en realidad se contrasta es si  $\ln A = 7.003$ .

En casi todos los casos el valor del estadístico del contraste es, en valor absoluto, menor que 1.96 por lo que, con un nivel de significación  $\alpha=0.05$  podemos aceptar la hipótesis nula de que las igualdades de la ecuación (3.9) son ciertas. Los contrastes para B en Málaga y C en Castellón dan valores ligeramente por encima del valor crítico, pero cabe suponer que esto se debe a la aleatoriedad de los datos y no a que B y C sean diferentes de los valores especificados. Así pues, podemos admitir que la función dada por la expresión:

$$G_{h,max}=1100(\text{sen}\gamma)^{1.05} \quad (3.11)$$

puede utilizarse para calcular los valores máximos de exposición global horaria de todas las localidades. También realizamos un contraste conjunto de las tres hipótesis contenidas en (3.9) con un estadístico chi-cuadrado con 3 grados de libertad. Nuevamente se aceptaba, con nivel de significación  $\alpha=0.05$ , la veracidad de la ecuación (3.9).

Si comparamos el método propuesto con el desarrollado por Caridad et al., /19/, observamos que el cálculo de las funciones de Fourier para cada hora, ver /39/, se ajusta mejor a los datos reales. Sin embargo, presenta el inconveniente de que para poder utilizarlo es necesario calcular un número excesivamente elevado de parámetros, por lo que hay que utilizar un gran número de datos para conseguir errores pequeños.

Respecto a la expresión propuesta por P. Brichambaut et al., /53/, puede observarse que hemos llegado a una expresión del mismo tipo, pero con los coeficientes obtenidos con datos de localidades españolas; esto se refleja en que el error y desviación estándar utilizando los estimadores propuestos en (3.9) es menor que si utilizamos los valores de P.Brighambaut.



### 3.4 ANÁLISIS DE LAS SERIES DE EXPOSICIÓN HORARIA

Si nos fijamos en la radiación que llega a la superficie de la tierra durante muchos días de un año, vemos que los valores que recibimos pueden variar ampliamente: desde recibir "muy poca intensidad" (cuando el cielo está totalmente cubierto), recibir una intensidad "algo mayor" (un cielo parcialmente cubierto), recibir una intensidad "elevada" (momentos en que el cielo está totalmente despejado), etc. Es decir, la posibilidad de situaciones es bastante amplia. Algunos autores, Suehrcke et. al., /60/, nos hablan de la "bimodalidad" de la radiación solar instantánea, en el sentido de cielo despejado/cubierto; posiblemente esta característica de bimodalidad sí se presente cuando hablamos de radiación global directa (radiación que llega en línea con el disco solar), pero, sin embargo, cuando hablamos de radiación global, no parece ser una clasificación adecuada.

Para corroborar nuestra hipótesis de que para radiación global no se puede mantener la clasificación única de cielo cubierto/despejado, vamos a utilizar una clasificación de "días tipo" realizada en el Laboratorio del Instituto de Energías Renovables (Laboratorio de Energía Solar Fotovoltaica). Utilizando medidas instantáneas registradas durante un año, se clasificaron la mayoría de los días utilizando ocho tipos distintos de días. En el Anexo I se muestra esta clasificación, con el porcentaje de días que se ha repetido cada uno de los "días tipo" a lo largo del año de medidas. En las figuras de este anexo se pueden observar las variaciones que se pueden producir dentro de un día en los niveles de radiación. Puesto que la variación de los componentes atmosféricos para un mismo día no es significativa, lo que en realidad se refleja en estas gráficas es la influencia (presencia) de nubes.

Como podemos observar, las variaciones en los valores de radiación entre dos instantes seguidos pueden ser desde muy significativas a poco o nada apreciables.

---

Esto nos lleva a rechazar, para la radiación global, la hipótesis de "bimodalidad de la radiación solar instantánea". Otra conclusión que podemos sacar de estas observaciones es que la forma de estos "días tipo" parece indicar que las variaciones que se producen en los niveles máximos de radiación para una hora no son independientes de la hora que consideremos.

Una vez hechas estas consideraciones generales sobre la naturaleza de la radiación solar, vamos a centrarnos en las características de las series de exposición horaria de radiación global.

Como observamos en el capítulo anterior, en estas series nos encontramos con una evolutividad en la media, debida a dos factores, a saber: la sucesión día-noche y la variación en la posición relativa sol-tierra a lo largo del año. En ambos casos la evolutividad que observamos es cíclica. La tendencia diaria observada se debe, como ya hemos comentado, a la variación de la masa atmosférica o camino óptico, para las distintas horas de un día. En la figura 2.1 se puede observar este hecho. La tendencia estacional se puede observar en la figura 2.2. La forma general de estas series, en lo que a tendencia se refiere, se repite para las restantes horas del día, en todas las localidades.

Así pues, al no ser estacionarias estas series, habrá que transformarlas para que se puedan considerar generadas por procesos estacionarios. Para tratar la evolutividad en la media vamos a utilizar el método de ajuste de funciones en la variable tiempo. Su justificación consiste en suponer que la serie original, la de valores de exposición global horaria,  $G_h$ , es una función matemática en la variable tiempo,  $f(h)$ , multiplicada por una componente estacionaria,  $Y_h$ . Esto es:

$$G_h = f(h) Y_h \quad (3.12)$$

De los métodos descritos en el capítulo anterior, en este trabajo vamos a utilizar la función que nos da el valor de exposición horaria máxima de radiación

global en función de la altura solar. Esta función es la que se recoge en la expresión (3.11).

El índice que vamos a utilizar en el análisis temporal que desarrollaremos en el siguiente capítulo, es el que resulta de dividir los valores de exposición horaria ( $G_h$ ) por los de exposición horaria máxima ( $G_{h,max}$ ) según la expresión:

$$X_h = \frac{G_h}{G_{h,max}} \quad (3.14)$$

donde  $G_{h,max}$  se calcula como se describió en el apartado anterior. Vamos a describir en los siguientes apartados las propiedades de las series de este índice para cada una de las localidades. Estudiaremos los valores medios, las varianzas y las funciones de distribución de probabilidades de ocurrencia. El discretizado que vamos a utilizar para estimar estas funciones es de 0.1.

En este capítulo el estudio será descriptivo. Nuestro objetivo será determinar si existe o no relación entre los distintos parámetros de la serie, como son: valor medio del índice, mes o época del año y función de distribución de probabilidad de ocurrencia de los valores del índice  $X_h$ . Basándonos en las observaciones que vamos a hacer en los siguientes apartados, en el capítulo 6 haremos un análisis formal de las variables que vamos a describir.

#### 3.4.1 Descripción de las series del índice $X_h$ de Badajoz

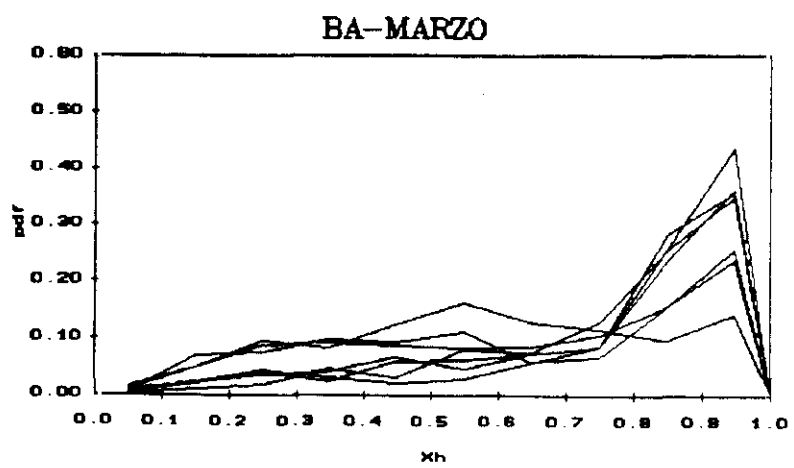
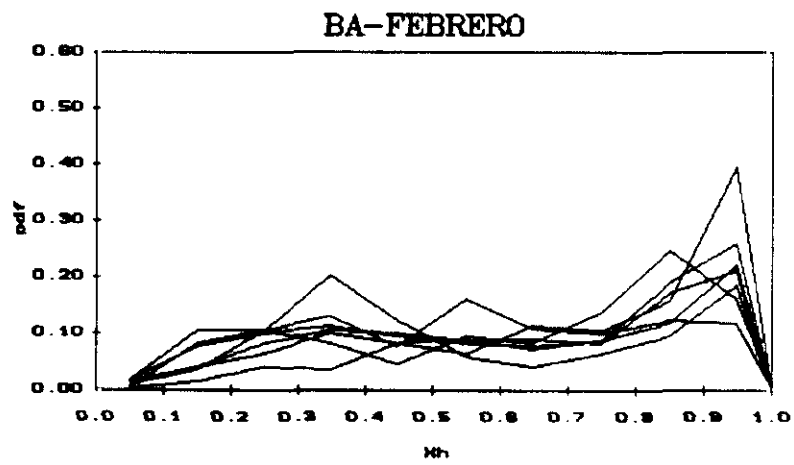
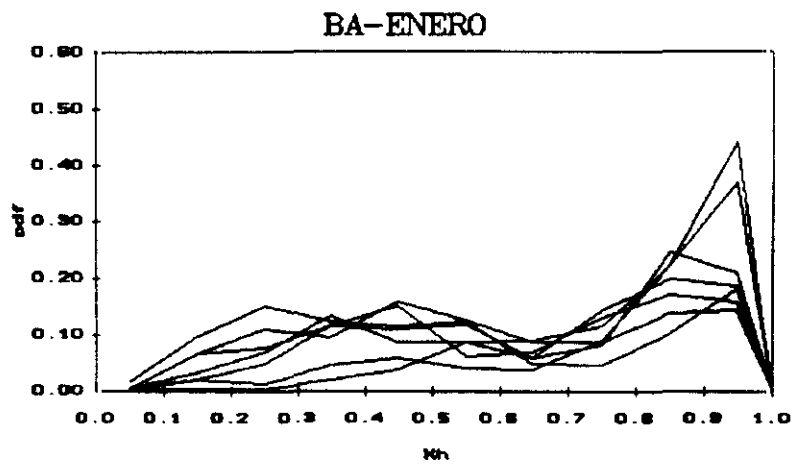
Para todas las series mensuales del índice  $X_h$  hemos calculado su valor medio mensual y varianza. En la Tabla 3.V se muestran los resultados obtenidos. En esta tabla, para cada año, en la primera línea se muestra el valor medio de cada serie mensual y en la segunda su varianza. Para los meses en los que aparece un cero para estos dos valores, es porque, o bien, están incompletos, o bien no hay datos. El mismo tipo de tabla es el que utilizaremos para las demás localidades.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1976	0.603 0.025	0.512 0.041	0.593 0.034	0.490 0.048	0.552 0.049	0.603 0.032	0.646 0.011	0.566 0.029	0.598 0.032	0.512 0.057	0.561 0.034	0.359 0.030
1977	0.403 0.044	0.432 0.038	0.610 0.024	0.602 0.032	0.574 0.039	0.597 0.031	0.643 0.019	0.662 0.011	0.601 0.017	0.500 0.036	0.493 0.043	0.380 0.044
1978	0.456 0.038	0.495 0.042	0.598 0.029	0.501 0.048	0.544 0.043	0.515 0.040	0.677 0.005	0.632 0.010	0.574 0.014	0.563 0.027	0.551 0.033	0.399 0.031
1979	0.427 0.036	0.411 0.049	0.505 0.048	0.595 0.039	0.615 0.024	0.628 0.019	0.607 0.018	0.651 0.007	0.600 0.022	0.440 0.050	0.592 0.032	0.463 0.037
1980	0.495 0.032	0.462 0.049	0.503 0.045	0.594 0.036	0.550 0.036	0.638 0.023	0.685 0.008	0.578 0.035	0.584 0.018	0.517 0.031	0.519 0.037	0.592 0.020
1981	0.614 0.017	0.571 0.028	0.460 0.037	0.498 0.047	0.611 0.029	0.636 0.018	0.664 0.014	0.627 0.016	0.576 0.033	0.576 0.024	0.590 0.015	0.402 0.042
1982	0.000 0.000	0.501 0.039	0.627 0.028	0.590 0.035	0.624 0.025	0.629 0.027	0.640 0.022	0.616 0.019	0.562 0.035	0.569 0.023	0.482 0.036	0.000 0.000
1983	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.625 0.018	0.609 0.021	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000

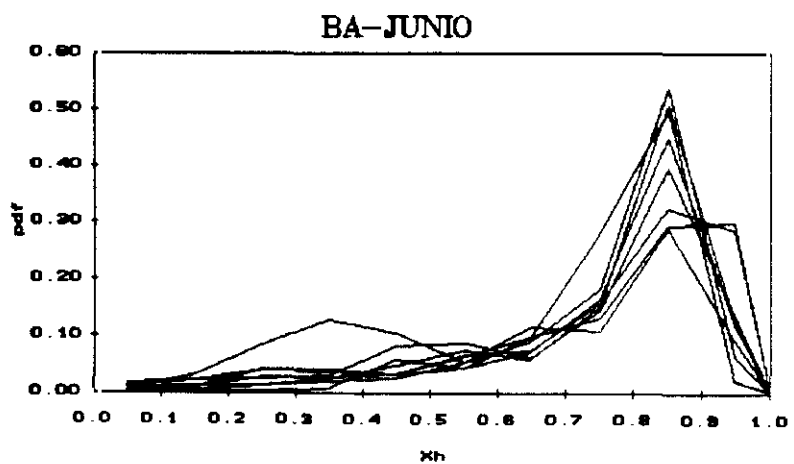
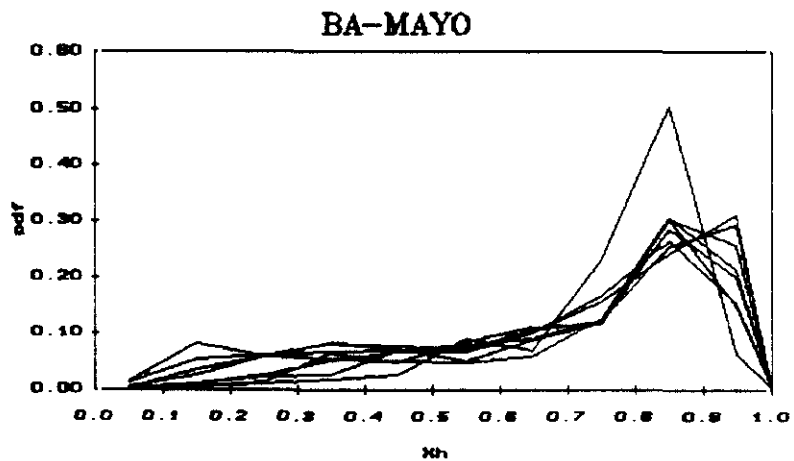
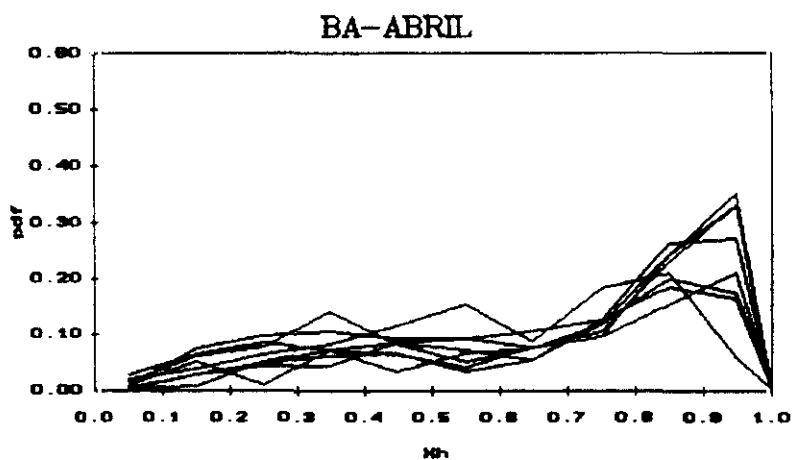
Tabla 3.V

Como puede observarse, los valores medios mensuales del índice presentan, en general, una relación con el mes del año. Para un mismo mes los valores suelen ser similares en la mayoría de los años, aunque siempre hay algún año en el cual el valor es sensiblemente diferente. Esto es lógico si pensamos en lo que representan, ya que estas series reflejan la disminución que se produce en los valores máximos de radiación esperables, sobre todo por el efecto de las nubes; esta variación guarda, en general, una fuerte dependencia con la época del año en que nos encontramos. Así, lo "normal" es que los meses de invierno presenten valores más pequeños porque en ellos suele haber más días nublados.

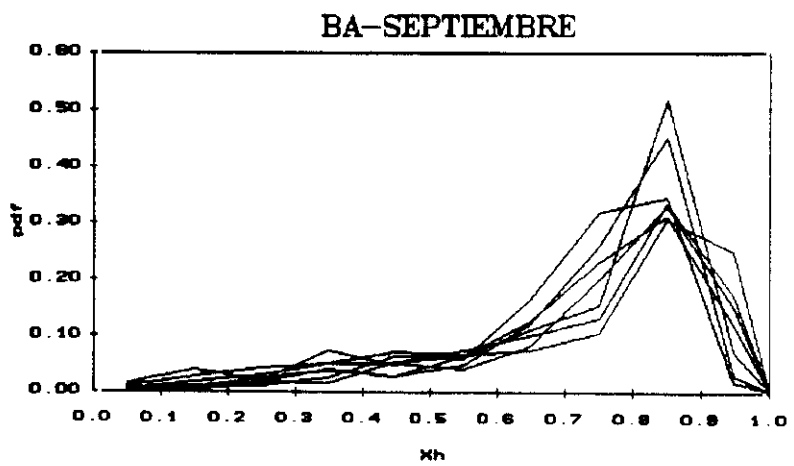
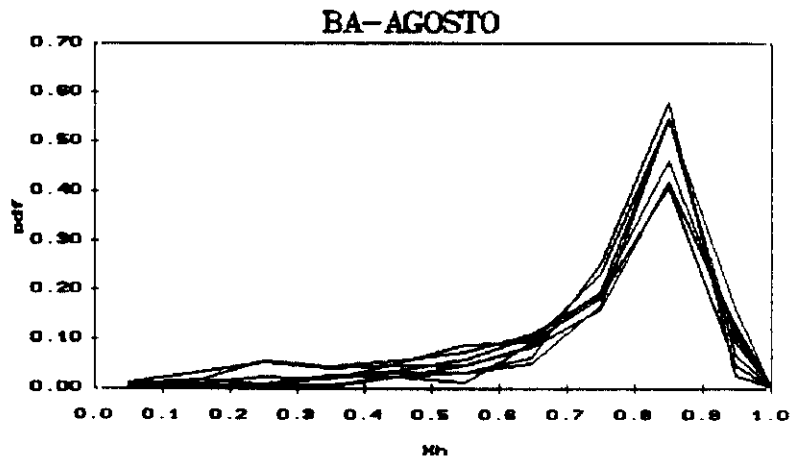
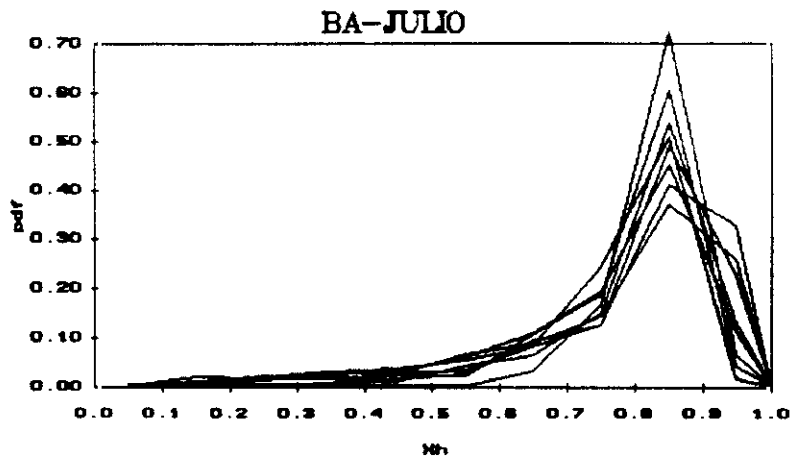
Vamos a representar también las funciones de distribución para cada uno de los meses del año (Figuras 3.3 a 3.14). En cada gráfica incluimos las funciones de distribución de las series de los distintos años para ese mes, normalizadas a la unidad (Años utilizados: 1976/83).



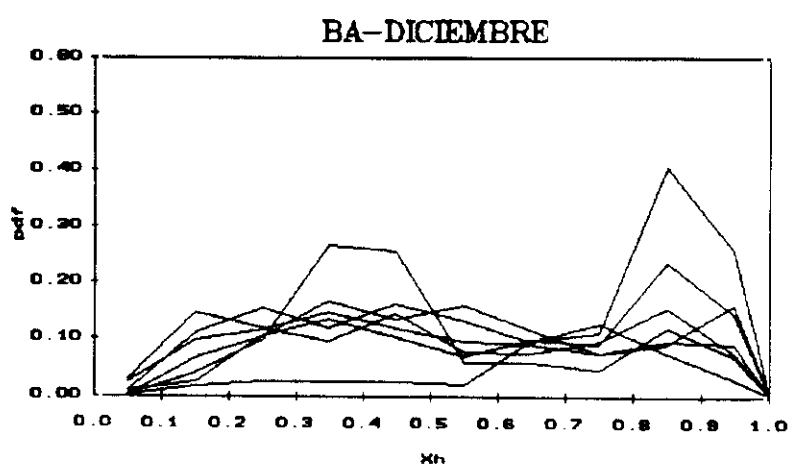
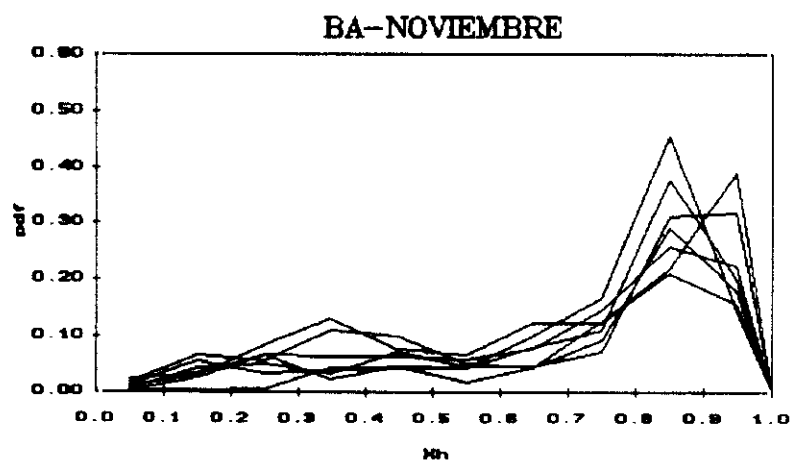
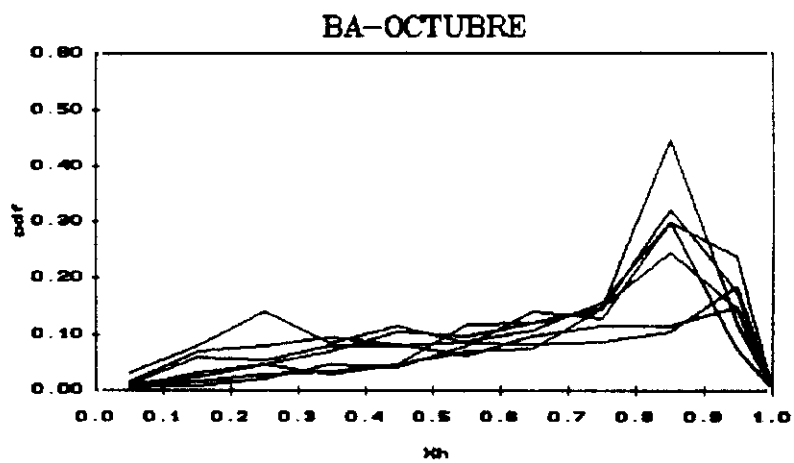
Figuras 3.3, 3.4 y 3.5



Figuras 3.6, 3.7 y 3.8



Figuras 3.9, 3.10 y 3.11



Figuras 3.12, 3.13 y 3.14



---

Si observamos estas figuras vemos que las funciones de probabilidad de ocurrencia para el índice  $X_h$  en las distintas series mensuales se pueden agrupar, fundamentalmente, en tres tipos que guardan relación con los meses del año:

- Series con un pico en torno a valores del índice de 0.75 a 0.95, que corresponden en general a los meses de verano y próximos a esta estación. La altura del pico varía entre 0.3 y 0.6 (esto es, hay un 30% o un 60% de horas, respectivamente, en los que se da ese valor).
- Series en las que hay dos valores, en general seguidos, de  $X_h$  con probabilidad de ocurrencia alta (30% o 40%). Los valores de  $X_h$  con esta probabilidad de ocurrencia están casi siempre en torno a 0.7 y 0.9, por lo que utilizando otro discretizado es probable que este tipo de funciones fuera similar al anterior. Esta forma de las funciones de distribución se repite para la mayoría de los meses de primavera y otoño.
- Series en las que hay varios valores de  $X_h$  con probabilidad de ocurrencia entre el 15% y el 25%. Es decir no hay ningún valor predominante en las series. Esta forma de las funciones de distribución se da, sobre todo, para las series de los meses de invierno.

Hay además, un cuarto tipo de funciones de distribución que, aunque se presenta con menos frecuencia, también aparece, sobre todo en los meses de invierno: son las funciones de tipo bimodal, con un valor más probable en el rango alto de valores y otro menos probable que éste, pero más que el resto, en el rango de valores medios o bajos.

Para cada mes hay siempre una forma de la función de distribución predominante (los distintos valores del índice ocurren, en los distintos años, aproximadamente con la misma probabilidad). En general, estas funciones, que

son similares para cada mes, corresponden a los años para los que también son similares los valores medios del índice  $X_b$ , según se indica en la Tabla 3.V. También para casi todos los meses siempre existe algún año para el que la distribución de los valores del índice no es similar a la de los demás años; estos años son, en general, los que corresponden a valores medios mensuales del índice muy distintos a los del resto de los años. Antes de sacar conclusiones más generales, vamos a analizar qué es lo que pasa para las demás localidades.

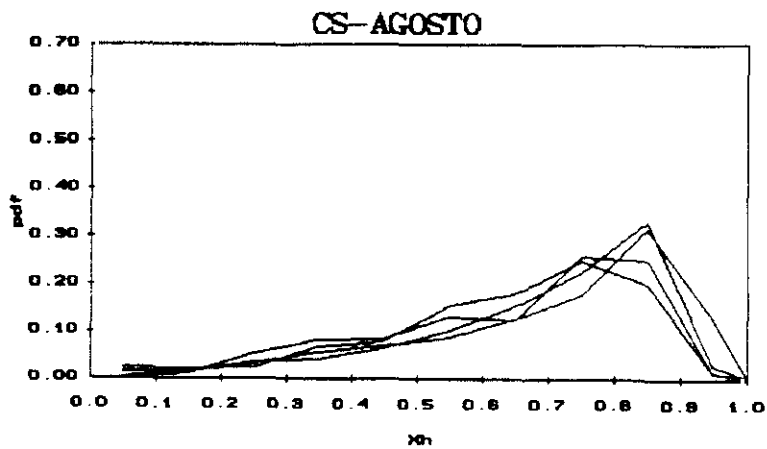
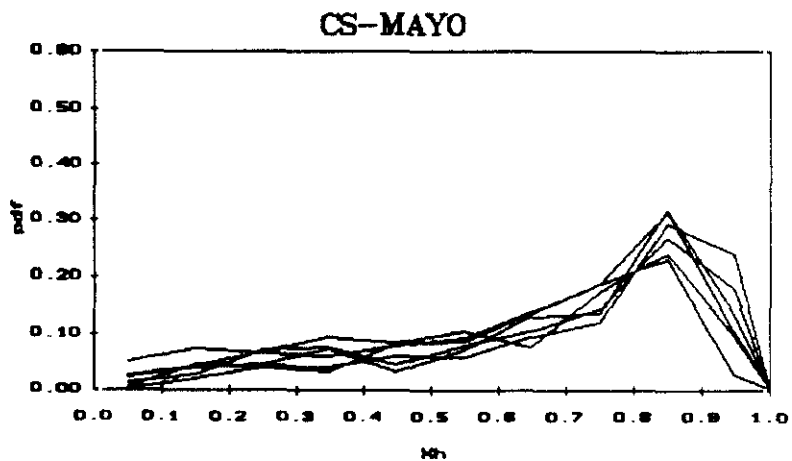
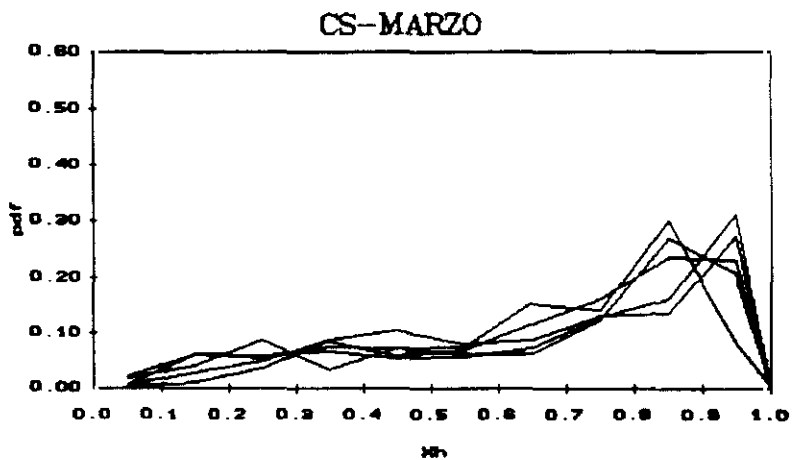
### 3.4.2 Descripción de las series del índice $X_b$ de Castellón

La Tabla 3.VI muestra los valores medios mensuales de las series del índice  $X_b$  y sus varianzas. En las figuras 3.15, 3.16, 3.17 y 3.18 se muestran, a modo de ejemplo, las curvas de probabilidad de ocurrencia para algunos meses, de los años 1979/84.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1979	0.390 0.060	0.488 0.048	0.542 0.039	0.587 0.041	0.554 0.039	0.549 0.037	0.588 0.020	0.575 0.027	0.470 0.040	0.000 0.000	0.589 0.019	0.522 0.029
1980	0.525 0.045	0.469 0.052	0.536 0.042	0.000 0.000	0.506 0.042	0.548 0.043	0.570 0.034	0.523 0.026	0.514 0.028	0.555 0.031	0.503 0.038	0.546 0.034
1981	0.594 0.032	0.527 0.049	0.561 0.029	0.468 0.050	0.587 0.037	0.582 0.031	0.646 0.014	0.553 0.026	0.572 0.027	0.584 0.032	0.541 0.021	0.527 0.037
1982	0.470 0.050	0.421 0.046	0.539 0.048	0.555 0.039	0.550 0.040	0.602 0.020	0.551 0.014	0.512 0.025	0.560 0.017	0.475 0.052	0.472 0.040	0.518 0.029
1983	0.570 0.011	0.511 0.036	0.510 0.036	0.586 0.026	0.570 0.027	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.466 0.028
1984	0.000 0.000	0.000 0.000	0.000 0.000	0.484 0.048	0.477 0.045	0.580 0.025	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000

Tabla 3.VI

Como se puede observar, los valores medios mensuales presentan, en general, una pequeña tendencia anual.



Figuras 3.15, 3.16 y 3.17

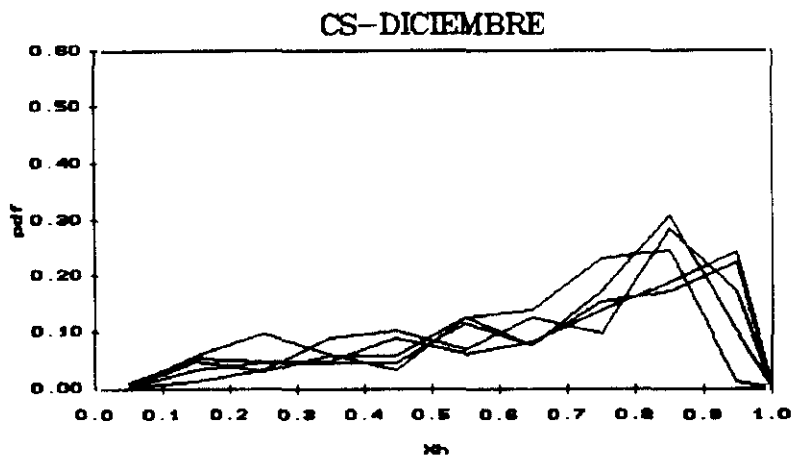


Figura 3.18

Las funciones de distribución presentan unas características generales que se repiten para los distintos meses:

- Hay siempre para cada mes una forma predominante de las funciones de los distintos años. Es decir, para cada mes existen, en general, unos valores del índice  $X_h$  que se repiten con la misma probabilidad en los distintos años. Cuando aparece en las gráficas de un mes alguna curva distinta sensiblemente de las demás, coincide siempre que para ese año se da un valor medio mensual del índice muy distinto al de los demás. Este hecho ya lo observamos para las series de Badajoz.
- En esta localidad los picos de las funciones de distribución son más pequeños que para la localidad de Badajoz. En general, nunca sobrepasan el valor de 0.3, es decir, ningún valor del índice se repite más de un 30% de las veces para un mes.

Los tipos de curvas que nos encontramos son similares a los descritos para la localidad de Badajoz.

### 3.4.3 Descripción de las series del índice $X_h$ de Logroño

En la Tabla 3.VII se muestran los valores medios mensuales del índice  $X_h$  y sus varianzas para los distintos meses y años analizados. En las figuras 3.19, 3.20, 3.21 y 3.22 se presentan las funciones de distribución de probabilidad de ocurrencia de estas series, durante los años 1981/84.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1981	0.406 0.050	0.498 0.043	0.479 0.041	0.459 0.059	0.532 0.048	0.634 0.035	0.605 0.041	0.554 0.039	0.567 0.035	0.502 0.041	0.535 0.037	0.393 0.044
1982	0.422 0.045	0.403 0.051	0.547 0.040	0.583 0.046	0.530 0.053	0.570 0.039	0.595 0.036	0.544 0.044	0.547 0.046	0.450 0.051	0.367 0.047	0.362 0.041
1983	0.496 0.038	0.479 0.042	0.503 0.046	0.511 0.046	0.498 0.052	0.595 0.039	0.585 0.034	0.503 0.049	0.621 0.023	0.569 0.036	0.394 0.046	0.499 0.038
1984	0.455 0.045	0.495 0.049	0.539 0.054	0.543 0.054	0.412 0.046	0.591 0.037	0.661 0.022	0.536 0.040	0.548 0.038	0.520 0.044	0.408 0.051	0.381 0.050

Tabla 3.VII

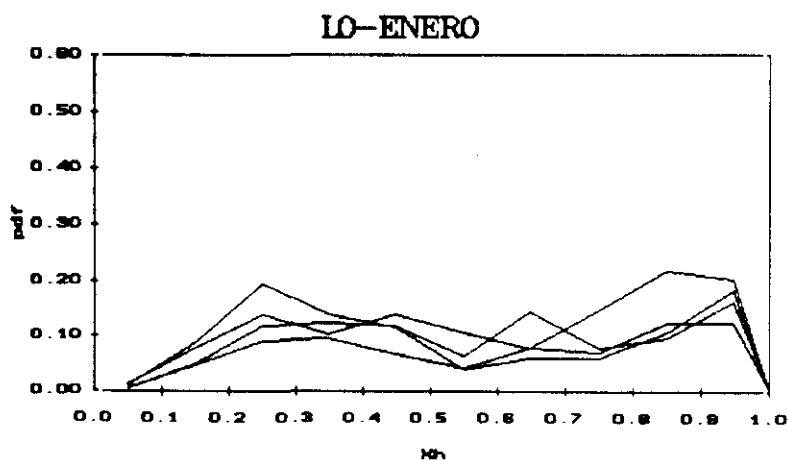
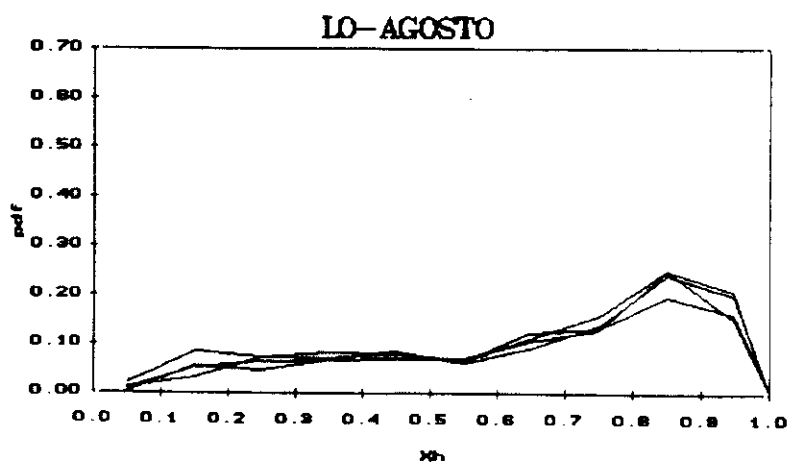
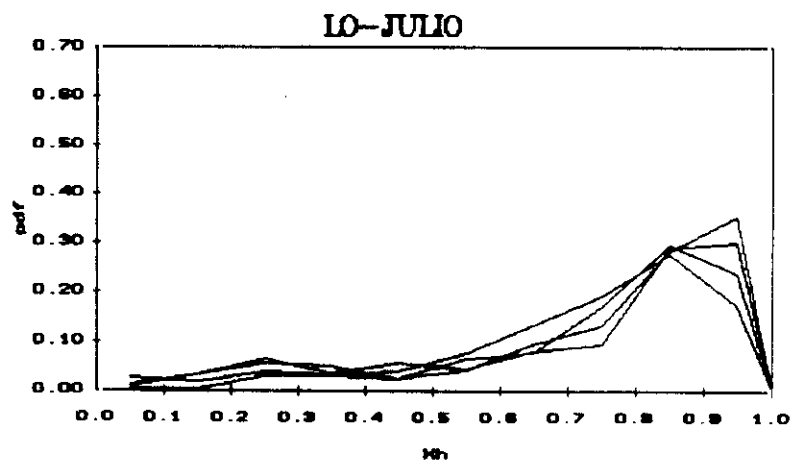
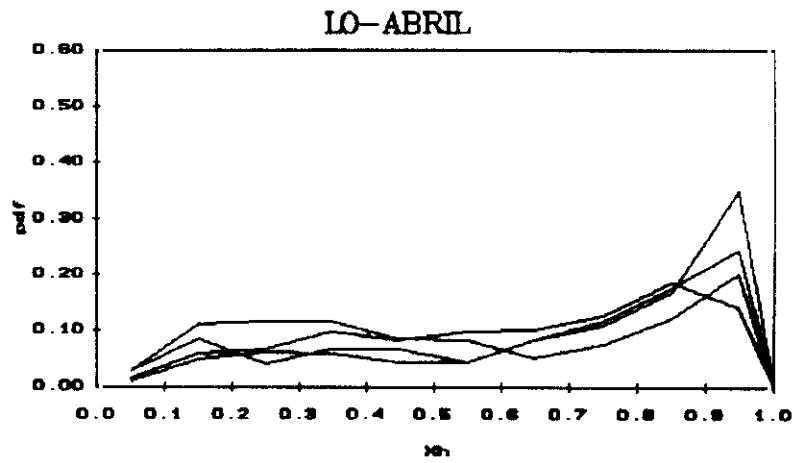


Figura 3.19



Figuras 3.20, 3.21 y 3.22

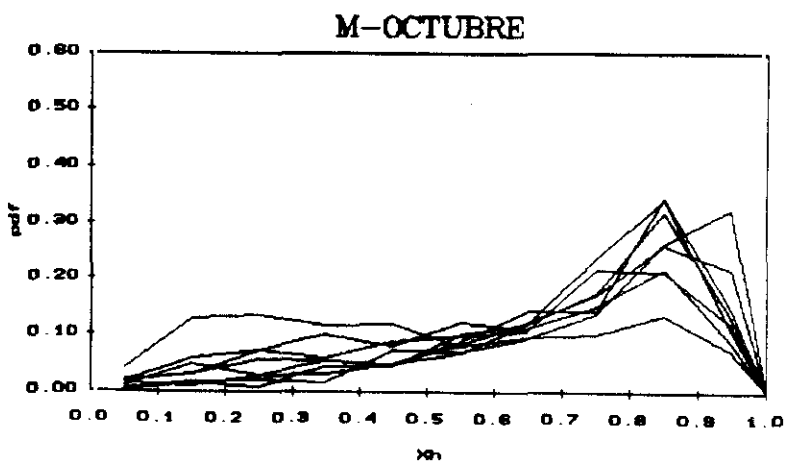
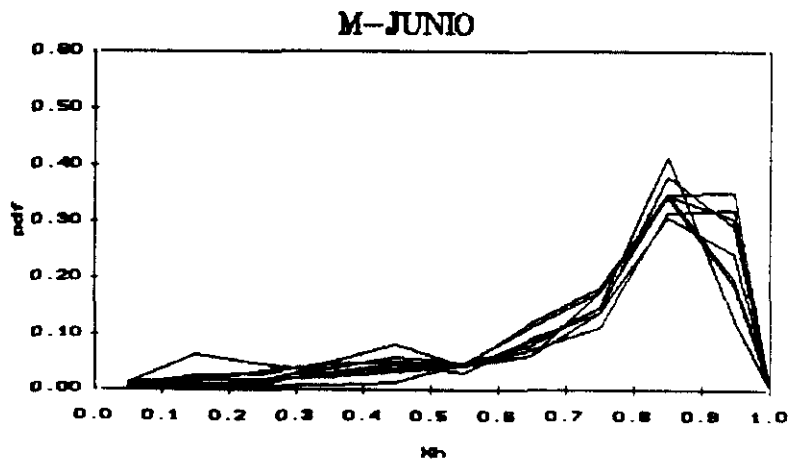
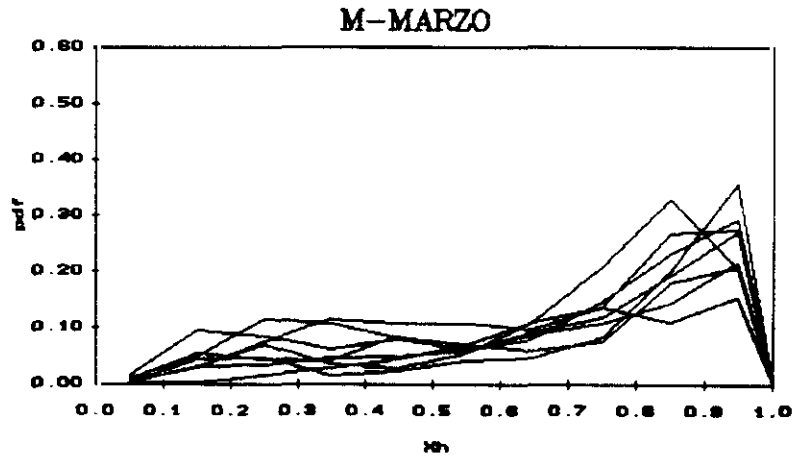
Como se puede observar, en las curvas de probabilidad de ocurrencia sólo aparecen, en general, dos tipos de formas: para casi ningún mes aparecen las curvas en las que había un único valor mucho más probable que los demás (un pico en la curva). El mayor porcentaje de ocurrencia de valores de  $X_h$  está cercano siempre al 30%. Nuevamente encontramos una clara relación entre los valores medios de las series y las formas de estas curvas: a valores medios mensuales similares les corresponden curvas similares.

#### 3.4.4 Descripción de las series del índice $X_h$ de Madrid

La Tabla 3.VIII muestra los valores medios mensuales de las series del índice  $X_h$  para los distintos meses y sus varianzas. En las figuras 3.23, 3.24, 3.25 y 3.26 se muestran los valores de la función de probabilidad de ocurrencia de estas series para los años 1979/86.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1979	0.369 0.054	0.437 0.060	0.502 0.051	0.592 0.044	0.612 0.036	0.630 0.029	0.607 0.024	0.665 0.009	0.567 0.024	0.388 0.044	0.572 0.020	0.464 0.039
1980	0.435 0.040	0.451 0.048	0.561 0.041	0.611 0.045	0.537 0.045	0.650 0.023	0.677 0.015	0.636 0.021	0.625 0.013	0.554 0.032	0.492 0.043	0.544 0.029
1981	0.612 0.017	0.536 0.049	0.501 0.041	0.486 0.060	0.599 0.043	0.652 0.020	0.680 0.018	0.627 0.021	0.637 0.022	0.593 0.024	0.578 0.020	0.401 0.052
1982	0.441 0.046	0.489 0.048	0.632 0.036	0.592 0.041	0.593 0.034	0.592 0.031	0.638 0.022	0.610 0.023	0.570 0.034	0.501 0.034	0.444 0.032	0.441 0.036
1983	0.550 0.011	0.475 0.037	0.598 0.017	0.573 0.035	0.580 0.038	0.621 0.021	0.677 0.008	0.576 0.028	0.648 0.010	0.555 0.026	0.341 0.037	0.439 0.037
1984	0.446 0.044	0.582 0.028	0.503 0.052	0.575 0.045	0.429 0.045	0.588 0.042	0.693 0.007	0.644 0.017	0.656 0.013	0.530 0.033	0.378 0.050	0.433 0.036
1985	0.441 0.049	0.422 0.036	0.587 0.033	0.518 0.048	0.552 0.040	0.607 0.029	0.652 0.013	0.669 0.011	0.631 0.009	0.561 0.019	0.456 0.035	0.384 0.034
1986	0.467 0.037	0.405 0.047	0.576 0.033	0.526 0.050	0.607 0.032	0.680 0.012	0.658 0.017	0.658 0.018	0.532 0.035	0.470 0.037	0.509 0.029	0.489 0.030

Tabla 3.VIII



Figuras 3.23, 3.24 y 3.25



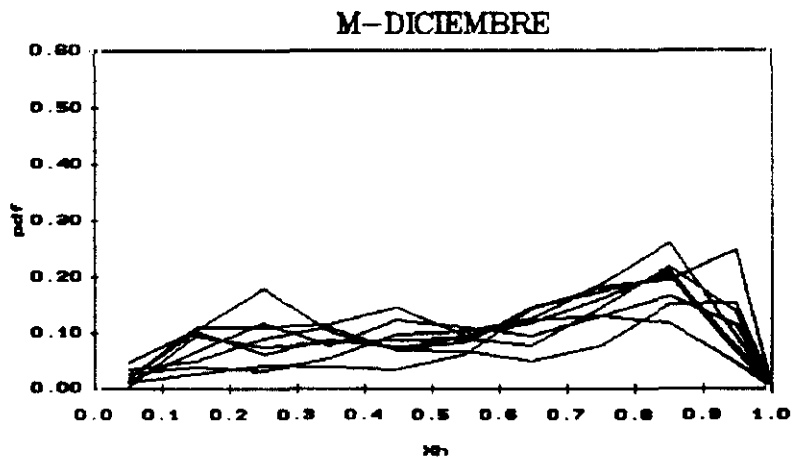


Figura 3.26

Las características generales de las funciones son las mismas que las comentadas para las localidades analizadas anteriormente. Como único hecho específico de esta localidad es que en muchos meses aparecen mezcladas curvas con un pico y curvas con dos valores mucho más probables que los demás (o picos). Además los valores de la función en éstos son algo más elevados (mayor porcentaje de veces que se presenta ese valor de  $X_b$ ).

#### 3.4.5 Descripción de las series del índice $X_b$ de Málaga

En la Tabla 3.IX se muestran los valores medios mensuales y varianzas de las series del índice  $X_b$  para la localidad de Málaga.

En las figuras 3.27, 3.28, 3.29 y 3.30 aparecen las funciones de distribución de probabilidad de ocurrencia de los valores del índice  $X_b$  en las series de datos de los años 1977/84.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1977	0.410 0.050	0.538 0.033	0.608 0.021	0.615 0.025	0.608 0.022	0.634 0.020	0.623 0.020	0.620 0.014	0.600 0.014	0.475 0.039	0.546 0.040	0.407 0.049
1978	0.496 0.042	0.571 0.034	0.629 0.020	0.536 0.037	0.575 0.028	0.562 0.039	0.635 0.012	0.583 0.013	0.000 0.000	0.000 0.000	0.549 0.036	0.435 0.039
1979	0.394 0.049	0.490 0.043	0.558 0.037	0.618 0.018	0.593 0.027	0.608 0.017	0.574 0.020	0.589 0.015	0.554 0.026	0.424 0.050	0.560 0.023	0.483 0.030
1980	0.484 0.038	0.524 0.030	0.507 0.043	0.547 0.042	0.564 0.037	0.635 0.027	0.663 0.012	0.562 0.025	0.566 0.021	0.564 0.026	0.503 0.043	0.565 0.027
1981	0.560 0.029	0.577 0.037	0.568 0.034	0.496 0.051	0.641 0.026	0.629 0.019	0.651 0.014	0.601 0.020	0.619 0.019	0.552 0.028	0.525 0.019	0.494 0.038
1982	0.515 0.037	0.520 0.046	0.566 0.035	0.497 0.047	0.580 0.035	0.647 0.011	0.581 0.032	0.569 0.027	0.574 0.019	0.588 0.023	0.507 0.036	0.560 0.026
1983	0.552 0.020	0.487 0.043	0.563 0.029	0.532 0.038	0.603 0.033	0.605 0.023	0.633 0.010	0.583 0.022	0.598 0.013	0.499 0.031	0.375 0.038	0.492 0.039
1984	0.547 0.023	0.552 0.035	0.000 0.000	0.497 0.053	0.504 0.046	0.622 0.020	0.594 0.024	0.596 0.023	0.611 0.014	0.563 0.026	0.419 0.052	0.449 0.032

Tabla 3.IX

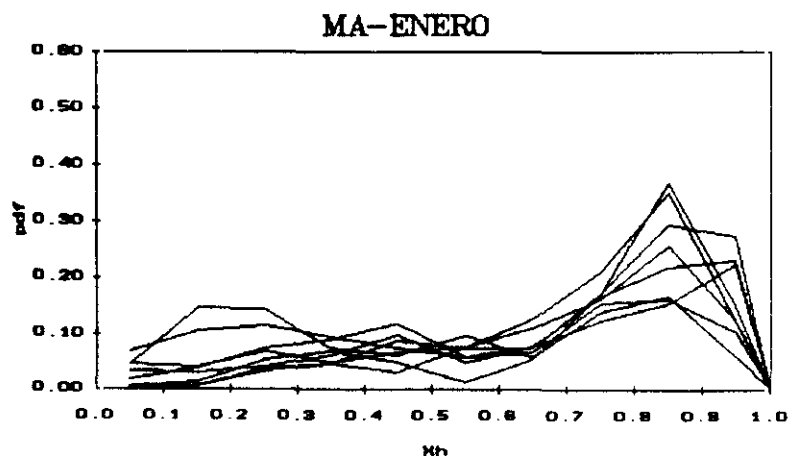


Figura 3.27

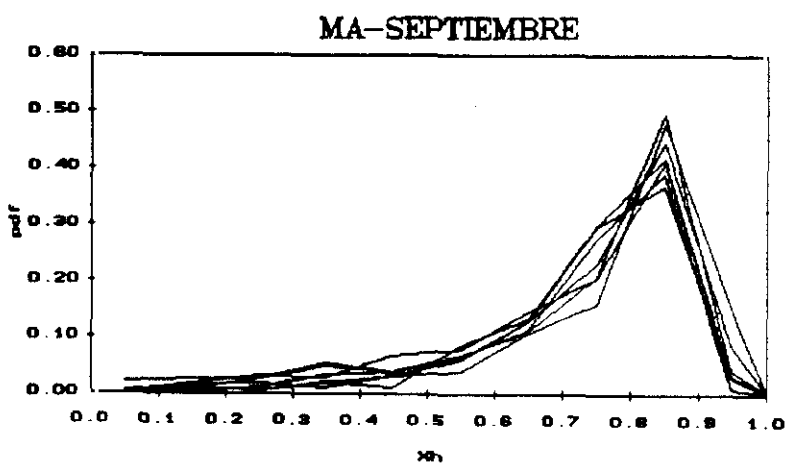
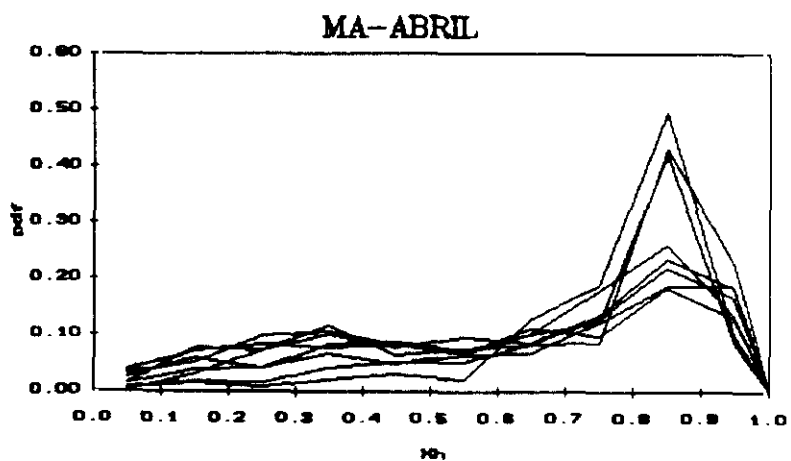
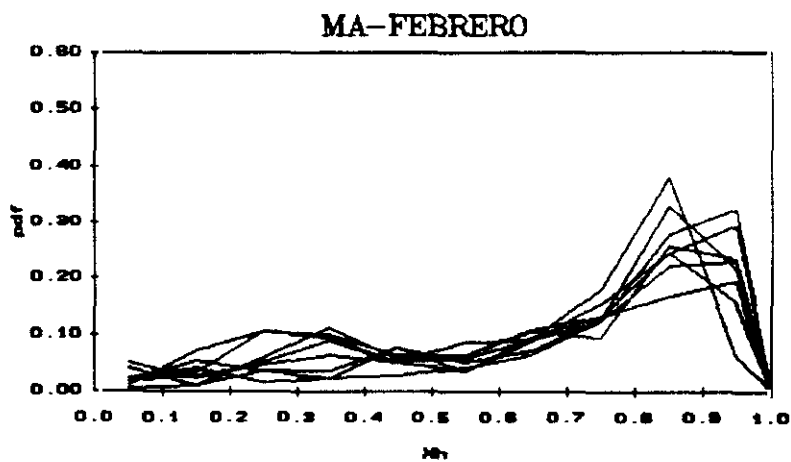


Figura 3.28, 3.29 y 3.30

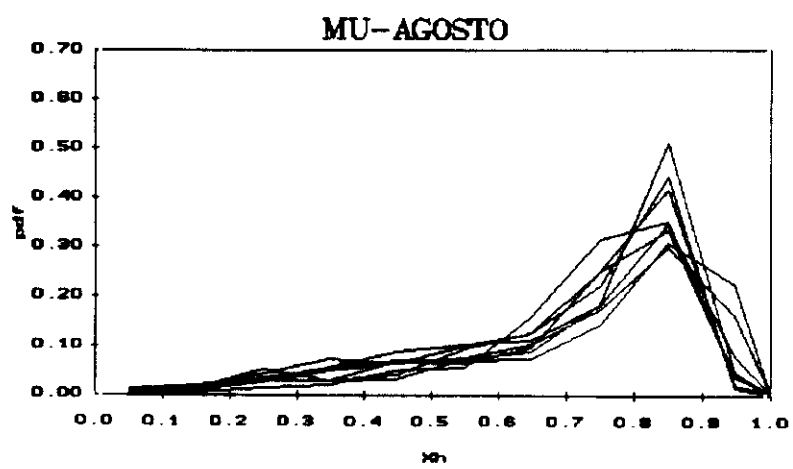
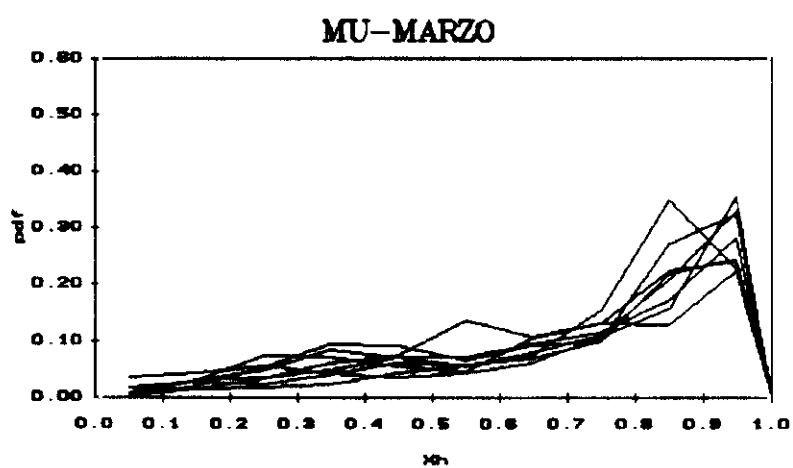
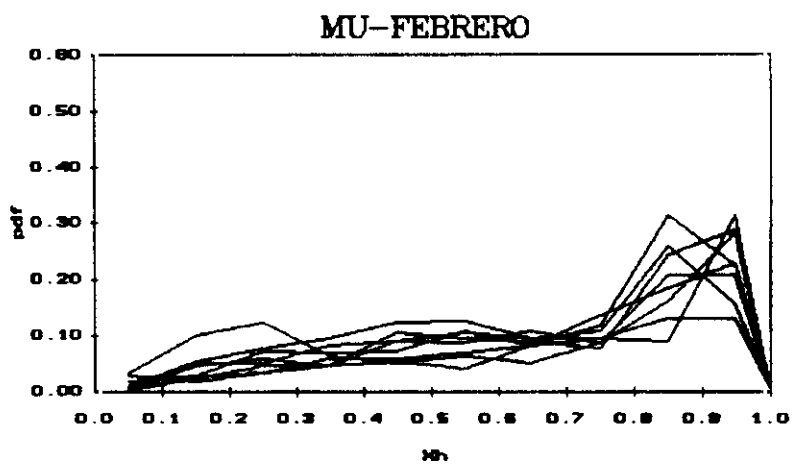
Para esta localidad, en general, nos encontramos sólo con dos formas de curvas de la función de distribución de probabilidades de ocurrencia: la que tiene un valor con mucha más probabilidad que los demás (pico), y la que tiene dos valores con más probabilidad que los demás, del tipo que llamamos bimodal. Además la frecuencia con que se dan este/estos valores es más elevada que para las localidades anteriormente analizadas. Otro hecho significativo es que ni siquiera en los meses de invierno hay, en general, funciones de distribución con muchos valores con probabilidad de ocurrencia similar.

#### 3.4.6 Descripción de las series del índice $X_h$ de Murcia

En la Tabla 3.X se recogen los valores medios mensuales y varianzas del índice  $X_h$  para todos los meses con datos de la localidad de Murcia. En las gráficas 3.31, 3.32, 3.33 y 3.34 se muestran las funciones de distribución de probabilidad de ocurrencia de los valores del índice  $X_h$  en las series de los años 1977/84 para esta localidad.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1977	0.374 0.047	0.506 0.035	0.601 0.026	0.572 0.036	0.534 0.044	0.631 0.023	0.596 0.023	0.601 0.022	0.587 0.017	0.472 0.033	0.541 0.037	0.450 0.037
1978	0.468 0.036	0.546 0.031	0.586 0.033	0.546 0.036	0.549 0.033	0.600 0.027	0.652 0.007	0.572 0.024	0.582 0.013	0.518 0.027	0.492 0.034	0.447 0.033
1979	0.384 0.054	0.528 0.036	0.553 0.039	0.601 0.028	0.568 0.036	0.589 0.023	0.578 0.020	0.592 0.015	0.560 0.027	0.450 0.041	0.589 0.027	0.523 0.027
1980	0.506 0.046	0.473 0.056	0.550 0.039	0.546 0.044	0.596 0.045	0.643 0.025	0.669 0.017	0.562 0.028	0.585 0.021	0.591 0.031	0.541 0.042	0.553 0.051
1981	0.625 0.041	0.588 0.044	0.566 0.034	0.494 0.054	0.629 0.039	0.627 0.032	0.693 0.012	0.586 0.034	0.626 0.022	0.577 0.036	0.574 0.020	0.569 0.039
1982	0.539 0.044	0.480 0.042	0.575 0.048	0.561 0.044	0.536 0.040	0.629 0.019	0.629 0.021	0.573 0.031	0.562 0.038	0.556 0.038	0.491 0.045	0.564 0.033
1983	0.594 0.020	0.555 0.038	0.588 0.035	0.611 0.031	0.610 0.033	0.563 0.030	0.613 0.021	0.559 0.024	0.612 0.015	0.532 0.029	0.466 0.040	0.530 0.027
1984	0.549 0.025	0.549 0.033	0.547 0.037	0.578 0.036	0.471 0.048	0.613 0.028	0.617 0.012	0.583 0.021	0.586 0.020	0.535 0.027	0.452 0.036	0.465 0.033

Tabla 3.X



Figuras 3.31, 3.32 y 3.33

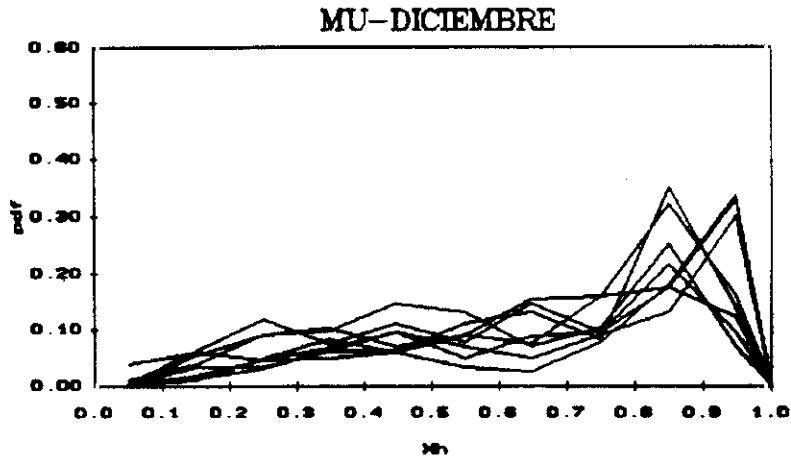


Figura 3.35

Las formas de las funciones de estas gráficas son, para la mayoría de los meses, solamente de dos tipos: con uno o dos picos (bimodal). Sólo hay algunos meses para los que se presenta siempre alguna curva del tipo de varios valores con probabilidad de ocurrencia similar. Las características de estas curvas para todos los meses son similares a las descritas para la localidad de Málaga.

#### 3.4.7 Descripción de las series del índice $X_n$ de Oviedo

En la Tabla 3.XI aparecen los valores medios mensuales y varianzas del índice  $X_n$  para la localidad de Oviedo.

En las gráficas 3.35, 3.36, 3.37 y 3.38 se muestran las funciones de distribución de probabilidad de ocurrencia para algunos meses de los años 1977/84.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1977	0.385 0.049	0.425 0.042	0.441 0.041	0.436 0.051	0.350 0.057	0.383 0.053	0.358 0.053	0.379 0.050	0.441 0.043	0.403 0.042	0.422 0.041	0.381 0.043
1978	0.325 0.038	0.393 0.045	0.466 0.048	0.374 0.061	0.362 0.052	0.376 0.057	0.424 0.051	0.451 0.049	0.513 0.055	0.490 0.050	0.394 0.037	0.386 0.042
1979	0.311 0.048	0.376 0.060	0.388 0.043	0.441 0.041	0.405 0.053	0.407 0.057	0.406 0.062	0.356 0.043	0.335 0.037	0.350 0.038	0.377 0.042	0.333 0.035
1980	0.352 0.047	0.366 0.044	0.352 0.045	0.435 0.066	0.318 0.038	0.421 0.046	0.434 0.050	0.440 0.040	0.390 0.044	0.360 0.030	0.346 0.036	0.390 0.036
1981	0.374 0.036	0.438 0.042	0.425 0.047	0.405 0.050	0.440 0.042	0.461 0.055	0.438 0.054	0.378 0.062	0.477 0.034	0.336 0.032	0.474 0.022	0.307 0.032
1982	0.378 0.026	0.389 0.039	0.453 0.041	0.527 0.035	0.365 0.049	0.384 0.045	0.363 0.042	0.375 0.053	0.436 0.041	0.369 0.035	0.386 0.031	0.270 0.027
1983	0.437 0.021	0.364 0.032	0.392 0.036	0.418 0.041	0.415 0.039	0.449 0.050	0.355 0.045	0.337 0.042	0.507 0.037	0.389 0.045	0.349 0.031	0.433 0.027
1984	0.324 0.039	0.395 0.050	0.431 0.045	0.479 0.045	0.265 0.032	0.410 0.051	0.458 0.045	0.454 0.038	0.453 0.038	0.432 0.044	0.350 0.035	0.385 0.032

Tabla 3.XI

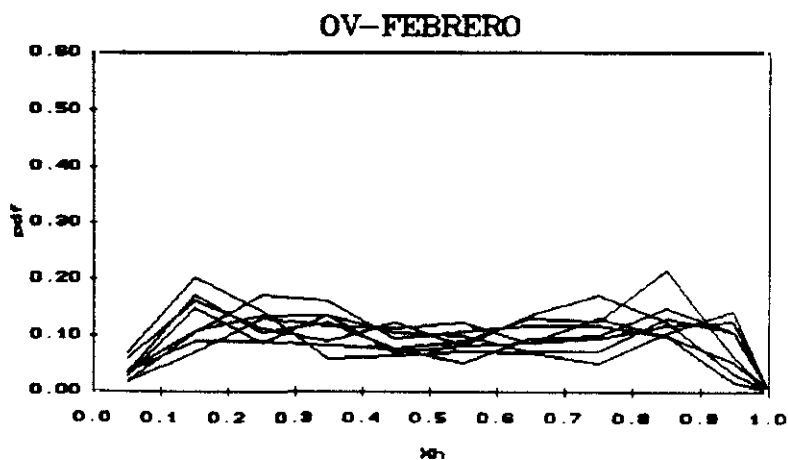
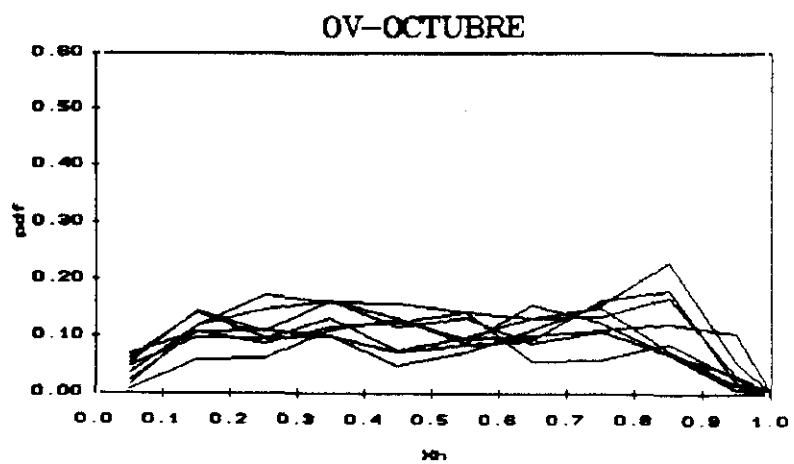
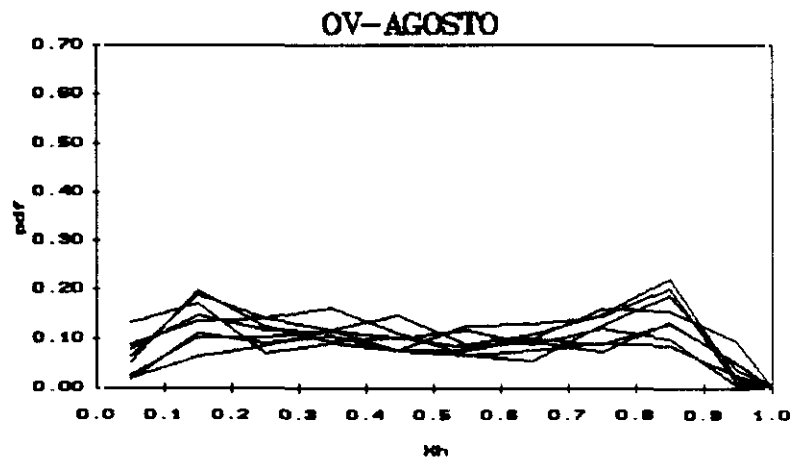
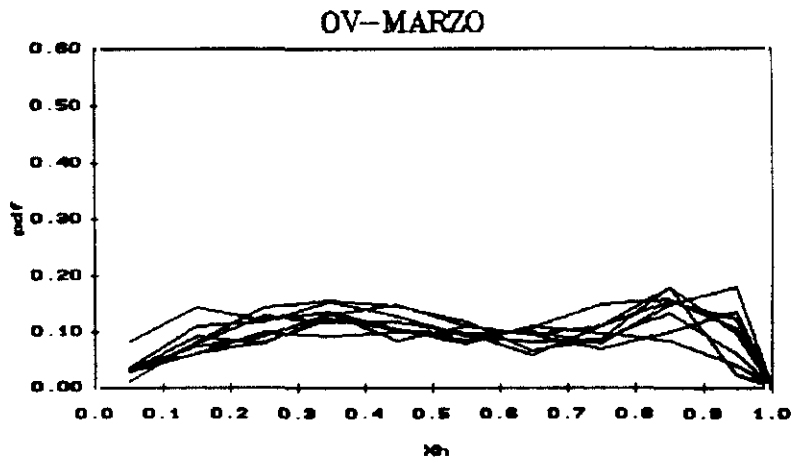


Figura 3.35



Figuras 3.36, 3.37 y 3.38



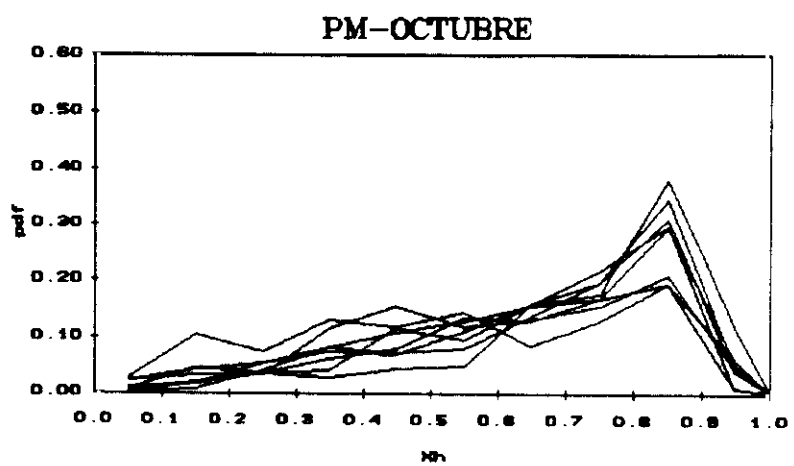
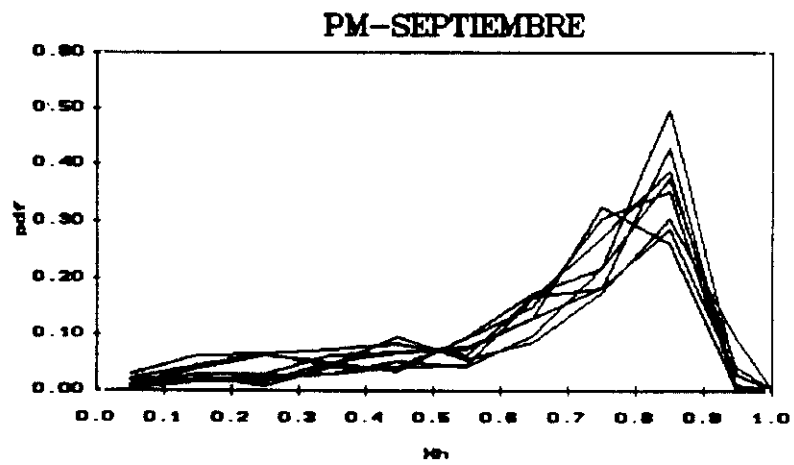
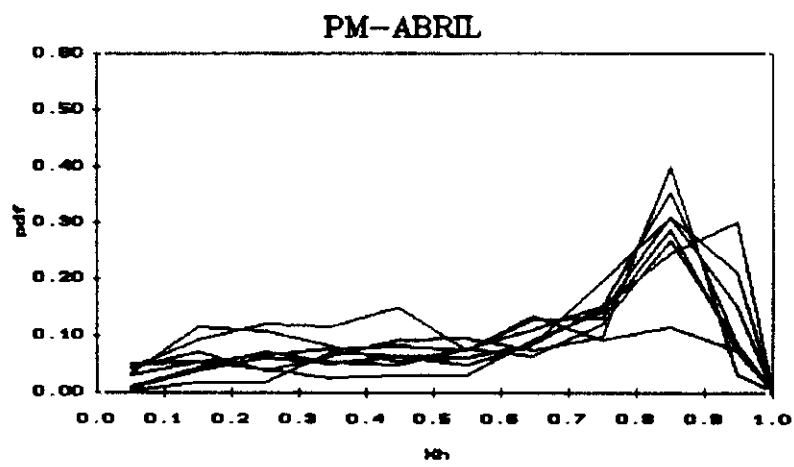
En la mayoría de los meses y años nos encontramos con que hay varios valores con probabilidad de ocurrencia similar, y, naturalmente, no muy elevada. Casi todas las curvas son de este tipo, con un valor un poco más probable para algún año/mes pero nunca con probabilidad mayor que el 30%. Sólo para alguna serie mensual aparece una curva con un pico.

### 3.4.8 Descripción de las series del índice $X_n$ de Palma de Mallorca

En la Tabla 3.XII aparecen los valores medios mensuales del índice  $X_n$  y sus varianzas para la localidad de Mallorca. En las gráficas 3.39, 3.40, 3.41 y 3.42 se muestran los valores de las curvas de distribución de probabilidades de ocurrencia de los valores del índice  $X_n$  para las distintas series mensuales.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1977	0.438 0.033	0.479 0.026	0.537 0.029	0.502 0.044	0.488 0.038	0.561 0.030	0.569 0.021	0.561 0.028	0.545 0.027	0.465 0.031	0.473 0.037	0.457 0.038
1978	0.416 0.037	0.463 0.036	0.508 0.029	0.499 0.042	0.509 0.033	0.549 0.030	0.605 0.012	0.544 0.021	0.561 0.017	0.501 0.026	0.505 0.030	0.438 0.028
1979	0.372 0.037	0.441 0.038	0.499 0.044	0.537 0.036	0.565 0.022	0.567 0.018	0.560 0.018	0.545 0.015	0.518 0.028	0.406 0.037	0.493 0.033	0.465 0.029
1980	0.434 0.037	0.427 0.043	0.555 0.035	0.464 0.052	0.543 0.035	0.619 0.012	0.000 0.000	0.588 0.010	0.560 0.022	0.518 0.024	0.485 0.026	0.435 0.045
1981	0.437 0.038	0.477 0.048	0.519 0.029	0.411 0.044	0.573 0.037	0.603 0.029	0.620 0.025	0.589 0.019	0.571 0.018	0.569 0.022	0.535 0.023	0.462 0.026
1982	0.497 0.024	0.431 0.039	0.513 0.045	0.557 0.036	0.549 0.041	0.611 0.015	0.578 0.016	0.566 0.017	0.502 0.034	0.469 0.033	0.438 0.034	0.480 0.029
1983	0.582 0.015	0.462 0.035	0.548 0.034	0.597 0.028	0.623 0.026	0.585 0.025	0.619 0.012	0.577 0.022	0.590 0.021	0.534 0.025	0.447 0.033	0.473 0.031
1984	0.487 0.026	0.468 0.036	0.518 0.047	0.565 0.047	0.475 0.048	0.619 0.023	0.000 0.000	0.589 0.024	0.517 0.043	0.504 0.030	0.451 0.034	0.449 0.031

Tabla 3.XII



Figuras 3.39, 3.40 y 3.41

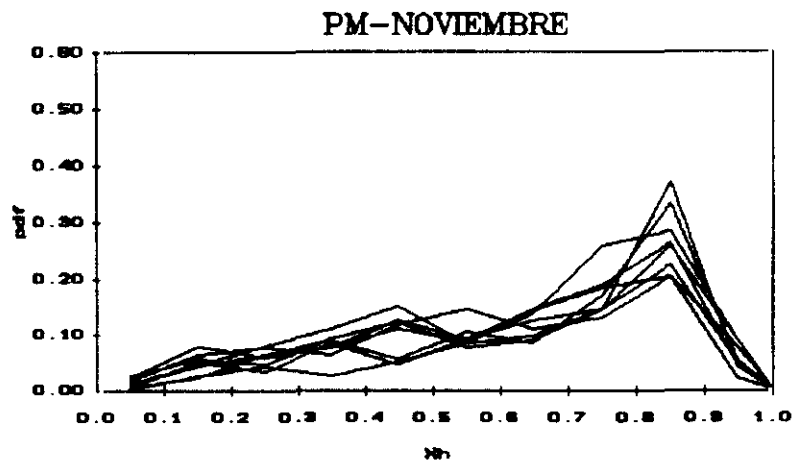


Figura 3.42

La forma de las funciones para esta localidad es similar a la observada para las localidades de Murcia y Málaga. Para la mayoría de los meses aparecen curvas con uno o dos valores con probabilidad de ocurrencia más elevada que la de los demás. Los picos de estas funciones son más elevados para los meses de verano que para los de invierno.

#### 3.4.9 Descripción de las series del índice $X_n$ de Sevilla

En la tabla 3.XIII se muestran los valores medios mensuales y varianzas del índice  $X_n$  para Sevilla.

En las gráficas 3.43, 3.44, 3.45 y 3.46 hemos representado la probabilidad de ocurrencia (en tanto por ciento) de los distintos valores del índice  $X_n$  para cada mes de cada año con datos.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1977	0.386 0.039	0.391 0.046	0.594 0.018	0.579 0.027	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000
1978	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.586 0.037	0.422 0.046
1979	0.463 0.050	0.454 0.069	0.558 0.048	0.622 0.040	0.634 0.031	0.646 0.021	0.600 0.025	0.674 0.014	0.645 0.019	0.464 0.052	0.621 0.035	0.522 0.037
1980	0.530 0.032	0.542 0.044	0.555 0.043	0.593 0.042	0.589 0.035	0.673 0.016	0.699 0.006	0.608 0.027	0.590 0.022	0.572 0.031	0.576 0.041	0.624 0.023
1981	0.639 0.017	0.603 0.039	0.551 0.039	0.510 0.050	0.662 0.026	0.666 0.016	0.687 0.011	0.622 0.016	0.622 0.018	0.585 0.030	0.611 0.018	0.459 0.052
1982	0.603 0.058	0.588 0.041	0.632 0.033	0.546 0.041	0.625 0.022	0.655 0.016	0.650 0.017	0.600 0.031	0.621 0.021	0.639 0.030	0.589 0.051	0.616 0.044
1983	0.620 0.017	0.565 0.039	0.639 0.032	0.595 0.037	0.625 0.035	0.656 0.021	0.718 0.015	0.660 0.031	0.616 0.015	0.551 0.024	0.411 0.042	0.521 0.028
1984	0.484 0.025	0.555 0.027	0.527 0.031	0.521 0.040	0.512 0.031	0.582 0.030	0.613 0.009	0.595 0.009	0.595 0.014	0.582 0.025	0.421 0.039	0.496 0.029

Tabla 3.XIII

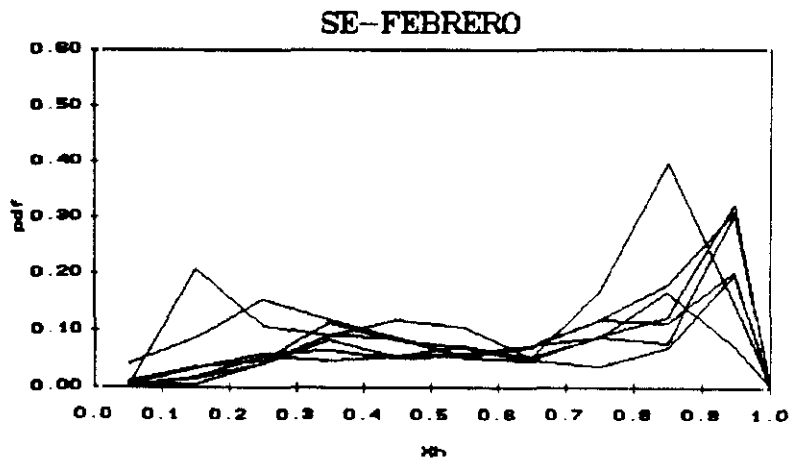
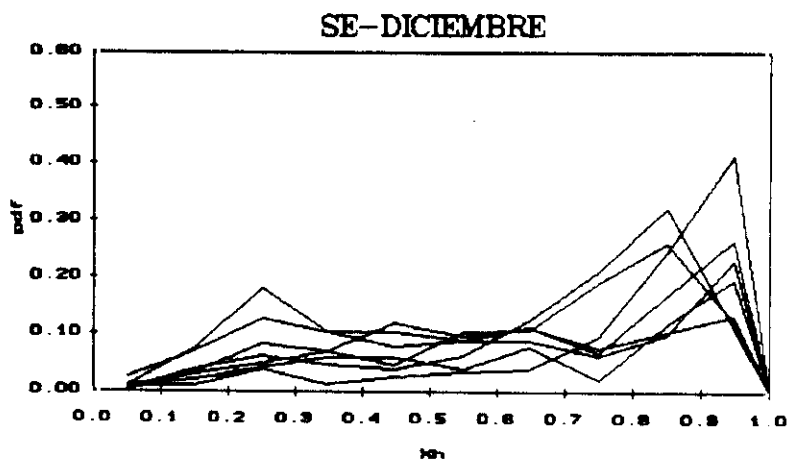
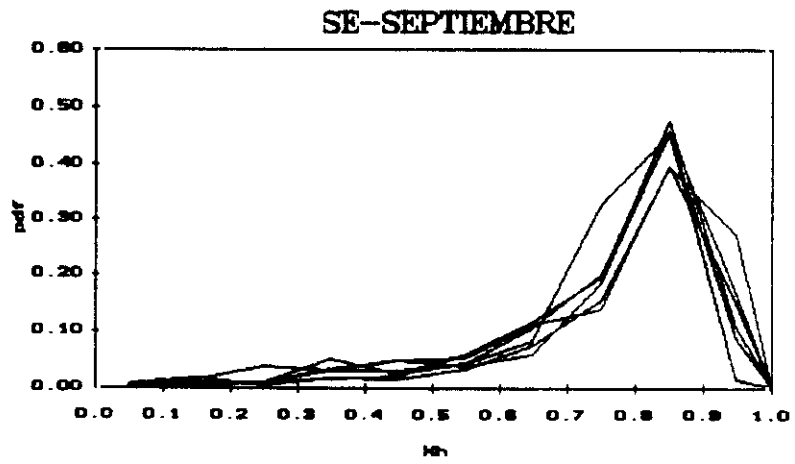
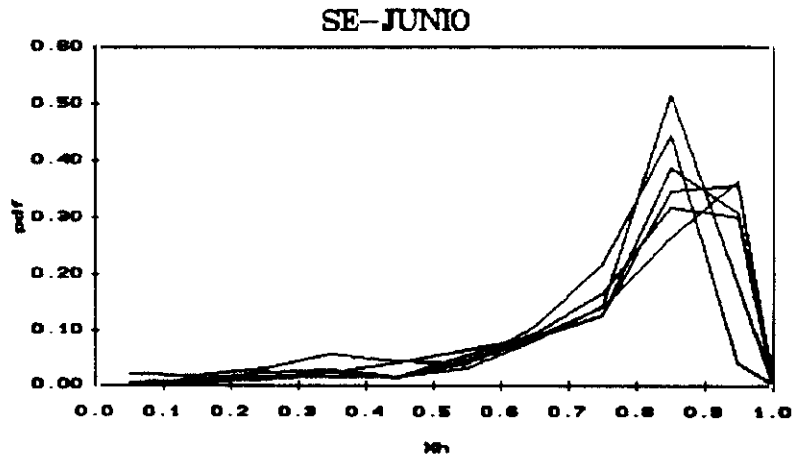


Figura 3.43



Figuras 3.44, 3.45 y 3.46

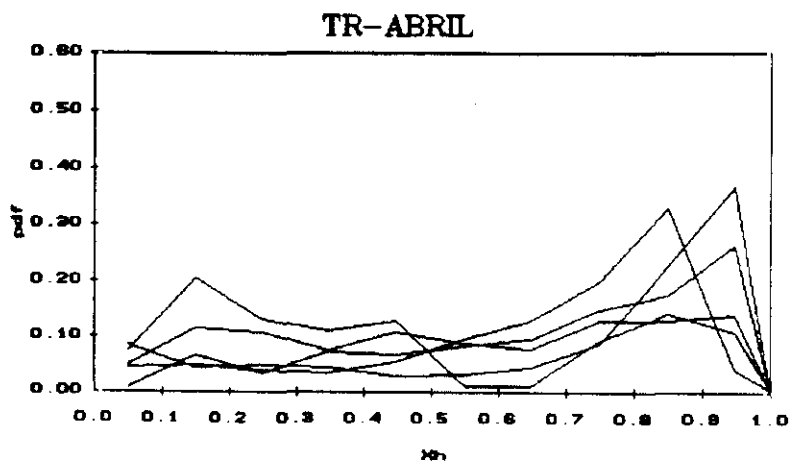
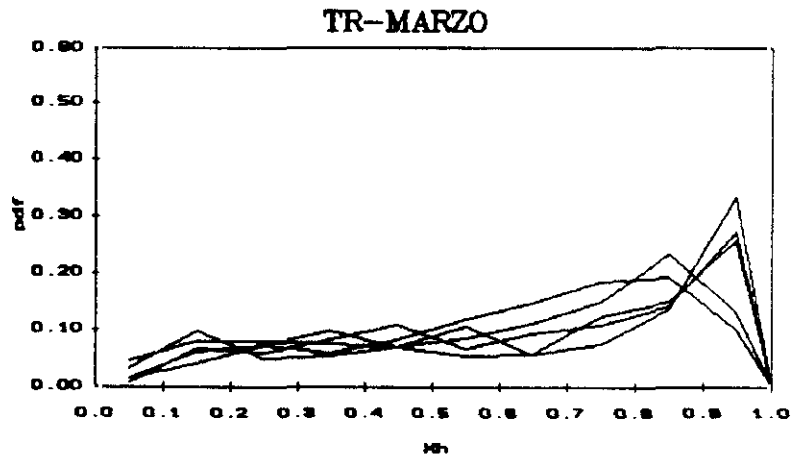
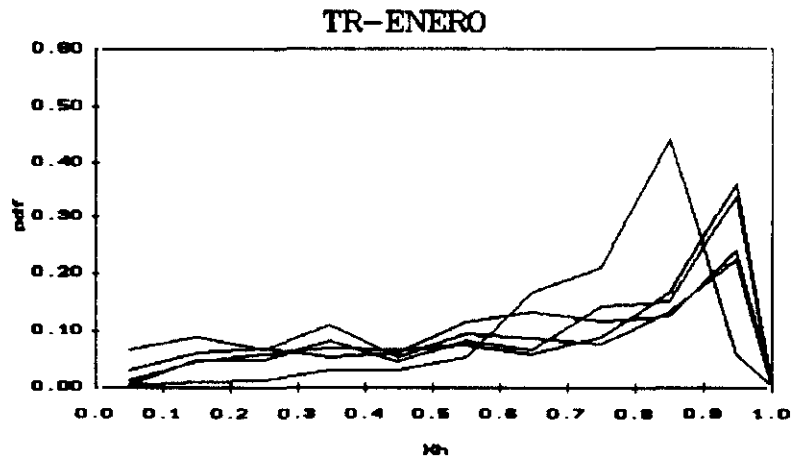
Para los meses de invierno nos encontramos con varias formas de distribución de los valores del índice  $X_b$ . Aunque hay muchos meses que presentan una curva con un único pico, también nos encontramos, con más frecuencia que para las localidades anteriores, funciones con varios valores de probabilidad similar. Para los meses de verano nos encontramos, **sobre todo**, con funciones con uno o dos picos. El valor de probabilidad de **ocurrencia** de estos valores no es, sin embargo, muy elevado (30-40%).

#### 3.4.10 Descripción de las series del índice $X_b$ de Tortosa

En la Tabla 3.XIV se muestran los valores medios mensuales y varianzas de las series del índice  $X_b$ , para la localidad de Tortosa. En las figuras 3.47, 3.48, 3.49 y 3.50 se muestran las funciones de distribución de probabilidades de ocurrencia de valores del índice  $X_b$  para algunos meses.

A/M	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
1980	0.485 0.041	0.450 0.056	0.518 0.047	0.564 0.062	0.479 0.050	0.549 0.044	0.599 0.032	0.541 0.031	0.526 0.028	0.512 0.038	0.476 0.044	0.554 0.028
1981	0.531 0.040	0.483 0.052	0.507 0.048	0.439 0.055	0.522 0.054	0.586 0.029	0.616 0.024	0.511 0.033	0.501 0.035	0.542 0.031	0.516 0.033	0.440 0.049
1982	0.443 0.053	0.428 0.048	0.494 0.034	0.522 0.037	0.517 0.037	0.553 0.022	0.547 0.026	0.517 0.030	0.553 0.025	0.486 0.038	0.479 0.035	0.486 0.027
1983	0.558 0.013	0.469 0.035	0.485 0.046	0.546 0.043	0.546 0.041	0.559 0.041	0.565 0.026	0.501 0.040	0.605 0.023	0.547 0.031	0.398 0.046	0.525 0.039
1984	0.532 0.042	0.492 0.055	0.507 0.060	0.366 0.061	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000

Tabla 3.XIV



Figuras 3.47, 3.48 y 3.49

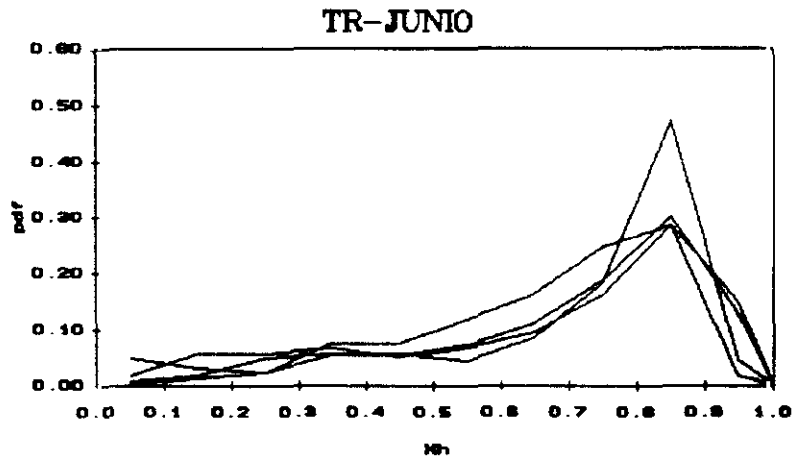


Figura 3.50

Como se puede observar, para la mayoría de los valores de  $X_n$  (utilizando diez intervalos) aparecen probabilidades de ocurrencia en torno al 10%, excepto para el valor 0.8-0.9 que es algo mayor. En general, tienen uno o dos picos pero con poca altura. El tipo de función con muchos valores de  $X_n$  con probabilidad similar (y, por tanto, ningún pico) sólo aparece en los meses de Abril y Noviembre de algunos años.



### 3.5 CONCLUSIONES

En este capítulo hemos propuesto una forma de eliminar las tendencias observadas en las series de exposición horaria de radiación global, que se basa en la utilización de los valores máximos de esta variable. Con ello conseguimos que en los valores de las nuevas series se refleje, principalmente el efecto de las nubes y desaparezcan las tendencias (de tipo determinista) observadas en las series de exposición horaria.

Una aportación muy importante del capítulo es que se propone una expresión general (válida para todas las localidades analizadas) para calcular los valores máximos de exposición horaria de radiación global. Ésta sólo depende del valor, para cada hora, de la altura solar.

Utilizando esta expresión hemos definido un índice en el que ya no aparecen las tendencias estacionales observadas en las series originales. Para las series de este índice (que hemos llamado  $X_h$ ) hemos analizado, de forma descriptiva, sus funciones de distribución de probabilidades de ocurrencia y valores medios.

Las conclusiones más importantes que podemos extraer de este estudio son las siguientes:

- Las funciones de distribución de probabilidad de ocurrencia de los valores del índice  $X_h$  parecen guardar relación con la época del año. De hecho nos encontramos, básicamente, sólo con tres formas de curvas (en algunas localidades sólo dos). Una primera forma con un pico, esto es, con un valor que se da muchas más veces que los demás; otra forma típica con dos picos, uno mayor y otro más pequeño, esto es, de tipo bimodal; y

---

una tercera forma de curva en la que aparecen varios valores con probabilidad similar. Estos resultados están de acuerdo con los que presentan Aguiar et al., /2/, para series del índice de transparencia atmosférico horario, aunque estos autores analizan por separado las series de cada hora.

- El comportamiento de las series de Oviedo es **sensiblemente diferente** al de las series de otras localidades. La forma de las funciones de distribución parece ser independiente del mes que consideremos; en general, presentan varios valores con probabilidad de ocurrencia similar (curvas sin ningún pico).

Respecto a los valores medios de las series es necesario hacer un análisis formal para poder extraer conclusiones. Aunque parece haber en la mayoría de las localidades una relación entre mes del año y media de la serie  $\{X_b\}$ , es necesario emplear un análisis de regresión que nos permita ver cuál es el grado de variabilidad de la media que realmente puede explicarse con el mes del año. Este análisis lo haremos en el Capítulo 6, que dedicaremos a estudiar cuáles son los parámetros mínimos necesarios para caracterizar las series del índice definido.

## **4. UTILIZACIÓN DE LA TEORÍA DE PROCESOS ESTOCÁSTICOS**

La teoría de procesos estocásticos estacionarios nos proporciona un esquema general para estudiar una sucesión infinita de variables aleatorias correlacionadas entre sí, en función de su proximidad en el tiempo. El método de trabajo que normalmente se sigue se muestra en la figura 4.1, tomada de Peña, /51/.

En esta figura se muestran las distintas etapas en la construcción de un modelo: se trata, primero, de identificar el modelo al que responden las series, después estimar sus parámetros, validarlo y por último utilizarlo en la simulación de nuevas series. Una condición necesaria para que pueda utilizarse la teoría de procesos estocásticos estacionarios a una serie de datos, es que ésta sea estacionaria tanto en la media como en la varianza.

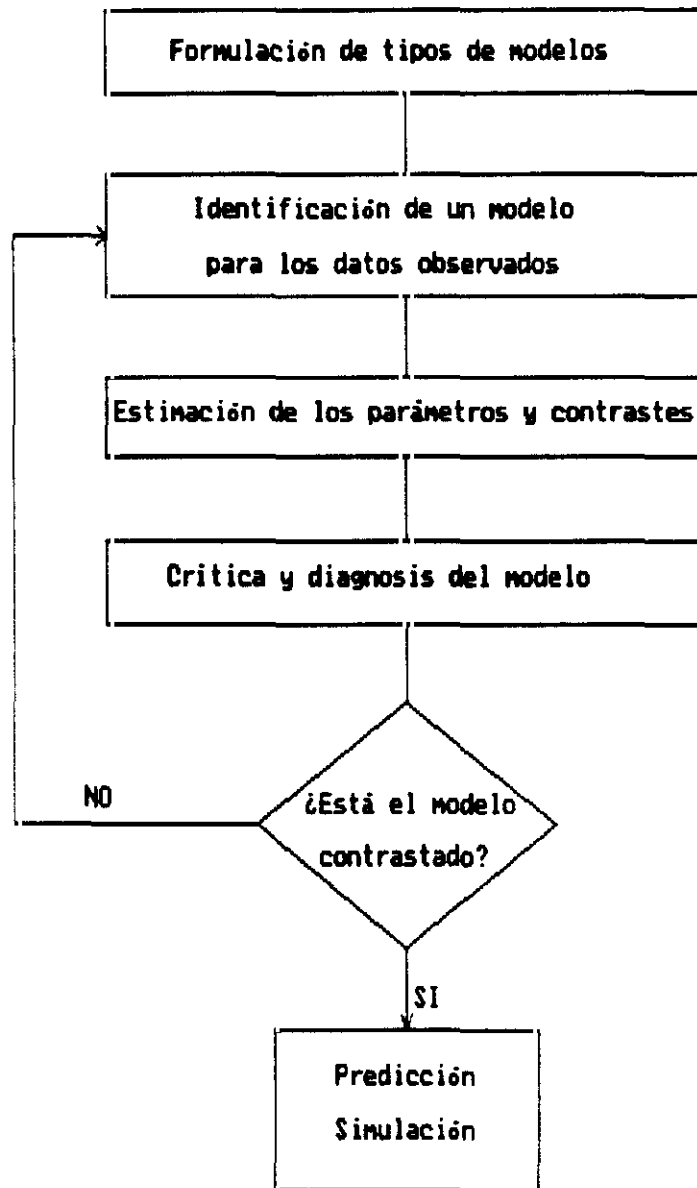


Figura 4.1 Etapas en la construcción de un modelo de series temporales

Las series de exposición global horaria de la radiación solar, sin embargo, ya hemos visto que no son estacionarias. Por ello, en el siguiente apartado

proponemos una serie de modificaciones de la serie original para conseguir una serie estacionaria.

#### 4.1 DESCRIPCIÓN DE LAS SERIES UTILIZADAS

La función matemática que vamos a utilizar para caracterizar la componente estacional de las series de exposición horaria de la radiación global es la expresión de exposición horaria máxima de radiación global. En esta expresión, la variable tiempo viene recogida en el ángulo de la altura solar. Así la función  $f(h)$  será:

$$f(h) = G_{h,\max} = A(\text{sen}\gamma)^B \quad (4.1)$$

Vamos a utilizar esta función porque, como ya vimos en un apartado anterior, la componente estacional de las series de exposición horaria es debida a la variación en la radiación que llega a la parte exterior de la atmósfera y a la variación de la masa atmosférica que atraviesa la radiación. Al utilizar la radiación máxima que podría llegar a la superficie de la tierra para cada hora eliminamos estos dos efectos; en las nuevas series lo que se recogerá será la variación en los distintos componentes atmosféricos (sobre todo, presencia de nubes). Vamos a evaluar en este apartado si la componente  $X_h$ , dada por la expresión:

$$X_h = \frac{G_h}{G_{h,\max}} \quad (4.2)$$

es estacionaria. Para definir la estacionariedad de un proceso se puede utilizar bien las funciones de distribución o alternativamente los momentos (son los conceptos conocidos como estacionariedad en sentido estricto y en sentido amplio, respectivamente). En este estudio trabajaremos siempre con este último concepto, menos restrictivo, y a él nos referiremos simplemente como "estacionariedad".

Se dice que un proceso es estacionario si se verifican las tres condiciones siguientes:

- (1ª) La media es finita y constante a lo largo del tiempo, es decir:

$$E(Y_t) = \mu \quad \forall t \quad (4.3)$$

*(Estacionario en media)*

- (2ª) La varianza es finita y permanece constante a lo largo del tiempo, es decir:

$$E(Y_t - \mu)^2 = \sigma^2 < \infty \quad \forall t \quad (4.4)$$

*(Estacionario en varianza)*

- (3ª) La autocovarianza entre dos períodos distintos de tiempo únicamente viene afectada por el lapso de tiempo transcurrido entre esos dos periodos. Así:

$$E[(Y_{t+k} - \mu)(Y_t - \mu)] = \gamma_k \quad \forall t, \forall k \in \mathbb{N} \quad (4.5)$$

que sería una covarianza de orden  $k$ , por ser éste el lapso de tiempo que separa  $Y_t$  de  $Y_{t+k}$ . Su valor,  $\gamma_k$ , es independiente del período  $t$  que se considere.

Además de la estacionariedad, es necesario que el proceso estocástico goce de la propiedad de ergodicidad, con objeto de que el proceso de inferencia pueda realizarse de una forma adecuada. Comentaremos, sólo de forma intuitiva, el concepto de ergodicidad. Cuando valores de la serie temporal alejados en el tiempo están muy correlacionados sucederá que al aumentar el tamaño de la muestra se añade poca información nueva, en comparación con el número de parámetros que se deberán estimar por haber aumentado el tamaño de la muestra. La consecuencia de este hecho en el plano estadístico será que los estimadores obtenidos no son consistentes, ya que el aumento del tamaño de la muestra no tendrá una especial utilidad, puesto que se tendrá que calcular un mayor número

---

de autocovarianzas para caracterizar adecuadamente el proceso. Es necesario, pues, exigir que la correlación entre observaciones vaya disminuyendo conforme vayan distanciándose éstas.

Cuando se verifica la propiedad de ergodicidad, sí se pueden obtener estimadores consistentes. Una condición necesaria, aunque no suficiente, para la ergodicidad es que las autocorrelaciones tiendan a cero cuando los intervalos de tiempo o desfases considerados tiendan a infinito. Como veremos en el apartado siguiente, en el que se calculan los correlogramas simple y parcial de las series, los valores de autocorrelación tienden a cero cuando el desfase tiende a cero, por lo que se cumple la condición anterior.

De una manera gráfica vamos a comprobar si se cumplen o no las propiedades de estacionariedad en las series del índice  $X_h$  propuesto.

Estas series de índices  $X_h$  serán estacionarias si son estables la media, varianza y covarianzas. Para comprobar la estacionariedad en media y varianza utilizaremos las gráficas de estas series. Para poder seleccionar, de forma precisa, periodos temporales en los que la varianza y la media permanezcan constantes vamos, primero, a analizar las series que se obtienen en función de que la serie de datos corresponda a un mes, un cuatrimestre, un año o un "mes global".

De forma genérica utilizaremos la expresión "mes global" para referirnos a series de datos para cada mes en las que se incluyen todos los datos disponibles para ese mes, para todos los años, o para varios años. Este tipo de series es ya utilizada por Gordon y Reddy, /28/, para series de valores diarios del índice de transparencia atmosférico.

Intuitivamente se puede justificar la utilización de estas series debido a que, en principio, en ellas se incluirá toda la "información estadística" contenida

en distintos años. Nos parece más adecuado el uso de estas series que el uso de series de valores medios puesto que en estas últimas, naturalmente, se pierde parte de la información de las series originales.

Si consideramos periodos temporales de un año, tendremos series de más de tres mil datos. Para poder ver si existe evolución en la media o en la varianza, vamos a representar únicamente los valores para una hora del día y observar si hay en éstos una variación a lo largo del año. En la figura 4.2 se muestran los valores de los índices  $X_h$ , calculados para los datos de exposición global horaria de Málaga para las 12 horas a lo largo del año 1977.

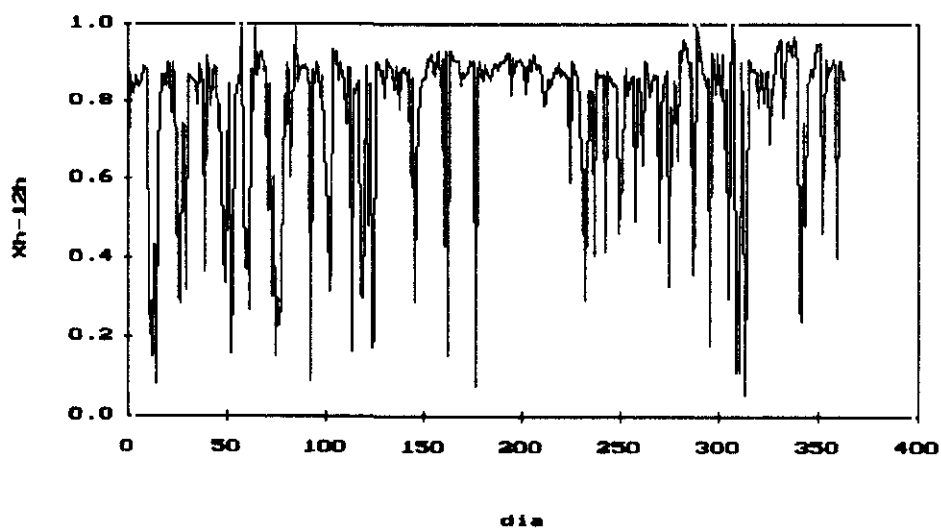


Figura 4.2. Índice  $X_h$  a las 12 horas, Málaga-1977.

Las series para las demás horas del día presentan características similares.

Como se puede observar en esta figura persiste una pequeña evolutividad en la media, menor que la que presentaban las series de radiación; lo que se observa claramente es una evolutividad en la varianza, es decir no podemos considerar que se trate de series estacionarias.



Con el objeto de evitar este hecho vamos a considerar periodos temporales más cortos.

Una posibilidad es la de dividir el año en tres periodos: es decir, utilizar series cuatrimestrales. El primer cuatrimestre lo formarían los meses de Noviembre, Diciembre, Enero y Febrero; el segundo los meses de Marzo, Abril, Mayo y Junio; y el tercero los meses de Julio, Agosto, Septiembre y Octubre.

Para analizar estas series vamos también a utilizar las gráficas de las mismas. En las figuras 4.3 y 4.4 se muestran los valores de estas series cuatrimestrales para dos cuatrimestres de la localidad de Palma de Mallorca.

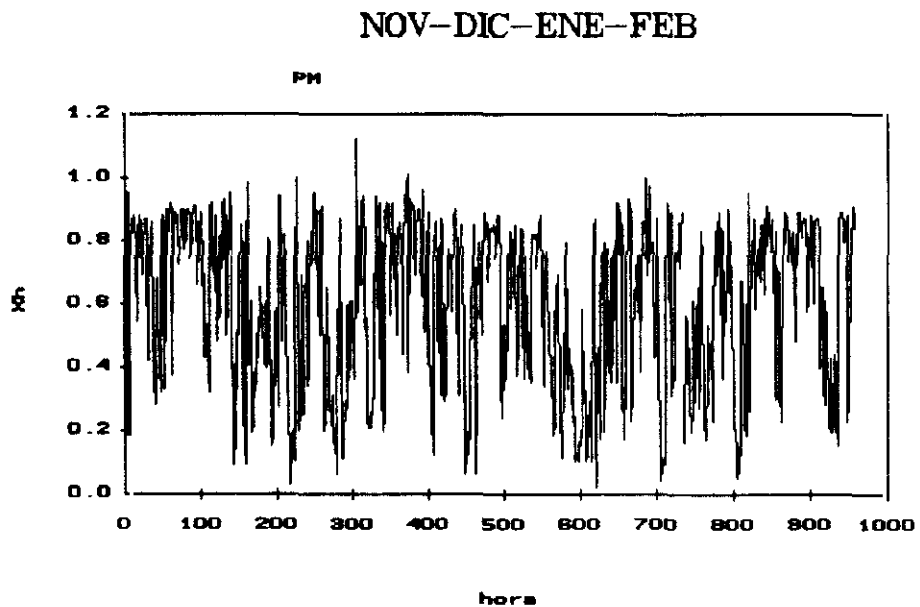


Figura 4.3 Valores del índice  $X_t$  para los meses de Noviembre, Diciembre y Enero y Febrero de 1977/78 para la localidad de Palma de Mallorca.

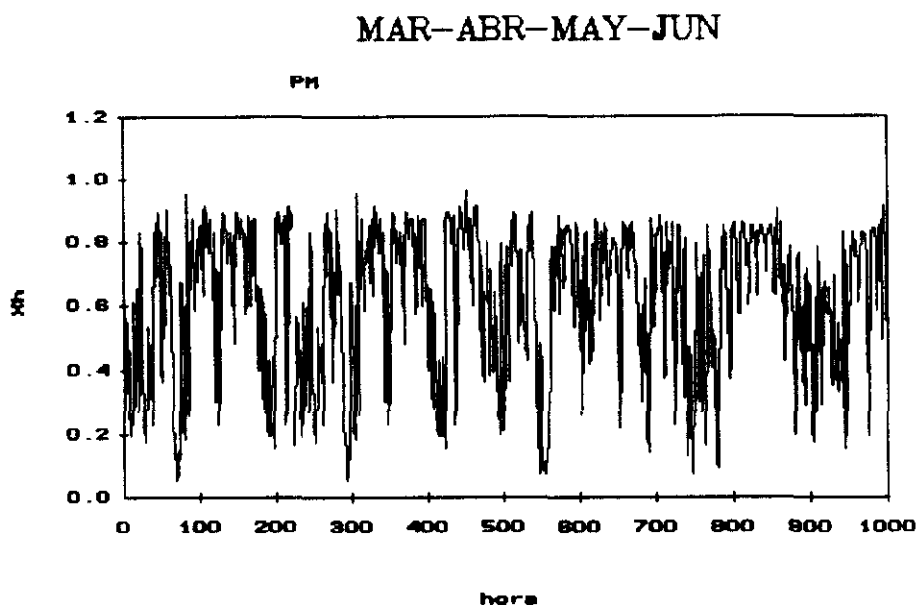


Figura 4.4 Valores del índice  $X_h$ , para los meses de Marzo, Abril, Mayo y Junio de 1977/78, para la localidad de Palma de Mallorca.

En el eje de ordenadas se han representado las horas de cada uno de los meses que compone el cuatrimestre. A efectos de numeración de las horas se han concatenado los distintos meses.

En estas gráficas observamos que, si bien hay cuatrimestres en los que pueden considerarse constantes la media y la varianza, hay otros en los que esto no es así. En general se trata, por tanto, de series no estacionarias.

Estas series cuatrimestrales, por tanto, tampoco pueden considerarse estacionarias, debido a la evolutividad en la media y/o varianza que presentan. Desde el punto de vista físico, es razonable que existan estas evolutividades, pues el clima, que será el que condicione los valores que alcanzan nuestras series, no permanece constante durante periodos tan grandes de tiempo.

Vamos ahora a analizar las series de "meses globales". En las figuras 4.5, 4.6, 4.7, 4.8, 4.9 y 4.10 se han representado varios de estos meses, para las localidades de Oviedo y Palma de Mallorca.

Como se ha explicado, la forma de "construirlos" ha sido ir concatenando cada mes de un año con el mismo mes del año siguiente y así sucesivamente hasta formar la serie. Aunque en estas gráficas sólo se muestra parte de las series (la mayoría tiene más de tres mil datos), para nuestro estudio es suficiente.

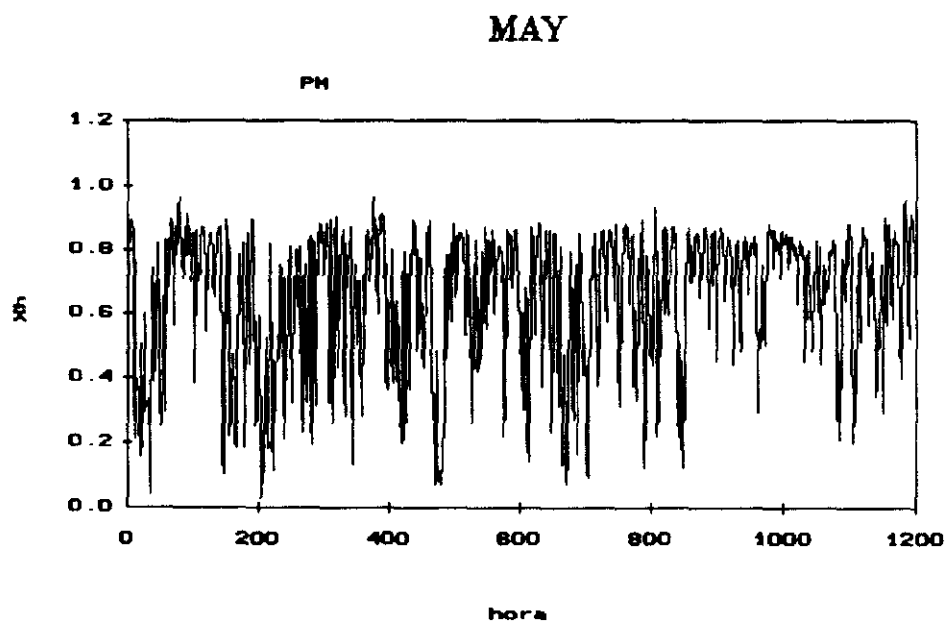


Figura 4.5 Serie de un "mes global", para el mes de Mayo, de la localidad de Palma de Mallorca.

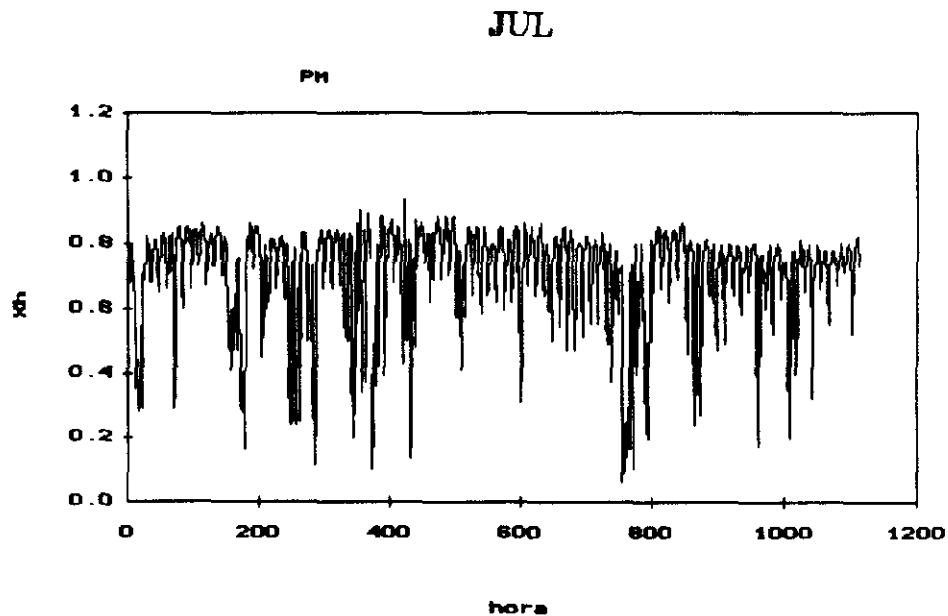


Figura 4.6 Serie de un "mes global", para el mes de Julio, de la localidad de Palma de Mallorca.

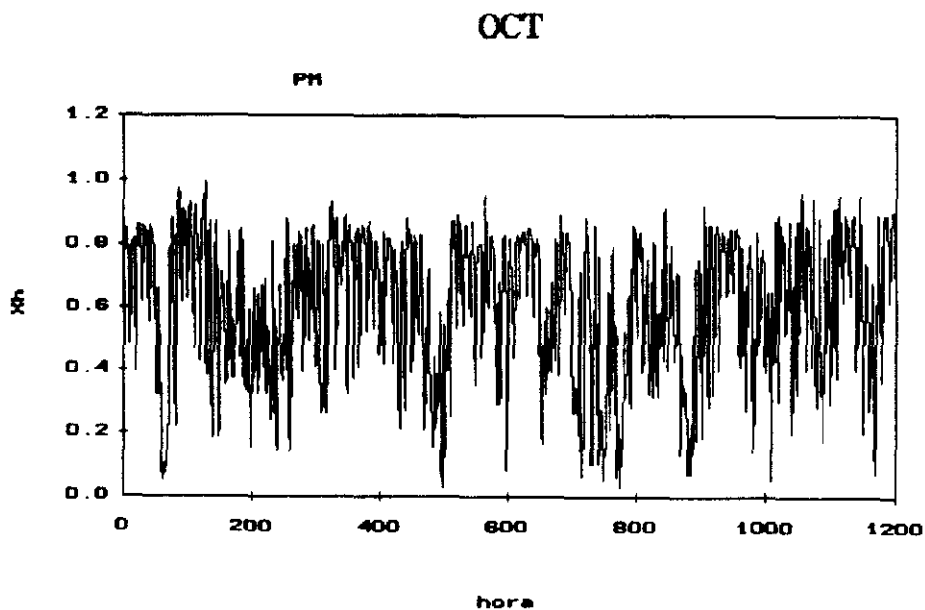


Figura 4.7 Serie de un "mes global", para el mes de Octubre, de la localidad de Palma de Mallorca.

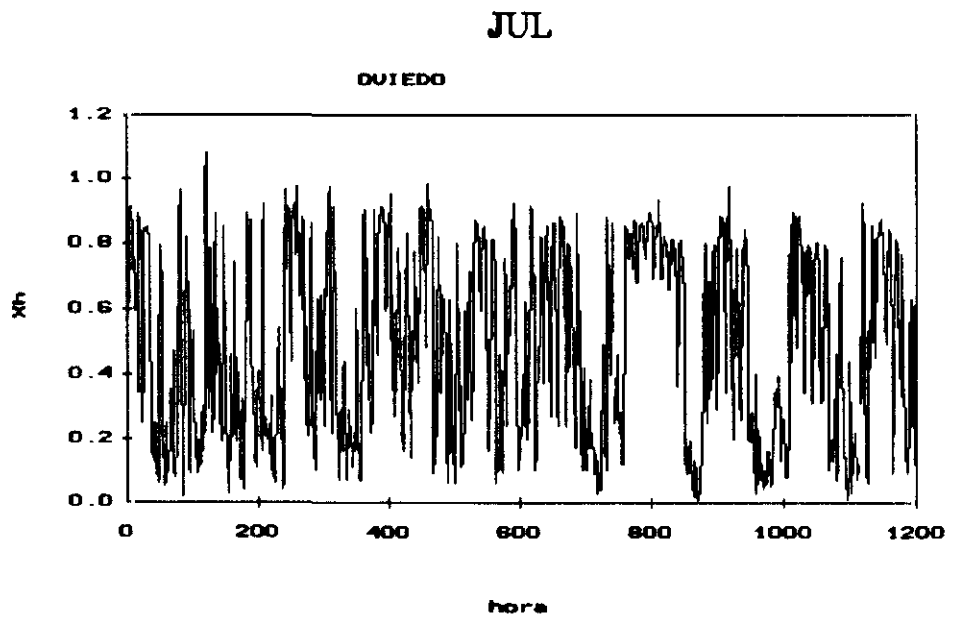


Figura 4.8 Serie de un "mes global", para el mes de Julio, para la localidad de Oviedo.

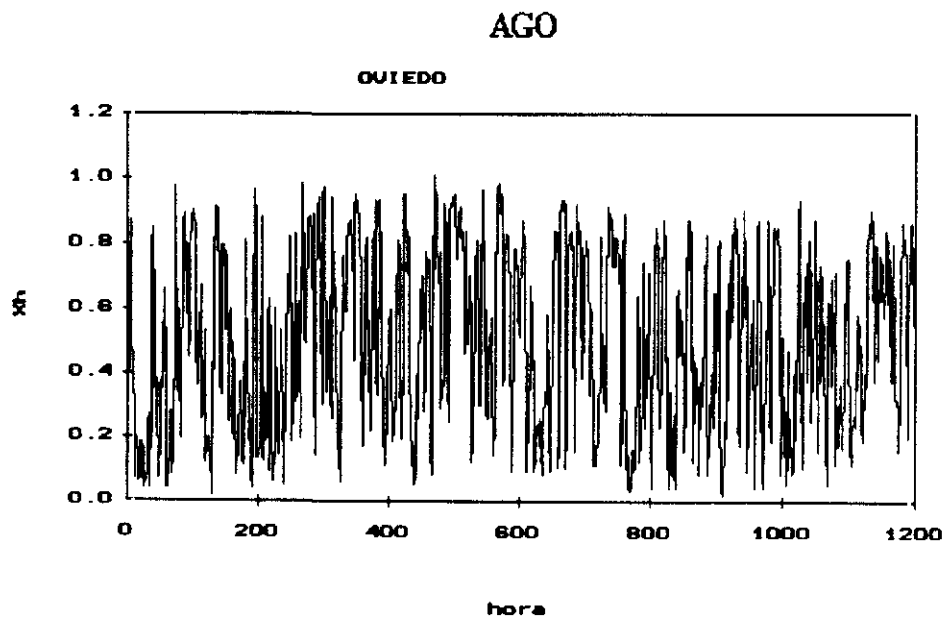


Figura 4.9 Serie de un "mes global", para el mes de Agosto, de la localidad de Oviedo.

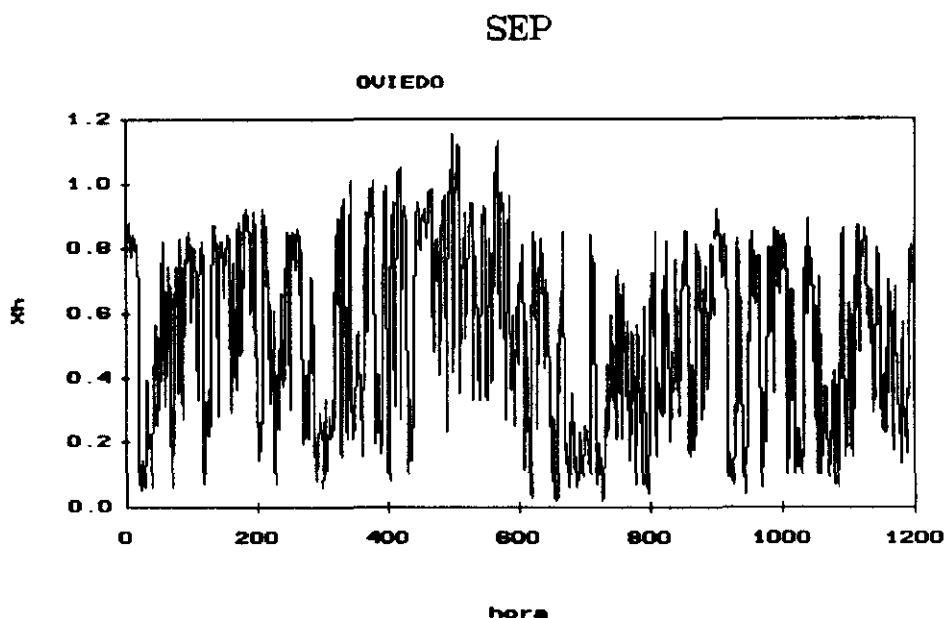


Figura 4.10 Serie de un "mes global", para el mes de Septiembre, de la localidad de Oviedo.

Como se puede observar en estas gráficas, para algunos meses tipo no se pueden considerar constantes la varianza y/o la media. Por ejemplo, en el gráfico correspondiente a Julio (Palma de Mallorca) se observa que el periodo [200,400] tiene media inferior y varianza superior a la que tiene el periodo [400,600].

Desde el punto de vista del fenómeno que representan las series, este hecho es totalmente explicable. No olvidemos que estamos considerando "iguales" cada uno de los meses del año, a lo largo de varios años. Es decir suponemos, por ejemplo, que el mes de Enero de todos los años va a ser igual siempre. Naturalmente, en la realidad, esto no es así.

Si bien podemos encontrar que, en general, los fenómenos climáticos son cíclicos, existen periodos de tiempo bastante diferentes de lo esperable o normal. Este hecho ya lo observamos cuando analizamos las funciones de distribución de

probabilidades de ocurrencia de la serie de índice  $X_b$  en el capítulo anterior. Para eliminar la no estacionariedad observada en la media y/o varianza podríamos modificar las series mensuales y trabajar con series normalizadas. Esto supondría, sin embargo, que necesitaríamos conocer los valores medios y varianzas de las series. Estos datos no están normalmente disponibles. Es decir, estaríamos incrementando el número de parámetros necesarios para caracterizar las series de exposición horaria de radiación global.

Para evitar los inconvenientes de los tres periodos de tiempo estudiados, nos parece conveniente analizar qué sucede cuando consideramos series mensuales.

En las figuras 4.11 y 4.12 hemos representado dos de estas series para el mes de Julio de 1978 de la localidad de Palma de Mallorca y Julio de 1977 de Murcia.

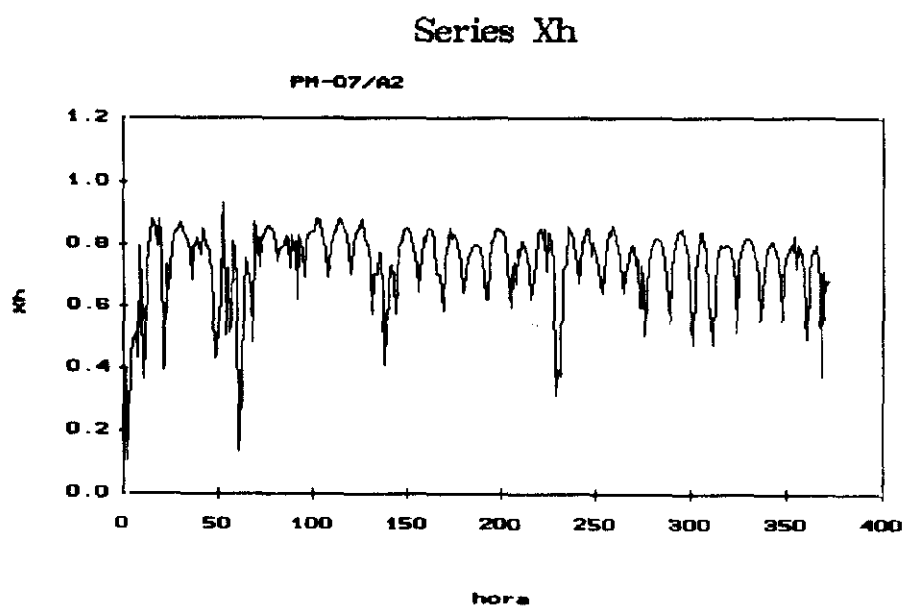


Figura 4.11 Serie mensual del índice  $X_b$ , para el mes de Julio de 1978, para la localidad de Palma de Mallorca.

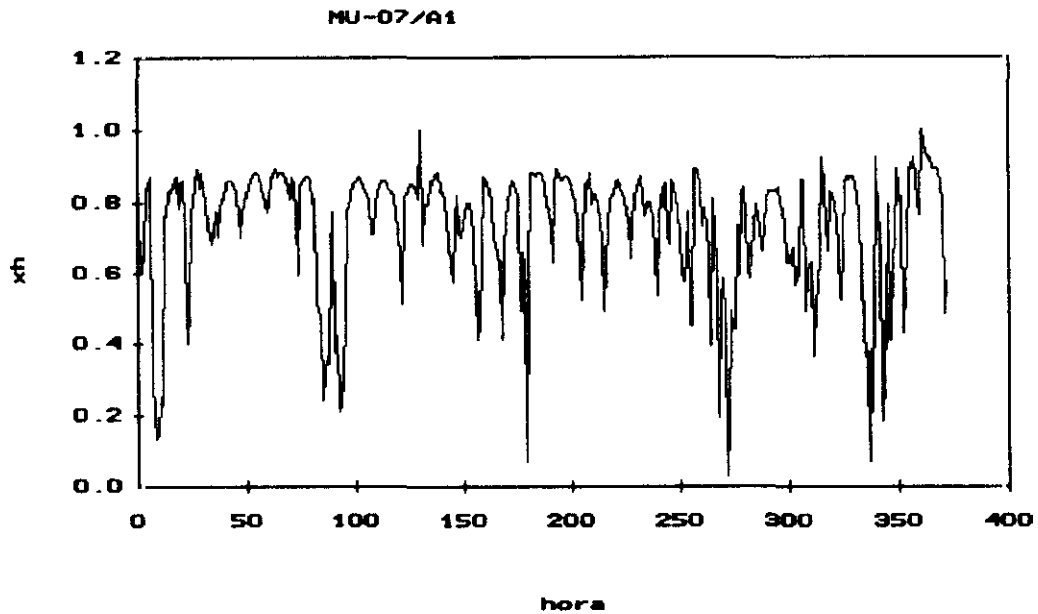


Figura 4.12 Serie mensual del índice  $X_h$ , para el mes de Julio de 1977, para la localidad de Murcia.

Como se puede observar en estas gráficas, al haber considerado un periodo temporal más pequeño hemos conseguido que la mayoría de las series sean estacionarias en varianza. Lógicamente, las series no son estacionarias en media, porque, por ejemplo, la media de las observaciones correspondientes a las primeras horas de cada día no coincide con la media de las observaciones correspondientes a las horas centrales.

Aún más, si nos fijamos en las series mensuales propuestas observamos que para algunos meses, sobre todo los que tienen muchos "días claros", todavía hay una cierta evolutividad diaria. Por ejemplo, en las figuras 4.13 y 4.14 se muestran los valores de  $X_h$  para las distintas horas de los días 14 a 19 de Enero y Julio, respectivamente, para el año 1977 calculados con los datos de exposición global horaria de Málaga.



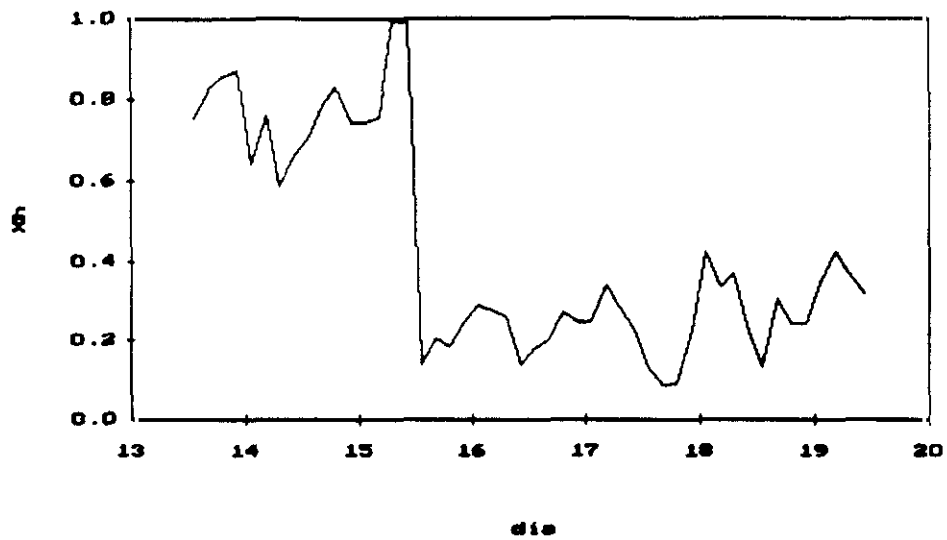


Figura 4.13 Valores del índice  $X_h$  para los días 14 a 19 de Enero de 1977. Málaga.

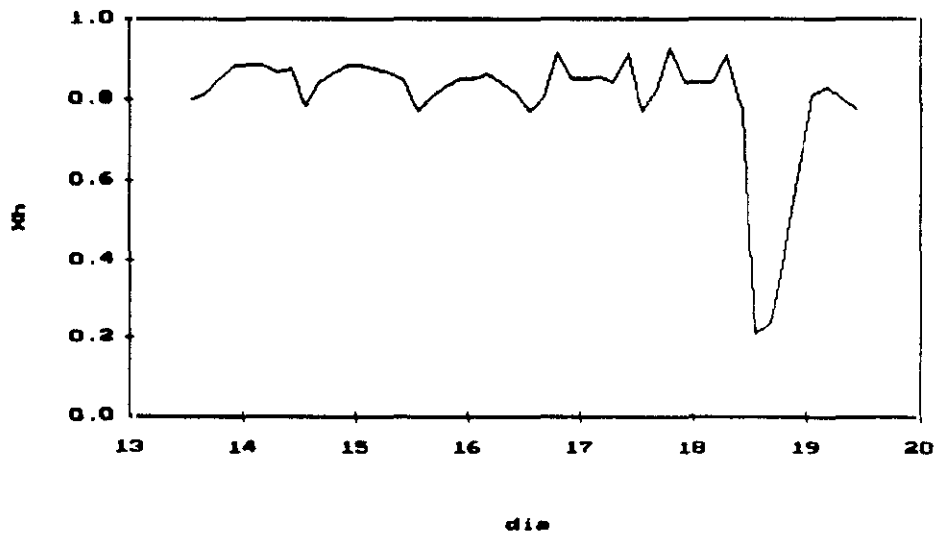


Figura 4.14 Valores del índice  $X_b$  para los días 14 a 19 de Julio de 1977. Málaga.

Como comentamos en el capítulo 2, existen varias técnicas con las que se puede eliminar esta evolutividad. Esto supone, para cada mes, calcular el valor medio y la varianza correspondiente a cada hora (las tendencias observadas son diarias), y obtener una nueva serie de índices de acuerdo con la expresión:

$$Y_h = \frac{X_h - X_{h,m}}{\sigma_{h,m}}$$

De nuevo nos encontramos con el inconveniente, ya comentado, de que los valores medios para cada hora y, sobre todo, la varianza de las series no están, normalmente, disponibles. Por ello, y aunque no afirmamos que no sea una técnica válida, no vamos a utilizarla en este trabajo.

Otra posibilidad es utilizar el operador diferencias. El orden de este operador coincidirá, para cada mes, con el número de horas por día considerados (de forma general le llamaremos de orden  $s$ ). Este método no se ha utilizado hasta ahora para series horarias. Se trata, sin embargo, de una técnica muy sencilla, que, como veremos, no necesita nuevos parámetros para su utilización.

Nuestro objetivo es obtener una serie estacionaria. La hipótesis de que al aplicar un operador diferencias de orden  $s$  a nuestra serie se va a obtener una serie estacionaria implica, fundamentalmente, que la diferencia entre observaciones separadas  $s$  intervalos es una variable aleatoria cuya media y varianza son constantes a lo largo del tiempo. Este supuesto es aceptable en el índice que utilizamos puesto que en él se refleja principalmente el efecto de las nubes y estamos considerando periodos de un mes. Este operador se puede expresar como:

$$D_s X_t = X_t - X_{t-s} \quad (4.6)$$

donde  $s$  toma los valores 8 (Enero, Febrero, Noviembre y Diciembre), 10 (Marzo, Abril, Septiembre y Octubre) y 12 (Mayo, Junio, Julio y Agosto).

Al aplicar este operador a cada serie mensual conseguimos eliminar la evolutividad que todavía había en las series. Las nuevas series son también estables en la media y la varianza y, por tanto, podemos admitir, en principio, que se trata de procesos estacionarios (falta por comprobar el tercero de los supuestos exigidos en la definición de serie estacionaria; este supuesto se contrastará en el proceso de validación de la serie).

Para comprobar la estacionariedad en media y varianza de las series, mostramos como ejemplo en las figuras 4.15 y 4.16 las series diferenciadas del índice  $X_h$  para los meses de Julio ( $s=12$ ) y Octubre ( $s=10$ ) del año 1977, para la localidad de Palma de Mallorca.

Estas series diferenciadas serán las que utilizaremos, y ya que el intervalo considerado es una hora, las llamaremos  $\{D_s X_h\}$ .

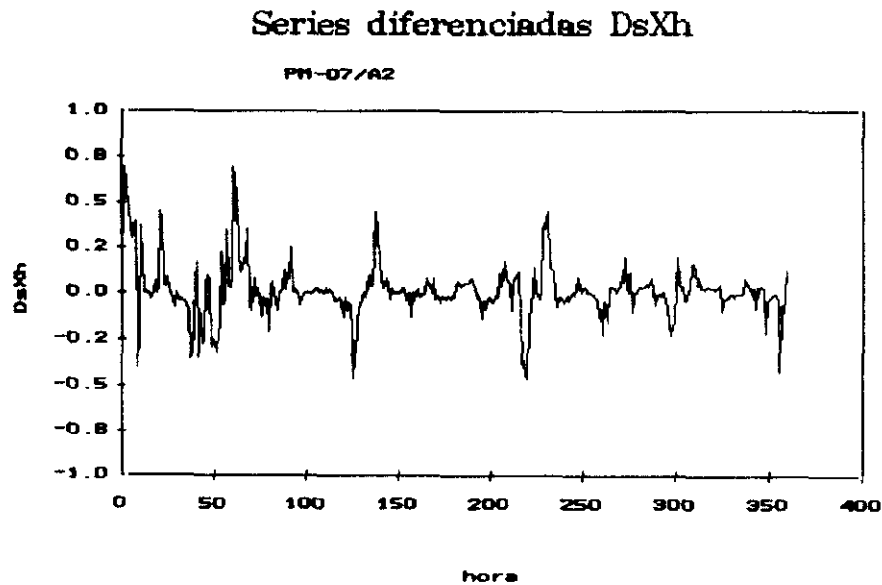


Figura 4.15 Serie diferenciada para el mes de Julio de 1977, de la localidad de Palma de Mallorca.

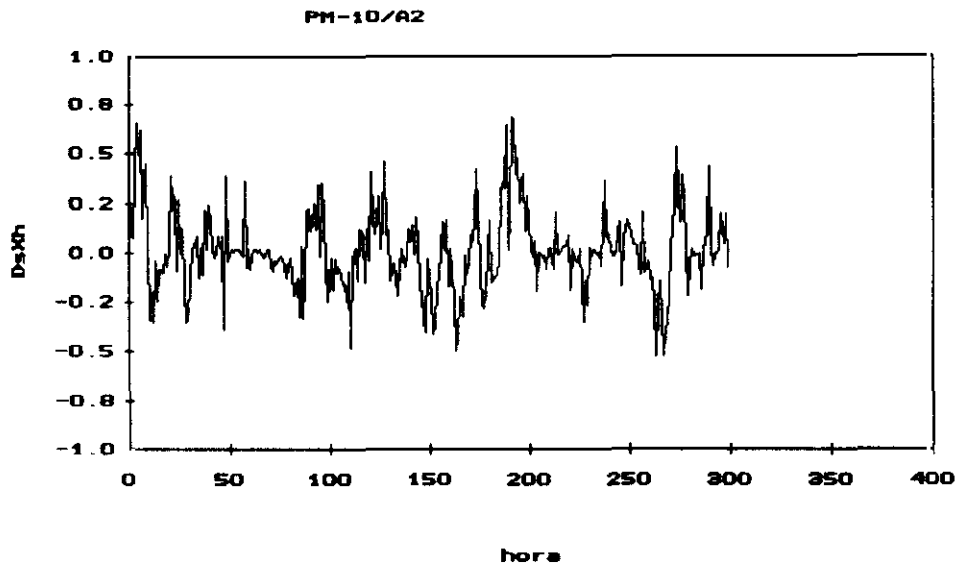


Figura 4.16 Serie diferenciada para el mes de Octubre de 1977, de la localidad de Palma de Mallorca.

Las funciones de distribución de estas series diferenciadas se pueden identificar, para la mayoría de las series mensuales con procesos gaussianos. Para los meses en los que esta función no es la correspondiente a un proceso gaussiano, sí podremos considerar, sin embargo, que los resultados que obtengamos (estimación de parámetros para el modelo subyacente) serán asintóticos, por trabajar con muchos datos (más de 240 para todos los meses). Además serán los resultados de la validación del modelo que propongamos los que nos harán aceptar o no la hipótesis de que las funciones de distribución son gaussianas.

En las figuras 17, 18 y 19 se muestran las funciones de distribución de probabilidades de ocurrencia de las series diferenciadas correspondientes a los meses de Enero, Mayo y Septiembre de 1979, de la localidad de Oviedo.

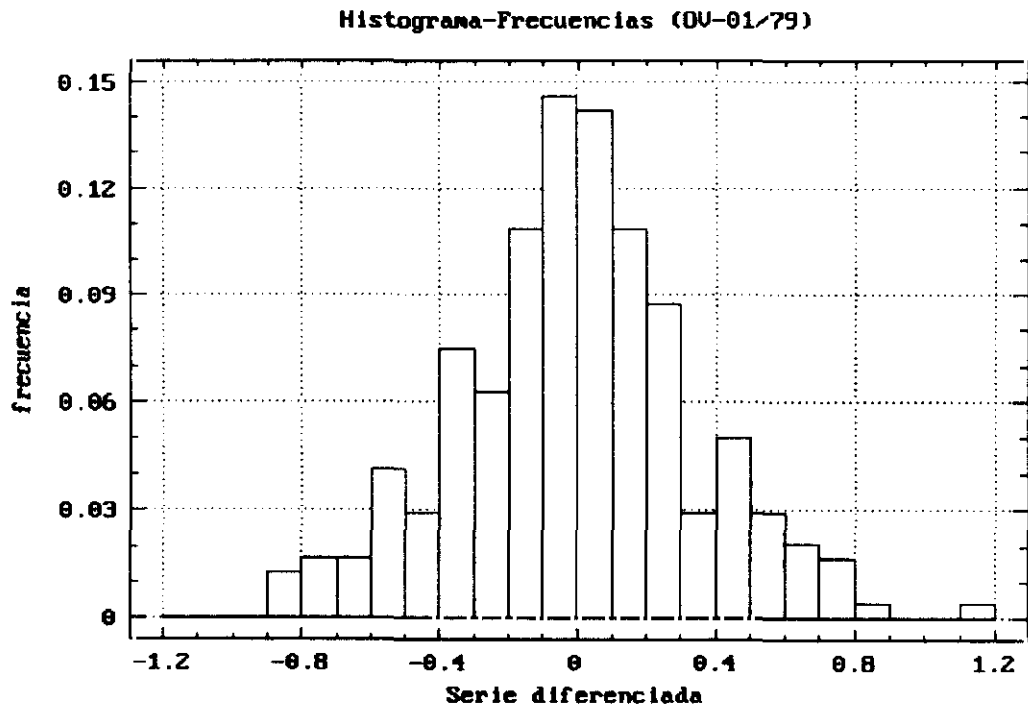


Figura 4.17 Distribución de probabilidades de ocurrencia de los valores de la serie diferenciada del mes de Enero-1979. Oviedo.

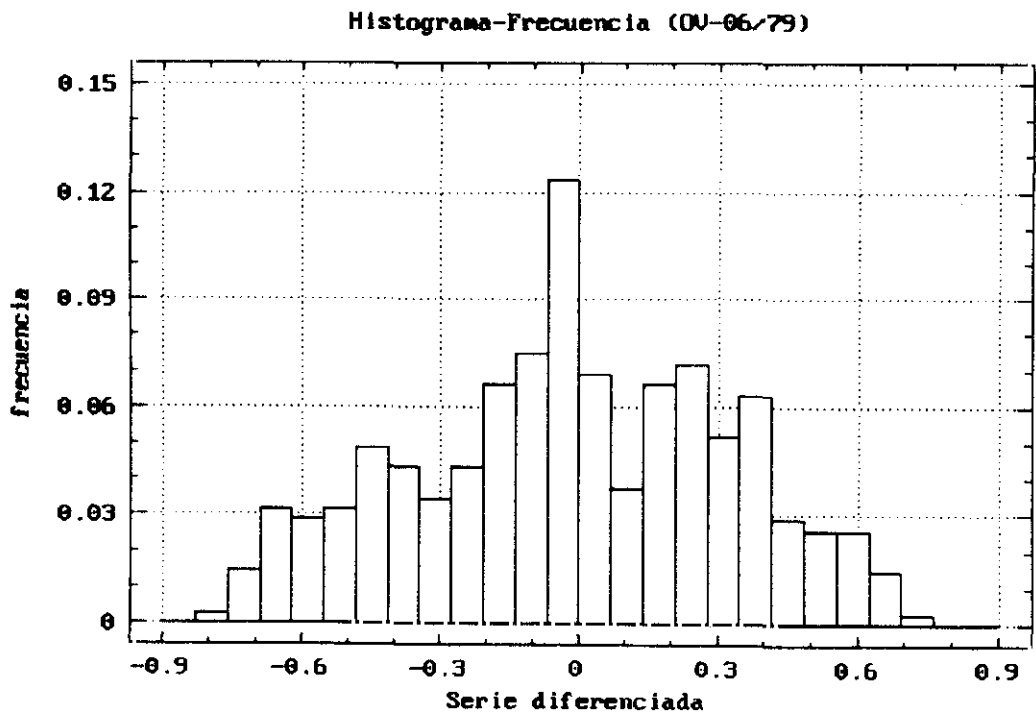


Figura 4.18 Distribución de probabilidades de ocurrencia de los valores de la serie diferenciada del mes de Junio-1979. Oviedo.

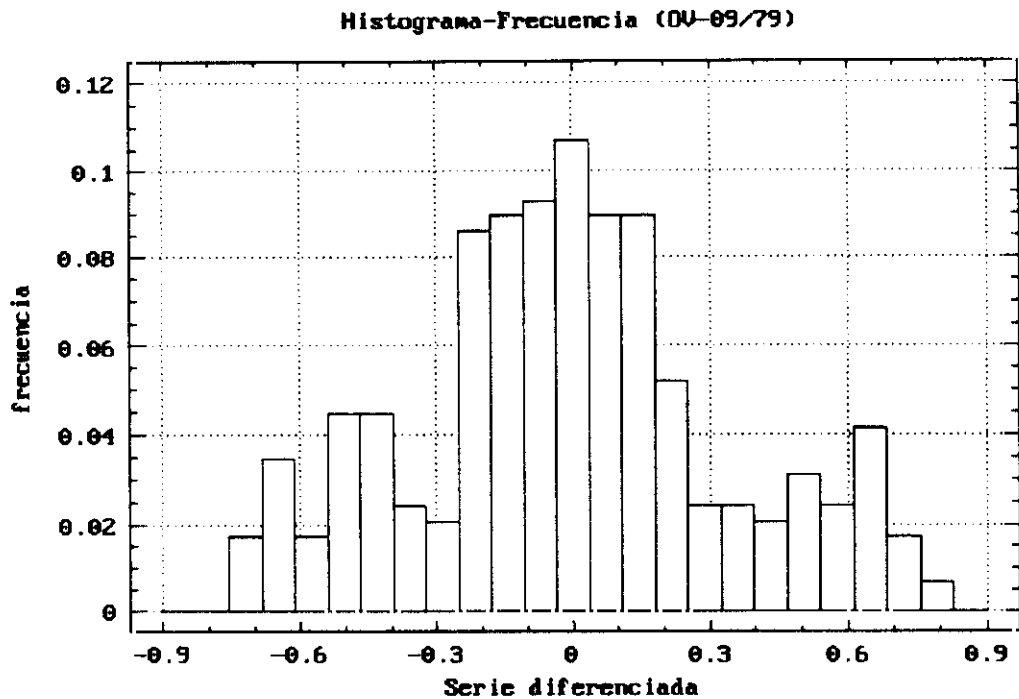


Figura 4.19 Distribución de probabilidades de ocurrencia de los valores de la serie diferenciada del mes de Septiembre-1979. Oviedo.

## 4.2 IDENTIFICACIÓN DEL MODELO

La identificación del modelo subyacente la haremos a partir de gráficos y correlogramas. Para encontrar el modelo con el que se pueden identificar las series  $\{D, X_n\}$  vamos a utilizar las funciones de autocorrelación y autocorrelación parcial.

Para un proceso estocástico los valores de la función de autocorrelación se calculan a partir de los valores de la autocovarianza. La función de autocovarianzas de un proceso es una función que describe las covarianzas en dos instantes cualesquiera:

donde  $\{z_t\}$  son las variables del proceso considerado (serie temporal) y  $\mu_t$

$$\text{Cov}(t, t+j) = \text{Cov}(z_t, z_{t+j}) = E[(z_t - \mu_t)(z_{t+j} - \mu_{t+j})] \quad (4.7)$$

representa la esperanza de  $z_t$ ,  $t$  es el instante inicial y  $j$  el intervalo entre observaciones.

La función de autocorrelación es la estandarización de la función de autocovarianzas:

$$\rho(t, t+j) = \frac{\text{Cov}(t, t+j)}{\sigma_t \sigma_{t+j}} \quad (4.8)$$

donde  $\sigma_t$  es la desviación típica de  $z_t$ .

Como en todo proceso estacionario se cumple que son estables la media, la varianza y las covarianzas, es decir

$$\mu_t = \mu_{t+j} = \mu \quad (4.9)$$

$$\sigma_t^2 = \sigma_{t+j}^2 = \sigma^2 = \gamma_0 \quad (4.10)$$

$$\text{Cov}(t, t+k) = \text{Cov}(t, t-k) = \gamma_k \quad k=0, \pm 1, \pm 2, \dots \quad (4.11)$$

la función autocorrelación (4.8) quedará:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (4.12)$$

Esta función nos proporciona información sobre la dependencia lineal entre los valores de la serie temporal para distintos desfases o retardos.

En los procesos estacionarios además de la función de autocorrelación (4.8) tiene también interés la llamada función de autocorrelación parcial. Ésta nos da una medida de la relación entre observaciones para distintos desfases con

independencia de los valores intermedios registrados entre dichas observaciones. El valor  $\phi_k$  que mide la autocorrelación parcial de orden  $k$  ( $k \geq 1$ ,  $k \in \mathbb{N}$ ) puede obtenerse como sigue:

El valor de  $\phi_k$  es el coeficiente de correlación simple entre  $U_t$  y  $V_t$ , siendo:  $U_t$  el término de error de una regresión con  $z_t$  como variable dependiente y  $z_{t-1}, \dots, z_{t-k+1}$  como variables independientes; y  $V_t$  el término de error de una regresión con  $z_{t-k}$  como variable dependiente y  $z_{t-1}, \dots, z_{t-k+1}$  como variables independientes (por la estacionariedad del proceso, el valor de  $t$  que se tome no afecta en la determinación de  $\phi_k$ ).

Las funciones de autocorrelación simple y parcial se estiman mediante los llamados correlogramas simple y parcial, respectivamente. El examen de estos correlogramas permite caracterizar el modelo con el que se pueden identificar las series diferenciadas del índice  $X_t$ .

Calcularemos los correlogramas simple y parcial de todas las series mensuales de que disponemos. Como se especifica en la Tabla 1.1 el número total de meses que se pueden utilizar es de 793.

La forma que presentan estos correlogramas es bastante similar para todos los meses considerados. En el correlograma simple observamos, de manera sistemática, que aparecen varios valores significativamente distintos de cero para desfases, a veces, de hasta orden cinco. La magnitud de estas correlaciones decrece hasta llegar al cuarto o quinto desfase, cuando puede admitirse ya que la correlación existente es nula. Además también aparece siempre un valor negativo significativamente distinto de cero para desfases que coinciden con el número de horas al día que hemos considerado (que llamamos  $s$ , este valor es además el mismo que utilizamos para diferenciar las series originales del índice  $X_t$ ). También a la izquierda y a la derecha de este valor (desfase igual a  $s-1$  y  $s+1$ )



aparecen valores, algo menores, pero significativamente distintos de cero.

Respecto al correlograma parcial de las series observamos que para desfases pequeños sólo hay un valor significativamente distinto de cero que se da para desfase igual a uno. Además, en los alrededores del número "s", horas al día consideradas, aparecen varios valores significativamente distintos de cero. El valor del coeficiente de autocorrelación parcial para el desfase igual al número de horas al día consideradas es negativo. A su izquierda aparece también otro valor negativo, algo menor, y a su derecha uno positivo, ambos significativamente distintos de cero. Estos tres valores se repiten para desfases que coinciden con múltiplos del número de horas al día consideradas, aunque su valor absoluto va disminuyendo conforme aumenta el desfase o distancia entre intervalos. A modo de ejemplo en las figuras 4.20, 4.21, 4.22 y 4.23 se muestran estas funciones para el mes de Julio de 1980 para Málaga y para el mes de Febrero de 1977 para Oviedo.

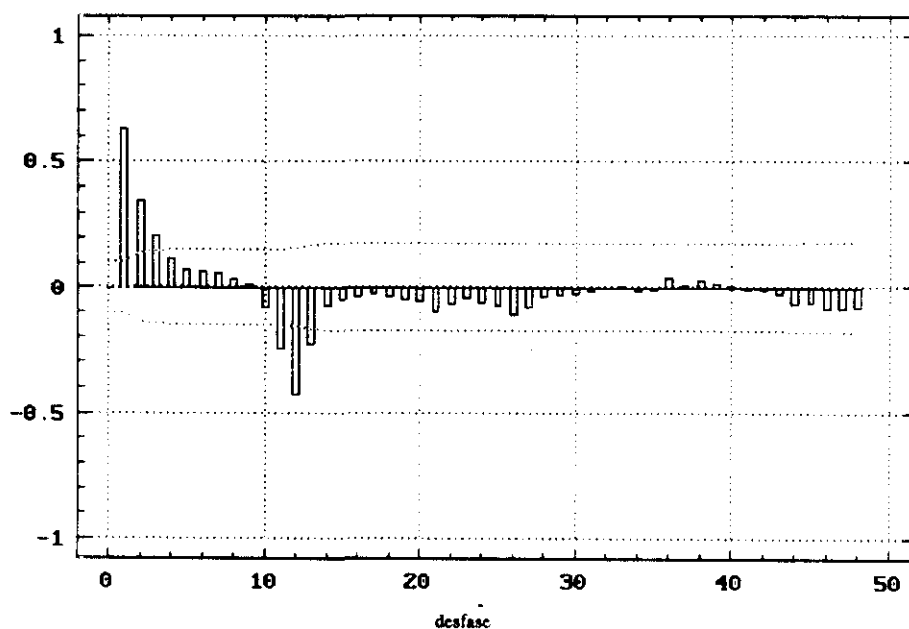


Figura 4.20 Correlograma simple. Julio, 1980. Málaga

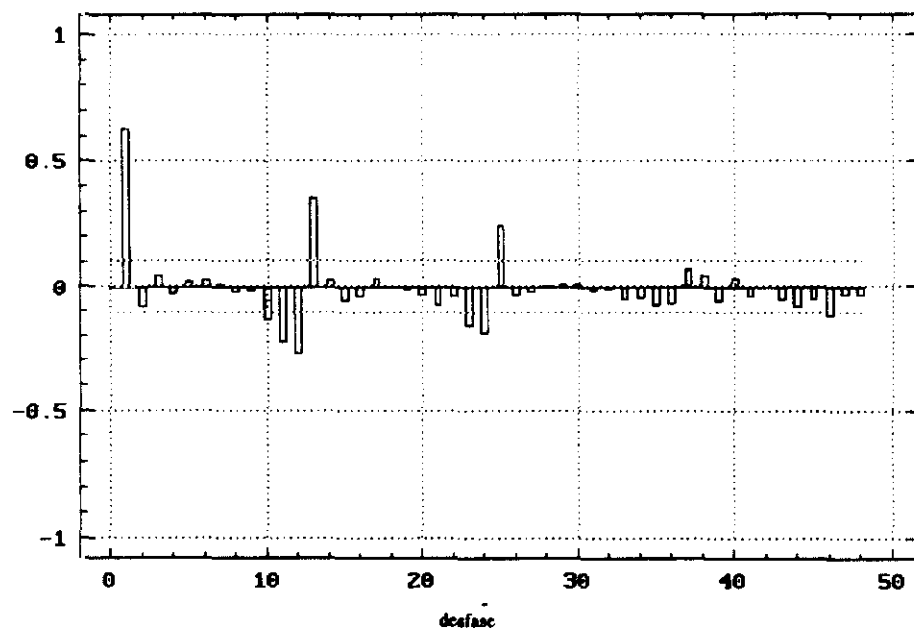


Figura 4.21 Correlograma parcial. Julio, 1980. Málaga

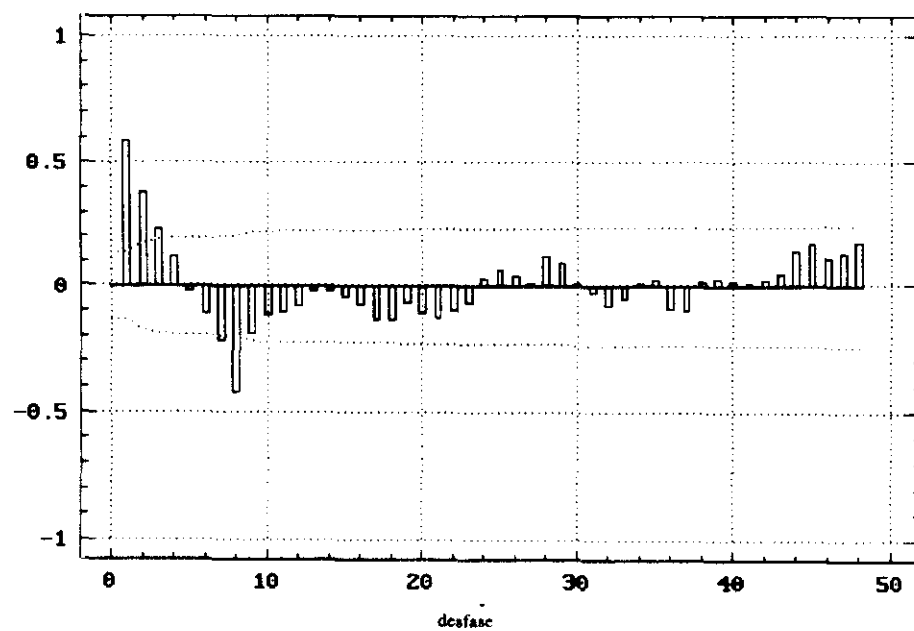


Figura 4.22 Correlograma simple. Febrero, 1977. Oviedo.

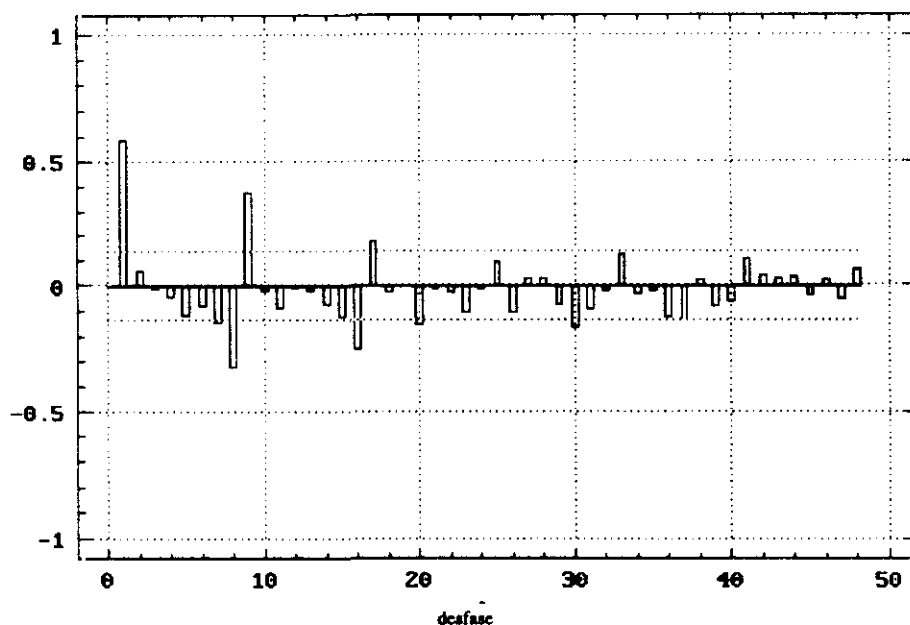


Figura 4.23 Correlograma parcial. Febrero, 1977. Oviedo.

En estas gráficas se pueden observar las características de los correlogramas simple y parcial que hemos comentado. Estas características se repiten para la mayoría de los meses analizados.

Siguiendo la metodología habitual en el estudio de series temporales estacionarias (ver, por ejemplo, Peña, /51/, cap.5 o Box-Jenkins, /16/), propondremos un modelo ARMA multiplicativo, es decir: analizaremos por separado las partes regular y estacional de la serie (tal y como se definen a continuación) y el modelo resultante será el producto de los modelos aplicables a cada parte.

La parte regular de la serie representa la relación entre observaciones consecutivas. La parte estacional de la serie representa la relación entre observaciones de la misma estación (en este caso, misma estación quiere decir

misma hora del día). Así pues, la parte regular la caracterizaremos utilizando los primeros valores de los correlogramas simple y parcial y la parte estacional la caracterizaremos a partir de los valores que se dan para el desfase que coincide con el número de horas al día consideradas y con los múltiplos de este valor. Vamos a analizar las características de cada una de ellas por separado.

Parte regular: el correlograma simple presenta varios valores significativamente distintos de cero, esto es presenta un decrecimiento lento. El correlograma parcial presenta un sólo valor significativo que se da para desfase uno. Estas características son las típicas de un modelo autorregresivo de orden uno (ARMA (1,0)), que es por tanto el que proponemos para la parte regular. Un proceso  $z_t$  es un proceso estacionario y ergódico autorregresivo de primer orden si puede expresarse como:

$$z_t = \phi z_{t-1} + a_t \quad \forall t \quad (4.13)$$

siendo el parámetro  $\phi$  menor que uno en valor absoluto y el término de error  $\{a_t\}$  un ruido blanco.

Por su parte un proceso ruido blanco  $\{a_t\}$  es un proceso estacionario con media cero, en el que la correlación entre  $a_t$  y  $a_s$  es cero siempre que  $t$  sea distinto de  $s$ .

Dado un valor inicial  $z_0$ , un proceso ARMA(1,0) queda caracterizado por el valor de  $\phi$  y la varianza del término de error  $\sigma^2$ .

Desde el punto de vista físico, en nuestras series proponer un proceso ARMA(1,0) significará que en los valores de un mes el valor de una hora queda determinado por el valor de la hora anterior más un término de error impredecible. Es decir, puesto que nuestras series originales estaban normalizadas al valor máximo de radiación esperable para una hora, la variación que los

componentes atmosféricos introducen en la parte regular de nuestra serie es sólo predecible a partir de la variación que hubo en la hora anterior.

Parte estacional: el correlograma simple de la parte estacional (valores múltiplos de  $s$ ) presenta sólo un valor significativamente distinto de cero: el primero, correspondiente al número de horas al día consideradas ( $s$ ). El correlograma parcial presenta varios valores significativamente distintos de cero para desfases que coinciden con el número de horas al día consideradas y múltiplos de este valor. Es decir, para la parte estacional el correlograma simple presenta un decrecimiento brusco (un sólo valor significativo) y el parcial presenta un decrecimiento lento.

Estas características son las que corresponden a un modelo de medias móviles de orden uno, que es el que proponemos para la parte estacional. Un proceso  $v_t$  es un proceso estacionario y ergódico de medias móviles de primer orden (ARMA(0,1)) si puede expresarse como:

$$v_t = \theta a_{t-1} + a_t \quad \forall t \quad (4.14)$$

siendo el parámetro  $\theta$  menor que uno en valor absoluto, y el término de error  $a_t$  un proceso ruido blanco. En nuestro caso, como lo aplicamos a la parte estacional escribimos  $(0,1)_s$ , y  $v_t$  será el proceso obtenido con múltiplos de orden  $s$  del proceso original.

Desde un punto de vista físico, el modelo de medias móviles de primer orden propuesto para la parte estacional nos indica que las perturbaciones introducidas a una hora del día están relacionadas con las perturbaciones que tendrán lugar a la misma hora del día siguiente. Esto está de acuerdo con las observaciones que hicimos en un apartado del capítulo anterior cuando analizamos la clasificación de tipos de días para la localidad de Madrid.

---

Estas características nos llevan a proponer la utilización de un **modelo ARMA multiplicativo**, en el que la parte regular sigue un **modelo autorregresivo de orden uno (AR1)** y la parte estacional responde a un **modelo de medias móviles de orden uno (MA1)**. Se representa como  $ARMA(1,0) \times ARMA(0,1)$ . El carácter multiplicativo del modelo es lo que explica que aparezcan en el correlograma parcial valores significativos para desfases de orden  $s-1$  y  $s+1$ .

El hecho de que en las series haya aparecido una parte estacional hace que no podamos identificar las series con modelos de Markov de primer orden. Estos modelos han sido propuestos para las series de índices  $X_n$  en los trabajos de E. Palomo, /46/, y Ll. Mora et al., /39/. En estos trabajos no se menciona la presencia en las series de una parte estacional. Esto puede ser debido a dos razones: primero porque no se analizan las funciones de autocorrelación simple y parcial para desfases suficientemente grandes; y segundo porque no se utilicen, para cada mes, días con igual número de horas o valores. Respecto a la primera hipótesis, si observamos los dos trabajos anteriormente citados, en las gráficas presentadas en el trabajo de E. Palomo, /46/, de estas dos funciones aparecen valores significativamente distintos de cero alrededor del desfase 9 (cuando la serie corresponde al mes de Diciembre) y alrededor de 16 (cuando la serie corresponde al mes de Junio). Estos valores no pueden ser debidos a la concatenación artificial que se hace en las series (suponer que el primer valor de un día es el siguiente, en la serie, al último del día anterior), ya que esto se reflejaría en la parte regular que es la que nos da cuenta del tipo de relación entre dos valores seguidos.

Evaluaremos también si la presencia de esta parte estacional es debida a la diferenciación de las series. Si esto fuera así, el coeficiente de medias móviles de la parte estacional debería ser igual a uno; este aspecto lo analizaremos cuando contrastemos la hipótesis de si este coeficiente es o no igual a uno. En cualquier

caso, y como ya hemos comentado, las series originales de índices  $X_b$  (antes de diferenciar) no eran estacionarias.

Una vez identificado el modelo, el siguiente paso será la estimación de los parámetros del mismo.

### 4.3 ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO PROPUESTO

Como hemos visto, para la mayoría de las series mensuales disponibles, la forma de los correlogramas simple y parcial nos ha sugerido la utilización de un modelo ARMA multiplicativo. Estos modelos, de forma general, se pueden representar mediante la expresión:

$$\phi_p(L)\tilde{\phi}_q(L)w_t = \theta_q(L)\tilde{\theta}_q(L)a_t \quad (4.15)$$

donde  $\phi_p(L)$  es un polinomio en retardos que representa el modelo autorregresivo de la parte regular,  $\tilde{\phi}_q(L)$  el de la parte estacional,  $\theta_q(L)$  es un polinomio en retardos que representa el modelo de medias móviles de la parte regular,  $\tilde{\theta}_q(L)$  el de la parte estacional y  $a_t$  es un ruido blanco.

Como en nuestro caso para la parte regular hemos propuesto un modelo AR1 y para la estacional un MA1, la expresión anterior quedará:

$$(1 - \phi_1 L)w_t = (1 - \tilde{\theta}_1 L^s)a_t \quad (4.16)$$

es decir:

$$w_t = \phi_1 w_{t-1} - \tilde{\theta}_1 a_{t-s} + a_t \quad (4.17)$$

Los parámetros que debemos estimar son, por tanto,  $\theta_1$ ,  $\phi_1$  y la varianza del ruido blanco (supondremos que éste sigue una distribución normal con media 0 y varianza finita en cada una de las series que estamos considerando). El método de estimación utilizado es el de máxima verosimilitud. Es un procedimiento de estimación que requiere optimización. Con este método se selecciona, de todos los valores posibles del parámetro, aquél que hace que el resultado obtenido sea más probable.

Utilizando esta metodología hemos calculado los estimadores de máxima verosimilitud de  $\theta_1$ ,  $\phi_1$ , y  $\sigma^2$  (que representaremos como  $\hat{\theta}_1$ ,  $\hat{\phi}_1$ , y  $\hat{\sigma}^2$ ). Hemos estimado la matriz de varianzas y covarianzas de:

$$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\theta}_1 \end{pmatrix} \quad (4.18)$$

y hemos calculado el valor del estadístico Box-Ljung que se utiliza para contrastar la hipótesis nula de que  $a_t$  es ruido blanco. Este estadístico se describe en el apartado siguiente. Los valores de los estimadores obtenidos para cada una de las localidades con datos se muestran en el Apéndice II.

#### 4.4 DIAGNOSIS DEL MODELO. DESCRIPCIÓN DE LOS CONTRASTES

Para evaluar la validez del modelo propuesto vamos a realizar una serie de contrastes sobre las estimaciones realizadas, a saber:

- Contraste sobre la significatividad de los parámetros que son de interés
- Contraste sobre el término de error
- Contraste sobre estacionariedad

Los contrastes que hemos utilizado son los siguientes:



*Análisis de la estimación*

- Contraste respecto al orden del modelo autorregresivo: analizaremos si el coeficiente  $\phi_1$  es significativamente distinto de cero. La hipótesis nula,  $H_0$ , que proponemos es:

$$\phi_1=0 \quad (4.19)$$

Para contrastar esta hipótesis calculamos el estadístico t, según la expresión:

$$\frac{\hat{\phi}_1}{S.E.(\hat{\phi}_1)} \quad (4.20)$$

Si la hipótesis nula es cierta este estadístico sigue, aproximadamente, una distribución normal, con media cero y varianza uno. Tomaremos como nivel de significación  $\alpha=0.05$ . Por tanto, si el estadístico t es mayor que 1.96, en valor absoluto, el p-valor (que es la salida que nos da el programa de análisis utilizado) será menor que 0.05 y habremos de rechazar la hipótesis nula y por tanto el modelo será de orden mayor que cero. Si por el contrario el p-valor de este contraste es mayor que 0.05 aceptaremos la hipótesis nula.

También haremos un contraste respecto a que este parámetro sea o no uno. La hipótesis nula será:

$$\phi_1=1 \quad (4.21)$$

Para contrastar esta hipótesis calculamos:

$$\frac{\hat{\phi}_1 - 1}{S.E.(\hat{\phi}_1)} \quad (4.22)$$

Los criterios que utilizaremos son los mismos que los vistos para la hipótesis nula de que el modelo sea de orden cero. Si aceptáramos la hipótesis nula habría evidencia de que la serie que se analiza no es estacionaria (recuérdese que, si lo es, el verdadero valor  $\phi$  es menor que uno).

- Contraste respecto al orden del modelo de medias móviles: seguiremos los mismos pasos detallados para el modelo autorregresivo, es decir, haremos los mismos contrastes anteriores para  $\theta_1$ .

#### *Análisis de los residuos*

- Contraste sobre si las perturbaciones son o no ruido blanco: la hipótesis nula que proponemos es que el término de error es un ruido blanco. El estadístico que se utiliza para este contraste, conocido como estadístico Box-Ljung es:

$$Q(h) = T(T+2) \sum_{j=1}^h \frac{r_a^2(j)}{T-j} \quad (4.23)$$

siendo  $r_a(j)$  el coeficiente de correlación muestral entre  $\hat{a}_t$  y  $\hat{a}_{t+j}$ , es decir,

$$r_a(j) = \frac{\sum_{t=1}^{T-j} (\hat{a}_t - \bar{a})(\hat{a}_{t+j} - \bar{a})}{\sum_{t=1}^T (\hat{a}_t - \bar{a})^2} \quad (4.24)$$

El estadístico  $Q(h)$  se utiliza para contrastar si hay correlación entre los  $h$  primeros retardos de  $a_t$ . Si la hipótesis nula de que el término de error es ruido blanco es cierta, entonces no hay correlación entre los  $h$  primeros retardos y  $Q(h)$  sigue una distribución  $\chi^2$  con  $h-f$  grados de libertad, siendo  $f$  el número de parámetros estimados en el modelo. Hemos realizado este contraste con  $h=20$  ( en nuestro caso  $f=2$ ) y obtenido el correspondiente p-valor. Cuanto mayor sea el p-valor del contraste más evidencia habrá a favor de  $H_0$ . Tomaremos como nivel de significación  $\alpha=0.05$ . En el caso de que el p-valor sea menor que 0.05 rechazaremos la hipótesis nula, y por tanto no podremos aceptar que el término de error sea un ruido blanco.

También analizaremos la normalidad a partir de la gráfica: en ella, aproximadamente, un 50% de los valores deben ser mayor que cero y, aproximadamente, un 2.5% debe ser mayor que  $1.96\hat{\sigma}$  y un 2.5% menor que  $1.96\hat{\sigma}$ .

Asimismo contrastaremos la hipótesis de normalidad utilizando los coeficientes de asimetría y curtosis obtenidos con los residuos. El estadístico que se utiliza para este contraste (conocido como estadístico de normalidad Bera-Jarque, BJ) es:

$$BJ = \left( \frac{s^2}{6} + \frac{(E-3)^2}{24} \right) T \quad (4.25)$$

siendo  $T$  el número de observaciones y  $S, E$  los coeficientes muestrales de asimetría y curtosis de los residuos, respectivamente. Si la hipótesis nula de normalidad del término de error es cierta, entonces  $BJ$  sigue, aproximadamente una distribución  $\chi^2$  con dos grados de libertad, lo que permite determinar el  $p$ -valor del contraste.

Completaremos el análisis de los residuos, por último, con un examen del correlograma residual: si el modelo está correctamente especificado, únicamente un 5% de los valores del correlograma, aproximadamente, serán significativos, cuando se toma como nivel de significación 0.05.

## 4.5 RESULTADOS DE LA DIAGNOSIS

### *Análisis de la estimación:*

En todos los meses analizados se obtiene un  $p$ -valor igual a cero, tanto para la hipótesis de que el modelo autorregresivo es de orden cero como para la de que es cero el modelo de medias móviles. Esto nos lleva a rechazar ambas, es decir el coeficiente  $\phi_1$  del modelo autorregresivo de la parte regular y el coeficiente  $\tilde{\theta}_1$  del modelo de medias móviles de la parte estacional son significativamente distintos de cero.

Del mismo modo, para la hipótesis de que ambos coeficientes son igual a uno, para casi todos los meses se cumple que el  $p$ -valor en ambos

casos es menor que 0.05 por lo que rechazaremos la hipótesis nula de que ambos sean uno. Sólo para algunos meses (menos del 4%), obtenemos un valor superior a 0.05. De acuerdo con lo que comentamos en un apartado anterior, la parte estacional que aparece en estas series sí podría ser debida a la diferenciación (aunque también este porcentaje lo podemos explicar por la propia naturaleza estadística del fenómeno). Para las demás series podemos afirmar que la diferenciación no hace que aparezca una parte estacional.

*Análisis de los residuos:*

El primer contraste que hemos analizado es el de que son un ruido blanco. Como ya comentamos, si esta hipótesis es cierta el estadístico Box-Ljung ha de seguir una distribución  $\chi^2$ . En la Tabla 4.I se recogen para todas las localidades con datos el número de meses en los que el p-valor del contraste es mayor que 0.05 (columna denominada RB) y el porcentaje que representan respecto al total.

LOCALIDAD	MESES UTIL.	RB	%
BADAJOS	84	69	82.1
CASTELLON	55	44	80.0
LOGROÑO	48	36	75.0
MADRID	96	78	81.3
MÁLAGA	94	82	87.2
MURCIA	96	83	86.5
OVIEDO	96	85	88.5
P.MALLORCA	94	78	83.0
SEVILLA	78	59	75.6
TORTOSA	52	45	86.5
Total	793	659	83.1

Tabla 4.I

---

En, aproximadamente, un 83% de los meses analizados puede aceptarse con el contraste de Box-Ljung la hipótesis nula de que el término de error es un ruido blanco.

Como el nivel de significación tomado es  $\alpha=0.05$ , este porcentaje es algo mayor que el que cabe esperar si la hipótesis nula fuera cierta en todos los casos. Con el resto de contrastes descritos se obtienen buenos resultados, en el sentido de que en aproximadamente un 95% de los casos se acepta la hipótesis nula. Así pues, el análisis de residuos confirma la validez de los modelos propuestos para algo más del 83% de los meses, aunque no para el 100% (además del 83% de meses para los que se acepta  $H_0$  con el contraste Box-Ljung, en un 5% más de casos podría admitirse que el modelo es correcto y que el resultado negativo obtenido se debe a la aleatoriedad). A modo de ejemplo en las figuras 4.24, 4.25, 4.26 y 4.27 se muestran los gráficos de las funciones de distribución acumuladas de los residuos para los siguientes meses y localidades: Abril-1981 y Septiembre-1983, Málaga; Noviembre-1979, Oviedo; Febrero-1978, Murcia.

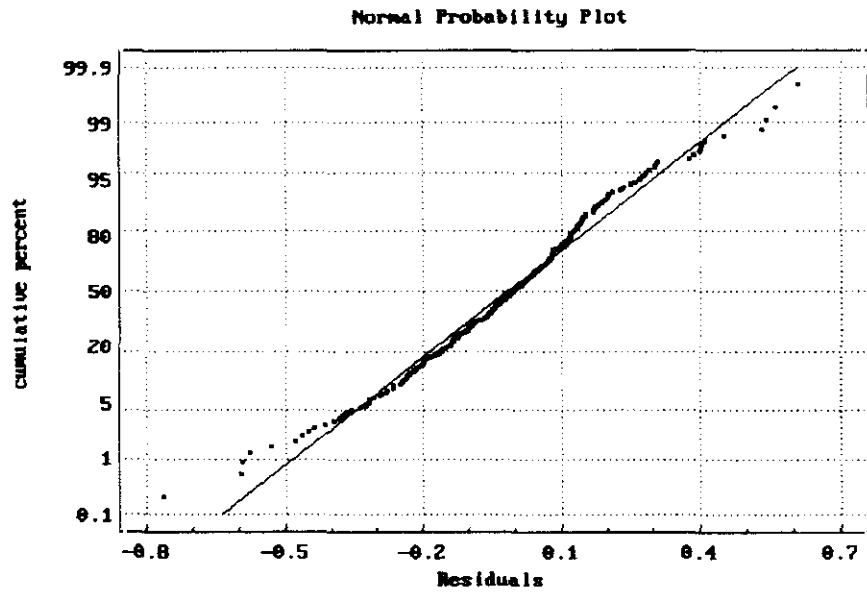


Figura 4.24 Serie residual, Abril-1981. Málaga

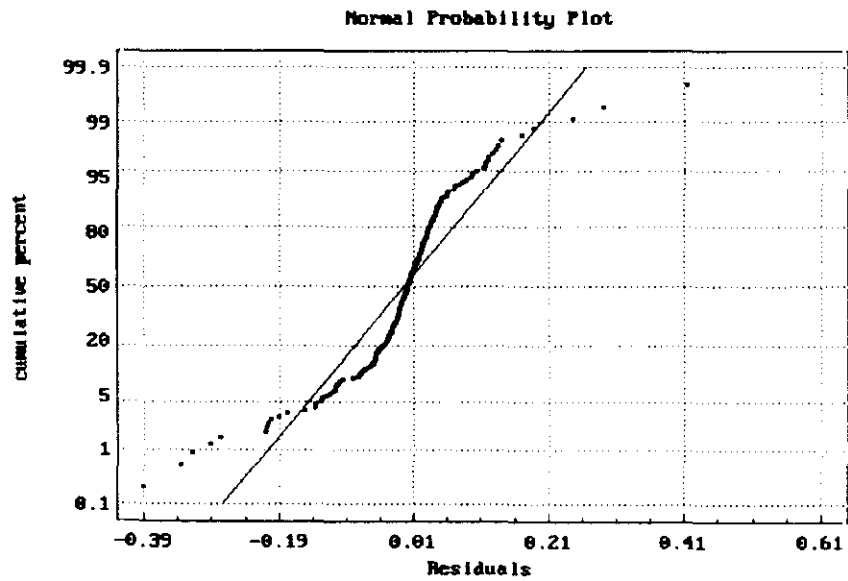


Figura 4.25 Serie residual, Septiembre-1983. Málaga

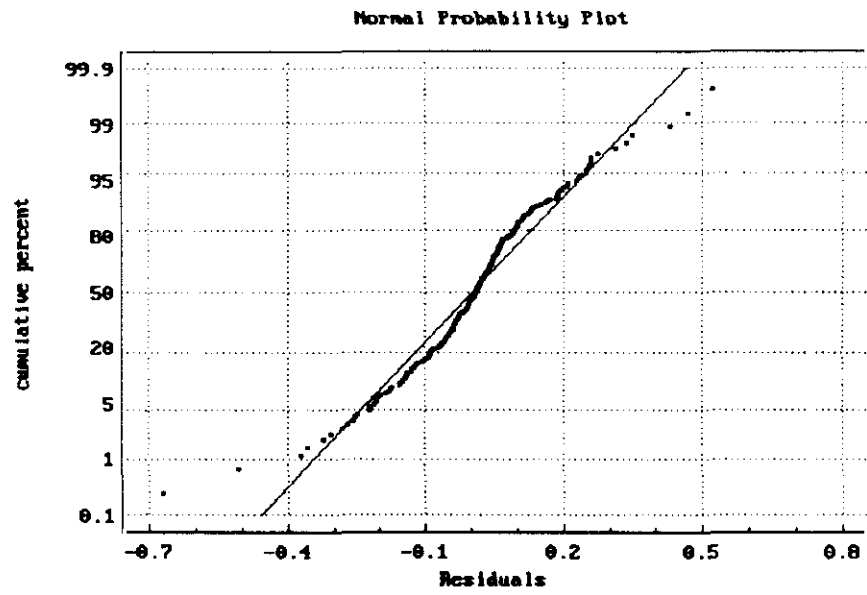


Figura 4.26 Serie residual, Noviembre-1979, Oviedo.

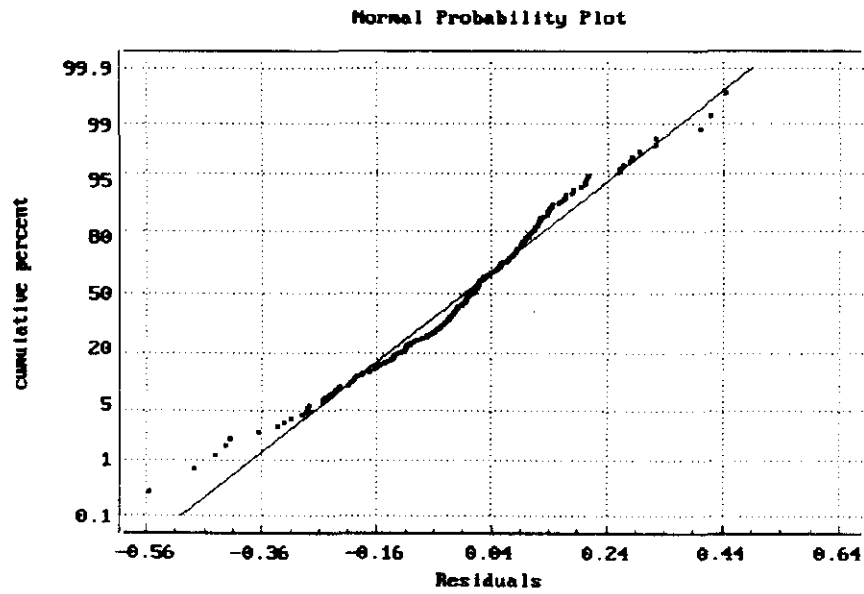


Figura 4.27 Serie residual, Febrero-1978, Murcia.



Como se puede observar en todas ellas se cumple la condición especificada anteriormente: aproximadamente un 50% de los valores es mayor que cero, y un 2.5% es mayor que  $1.96\hat{\sigma}$  y un 2.5% menor que este valor.

El correlograma residual de la gran mayoría de las series para las que se acepta  $H_0$  en el contraste de Box-Ljung presenta a lo sumo un 5% de valores significativos.

Como ejemplo, en las figuras 4.28 y 4.29 se muestran los correlogramas residuales para los meses de Septiembre-1983 y Febrero-1979 de las localidades de Málaga y Murcia respectivamente.

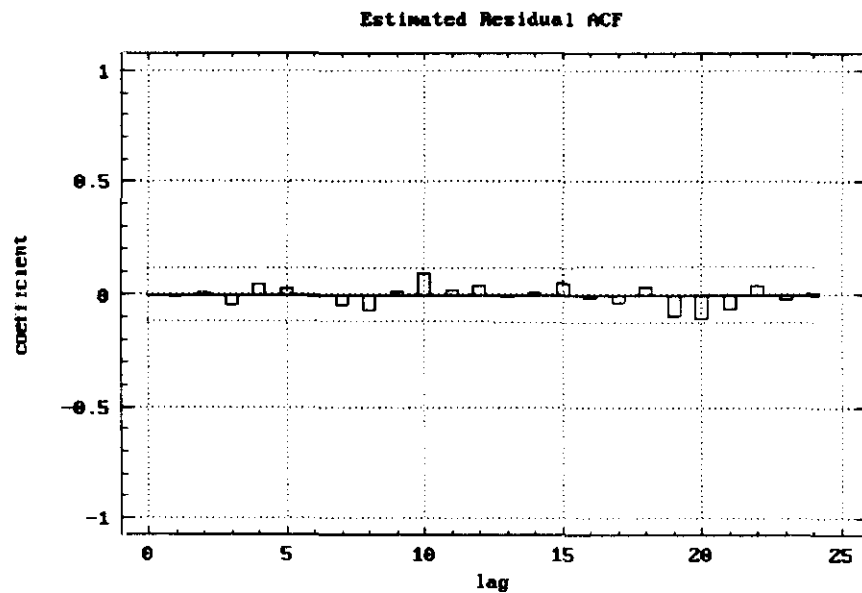


Figura 4.28 Correlograma serie residual. Septiembre-1983. Málaga

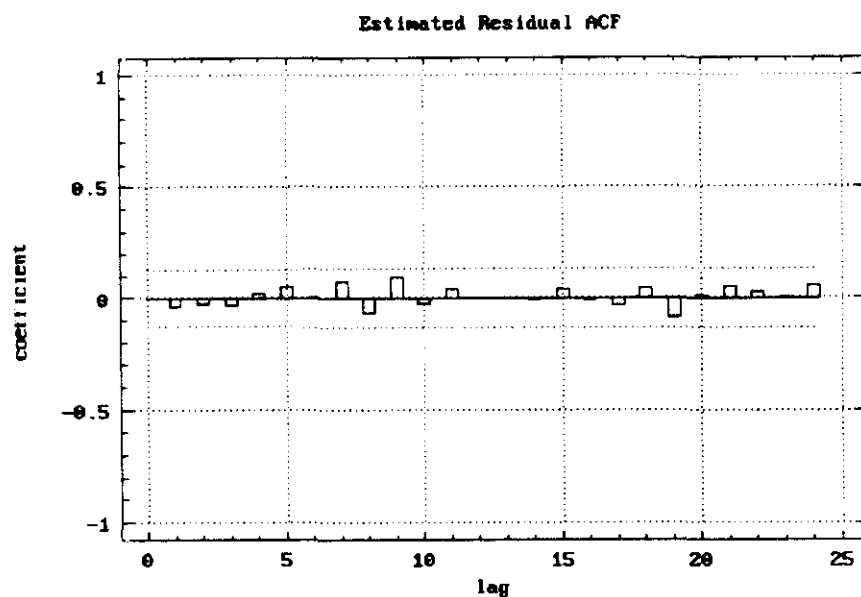


Figura 4.29 Correlograma serie residual, Febrero-1979. Murcia

A la vista de estos resultados surgen, al menos, dos cuestiones de interés:

- La primera se refiere a cómo se pueden generalizar los resultados obtenidos para los meses en los que el modelo ARMA multiplicativo ha resultado ser el adecuado. Nuestro intento de generalización será en el sentido de encontrar parámetros generales que definan las distintas localidades. También analizaremos las relaciones que existen entre los parámetros estimados en el modelo propuesto y otras características de las series.
- La segunda está relacionada con los meses en los que el modelo propuesto no ha funcionado; es decir, aquellos meses en los que no puede aceptarse una especificación  $ARMA(1,0) \times (0,1)$ , porque se rechaza con esta especificación que los residuos sean ruido blanco.

Aunque la mayoría de los meses se puedan identificar con modelos ARMA multiplicativos, el hecho de que para más de un 5% el modelo no funcione (aproximadamente el 15%) nos obliga a analizar con más detalle estos meses.

De la primera de las cuestiones planteadas nos ocuparemos en el capítulo 6. La segunda será objeto de análisis en el apartado siguiente.

#### **4.6 ANÁLISIS DE LOS MESES PARA LOS QUE NO FUNCIONA EL MODELO PROPUESTO**

Para el análisis de los meses en los que el modelo ARMA multiplicativo no ha resultado ser el adecuado vamos a utilizar las gráficas de estas series. De la observación de éstas podemos clasificar estos meses en tres tipos:

- Meses en los que el valor medio y la varianza no permanecen constantes: esto es, no se trata de series estacionarias a las que se pueda aplicar el análisis propuesto. Este fenómeno lo hemos encontrado, aproximadamente en un 35% de los meses en los que no ha funcionado el modelo propuesto (un 5.9% del total), principalmente en los meses de Mayo, Junio y Agosto. Desde el punto de vista del fenómeno físico que representan las series analizadas, son los meses en los que se produce un cambio climático brusco. Podemos observar cómo hay una parte de los mismos en la que la serie es similar a la del mes anterior y otra parte que es similar a la del mes siguiente. En general, corresponden a meses y localidades en las que se pasa de pronto del "frío" al "calor" o viceversa, procesos, por otro lado, típicos de un clima continental. Para caracterizar estas series con el

modelo propuesto deberíamos modificar el periodo de estudio, bien tomando series temporales más cortas, o bien incluyendo los periodos correspondientes en los meses anterior y posterior. En las figuras 4.30 y 4.31 se muestran dos ejemplos de este tipo de series. La primera corresponde al mes de Agosto de 1976 y la segunda al de Abril de 1979, ambas para la localidad de Badajoz.

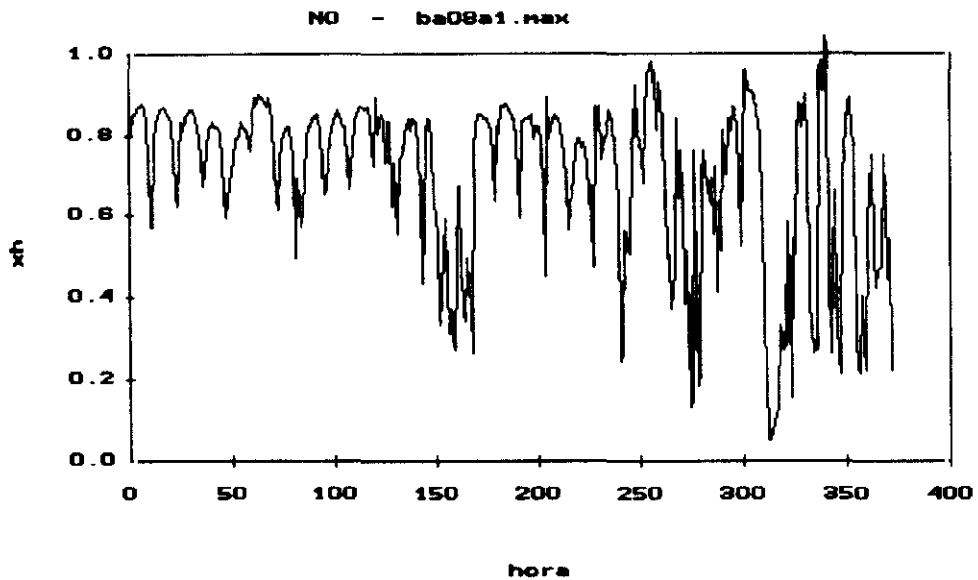


Figura 4.30. Valores del índice  $X_n$ , para el mes de Agosto de 1976, Badajoz.

Debido a que las secuencias que se dan en estas series son similares a las que se dan los meses anterior y posterior, el no tener en cuenta estos meses no afectará de manera significativa a las series que generemos mediante el procedimiento descrito en un apartado posterior.

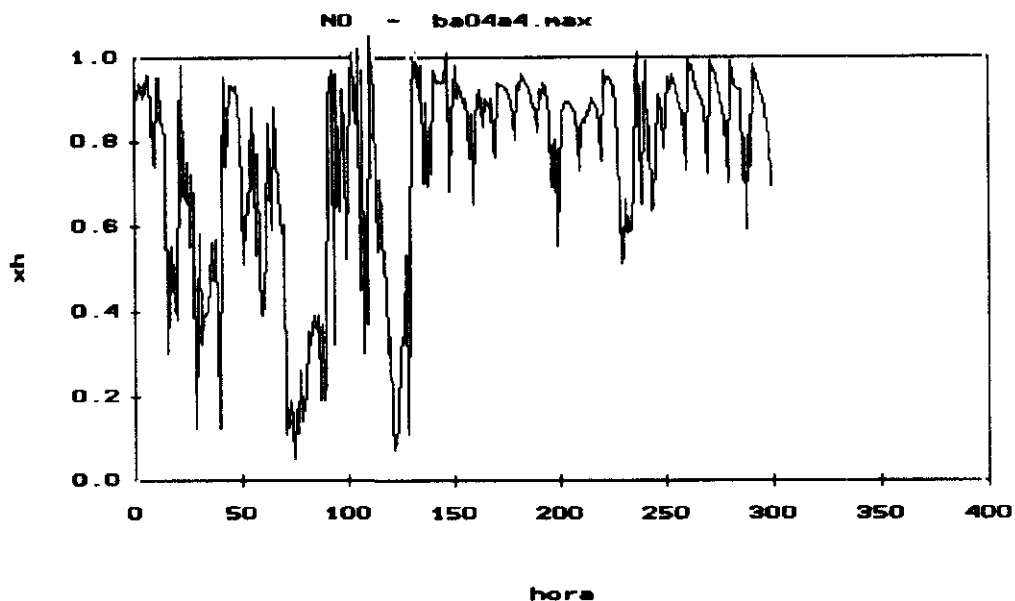


Figura 4.31 Valores del índice  $X_h$ , para el mes de Abril de 1979, Badajoz.

- Meses en los que el correlograma simple y parcial de los residuos no corresponde al de un ruido blanco, ya que se observa algún pico significativo. En estos correlogramas hay algún/os datos significativamente distintos de cero. Puesto que todavía había una relación entre las observaciones del mes hemos probado la utilización de modelos ARMA en los que la parte autorregresiva es de orden mayor que uno. El resultado es que, en general, estos meses se pueden identificar con modelos  $ARMA(h,0) \times (0,1)$ , donde  $h$  puede ser hasta de orden 5. Entre las series a que corresponden estos meses no se observa ninguna característica diferente a las de las series en las que sí ha funcionado el modelo ARMA. El número total de meses para los que se ha identificado un modelo ARMA de orden superior es de un 25% del total de meses en los que no ha funcionado el modelo propuesto (un 4.25%).

- Meses para los que, aún presentando características similares a los otros, el modelo ARMA propuesto no ha sido válido; representan un 40% de los meses en los que no ha funcionado el modelo (un 6.8% del total). Este porcentaje se explica por la aleatoriedad del estimador del contraste.

## 4.7 INTERPRETACIÓN DE LOS RESULTADOS. CONCLUSIONES

En este capítulo hemos analizado varios periodos temporales que se pueden utilizar para el estudio de las series anuales de exposición horaria. De este análisis podemos sacar dos conclusiones: la primera es que cuando consideramos periodos mayores que un mes nos "arriesgamos" a que las series que utilicemos ya no sean estacionarias; la segunda es que la utilización de "meses globales" en el sentido que proponen Gordon et al., /28/, (para valores diarios) no parece ser adecuada en el caso de las series de exposición horaria; como hemos visto, aunque, en general, las series de un mismo mes para distintos años son similares, a veces, aparecen meses lo suficientemente diferentes como para que las series así construidas dejen de ser estables en media y varianza. Además, en el caso de series de exposición horaria no se presenta el problema apuntado por los autores citados respecto a la longitud de la serie, por lo que el no poder concatenar series mensuales de distintos años no supone ningún inconveniente.

Otra aportación importante de este capítulo es la utilización del operador diferencias que nos permite eliminar la "pequeña" tendencia diaria que todavía persistía en las series del índice  $X_n$ . Si bien este operador ya había sido utilizado para series de exposición diaria de radiación global, ver Boileau, /15/, presentaba un inconveniente a la hora de abordar el problema de generación de nuevas series que, según el autor hacía desaconsejable su utilización. Como describiremos en

---

el apartado siguiente este inconveniente se puede resolver de manera sencilla.

Para las series diferenciadas del índice  $X_t$  hemos propuesto y parametrizado un modelo (para, aproximadamente, un 83% de las series). Se trata de un modelo ARMA multiplicativo en el que la parte regular sigue un modelo autorregresivo de orden uno y la parte estacional sigue un modelo de medias móviles de orden uno. Para proponer este modelo nos hemos basado en los correlogramas simple y parcial de las series.

Estos correlogramas nos indican dos hechos que tienen aplicabilidad en las series de exposición horaria de radiación global:

- El valor de exposición de radiación global que se da en una hora depende fuertemente del valor que se dió la hora anterior.
- La perturbación que se introduce en los valores de radiación en una hora está muy relacionada con la que se produjo el día anterior a esa misma hora, y no guarda gran relación con la que se produjo en días anteriores.

La segunda observación tiene relación con un hecho que ya apuntamos en el capítulo tres de este trabajo. Las perturbaciones o variaciones que se producen en los niveles de radiación en un día dependen de la hora del día considerada.

## **5. MÉTODO DE GENERACIÓN DE SERIES DE EXPOSICIÓN HORARIA**

Una vez determinado el modelo con el que se pueden identificar las series diferenciadas del índice  $X_h$  propuesto (obtenidas a partir de las series de exposición horaria de radiación global) y estimado sus parámetros, vamos a proponer un método que nos permita generar nuevas series teniendo en cuenta este modelo ARMA propuesto. El interés por conseguir un método adecuado y eficaz de generar nuevas series es doble:

- Primero porque con él podremos, utilizando sólo unos pocos parámetros de entrada, generar nuevas series con las mismas propiedades que las originales. Desde el punto de vista del diseño, dimensionado y simulación de instalaciones de aprovechamiento de la energía solar, es mucho más cómodo utilizar modelos que trabajar con series de datos experimentales, difíciles de obtener o inexistentes en muchos casos (necesarios si se quiere



tener en cuenta las propiedades estadísticas de la radiación a largo plazo).

- Segundo porque si conseguimos definir qué propiedades son las que caracterizan las series de exposición horaria de radiación global y qué relación guardan éstas con los parámetros del modelo ARMA propuesto (estudio que abordaremos en el capítulo sexto) podremos generar series de exposición horaria para localidades en las que no existan este tipo de medidas. Esto tiene una gran importancia a la hora de la extrapolación de resultados de los sistemas solares, de manera que se podría acometer el diseño de estos sistemas, para estas localidades sin necesidad de disponer de datos medidos de radiación, y lo que es más importante, hacerlo con una gran fiabilidad debido a que las series generadas contienen toda la información estadística significativa de las series reales.

Así, por ejemplo, el dimensionado de los sistemas de energía solar fotovoltaica se podrá realizar utilizando métodos numéricos, mucho más precisos que los métodos analíticos y simplificados, ver por ejemplo Gordon, /27/, o Egido, /21/. Esto es importante cuando se trata de sistemas de gran potencia, ya que mediante los métodos numéricos se puede realizar una optimización energética y/o económica de los sistemas, ver por ejemplo Sidrach, /58/ y /59/.

Por esto el objetivo de este apartado será encontrar un método que nos permita la generación de series de exposición horaria de radiación global.

Para la generación de nuevas series de exposición horaria de radiación global utilizaremos el siguiente procedimiento

- (i) Generar la serie  $D_r X_h$ , serie de valores diferenciados, a partir de los coeficientes del modelo ARMA, calculados para cada mes.
- (ii) Integrar las series  $D_r X_h$  generadas, para obtener las series  $X_h$ . En

la integración analizaremos la influencia de los valores horarios que se suponen para las primeras horas.

- (iii) Calcular las series  $G_h$  a partir de las series  $X_h$ ; para ello utilizaremos la expresión de  $G_{h,\max}$  propuesta.

De las tres tareas que será preciso realizar en la simulación o generación de nuevas series, la segunda será la más importante. Como ya veremos, en la integración de las series diferenciadas será necesario tener en cuenta la naturaleza de las series originales. En los apartados siguientes se detallan cada uno de los pasos a seguir.

## 5.1 GENERACIÓN DE LAS SERIES DEL ÍNDICE DsXh

Para generar la serie DsXh necesitamos conocer, para cada mes, el valor del coeficiente autorregresivo de la parte regular de la serie y el del coeficiente de medias móviles de la parte estacional, así como el valor de la varianza del término de error. En el apartado de estimación de los parámetros del modelo explicamos cómo se obtienen los mismos. En el Anexo II se muestran estos valores para todas las series mensuales en las que el modelo ARMA multiplicativo ha resultado ser el adecuado.

A partir de estos valores, y puesto que el modelo propuesto es un ARMA(1,0)xARMA(0,1), la expresión que vamos a utilizar es:

$$(1-\phi_1L)w_t=(1-\tilde{\theta}_1L^s)a_t \quad (5.1)$$

donde  $(1-\phi_1L)$  es el polinomio en retardos del modelo autorregresivo de la parte regular y  $(1-\tilde{\theta}_1L^s)$  es el polinomio en retardos del modelo de medias móviles de la parte estacional. Desarrollando esta expresión nos queda:

donde  $a_t$  son variables aleatorias independientes e idénticamente distribuidas

$$w_t = \phi_1 w_{t-1} - \tilde{\theta}_1 a_{t-s} + a_t \quad (5.2)$$

(i.i.d.) que siguen una distribución normal con media 0 y varianza la de la serie residual. De manera práctica, como se puede observar, para poder calcular la serie  $D_t X_h$  es necesario suponer un valor inicial de la misma  $\delta_0$ . En la simulación hemos utilizado varios valores, entre 0 y 1.

En las gráficas 5.1, 5.2 y 5.3 se representan las series  $D_t X_h$  obtenidas cuando se utiliza como valor inicial de la serie 0, 0.5 y 1, respectivamente.

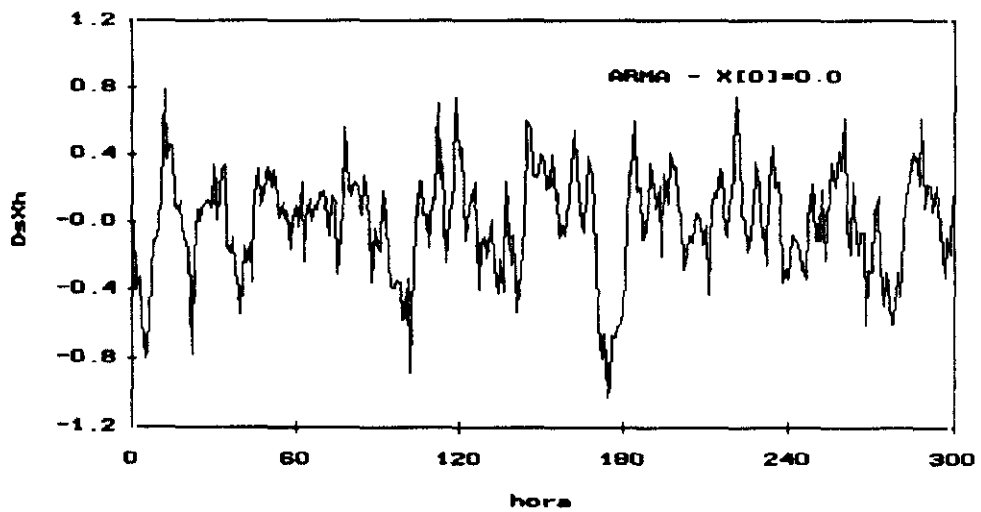


Figura 5.1 Serie  $D_t X_h$ , con  $\delta_0=0.0$ . Octubre-1977. Málaga

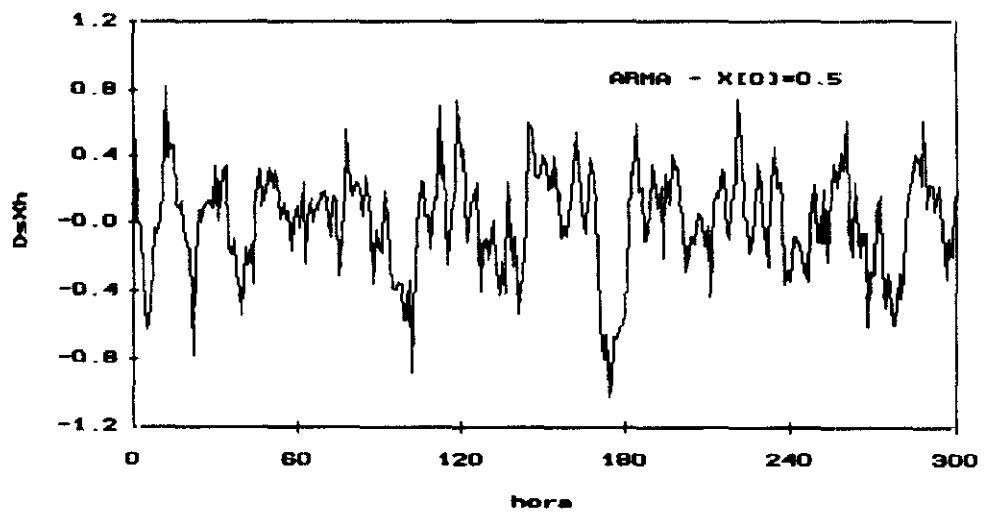


Figura 5.2 Serie  $D_s X_h$ , con  $\delta_0=0.5$ . Octubre, 1977. Málaga.

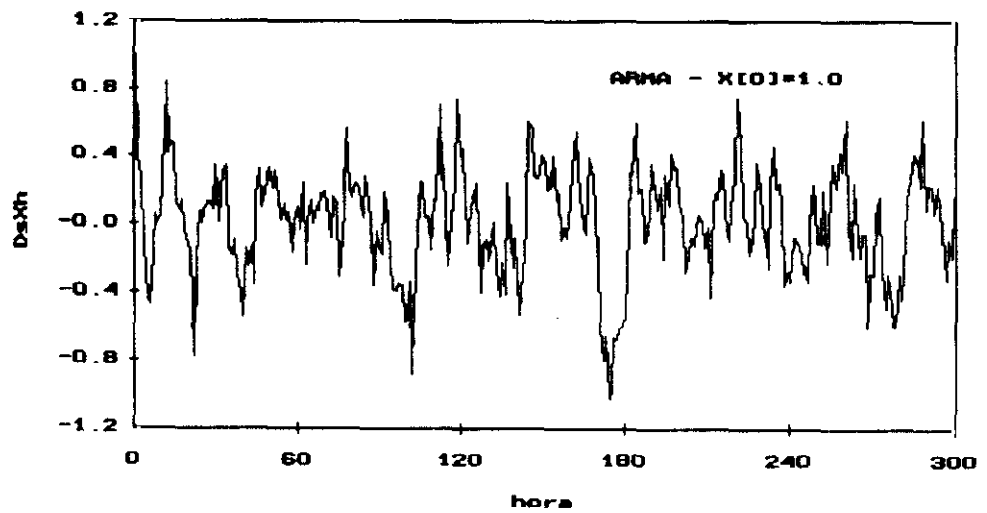


Figura 5.3 Serie  $D_s X_h$ , con  $\delta_0=1.0$ . Octubre, 1977. Málaga.

Como se puede observar en estas figuras, la serie obtenida es independiente del valor inicial que se utilice para la generación de la misma, en el sentido de que sólo en los primeros valores de la serie hay diferencias. De hecho si calculamos para cada una de estas series su valor medio y varianza, aplicando un contraste de igualdad de varianzas y medias obtenemos que las tres series tienen media cero e igual varianza (Estos contrastes se describen en el Anexo III). En concreto, y a modo de ejemplo, en la Tabla 5.I se dan estos dos valores para las tres series generadas.

	$X_0$	Xhm	Var
Serie 1	0.0	-0.003	0.096
Serie 2	0.5	0.010	0.096
Serie 3	1.0	0.003	0.094

Tabla 5.I Valor medio y varianza de las series  $\{D_s X_h\}$ , para distintos valores de  $\delta_0$ . (Parámetros del modelo ARMA para el mes de Octubre, 1977-Málaga)

## 5.2 INTEGRACIÓN

Para integrar las series  $D_s X_h$  obtenidas mediante el modelo ARMA multiplicativo propuesto, es necesario utilizar un total de  $s$  valores iniciales del índice  $X_h$ ; siendo  $s$  el número de horas por día consideradas para cada mes. Con estos primeros  $s$  valores de la serie  $X_h$  ya podemos utilizar la expresión:

$$X_h = D_s X_h + X_{h-s} \quad (5.3)$$

para calcular la serie  $X_h$ .

En las gráficas de las series  $D_s X_h$  del apartado anterior (figuras 5.1, 5.2 y 5.3), se puede observar que esta serie toma valores menores que cero en

---

muchos casos. Estos valores aparecen como consecuencia de aplicar el operador diferencias a las series originales. Los límites de oscilación de las mismas vendrán marcados por los parámetros del modelo ARMA. El hecho de que en esta serie aparezcan valores negativos hará que sí sea importante la forma en que se definan los primeros  $s$  valores de la serie  $X_n$ . De hecho, si utilizáramos, por ejemplo, los  $s$  primeros valores de la serie real, en muchos casos, obtendríamos valores negativos para el índice  $X_n$ , hecho totalmente inaceptable desde el punto de vista del índice que representan las series originales.

Este hecho ya fue observado por Boileau, /15/, para series diarias. El orden del operador diferencias que utilizaba era uno. Este autor, sin embargo, no presentaba ninguna solución para conseguir generar series simuladas que no presentaran valores negativos, por lo que no aconsejaba su utilización.

Por otra parte, las nuevas series deben presentar el mismo valor medio que las series originales.

Por estas dos razones, la forma de definir los  $s$  primeros valores de la serie será fundamental. En este trabajo se propone un método de integración de las series que resuelve los dos problemas anteriormente citados. Se trata de un procedimiento iterativo, en el que se parte de  $s$  valores que presentan una tendencia similar a la de los valores medios de las series reales (de tipo sinusoidal). A partir de éstos se crea una primera serie integrada. Para esta serie se comprueba que su valor medio sea similar al original (estadísticamente igual), así como que no haya ningún valor menor que cero. Si alguna de estas dos condiciones no se cumple se modifica la serie de los  $s$  primeros valores, teniendo en cuenta el valor medio de la que se ha generado y en qué horas se producen valores negativos.

El factor utilizado para incrementar o disminuir los  $s$  valores iniciales de

---

la serie original depende de la iteración en que nos encontremos. Inicialmente toma un valor igual a uno. En la segunda iteración a este factor se le suma (resta) un incremento (inicialmente este incremento toma el valor 0.1), si el valor medio obtenido en la primera iteración era menor que el real. Si el valor medio de la serie generada es menor que el de la serie real los  $s$  primeros valores se multiplican por este nuevo factor. En la tercera iteración se tienen en cuenta los valores medios obtenidos en las dos iteraciones anteriores. Si en las dos el valor medio mensual de la serie generada era menor (mayor) a los  $s$  primeros valores se les suma (resta) a factor el valor de incremento sin modificar. Si en la penúltima iteración el valor medio era mayor (menor) que el de la serie original y en la última el valor medio obtenido es menor (mayor) se le suma (resta) a factor el valor que tenía incremento dividido por dos (este pasa a ser el nuevo valor de incremento). Este proceso se repite hasta conseguir un valor medio similar al real (se puede admitir, por ejemplo, una variación entre ambos del cinco por ciento).

Además, los valores de la serie se distinguen en función de la hora del día a que corresponden. Así para la/s hora/s en que se haya presentado algún valor negativo a lo largo de la serie se incrementa el valor original que se tomó para esa hora (por ejemplo, en un diez por ciento)

En todos los casos analizados son muy pocos los valores negativos que aparecen en las series, y tras la primera o segunda iteración desaparecen. En el caso de que una vez alcanzado un valor medio similar al original, persistan valores negativos, se hacen éstos igual a cero (en ninguna de las series generadas con esta técnica se ha producido esta situación).

El procedimiento descrito es convergente puesto que el valor que se suma o resta va variando en función de los resultados previos. De hecho, en general, son suficientes de 3 a 4 iteraciones para generar cada serie mensual.

---

Por supuesto las nuevas series así generadas podrían volver a ser estacionarias sólo aplicándoles el operador diferencias de orden 's', y los parámetros que les corresponderían serían los mismos que los que se utilizaron para crearlas.

Con este método se consiguen, en definitiva, series  $X_n$  que presentan igual valor medio que las series originales.

Respecto a la varianza de estas series, en el Anexo III hemos desarrollado una expresión para calcularla a partir de los parámetros del modelo ARMA con que ha sido generada. Este valor hemos comprobado que es muy similar al estimador que se obtiene utilizando los valores de las series originales y generadas.

Una característica que presentaban la series originales era que para cada mes había una forma de función de distribución predominante. Estas funciones de distribución no eran, en general, normales. Pues bien, como se puede comprobar, con el método de integración que aquí proponemos se pueden generar funciones de distribución similares a las originales, ya que no existe "a priori", ninguna restricción respecto a la forma de ésta.

A modo de ejemplo en las gráficas 5.4 y 5.5 se muestran las series generada y real, respectivamente, para el mes de Febrero de 1977 para la localidad de Oviedo, y en las gráficas 5.6 y 5.7 se muestran estas mismas series para el mes de Julio de 1977, también para la localidad de Oviedo.



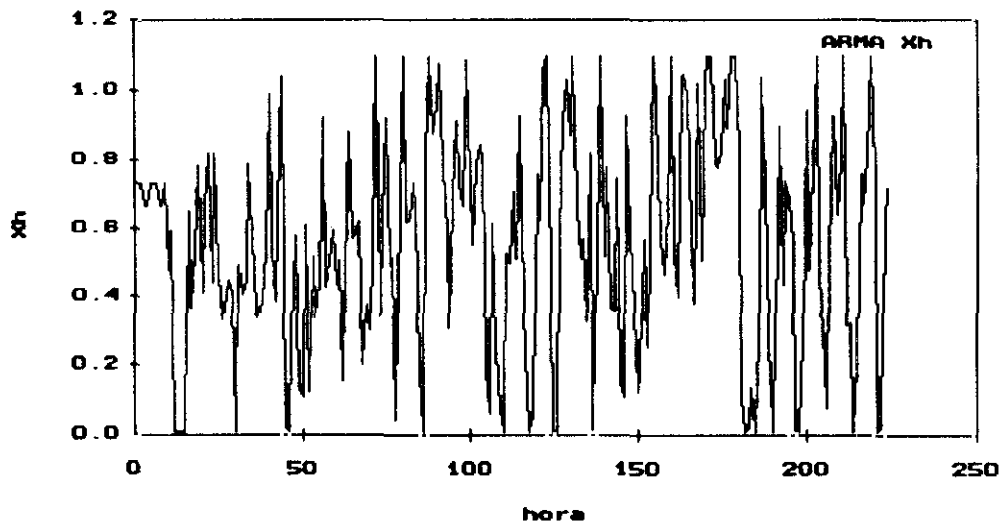


Figura 5.4 Serie Xh generada, para el mes de Febrero-1977. Oviedo.

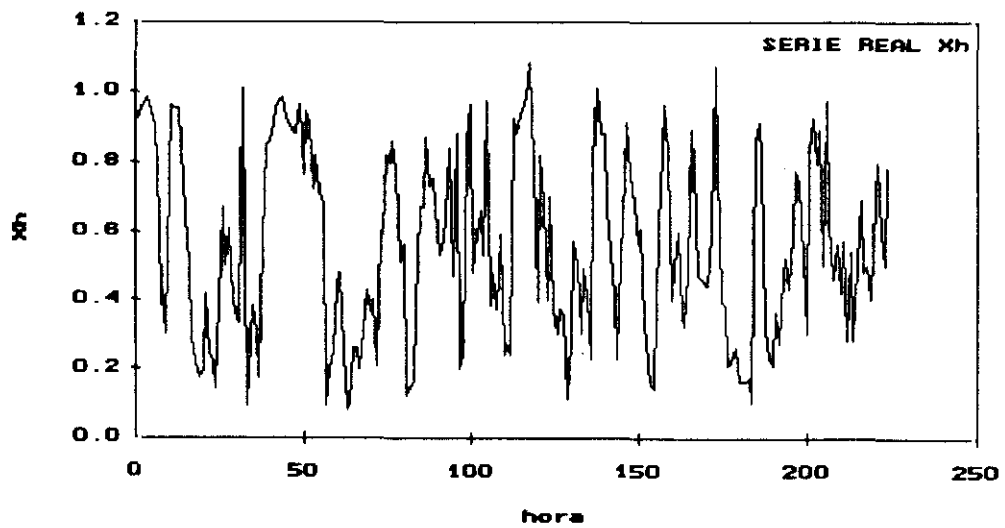


Figura 5.5 Serie Xh real, para el mes de Febrero-1977. Oviedo.

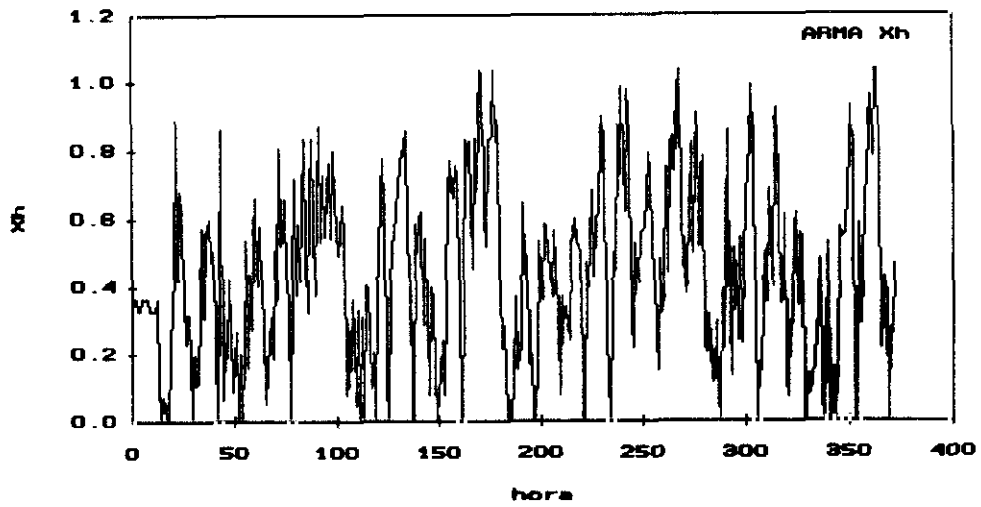


Figura 5.6 Serie  $X_h$  generada, para el mes de Julio-1977. Oviedo

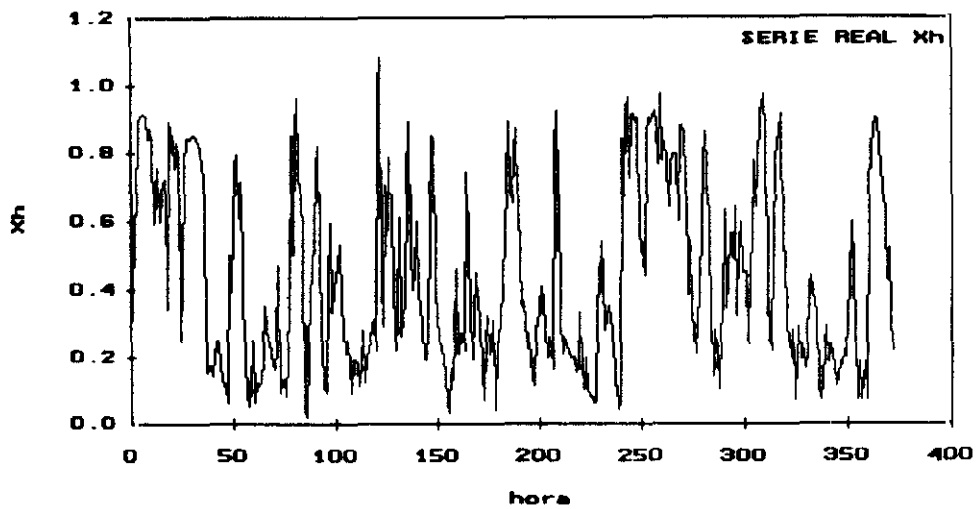


Figura 5.7 Serie  $X_h$  real, para el mes de Julio de 1977. Oviedo.

Respecto a las propiedades secuenciales de las series reales y generadas, también son similares. Por una parte, las funciones de distribución de probabilidades de ocurrencia de valores del índice  $X_h$  son en general similares; las funciones de autocorrelación simple y parcial, obtenidas tras diferenciar las

series (ya que las series diferenciadas son las estacionarias) son también similares.

### 5.3 CÁLCULO DE LAS SERIES DE EXPOSICIÓN HORARIA DE RADIACIÓN GLOBAL

A partir de la serie  $X_h$  generada mediante el modelo ARMA propuesto e integrada según el algoritmo descrito en el apartado anterior se genera la serie de valores de exposición horaria de radiación global. Para ello se utilizan los valores de exposición horaria máxima ya definidos. La expresión que nos da los valores de  $G_h$  es:

$$G_h = X_h G_{h,\max} \quad (5.4)$$

$$G_{h,\max} = A (\text{sen } \gamma)^B \quad (5.5)$$

En las gráficas 5.8, 5.9, 5.10 y 5.11 se muestran las series reales y las generadas mediante el modelo ARMA multiplicativo, de los valores de exposición horaria de radiación global para los meses de Agosto-79 y Febrero-79 para la localidad de Madrid. También para los mismos meses, en las figuras 5.12 y 5.13 se muestran los histogramas de estas series.

Los resultados que se muestran en estas gráficas se repiten para la mayoría de los meses y localidades analizados en los que el modelo ARMA multiplicativo ha funcionado. Naturalmente no es necesario que se reproduzcan en la serie generada la misma secuencia de la serie original; es suficiente, para la utilización que se hace de las series generadas, con que se reproduzcan todas las propiedades estadísticas, a saber: valor medio, varianza y propiedades secuenciales (ej.: número de veces que se dan, seguidos, distintos valores de exposición horaria, etc.). Este último hecho, se consigue ya que los parámetros que utilizamos en la simulación son los obtenidos para cada mes.

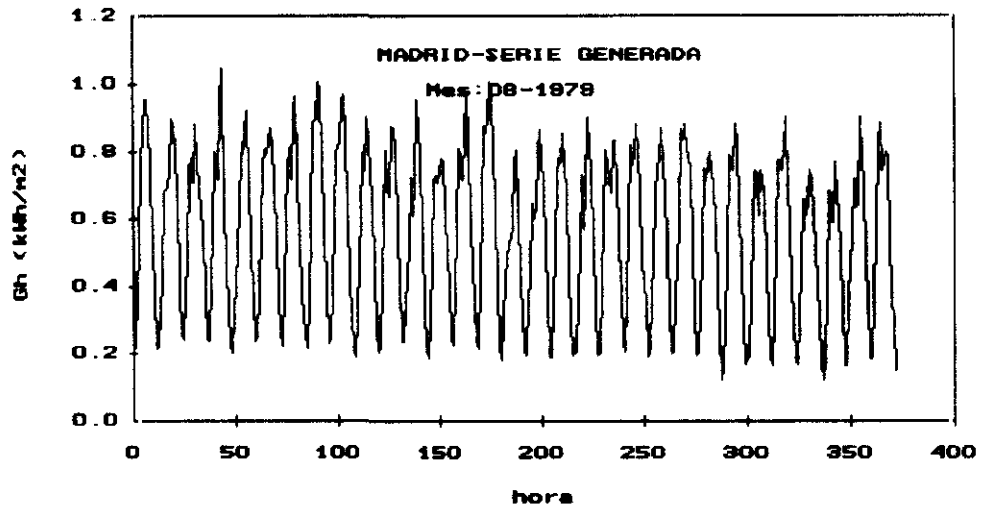


Figura 5.8 Valores de Gh generados, para el mes de Agosto-1979. Madrid.

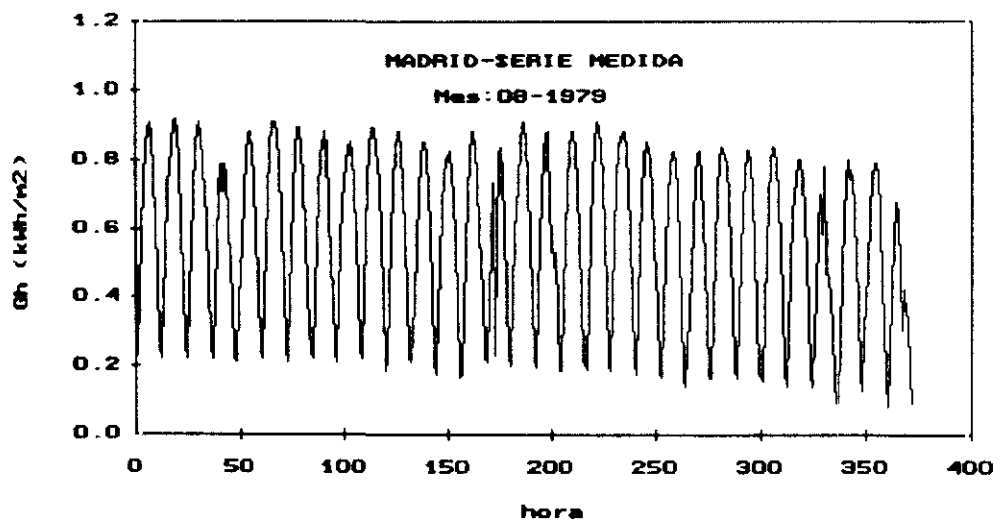


Figura 5.9 Valores de Gh reales, para el mes de Agosto de 1979. Madrid.

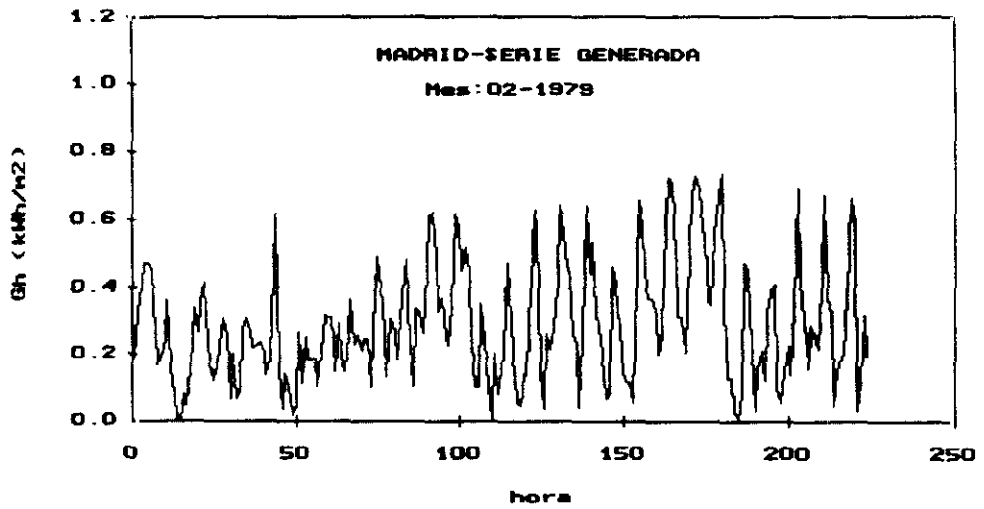


Figura 5.10 Valores de Gh generados, para el mes de Febrero de 1979. Madrid.

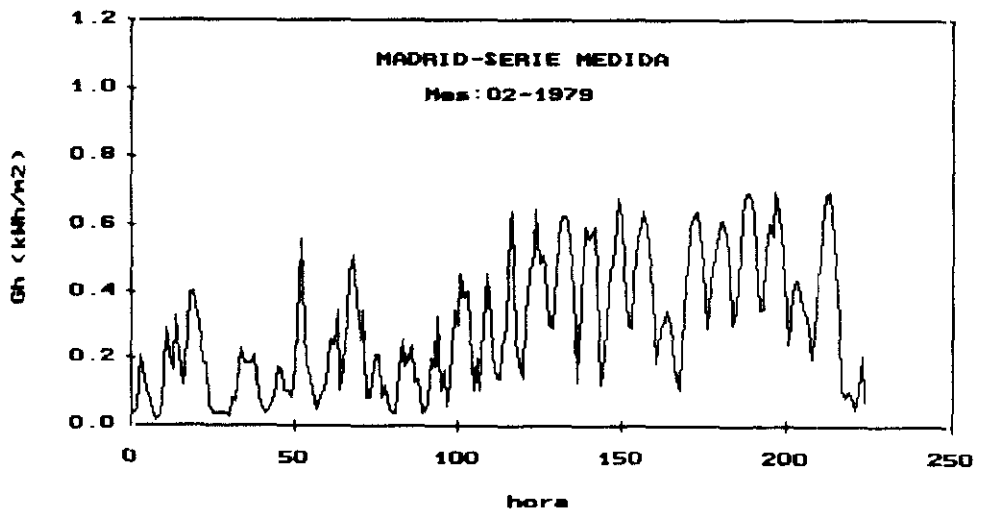


Figura 5.11 Valores de Gh reales, para el mes de Febrero-1979. Madrid.

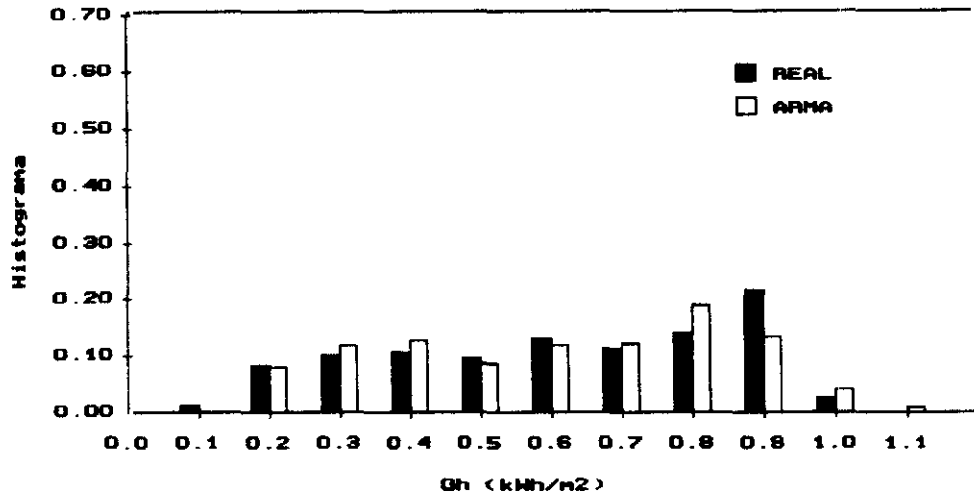


Figura 5.12 Histogramas de las series de exposición horaria de radiación global real y generada (ARMA), para el mes de Agosto-1979. Madrid.

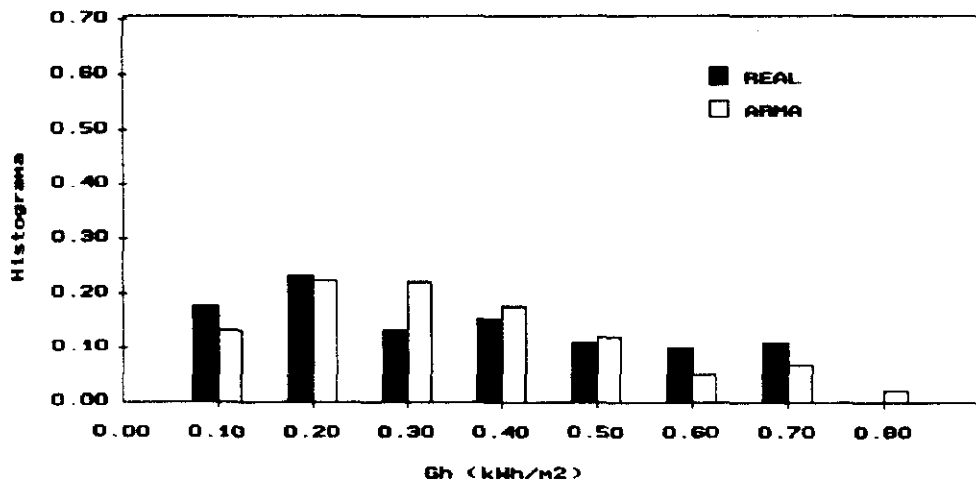


Figura 5.13 Histogramas de las series de exposición horaria de radiación global real y simulada, para el mes de Febrero-1979. Madrid

## 5.4 CONCLUSIONES

En este capítulo hemos visto cómo se pueden generar de una manera muy sencilla series de exposición horaria de radiación global, que tengan las mismas propiedades secuenciales que las originales, utilizando un modelo ARMA multiplicativo.

La utilización de modelos ARMA para la caracterización y generación de series de exposición horaria y diaria de radiación global estaba limitada porque con estos modelos sólo se pueden reproducir funciones de distribución de probabilidad de ocurrencia normales; sin embargo las series originales presentan funciones de distribución que, en general, no son normales. Esta dificultad ha sido comentada por diversos autores, ver por ejemplo Aguiar et al., /1/. Así para poder utilizar estos modelos era necesario transformar primero las series mediante técnicas gaussianas, ver por ejemplo Amato et al., /4/.

En este sentido una conclusión muy importante de este capítulo es que, al utilizar el operador diferencias, no es necesario que las series originales sean normales para poder utilizar modelos ARMA. Sí son normales una vez diferenciadas, y también las que corresponden a cada hora, resultado que coincide con el presentado por Aguiar et al., /2/.

Por otra parte, mediante el procedimiento de integración de series generadas con modelos ARMA, que aquí se describe, se consigue generar series cuya función de distribución no es normal.

Es decir, al utilizar el operador diferencias conseguimos, por una parte, eliminar la evolutividad diaria que había en las series, y por otra, no imponer la

condición de que las series originales sean normales. Además mediante la integración de las series generadas conseguimos funciones de distribución para las que no hay "a priori" ninguna restricción.

Otra aportación importante que destacamos es el método de integración de series diferenciadas que proponemos. Con éste método se resuelve el problema apuntado por Boileau, /15/, de la presencia de valores negativos en las series generadas.



## **6. PARÁMETROS QUE CARACTERIZAN LAS SERIES DE EXPOSICIÓN HORARIA DE RADIACIÓN GLOBAL**

En los apartados anteriores hemos encontrado y parametrizado el modelo con el que se pueden identificar las series de exposición horaria de radiación global. El periodo de análisis utilizado ha sido un mes, lo que nos ha llevado a tener que estimar todos los parámetros significativos, para cada mes de cada año de cada localidad, del modelo ARMA propuesto. Así, si bien hemos conseguido reproducir adecuadamente las distintas series originales, nos ha sido necesario determinar muchos estimadores para caracterizar las series de cada localidad. En este capítulo del trabajo vamos a ver qué tienen en común las series de los

distintos meses y localidades, con el ánimo de poder reproducir todas las series mensuales originales de cada localidad utilizando sólo los datos de que normalmente se puede disponer.

Los datos que con más frecuencia se pueden encontrar en los mapas o libros de radiación son valores medios de exposición diaria de radiación global. El índice de transparencia atmosférico se relaciona con estos valores a través de la radiación solar extraterrestre (fácilmente calculable). Por esta razón en los siguientes apartados analizaremos si existe relación entre éste y el valor medio mensual del índice  $X_h$ . Además analizaremos si existe relación entre los valores medios mensuales de  $X_h$  y otras características de las series  $X_h$ : época del año, coeficiente de orden uno del modelo autorregresivo de la parte regular, coeficiente de orden uno del modelo de medias móviles de la parte estacional y varianza residual.

Nuestro objetivo final es definir los parámetros mínimos que necesitamos conocer para generar series de exposición horaria de radiación global. Buscar y proponer relaciones entre distintas variables es sólo una forma de hallar herramientas que nos permitan conseguir este objetivo.

## **6.1 RELACIÓN ENTRE EL ÍNDICE DE TRANSPARENCIA ATMOSFÉRICO DIARIO Y EL ÍNDICE PROPUESTO $X_h$**

En primer lugar vamos a estudiar si existe algún tipo de relación entre el índice de transparencia atmosférica diario, en valor medio mensual ( $k_{d,m}$ ) y el índice propuesto, también en valor medio mensual ( $X_{h,m}$ ). Para ello vamos a representar en la figura 6.1 los valores medios mensuales de éstos para cada serie mensual de cada localidad. El valor medio mensual del índice de transparencia

diario lo calculamos como la media de los índices de transparencia diarios (éstos se calculan como el cociente entre exposición diaria de radiación global y radiación extraterrestre diaria).

### TODAS LAS LOCALIDADES

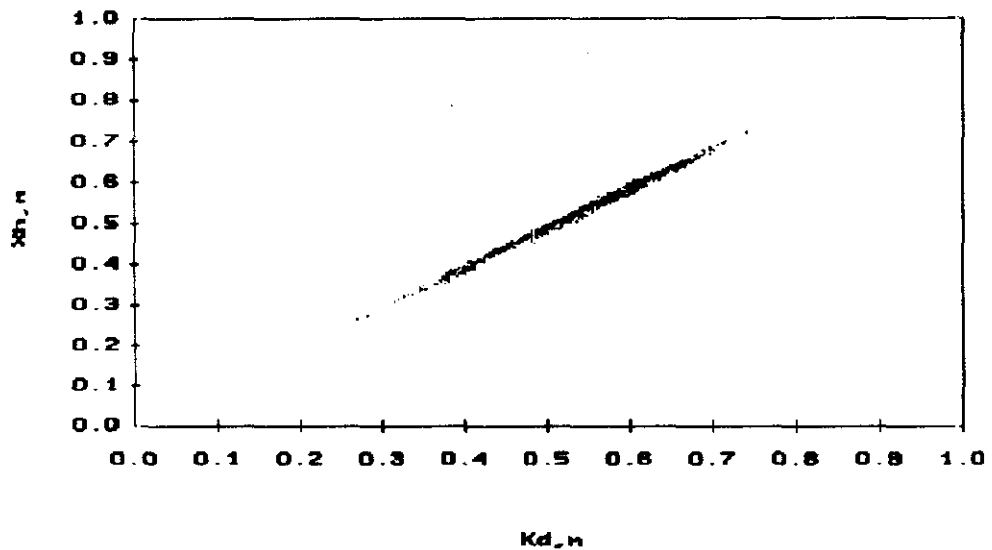


Fig. 6.1. Valores de  $K_{d,m}$  frente a los de  $X_{h,m}$

Como se puede observar en esta gráfica parece haber una clara relación lineal entre las dos variables representadas.

Basándonos en esta observación vamos a utilizar un modelo lineal para ajustar los datos. La expresión que utilizaremos será del tipo:

$$X_{h,m} = a + bK_{d,m} \quad (6.1)$$

Los coeficientes  $a$  y  $b$  los estimaremos mediante un análisis de regresión. En la tabla 6.I se muestran los resultados obtenidos para estos valores así como otros valores de interés del análisis efectuado.

Parámetro	Estimador	Error estándar
a	0.00125	0.00145
b	0.9677	0.00262

Nºobservaciones: 793  
 $R^2 = 99.45$

Coef.correlación: 0.997

Tabla 6.1

De acuerdo con estos resultados podemos afirmar que utilizando la *relación lineal propuesta en la ecuación (6.1)* y los valores, normalmente disponibles, del índice de transparencia atmosférico diario, en valor medio mensual, podemos predecir con gran precisión los valores medios mensuales de las series del índice  $X_h$ .

Desde el punto de vista físico esta relación tiene bastante significado. Pensemos, por ejemplo, en una situación extrema: valores muy altos del índice de transparencia diario medio mensual. Para los meses en que este índice tenga un valor muy elevado tendremos series en las que aparecerán muchos valores diarios, del índice de transparencia, muy altos; es decir, la atenuación de la atmósfera sobre la radiación extraterrestre será mínima y, podemos suponer, que habrá habido pocos días con nubes (como ya comentamos, los demás componentes sólo hacen que la atenuación en los valores de radiación recibidos sea como máximo de un 3 a un 5 por ciento mayor que para días claros). Este hecho, ausencia de nubes, hace, por otra parte, que la radiación recibida sea próxima o igual a la máxima esperable; es decir, si hay muchos días claros (valores elevados del índice de transparencia diario) será porque hay muchas horas (las de esos días) en las que se alcanzan valores próximos a los máximos. Este razonamiento podríamos hacerlo igual para valores bajos de estas dos variables.

## 6.2 RELACIÓN ENTRE EL ÍNDICE $X_{h,m}$ Y EL MES DEL AÑO

En este apartado vamos a analizar la relación que existe entre el valor medio mensual del índice  $X_h$ , que denominaremos  $X_{h,m}$ , y el mes del año (enero, febrero, etc.) a que corresponde. Vamos a utilizar para ello todos los datos disponibles de todas las localidades. Nuestro objetivo es ver en qué medida es biunívoca la relación entre las variables "mes del año" y "valor medio mensual del índice  $X_h$ ".

Estimaremos un modelo de regresión en el que tomaremos como variable dependiente las desviaciones de  $X_{h,m}$  de cada mes respecto a la media de todos los valores de  $X_{h,m}$ , es decir:

$$Y_t = X_{h,m}^t - X_{h,M} \quad (6.2)$$

donde  $t$  indica el mes, y  $M$  indica valor medio de los valores de  $X_{h,m}$ :

$$X_{h,M} = \frac{\sum_{t=1}^T X_{h,m}^t}{T} \quad (6.3)$$

siendo  $T$  el número total de meses.

Consideraremos doce variables independientes en este modelo de regresión:  $D_1, \dots, D_{12}$ , que serán variables ficticias (o "dummy"), definidas de manera que  $D_i$  es uno si la observación  $t$ -ésima corresponde al mes  $i$ -ésimo y cero en caso contrario. Así pues, el modelo de regresión que vamos a estimar es:

$$Y_t = \beta_1 D_{1t} + \dots + \beta_{12} D_{12t} + U_t \quad (6.4)$$

Los coeficientes del modelo nos indican si los valores medios del índice

$X_{h,m}$  son diferentes en cada mes del año. Por ejemplo, si fuese  $\beta_1 < 0$  ello querría decir que la media de  $X_{h,m}$  para las observaciones de Enero no es la misma que la media de  $X_{h,m}$  para todas las observaciones, sino inferior. Al mismo tiempo el coeficiente de determinación  $R^2$  de la regresión (6.4) nos dirá en qué medida la pertenencia de una observación a un mes o a otro determina el valor de  $X_{h,m}$ .

Así, si obtuviésemos que todos los coeficientes  $\beta$  de (6.4) son estadísticamente no significativos y que el valor de  $R^2$  es próximo a cero, ello querría decir que el mes a que pertenezca la observación no afecta al índice  $X_{h,m}$ .

Si, por el contrario, obtuviéramos que los coeficientes  $\beta$  son todos diferentes entre sí y que el valor de  $R^2$  es próximo a uno, ello querría decir que el mes a que pertenece la observación determina casi por completo el valor de  $X_{h,m}$ : dado un valor de  $X_{h,m}$  podríamos predecir con gran fiabilidad el mes a que pertenece la observación y viceversa.

Hemos estimado la ecuación (6.4) por mínimos cuadrados utilizando las 659 observaciones de las 10 localidades (Hemos eliminado aquellas observaciones para las que el modelo ARMA propuesto en el capítulo cuatro no funcionaba). Los resultados de la estimación se muestran en la tabla 6.II.

Observamos que existen tres tipos de comportamientos. Por una parte hay un grupo de meses (Noviembre, Diciembre, Enero y Febrero) que tienen un coeficiente que es menor que cero y significativamente diferente de cero (nivel de significación 0.05; el estadístico  $t$ , en valor absoluto, es mayor que 1.96); a este conjunto de meses nos referiremos, a partir de ahora, como grupo 1.

Por otra parte, los meses de Junio, Julio, Agosto y Septiembre tienen coeficiente significativamente mayor que cero; a estos meses los incluiremos dentro del denominado grupo 2. Además, se observa que los cuatro meses del

grupo 2 tienen coeficientes similares (excepto Julio), y lo mismo ocurre con los meses del grupo 1 (excepto Diciembre).

Los demás meses (Marzo, Abril, Mayo y Octubre) los clasificaremos dentro del grupo 3; en los meses de este último grupo puede aceptarse que el coeficiente correspondiente es cero con un nivel de significación  $\alpha=0.05$ .

Variable dependiente:  $X_{h,m}-0.522158$  (\*)

V.independientes	Coefficiente	Estadístico t
Enero	-0.045061	-4.8922
Febrero	-0.041097	-4.6035
Marzo	0.012383	1.3335
Abril	0.012315	1.2592
Mayo	0.012062	1.1760
Junio	0.047523	4.4922
Julio	0.069820	6.5293
Agosto	0.042182	4.2341
Septiembre	0.040646	4.1938
Octubre	-0.015831	-1.5740
Noviembre	-0.043267	-4.4243
Diciembre	-0.058640	-6.0505

Número de observaciones: 659

$R^2=0.2422$

(\*) Valor de  $X_{h,m}=0.522158$

Tabla 6.II

Para confirmar la validez de los resultados obtenidos y de la agrupación de meses propuesta a partir de las estimaciones hechas con todas las observaciones, hemos repetido este análisis para cada una de las diez localidades por separado. En todos los casos, excepto en Oviedo, se obtienen resultados

similares. Mostramos a continuación las estimaciones obtenidas para Madrid y Oviedo, tablas 6.III y 6.IV, respectivamente.

El diferente comportamiento de Oviedo hace pensar que se obtendrán mejores resultados si eliminamos las observaciones de esta variable. En efecto, si estimamos (6.4) utilizando las 574 observaciones restantes obtenemos el resultado que se recoge en la tabla 6.V

Variable dependiente:  $X_{h,m}-0.547962$  (\*) (Madrid)

V.independientes	Coefficiente	Estadístico t
Enero	-0.096629	-4.3783
Febrero	-0.073337	-3.8370
Marzo	0.017324	0.8478
Abril	0.016438	0.6799
Mayo	0.002705	0.1226
Junio	0.078871	3.5737
Julio	0.106705	4.8348
Agosto	0.087609	4.2877
Septiembre	0.060288	3.1543
Octubre	-0.034762	-1.4379
Noviembre	-0.058105	-2.8437
Diciembre	-0.097391	-4.7664

Número de observaciones: 78

$R^2=0.6674$

(\*) Valor de  $X_{h,M}=0.547962$ , considerando sólo los datos de Madrid

Tabla 6.III



Variable dependiente:  $X_{h,m}-0.399635$  (\*) (Oviedo)

V.independientes	Coefficiente	Estadístico t
Enero	-0.040778	-2.3552
Febrero	-0.006385	-0.3942
Marzo	0.022651	1.3082
Abril	0.042794	2.4716
Mayo	-0.031635	-1.6916
Junio	0.011740	0.7249
Julio	0.004651	0.2686
Agosto	-0.010135	-0.5419
Septiembre	0.044365	2.7393
Octubre	-0.005349	-0.3090
Noviembre	-0.012385	-0.7647
Diciembre	-0.035968	-1.9233

Número de observaciones: 85

 $R^2=0.2854$ (\*) Valor de  $X_{h,M}=0.399635$ , considerando sólo los datos de Oviedo

Tabla 6.IV

Variable dependiente:  $X_{h,m}-0.540301$  (\*)

V.independientes	Coefficiente	Estadístico t
Enero	-0.048156	-6.7029
Febrero	-0.047129	-6.7365
Marzo	0.008792	1.2125
Abril	0.007595	0.9876
Mayo	0.016585	2.0648
Junio	0.061853	7.2498
Julio	0.085366	10.0057
Agosto	0.046359	5.9650
Septiembre	0.042303	5.5008
Octubre	-0.016545	-2.0831
Noviembre	-0.045812	-5.8947
Diciembre	-0.064801	-8.6000

Número de observaciones: 574

 $R^2=0.43$ (\*) Valor de  $X_{h,M}=0.540301$ , sin considerar los datos de Oviedo

Tabla 6.V

Hay que observar que el coeficiente de determinación  $R^2$  ha pasado de un 24 a un 43%. En lo sucesivo todas las estimaciones serán, pues, con las 574 observaciones que corresponden a las restantes nueve localidades.

A la vista de los resultados de las tablas 6.II y 6.V cabe pensar que la agrupación de los 12 meses del año en 3 clases (tal y como se especificó más arriba) puede hacer que se obtengan resultados similares a los obtenidos en la estimación de la ecuación 1 utilizando únicamente tres variables independientes. Definamos, por tanto, las tres variables siguientes:

$$F_{1t} = \begin{cases} 1 & \text{si el mes } t \text{ es Enero, Febrero, Noviembre o Diciembre} \\ 0 & \text{en caso contrario} \end{cases}$$

$$F_{2t} = \begin{cases} 1 & \text{si el mes } t \text{ es Junio, Julio, Agosto o Septiembre} \\ 0 & \text{en caso contrario} \end{cases}$$

$$F_{3t} = \begin{cases} 1 & \text{si el mes } t \text{ es Marzo, Abril, Mayo u Octubre} \\ 0 & \text{en caso contrario} \end{cases}$$

El modelo estimado ahora es:

$$Y_t = \gamma_1 F_{1t} + \gamma_2 F_{2t} + \gamma_3 F_{3t} + U_t \quad (6.5)$$

Los resultados de la estimación (6.5) figuran en la tabla 6.VI

Variable dependiente:  $X_{h,m} - 0.540301$  (\*)

V.independientes	Coefficiente	Estadístico t
Grupo 1	-0.051311	-13.6950
Grupo 3	0.004317	1.0988
Grupo 2	0.057520	13.9344

Número de observaciones: 574  $R^2 = 0.4041$   
 (\*) Valor de  $X_{h,m}$ , sin considerar los datos de Oviedo

Tabla 6.VI

Observamos claramente que los meses de cada grupo tienen valores de  $X_{h,m}$  sensiblemente diferentes de los demás. Y, lo que es más importante, el  $R^2$  de esta regresión es del 40%, frente al 43% de la regresión con doce variables; es decir, con sólo estas tres variables se explica prácticamente el mismo porcentaje de variabilidad de  $X_{h,m}$  que con las doce variables anteriores.

Para comprobar que efectivamente los tres grupos determinados son homogéneos hemos estimado por separado ecuaciones de regresión similares a (6.4) pero utilizando sólo las observaciones correspondientes a cada uno de los grupos analizados, es decir, hemos estimado:

$$Y_t = \delta_1 D_{1t} + \delta_2 D_{2t} + \delta_{11} D_{11t} + \delta_{12} D_{12t} + U_t \quad \text{Grupo 1} \quad (6.6)$$

$$Y_t = \delta_6 D_{6t} + \delta_7 D_{7t} + \delta_8 D_{8t} + \delta_9 D_{9t} + U_t \quad \text{Grupo 2} \quad (6.7)$$

$$Y_t = \delta_3 D_{3t} + \delta_4 D_{4t} + \delta_5 D_{5t} + \delta_{10} D_{10t} + U_t \quad \text{Grupo 3} \quad (6.8)$$

donde en la estimación de la ecuación (6.6) utilizamos las 210 observaciones del grupo 1, en la ecuación (6.7) las 173 observaciones del grupo 2 y en la ecuación (6.8) las 191 estimaciones del grupo 3. Los resultados de estas estimaciones se recogen en las tablas 6.VII, 6.VIII y 6.IX.

Variable dependiente:  $X_{h,m}-0.488991$

V.independientes	Coficiente	Estadístico t
Diciembre	-0.013491	-0.2430
Noviembre	0.005498	0.0960
Enero	0.003154	0.0596
Febrero	0.004181	0.0811
Número de observaciones: 210		$R^2=0.0047$
(*) Valor de $X_{h,m}$ para meses Grupo 1		

Tabla 6.VII

Variable dependiente:  $X_{h,m}-0.597819$  (\*)

V.independientes	Coficiente	Estadístico t
Junio	0.004335	0.0539
Julio	0.027848	0.3465
Agosto	-0.011159	-0.1524
Septiembre	-0.015215	-0.2100
Número de observaciones: 173		$R^2=0.0002$
(*) Valor de $X_{h,M}$ , para meses Grupo 2		

Tabla 6.VIII

Variable dependiente:  $X_{h,m}-0.544618$  (\*)

V.independientes	Coficiente	Estadístico t
Marzo	0.004475	0.0735
Abril	0.003278	0.0508
Mayo	0.012268	0.1819
Octubre	-0.020862	-0.3128
Número de observaciones: 191		$R^2=0.0071$
(*) Valor de $X_{h,M}$ para meses Grupo 3		

Tabla 6.IX

Examinando estas tablas se observa que, dentro de cada uno de los grupos considerados, el mes a que pertenezca la observación es irrelevante para determinar el índice  $X_{h,m}$ , puede admitirse que todos los coeficientes son cero y todos los  $R^2$  son, también, aproximadamente, cero. Es decir, dentro de cada grupo el mes a que pertenezca la observación no explica en medida alguna la variabilidad del índice  $X_{h,m}$ .

En resumen, podemos decir que no es el mes a que pertenece una observación lo que es relevante a la hora de explicar el comportamiento de  $X_{b,m}$ , sino la clase a que pertenece la observación dentro de las tres antes consideradas (grupo 1/grupo 2/grupo 3): dentro de cada una de estas clases el comportamiento es similar. En cualquier caso, esta clasificación sólo explica aproximadamente, un 40% de la variabilidad del índice.

### **6.3 RELACIÓN ENTRE $K_{d,m}$ Y LOS PARÁMETROS DEL MODELO ARMA**

En este apartado vamos a analizar el tipo de relación que existe entre el valor medio mensual del índice de transparencia diario y el modelo ARMA de cada mes. El objetivo de este estudio es analizar, en qué medida, a partir del modelo ARMA de cada mes se puede determinar el valor medio del índice de  $K_d$  de ese mes. Puesto que los valores medios mensuales del índice  $X_d$  están normalmente disponibles, si encontramos qué tipo de relación existe entre estas variables, podremos directamente generar series del índice propuesto, y por tanto de exposición horaria, sólo con los valores del índice de transparencia diario.

Vamos a empezar analizando las gráficas del valor de  $K_{d,m}$  frente al coeficiente autorregresivo, al coeficiente de medias móviles y a la varianza residual. En las figuras 6.2, 6.3 y 6.4 se muestran estas gráficas.

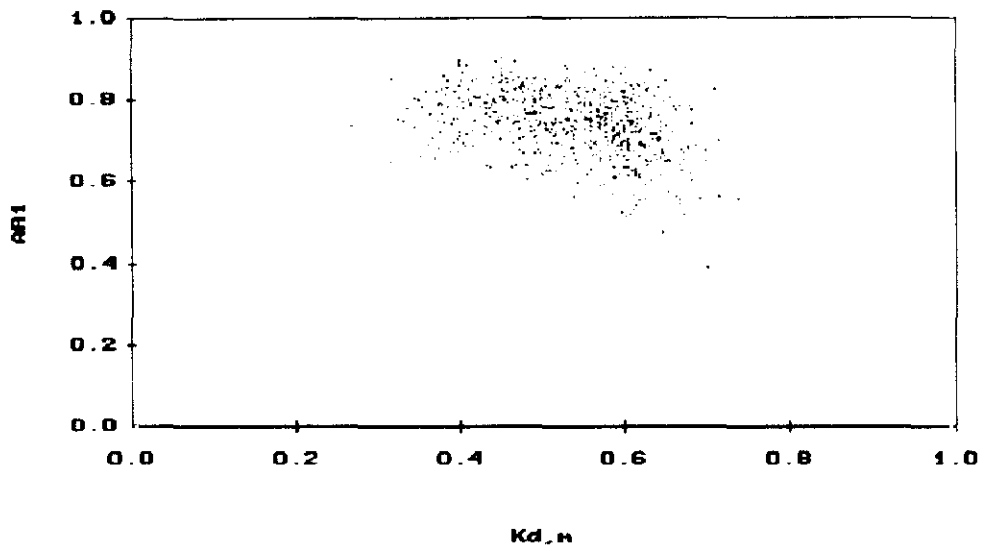


Figura 6.2. Valores medios mensuales del índice  $K_d$  frente a los valores del coeficiente autorregresivo.

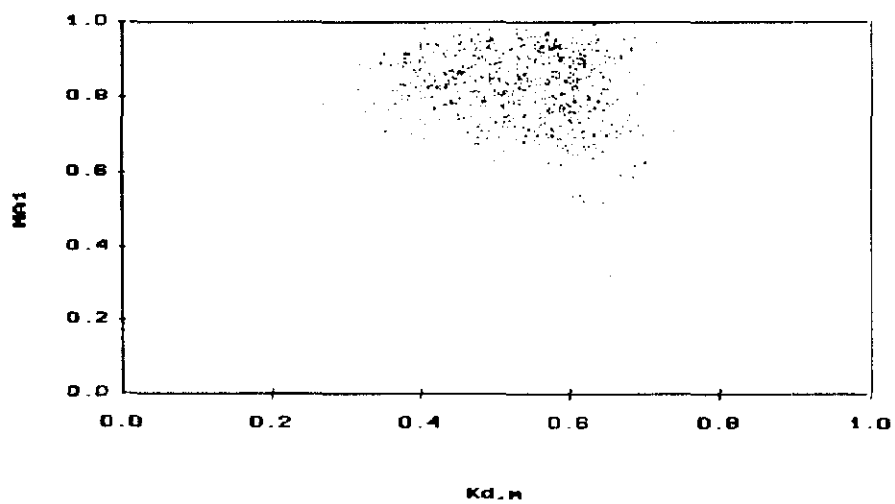


Figura 6.3. Valores medios mensuales del índice  $K_d$  frente a los valores del coeficiente de medias móviles.

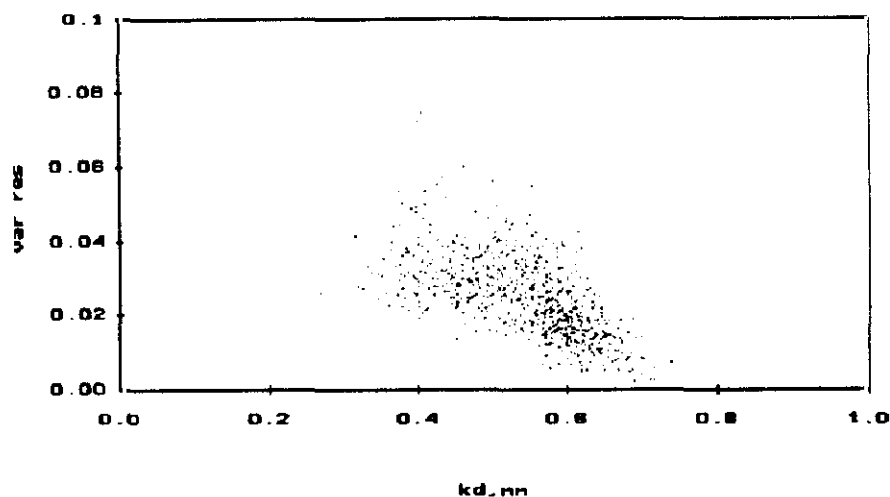


Figura 6.4. Valores medios mensuales del índice  $K_d$  frente a valores de la varianza residual.

De la observación de estas gráficas podemos concluir que el valor medio mensual del índice  $K_b$  no determina, por separado, cada una de las otras variables. Vamos por ello a utilizar un análisis de regresión conjunto de todas las variables (análisis multivariante).

El modelo que vamos a estimar es el siguiente:

$$Y_t = \beta_1 A_t + \beta_2 M_t + \beta_3 VAR_t + U_t \quad (6.9)$$

donde  $Y_t$ , variable dependiente, son las desviaciones del valor de  $K_{d,m}$  de cada mes respecto a la media de todos los valores de  $K_{d,m}$ ;  $A_t$  es el estimador del coeficiente AR menos la media de todos los coeficientes AR,  $M_t$  es el estimador del coeficiente MA menos la media de todos los coeficientes MA y  $VAR_t$  es el estimador de la varianza residual menos la media de todas las varianzas residuales. Obsérvese que todas las variables del modelo tienen media 0; esto no afectará al coeficiente  $R^2$  y permitirá una más fácil interpretación de los parámetros estimados. Al tener todas las variables media 0, no incluimos constante en el modelo de regresión.

Puesto que ya vimos en el apartado anterior que las observaciones de Oviedo tenían un comportamiento sensiblemente diferente al de las demás, vamos a estimar el modelo propuesto sin considerar los datos de esta localidad. El total de observaciones que vamos a utilizar será de 574. Con estas observaciones hemos obtenido los resultados que se muestran en la tabla 6.X

Variable dependiente:  $K_{d,m}-0.556$  (\*)

V.independientes	Coefficiente	Estadístico t
A	-0.293801	-10.689
M	-0.019041	-0.837
VAR	-4.512060	-24.411
Número de observaciones: 574		$R^2=0.56$

(\*) Valor de  $K_{d,m}$ , sin considerar los datos de Oviedo

Tabla 6.X

Como se puede observar el coeficiente autorregresivo y la varianza residual de las series son muy significativos; por el contrario el coeficiente de medias móviles no parece tener ninguna influencia.

Como el valor del coeficiente  $R^2$  no es muy alto, podemos también incluir en el modelo una agrupación por meses, puesto que ya hemos visto que éstos tenían influencia. Incluiremos las tres variables de agrupación definidas en el apartado anterior; el modelo que vamos a estimar ahora es:

$$Y_t = \beta_1 A_t + \beta_2 M_t + \beta_3 VAR_t + \beta_4 F_{1t} + \beta_5 F_{2t} + \beta_6 F_{3t} + U_t \quad (6.10)$$

donde  $A_t$ ,  $M_t$  y  $VAR_t$  son como en (6.9) y  $F_{it}$  ( $i=1,2,3$ ) son como en (6.5). Los resultados que hemos obtenido se muestran en la tabla 6.XI



Variable dependiente:  $K_{d,m}-0.556$  (\*)

V.independientes	Coefficiente	Estadístico t
A	-0.225721	-8.6780
M	0.003177	0.1559
VAR	-3.670299	-17.9163
F1 (Grupo 1)	-0.032679	-9.9532
F2 (Grupo 2)	0.020803	5.1953
F3 (Grupo 3)	0.017087	5.3112

Número de observaciones: 574

$R^2=0.65$

(\*) Valor de  $K_{d,m}$ , sin considerar los datos de Oviedo

Tabla 6.XI

Como cabía esperar las variables significativas son las mismas que obtuvimos en los anteriores análisis: grupo (época del año), coeficiente autorregresivo y varianza residual. El coeficiente de medias móviles sigue sin ser significativo. El coeficiente  $R^2$  se incrementa hasta un 65%.

Vamos ahora a analizar si estos resultados son homogéneos para todas las localidades.

En la tabla 6.XII se presenta el resultado de estimar las ecuaciones (6.9) y (6.10) utilizando cada vez los datos de una localidad.

Localidad	Nº de obs	R <sup>2</sup> de (6.9)	R <sup>2</sup> de (6.10)
Badajoz	69	0.69	0.76
Castellón	44	0.54	0.62
Logroño	36	0.67	0.80
Madrid	78	0.61	0.80
Málaga	82	0.67	0.72
Murcia	83	0.49	0.59
Oviedo	85	0.04	0.14
Palma de M.	78	0.62	0.75
Sevilla	59	0.52	0.56
Tortosa	45	0.57	0.63
Total	659	0.51	0.55
Total(*)	574	0.56	0.65

(\*) Sin considerar los datos de Oviedo

Tabla 6.XII

Como podemos observar existen algunas diferencias entre las distintas localidades. En cuanto a Oviedo, se observa que sólo un 14% de la variabilidad del índice  $K_{d,m}$  puede explicarse con las variables utilizadas en el modelo de regresión (grupo, coeficiente autorregresivo y coeficiente de medias móviles). Esto es, la relación entre las variables utilizadas y el valor medio de las series es muy pequeña. Por otra parte nos encontramos con otro grupo de localidades para las que la relación es significativa, con un valor de  $R^2$  en torno al 60% (Castellón, Murcia, Sevilla y Tortosa). Por último, para las demás localidades (Badajoz, Logroño, Málaga, Madrid y Palma de Mallorca) el valor de  $R^2$ , mayor que 0.7, nos indica que existe una relación muy significativa.

## 6.4 TIPOS DE MESES EN FUNCIÓN DEL MODELO ARMA ASOCIADO

En el apartado anterior hemos visto la relación que existe, para cada serie mensual del índice  $K_d$ , entre el valor medio de ésta, la época del año y su modelo ARMA asociado.

Aunque para todas las localidades, excepto Oviedo, hemos encontrado que existe una relación significativa entre estas variables, el número de parámetros con el que hemos trabajado ha sido muy grande; así, las características que hemos encontrado significativas para cada mes de cada localidad han sido: su media, la época del año a que corresponde la serie y dos de los tres parámetros del modelo ARMA asociado.

Puesto que nuestro objetivo es caracterizar las series de exposición horaria de radiación global (entre otras cosas, para ser capaces de generar nuevas series) utilizando las características definidas, vamos ahora a estudiar si se pueden reducir/agrupar los distintos modelos ARMA encontrados. El objetivo de este apartado será definir cuántos "tipos de modelos ARMA" necesitamos para describir todos los modelos ARMA calculados para cada todos los meses de todas las localidades. Es decir, vamos a intentar simplificar la información que tenemos de cada localidad. Para ello vamos a utilizar los contrastes de igualdad de varianzas y parámetros estimados en el modelo ARMA propuesto que se describen en el Anexo III.

El método de trabajo que vamos a seguir es buscar, mediante los contrastes estadísticos adecuados, los meses que tienen similar modelo ARMA asociado y proponer unos parámetros únicos para estos meses. Haremos primero una agrupación de los meses en función de las varianzas residuales del modelo

---

ARMA correspondiente. Fijándonos en éstas, vemos que, si exceptuamos los datos de Logroño, el rango de variación, para las demás localidades oscila entre los valores de 0.006 y 0.045.

Las series mensuales correspondientes a la localidad de Logroño presentan, en general, varianzas mayores; de hecho, hay 16 meses de un total de 36 que tienen una varianza residual mayor que 0.045. Puesto que este hecho sólo ocurre para esta localidad, para poder agrupar todas las series mensuales de la misma sería necesario incluir un rango de variación de la varianza residual mucho mayor, lo que nos llevaría a tener que proponer más tipos de modelos ARMA, pero calculados éstos utilizando muy pocos meses de datos. Esto es, estos tipos de modelos ARMA no serían representativos de muchos meses. Para poder proponer tipos de modelos que sean más generales, no agruparemos esos meses con varianza residual mayor que 0.045.

Vamos a dividir el rango de varianzas especificado en cinco intervalos. En cada intervalo vamos a definir una "marca" o varianza residual del intervalo, de manera que todas las varianzas que estén dentro del intervalo pueda admitirse que son iguales a esta "marca". Vamos a utilizar un nivel de significación del contraste de igualdad  $\alpha=0.05$  para la varianza residual, por ser éste el parámetro que mayor significación tenía en el modelo de regresión propuesto. El nivel de significación del contraste de igualdad que vamos a utilizar para los otros parámetros del modelo ARMA será de 0.01. En este caso vamos a utilizar un valor menor del nivel de significación porque lo que pretendemos es ver qué meses se parecen: no es necesario que sean exactamente iguales. Además por la naturaleza de las series (series que reflejan sobre todo presencia/ausencia de nubes) podemos considerar que tenemos más evidencia a favor de la hipótesis nula; es decir, más evidencia a favor de que las series sean similares. De hecho si tenemos en cuenta nuestra "experiencia diaria" no existen tantos tipos de meses distintos como hemos estimado en un capítulo anterior (659 en total).

---

Dentro de cada uno de estos intervalos analizaremos el rango de variación del coeficiente autorregresivo y del coeficiente de medias móviles. En función de éstos, y utilizando el contraste de igualdad de parámetros del modelo ARMA descrito en el Anexo III obtendremos tres grupos; en cada uno de ellos habrá, como para la varianza residual, una "marca del coeficiente autorregresivo" y una "marca del coeficiente de medias móviles", de forma que todos los coeficientes autorregresivos y de medias móviles de meses que pertenezcan a ese subintervalo podrá admitirse que son iguales a las marcas del subintervalo. Dentro de cada intervalo de varianzas las marcas de los coeficientes autorregresivo y de medias móviles (de orden uno) que definen cada uno de los grupos se han seleccionado de manera que sea máximo el número de meses que quede comprendido en alguno de los tres grupos. (Obsérvese que un mes dado puede quedar fuera de todos los grupos: si no se acepta la hipótesis de que sus coeficientes coinciden con los que definen cada uno de los grupos).

El proceso de optimización para determinar las marcas de los coeficientes autorregresivo y de medias móviles de cada grupo, se ha llevado a cabo por un procedimiento de búsqueda con rejilla ("grid search"). Para cada rango de varianza hemos utilizado cuarenta valores distintos para el coeficiente autorregresivo de la parte regular y otros tantos para el coeficiente de medias móviles de la parte estacional; es decir, un total de 1600 modelos diferentes. Los valores de los coeficientes los hemos obtenido partiendo del valor mínimo observado para los estimadores de ambas partes (primer valor) e incrementándolos en 0.01 para obtener cada nuevo valor.

De los 1600 modelos analizados hemos seleccionado tres, con el criterio de que el total de meses que sean similares a alguno de esos tres modelos sea mayor; es decir, hemos seleccionado los tres modelos con los que conseguimos "agrupar" el mayor número de series mensuales de todas las que tienen una varianza residual similar a la de la marca del intervalo. (Hay que hacer notar que

todas las observaciones o series mensuales, pertenecen a más de uno de los 1600 modelos propuestos).

El procedimiento lo hemos repetido para los cinco rangos de varianza señalados. El número de "tipos de modelos ARMA" que hemos obtenido de esta manera es 15. Tres grupos para cada uno de los cinco intervalos de varianza residual. Las marcas de cada uno se muestran en la tabla 6.XIII. En la primera columna aparece la varianza residual del modelo (VAR), en la segunda el coeficiente autorregresivo (AR) y en la tercera el coeficiente de medias móviles (MA).

	VAR	AR	MA
TIPO 1	0.0084	0.620	0.645
TIPO 2	0.0084	0.720	0.920
TIPO 3	0.0084	0.820	0.970
TIPO 4	0.0121	0.670	0.895
TIPO 5	0.0121	0.745	0.720
TIPO 6	0.0121	0.820	0.920
TIPO 7	0.0175	0.670	0.895
TIPO 8	0.0175	0.745	0.720
TIPO 9	0.0175	0.820	0.920
TIPO 10	0.0249	0.720	0.720
TIPO 11	0.0249	0.745	0.870
TIPO 12	0.0249	0.820	0.945
TIPO 13	0.0361	0.670	0.845
TIPO 14	0.0361	0.745	0.795
TIPO 15	0.0361	0.770	0.945

Tabla 6.XIII

Como antes se señaló el que un mes pertenezca a uno de estos grupos querrá decir que puede admitirse, con nivel de significación  $\alpha=0.01$  (0.05 para la varianza residual), que los parámetros del modelo ARMA estimado para ese mes coinciden con los que definen cada grupo. Por ejemplo, para el mes Diciembre de 1979 de la localidad de Málaga, se obtuvo:

$$\begin{aligned} AR &= 0.666 & stdAR &= 0.049 \\ MA &= 0.627 & stdMA &= 0.055 \\ \sigma^2 &= 0.027 \end{aligned}$$

Este mes queda dentro del grupo 10, porque con estos datos puede aceptarse la hipótesis nula en un contraste:

$$\begin{aligned} H_0: & \sigma^2 = 0.0249 \\ H_1: & \sigma^2 \neq 0.0249 \end{aligned}$$

con un nivel de significación  $\alpha=0.05$ , y puede también aceptarse, con un nivel de significación  $\alpha=0.01$ , la hipótesis nula del contraste:

$$\begin{aligned} H_0: & \phi_1 = 0.72 \quad \bar{\theta}_1 = 0.72 \\ H_1: & \phi_1 \neq 0.72 \quad \bar{\theta}_1 \neq 0.72 \end{aligned}$$

Utilizando estos tipos de modelos ARMA propuestos para las 659 series mensuales de todas las localidades consideradas, conseguimos agrupar 559 series (aproximadamente, un 85%); es decir, del total de observaciones analizadas, puede admitirse que 559 son iguales, al menos, a uno de los modelos propuestos. Las demás observaciones, o bien quedan fuera del rango de varianzas propuestas (49 observaciones), o bien no se puede aceptar la hipótesis nula del contraste para ningún modelo (51 observaciones).

En la tabla 6.XIV se recogen los resultados obtenidos para cada localidad. En la primera columna (N<sup>o</sup>M) aparece el número de meses analizados; en la

segunda (AGR) el número de meses que en los que la hipótesis nula del contraste ha resultado ser cierta para, al menos, uno de los modelos ARMA propuestos, esto es el número de meses que hemos conseguido "agrupar" en alguno de los modelos ARMA propuestos; en la tercera (NO-V) el número de series (o meses) con varianza residual fuera (mayor o menor) de los intervalos propuestos; en la cuarta (NO-A) el número de meses para los que la hipótesis nula del contraste no ha resultado ser cierta, esto es, meses que no se pueden identificar con ninguno de los quince modelos ARMA propuestos; y en la quinta columna (%AGR) el porcentaje de meses que son iguales (con nivel de significación  $\alpha=0.01$ ) a uno, al menos, de los modelos propuestos, respecto al total de series con varianza residual dentro de los intervalos analizados.

LOCALIDAD	NºM	AGR	NO-V	NO-A	%AGR
BADAJOS	69	59	4	6	90.8
CASTELLÓN	44	37	1	6	86.0
LOGROÑO	36	19	16	1	95.0
MÁLAGA	82	72	5	5	93.5
MADRID	78	67	7	5	93.1
MURCIA	83	76	2	5	93.8
OVIEDO	85	63	8	14	81.8
P.MALLORCA	78	75	1	2	97.4
SEVILLA	59	52	3	4	92.9
TORTOSA	45	40	2	3	93.1
TOTAL	659	559	49	51	91.6

Tabla 6.XIV

Como podemos observar en esta tabla, la misma información, desde el punto de vista estadístico, que tenemos en 559 series (de un total de 610 que tienen varianza residual dentro de alguno de los intervalos analizados) la podemos conseguir utilizando sólo quince modelos ARMA. Si nos fijamos en el porcentaje



---

que este número de observaciones (meses o series) representa respecto al total que teníamos en los intervalos de varianza considerados, vemos que la mayoría de las observaciones (más de un 91%) es igual a alguno de estos modelos. Las observaciones que no son similares a ninguno de los modelos ARMA es porque dentro de su intervalo de varianza, o bien el coeficiente de la parte regular o bien el de la parte estacional, o ambos, no puede aceptarse que sea(n) iguales a los de alguno de los modelos propuestos para ese intervalo.

Con esta "agrupación" hemos conseguido reducir notablemente el número de parámetros necesarios para caracterizar todas las series de cada localidad (cada modelo ARMA lo definimos, como ya hemos visto, a partir de tres parámetros).

Una vez seleccionados los tipos de modelos ARMA con los que podemos representar la mayoría de las series mensuales de datos disponibles, vamos ahora a dar un paso más en el propósito de la caracterización de las series que estamos utilizando. Para ello, vamos a plantearnos una pregunta que surge como consecuencia de lo estudiado hasta el momento: puesto que ahora podemos trabajar con sólo 15 modelos ARMA, ¿cómo se distribuyen éstos para cada localidad en las distintas épocas del año observadas? Otra cuestión que se encuentra incluida en ésta es ver si existirán unos tipos de modelos ARMA que se presenten de forma predominante en cada época del año, para una o más localidades. Vamos a dedicar el siguiente apartado de este capítulo a realizar estos análisis.

## 6.5 CARACTERIZACIÓN Y GENERACIÓN DE SERIES ANUALES DE EXPOSICIÓN HORARIA DE RADIACIÓN GLOBAL

Hasta ahora hemos analizado, por una parte, todas las series mensuales de cada localidad y propuesto para la mayoría de las mismas un modelo ARMA con el que identificarlas y poder reproducirlas. Además hemos visto que el número de modelos ARMA necesarios para caracterizar la mayor parte de las series puede reducirse a quince. Por otra parte hemos estudiado las relaciones entre las distintas características de las series mensuales: valor medio del índice de transparencia diario,  $K_{d,m}$ , valor medio de  $X_h$ , mes del año (grupo de meses) y modelo ARMA propuesto.

Como resultado de este segundo análisis hemos comprobado que, en función de la localidad considerada, mediante las variables mencionadas podemos explicar entre un 10 y un 80% de la variabilidad del índice  $K_{d,m}$  (índice de transparencia atmosférico diario).

Nuestro objetivo ahora es determinar cómo puede unirse toda esta información para lograr caracterizar las series de radiación a largo plazo. Esto es, definir qué es lo que "ha ocurrido y qué se puede esperar que ocurra" durante varios meses, años, en lo que se refiere a valores de exposición horaria recibidos sobre la superficie de la tierra.

Una forma de resolver este problema es mediante la caracterización (y posterior generación) de series de varios años. Se pretende que en las series (o en los parámetros que se den de las mismas) queden recogidos todos los fenómenos secuenciales que se hayan presentado. Así, por ejemplo, cuando se hace una simulación de una instalación de energía solar es deseable poder incluir

---

todas las situaciones posibles que se han dado y que pueden darse (número seguidos de horas o días nublados, etc.). Este tipo de información es necesaria cuando se quiere, por ejemplo, realizar el dimensionado de un sistema de energía solar fotovoltaica utilizando métodos numéricos, o cuando se quiere simular el comportamiento de un sistema mixto fotovoltaico-generador auxiliar, ver por ejemplo Sidrach et al., /42/.

Con un elevado número de parámetros de entrada podríamos generar series anuales para cualquiera de las localidades analizadas; sólo tendríamos que generar para cada mes una nueva serie utilizando los parámetros del modelo correspondiente y concatenando los meses de manera correlativa (primer mes de enero del primer año seguido de primer mes de febrero de primer año, etc.). Sin embargo, esto tiene una relativa utilidad, ya que es demasiada la información que necesitamos conocer.

Lo que pretendemos es utilizar un número de parámetros de entrada mínimo para poder generar series anuales.

Para las localidades que presentan para la regresión (6.10) un valor de  $R^2$  significativo (todas excepto Oviedo), podemos suponer que, conocido el valor medio mensual del índice de transparencia diario,  $K_{d,m}$ , correspondiente a un determinado mes del año, podemos determinar el modelo ARMA que tiene asociado. El grado de fiabilidad de esta hipótesis va a depender del valor de  $R^2$  que hemos obtenido para cada localidad. Para las localidades en las que el coeficiente de  $R^2$  es elevado, existe una fuerte relación entre la variable dependiente y las variables independientes de (6.10).

Así, para estas localidades sólo conociendo los valores medios mensuales del índice de transparencia atmosférico diario podemos llegar a caracterizar y generar series de exposición horaria. El proceso a seguir para generar las mismas

sería:

- (i) Partir de los valores medios mensuales del índice de transparencia atmosférico diario,  $K_{d,m}$ , normalmente disponibles.
- (ii) Utilizando la regresión (6.1) calcular el valor medio mensual del índice  $X_h$
- (iii) Determinar, en función del mes (época del año) y del valor medio del índice de transparencia diario, el modelo ARMA asociado.
- (iv) Generar la serie diferenciada:  $D_h X_h$ , utilizando la expresión (5.2)
- (v) Integrar este serie, utilizando el procedimiento descrito en 5.2.
- (vi) Calcular los valores de  $G_h$  utilizando la expresión (5.4).

Para poder determinar el modelo ARMA asociado a cada valor medio del índice  $K_{d,m}$  (paso iii) vamos a utilizar los coeficientes de la expresión (6.10). Utilizando estos coeficientes, la media de todas las observaciones, la media de los coeficientes de los modelos autorregresivo y de medias móviles y la media de las varianzas residuales, y la expresión (6.10) obtenemos la tabla 6.XVII. En esta tabla aparece, para cada valor medio de índice de transparencia atmosférico diario,  $K_{d,m}$ , y cada mes (grupo del año) el modelo ARMA que le corresponde. Con esto queda, en parte, especificado el paso (iii). Decimos que sólo en parte porque, como se puede observar, existen, para algunos valores medios del índice  $K_{d,m}$  y Grupo más de un modelo. Esto es debido, a que con las variables independientes utilizadas no se explicaban, totalmente la variabilidad de los valores medios del índice utilizado. Una forma de resolver este problema es utilizar simultáneamente una tabla que recoja la frecuencia con que han aparecido, para las observaciones utilizadas, los distintos tipos de modelos ARMA en cada uno de los grupos de meses definidos.

Modelo	F1(Grupo1)	F2(Grupo2)	F3(Grupo3)
Tipo 1	0.605	0.655	0.659
Tipo 2	0.584	0.633	0.637
Tipo 3	0.561	0.611	0.615
Tipo 4	0.581	0.631	0.635
Tipo 5	0.564	0.614	0.618
Tipo 6	0.548	0.597	0.601
Tipo 7	0.562	0.611	0.615
Tipo 8	0.544	0.594	0.598
Tipo 9	0.528	0.578	0.582
Tipo 10	0.523	0.572	0.576
Tipo 11	0.517	0.567	0.571
Tipo 12	0.501	0.550	0.554
Tipo 13	0.493	0.543	0.547
Tipo 14	0.476	0.526	0.530
Tipo 15	0.471	0.521	0.525

Tabla 6.XVII

Cuando a un valor medio de  $K_{d,m}$ , para un mismo grupo, le correspondan dos modelos distintos, la selección se hace teniendo en cuenta la frecuencia con que ha aparecido cada uno de esos modelos; de manera que para la serie de años que se está generando se reproduzca un porcentaje similar.

En la tabla 6.XVIII se muestran las probabilidades (%) con que se han presentado los distintos modelos para cada Grupo (época) definido.

Modelo	F1(Grupo1)	F2(Grupo2)	F3(Grupo3)
Tipo 1	0.0	5.9	0.0
Tipo 2	0.0	8.2	0.0
Tipo 3	0.0	2.2	0.0
Tipo 4	0.5	10.4	0.6
Tipo 5	2.1	17.0	1.8
Tipo 6	0.5	5.2	1.8
Tipo 7	1.6	11.1	8.5
Tipo 8	3.2	14.8	7.3
Tipo 9	8.0	2.2	9.1
Tipo 10	4.3	8.1	15.1
Tipo 11	13.3	5.9	13.9
Tipo 12	21.3	3.7	4.9
Tipo 13	2.7	3.0	18.2
Tipo 14	22.9	1.5	12.7
Tipo 15	19.7	0.7	6.1

Tabla 6.XVIII

La interpretación física de estas distribuciones, teniendo en cuenta el significado de la varianza residual en el modelo propuesto, y la naturaleza de las series (o fenómeno físico que representan), es que, en general, los grupos que incluyen los meses meses con mayor presencia de nubes (1 y 3) presentan mayores valores de varianza residual. La aparición de nubes provoca mayores perturbaciones en los valores de exposición horaria. Por el contrario, en los meses de verano, en general con menos nubes, las perturbaciones que se introducen son menores. Como ya hemos indicado este comportamiento lo presentan, en general, todas las localidades, excepto Oviedo: para esta localidad no existen esas diferencias tan claras entre los meses de los distintos Grupos.

---

Para la localidad de Oviedo, el análisis efectuado nos lleva a proponer otro tipo de tratamiento. Por una parte obtuvimos que no existía correlación significativa entre el mes del año y los valores medios mensuales del índice propuesto. Por otra, tampoco obtuvimos una correlación significativa entre el índice de transparencia atmosférico, medio mensual, y los parámetros de los modelos ARMA. Esto es, para esta localidad, no existen periodos distintos de otros, en lo que a los parámetros que estamos analizando se refiere. Por ello, para simular series de datos de exposición horaria para varios años, nos será suficiente, conocer, por una parte, la frecuencia con que aparecen los distintos valores medios del índice  $X_{h,m}$  a lo largo del año, y por otra, la frecuencia con que se presentan los distintos modelos ARMA. Las series generadas habrán de mantener estas mismas distribuciones.

Para esta localidad, se puede observar que del total de meses que podían identificarse con un modelo ARMA multiplicativo (85), mediante los tipos de modelos ARMA definidos sólo conseguimos clasificar 63 (8 quedaban fuera del rango de varianzas y 14, aún teniendo varianza similar a la de algún modelo, no tenían similar coeficiente autorregresivo y/o de medias móviles). Esto nos sugiere que, quizá, para esta localidad (y para localidades con climas similares) con otros tipos de modelos ARMA habríamos recogido mejor la información estadística de las series originales. El problema es que, sólo con los datos de esta localidad, los tipos de modelos ARMA que seleccionaríamos serían poco representativos. Sería, *por tanto, necesario utilizar datos de otras localidades con clima similar.*

El diferente comportamiento de Oviedo lo podemos explicar si tenemos en cuenta otro tipo de variables climatológicas. Siguiendo, por ejemplo, el trabajo de M. Pardo et al., /49/, en el que se determinan las distintas áreas climáticas de la península, en función de los valores de exposición solar, precipitación y humedad, Oviedo pertenece a la denominada zona A, mientras que el resto de las localidades analizadas pertenece a la zona C.

De acuerdo con las diferencias encontradas entre la localidad de Oviedo y las demás, podemos afirmar, que, si bien el modelo ARMA propuesto parece servir para todas las localidades, el método con el que se van a generar nuevas series, cuando se pretende utilizar unos pocos parámetros, no es universal. Es decir, va a depender del clima de la localidad en estudio. Una forma de poder decidir "a priori", si una localidad concreta (distinta de las estudiadas) presenta una correlación significativa para la regresión (6.10), será, por tanto, utilizar otro tipo de parámetros climáticos que nos digan si es o no similar a las estudiadas. Un ejemplo podría ser la utilización de la clasificación de M.Pardo, ya mencionada.



---

## 6.6 CONCLUSIONES

El objetivo que nos planteamos en este capítulo era determinar qué parámetros de todos los definidos para las series de exposición horaria son los que necesitamos utilizar para caracterizar suficientemente las mismas. Teniendo en cuenta este objetivo, las conclusiones que podemos obtener del análisis efectuado en este capítulo son las siguientes:

- El número de parámetros que necesitamos para caracterizar y generar series de exposición horaria para una localidad va a depender del "tipo de clima" de esa localidad. Para las localidades analizadas, hemos encontrado dos tipos: por una parte Oviedo y, por otra, las demás localidades. Para conocer a cuál de estos grupos pertenecería una localidad para la que no haya datos horarios, es necesario utilizar otro tipo de variables climatológicas. En este sentido, con la clasificación propuesta por M.Pardo, /49/, sólo conociendo la zona a la que pertenece la localidad considerada podríamos saber el número de parámetros que necesitamos. (La agrupación de localidades que se presenta en el citado trabajo explica la agrupación que nosotros hemos obtenido utilizando las propiedades de las series de radiación horarias, a saber, relación entre el índice de transparencia atmosférico diario, en valor medio mensual, y las demás variables de análisis utilizadas).

- Para las localidades que presentan una correlación significativa entre el valor medio mensual del índice de transparencia atmosférico diario y las otras variables (parámetros del modelo ARMA, época del año), utilizando únicamente el valor medio mensual del índice de transparencia atmosférico diario se tienen caracterizadas en gran medida tanto las series mensuales como las anuales (periodos de varios años) de la mayoría de las localidades. Esto se explica teniendo en cuenta los resultados obtenidos:

---

primero porque el valor medio mensual de las series del índice  $X_n$  definido se puede relacionar con el valor medio mensual del índice de transparencia atmosférico diario utilizando el modelo de regresión que proponemos en (6.1), con un coeficiente de correlación mayor del 0.99; segundo porque la variabilidad del valor medio del índice  $K_{a,m}$  queda explicada en un alto porcentaje utilizando las siguientes características de las series: época del año (grupo) y tipo de modelo ARMA asociado al mes. En cada localidad, cuanto más alto es el coeficiente  $R^2$  de la tabla 6.XVII, mejores resultados proporcionará este método, debido a que la relación entre estas variables es más estrecha.

- Para las localidades que no presentan una correlación significativa entre el índice de transparencia diario, medio mensual, y las demás variables utilizadas (únicamente Oviedo), conociendo la distribución del primero ( $K_{a,m}$ ) y la de los modelos ARMA, para varios años (independientemente del mes o grupo considerado) quedan caracterizadas las series de exposición horaria. Del mismo modo, teniendo en cuenta estas distribuciones se pueden generar series para varios años. Para este tipo de localidades será necesario utilizar más series para asegurar que la selección de los tipos de modelos ARMA es la mejor.

Aunque, como ya hemos comentado, harían falta más datos de localidades con clima similar al de Oviedo que nos permitieran contrastar mejor los resultados que aquí se presentan, nos parece interesante destacar la metodología que hemos empleado: tanto para la caracterización de las series como para la selección de los tipos de modelos ARMA (tipos de meses) que hemos propuestos.

En nuestro análisis no hemos utilizado nunca la varianza de las series, esto ha sido debido, como se recoge en el Apéndice III, a que este parámetro está relacionado con los parámetros del modelo ARMA propuesto.

## 7. CONCLUSIONES

Los dos objetivos fundamentales que nos planteamos al principio de este trabajo eran:

- Caracterizar las series de exposición horaria de radiación global utilizando un número mínimo de parámetros.
- *Encontrar un método que nos permitiera generar series de estos valores utilizando, sólo, datos normalmente disponibles.*

Respecto al primer objetivo, en las series de exposición horaria hemos encontrado dos propiedades fundamentales, que nos han servido para determinar el modelo con el que se pueden identificar. Estas dos características son:

- El valor de exposición para una hora está muy relacionado con el valor que se dió la hora anterior.
- La perturbación que se produce en una hora del día está relacionada con la que se produjo el día anterior a esa misma hora.

Para llegar a estas conclusiones hemos desarrollado y utilizado una serie de herramientas que describimos a continuación:

- Hemos caracterizado los valores de exposición horaria máxima observados en todas las localidades analizadas a partir de una única expresión que es

---

sólo función de la altura solar para cada hora.

- Hemos propuesto una técnica de modificación de las series originales para eliminar tendencias que ha resultado ser adecuada, ya que en las series modificadas (que hemos llamado  $X_{it}$ ) ya no aparecen las tendencias y componentes cíclicas observadas en las series originales.
- Para el análisis y caracterización de estas series modificadas hemos utilizado la teoría de procesos estocásticos, o series temporales. Hemos propuesto una sencilla modificación de las series descritas que nos ha permitido obtener, a partir de series con funciones de distribución de probabilidad no necesariamente normales, series que sí presentaban funciones de distribución normales. Esta modificación se basa en la utilización del operador diferencias. Con esto hemos resuelto el problema apuntado por algunos autores (ver por ejemplo Aguiar et al., /1/, Amato et al., /4/) que presentaba la utilización de modelos ARMA en la caracterización y generación de series de exposición horaria o diaria de radiación global. Los resultados de este análisis nos han servido para extraer las dos conclusiones que, respecto a la naturaleza de la radiación solar, hemos descrito anteriormente.

Basándonos en las dos características de las series de exposición horaria mencionadas hemos propuesto, para identificar y reproducir estas series, un modelo ARMA multiplicativo de tipo  $ARMA(1,0) \times (0,1)_s$ , en el que la parte regular sigue un modelo autorregresivo de orden uno y la parte estacional un modelo de medias móviles también de orden uno.

Hemos hecho también un estudio de qué periodo es el más adecuado para "dividir", o estudiar, las series anuales de exposición horaria de radiación global. Hemos obtenido que el periodo temporal adecuado para el análisis de estas series es un mes. Otra conclusión de este análisis es que no parece adecuada una

---

utilización de "meses globales" similar a la propuesta por Gordon et al., /28/, para valores diarios, puesto que sería necesario modificar previamente las series ya que no son estacionarias. Esta modificación supondría la utilización de más parámetros de entrada, normalmente no disponibles.

Basándonos en los resultados de los capítulos 3 y 4 hemos propuesto una forma de generar series mensuales que "a priori" no impone ninguna restricción sobre la forma de las funciones de distribución de probabilidad de ocurrencia de los valores de radiación. En este sentido, consideramos una aportación muy interesante el método de integración de series diferenciadas que proponemos en este trabajo. Con el mismo se resuelve el problema señalado por E.Boileau, /15/, de la presencia de valores negativos en las series generadas.

La información que hemos obtenido en los capítulos precedentes la hemos conseguido simplificar en el capítulo 6. Las aportaciones más importantes del mismo son:

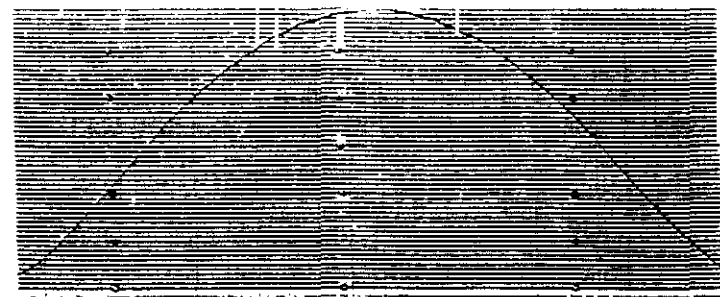
- Las características más significativas de las series de exposición horaria de radiación global han resultado ser: valor medio mensual del índice  $X_h$ , época del año (grupo) y tipo de modelo ARMA asociado.
- Hemos analizado la relación que existe entre estas características y los valores medios mensuales del índice de transparencia. Los resultados obtenidos nos muestran que no existe una relación universal, sino que, en función del tipo de clima, podemos distinguir dos tipos de localidades. Por una parte localidades en las que aparece como variable significativa la época del año (diferenciamos tres épocas). Estas localidades presentan una correlación significativa entre el valor medio mensual del índice de transparencia diario y los parámetros de los modelos ARMA y la época del año. Por otra parte, hemos encontrado localidades (sólo Oviedo) para las que esta relación no es significativa. Además, no parece tener influencia

**ANEXO I**

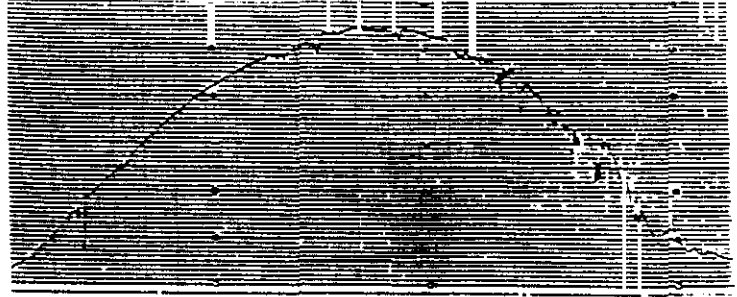
**CLASIFICACIÓN DE DÍAS TIPO EN FUNCIÓN DE LOS VALORES  
DE RADIACIÓN GLOBAL INSTANTÁNEOS RECIBIDOS.**

**(Datos del LIER-CIEMAT)**

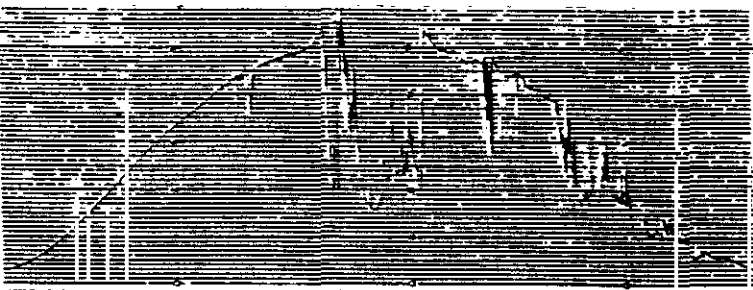
TYPICAL DAYS



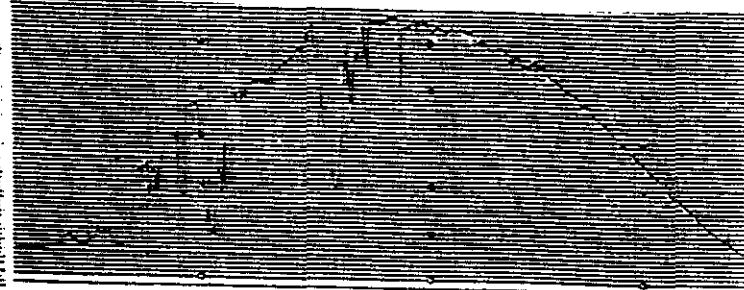
Number of days 62 (20%)



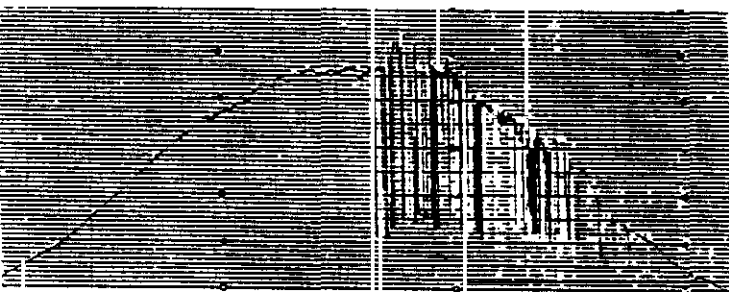
Number of days 58 (19%)



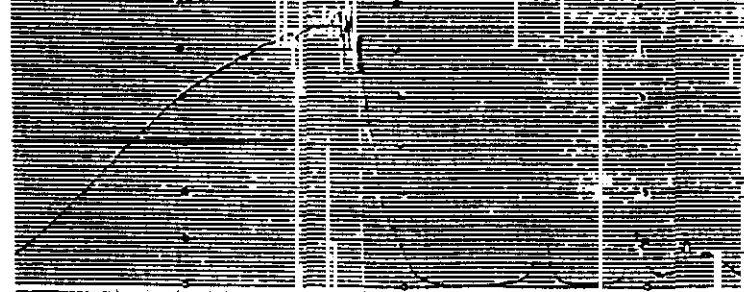
Number of days 33 (10%)



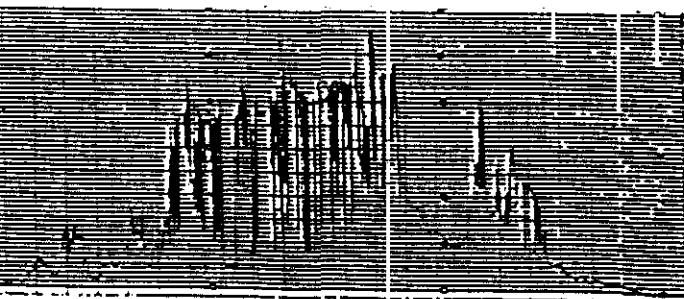
Number of days 14 (6%)



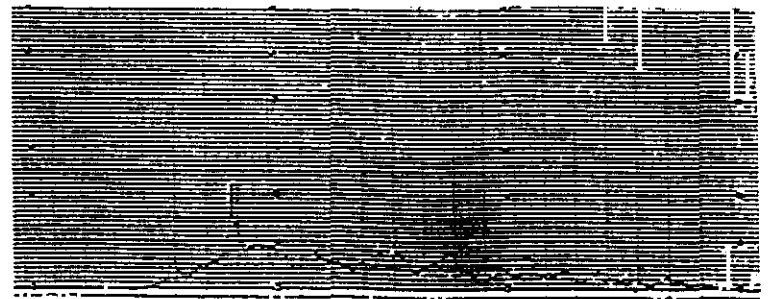
Number of days 35 (11%)



Number of days 16 (5%)



Number of days 58 (19%)



Number of days 35 (10%)

**ANEXO II**  
**TABLAS CON LOS PARÁMETROS ESTIMADOS DEL**  
**MODELO ARMA PROPUESTO**



En las siguientes tablas se muestran los parámetros estimados del modelo ARMA de todas las series mensuales utilizadas. Los datos los hemos agrupado por localidades y años. Para cada año hemos utilizado doce columnas, correspondientes a los doce meses del año. La información que se recoge, por filas, es, este orden, coeficiente autorregresivo de orden uno (AR1), coeficiente de medias móviles de orden uno (MA1) y varianza residual (res). Para los meses en los que no había datos, o el modelo ARMA propuesto no ha dado resultados aceptables, hemos utilizado dos rayas para indicar cualquiera de estas dos circunstancias.

BADAJOS

Año: 1976

AR1	0.822	0.691	0.757	0.779	--	0.736	0.717	--	--	0.625	--	0.653
MA1	0.758	0.733	0.887	0.920	--	0.940	0.902	--	--	0.787	--	0.885
res	0.012	0.037	0.026	0.033	--	0.020	0.007	--	--	0.045	--	0.033

Año: 1977

AR1	0.693	--	0.730	0.766	0.795	--	0.804	0.613	0.741	--	0.823	0.779
MA1	0.800	--	0.938	0.822	0.859	--	0.726	0.749	0.889	--	0.862	0.791
res	0.035	--	0.019	0.023	0.018	--	0.012	0.008	0.011	--	0.025	0.037

Año: 1978

AR1	0.738	0.615	0.698	0.810	0.792	--	0.789	0.644	0.608	--	0.817	0.712
MA1	0.777	0.731	0.646	0.807	0.831	--	0.805	0.782	0.607	--	0.899	0.825
res	0.038	0.041	0.024	0.031	0.026	--	0.002	0.005	0.011	--	0.020	0.031

Año: 1979

AR1	0.757	0.805	0.778	--	0.752	--	0.688	0.541	0.621	0.751	0.806	0.823
MA1	0.832	0.841	0.741	--	0.924	--	0.776	0.701	0.773	0.741	0.761	0.931
res	0.028	0.030	0.033	--	0.013	--	0.014	0.005	0.016	0.032	0.018	0.025

BADAJOSZ (...)

Año: 1980

AR1	0.625	0.890	0.723	0.842	0.733	--	--	0.879	0.706	0.820	0.750	0.742
MA1	0.809	0.948	0.718	0.729	0.799	--	--	0.941	0.929	0.978	0.898	0.757
res	0.037	0.026	0.036	0.016	0.023	--	--	0.011	0.014	0.018	0.029	0.015

Año: 1981

AR1	0.556	0.803	0.701	0.745	--	--	0.645	0.590	0.726	0.684	0.650	0.786
MA1	0.854	0.928	0.817	0.805	--	--	0.585	0.318	0.903	0.565	0.820	0.796
res	0.015	0.019	0.035	0.041	--	--	0.009	0.010	0.026	0.016	0.012	0.027

Año: 1982

AR1	--	0.725	0.721	0.747	0.715	0.728	0.730	0.762	0.804	0.670	0.797	--
MA1	--	0.773	0.516	0.753	0.821	0.735	0.791	0.863	0.935	0.903	0.933	--
res	--	0.035	0.013	0.026	0.013	0.020	0.013	0.009	0.021	0.018	0.026	--

Año: 1983

AR1	--	--	--	--	--	--	0.741	0.644	--	--	--	--
MA1	--	--	--	--	--	--	0.852	0.769	--	--	--	--
res	--	--	--	--	--	--	0.009	0.016	--	--	--	--

CASTELLON

Año: 1979

AR1	0.834	0.723	0.679	--	0.815	0.755	--	0.672	--	--	--	0.735
MA1	0.828	0.850	0.744	--	0.854	0.655	--	0.875	--	--	--	0.952
res	0.039	0.038	0.038	--	0.019	0.024	--	0.016	--	--	--	0.025

Año: 1980

AR1	0.793	0.854	0.771	--	0.743	0.859	--	0.724	--	0.772	0.743	--
MA1	0.788	0.881	0.849	--	0.786	0.966	--	0.681	--	0.967	0.730	--
res	0.032	0.025	0.033	--	0.029	0.018	--	0.015	--	0.020	0.031	--

Año: 1981

AR1	0.790	0.854	0.750	0.814	0.864	0.815	0.774	0.762	0.709	0.760	--	--
MA1	0.757	0.822	0.798	0.811	0.847	0.836	0.709	0.695	0.777	0.825	--	--
res	0.023	0.024	0.021	0.029	0.016	0.014	0.008	0.016	0.021	0.021	--	--

Año: 1982

AR1	0.826	0.789	0.781	0.825	--	0.684	0.652	0.699	0.668	0.851	0.749	0.715
MA1	0.852	0.858	0.873	0.886	--	0.887	0.766	0.820	0.837	0.954	0.920	0.847
res	0.028	0.039	0.027	0.020	--	0.013	0.009	0.016	0.012	0.022	0.029	0.024

CASTELLON (...)

Año: 1983

AR1	0.631	0.832	0.845	0.714	0.736	--	--	--	--	--	--	0.742
MA1	0.619	0.927	0.687	0.839	0.888	--	--	--	--	--	--	0.795
res	0.007	0.020	0.019	0.020	0.016	--	--	--	--	--	--	0.021

Año: 1984

AR1	--	--	--	--	0.820	0.800	--	--	--	--	--	--
MA1	--	--	--	--	0.848	0.847	--	--	--	--	--	--
res	--	--	--	--	0.022	0.012	--	--	--	--	--	--

LOGROÑO

Año: 1981

AR1	0.712	0.719	0.642	--	0.647	--	--	0.719	0.672	0.629	--	0.470
MA1	0.848	0.911	0.822	--	0.820	--	--	0.764	0.840	0.792	--	0.831
res	0.053	0.040	0.046	--	0.047	--	--	0.025	0.030	0.043	--	0.072

Año: 1982

AR1	0.588	0.763	0.559	0.748	--	0.646	--	0.727	0.669	0.646	--	0.659
MA1	0.737	0.873	0.896	0.674	--	0.742	--	0.796	0.809	0.875	--	0.905
res	0.055	0.050	0.046	0.040	--	0.038	--	0.036	0.042	0.060	--	0.054

Año: 1983

AR1	0.726	0.732	--	--	--	0.605	0.755	0.644	--	0.694	0.801	0.621
MA1	0.821	0.941	--	--	--	0.860	0.936	0.801	--	0.906	0.868	0.836
res	0.032	0.036	--	--	--	0.037	0.024	0.045	--	0.033	0.034	0.040

Año: 1984

AR1	0.524	0.640	0.686	0.763	0.593	--	0.704	0.760	0.678	0.715	0.567	0.651
MA1	0.861	0.700	0.781	0.823	0.764	--	0.786	0.910	0.935	0.697	0.805	0.742
res	0.067	0.056	0.054	0.038	0.055	--	0.014	0.029	0.034	0.038	0.074	0.056

## MADRID

Año: 1979

AR1	0.786	0.799	0.751	--	--	0.700	0.807	0.737	0.694	0.778	0.764	0.828
MA1	0.742	0.852	0.867	--	--	0.765	0.666	0.930	0.933	0.891	0.850	0.939
res	0.045	0.035	0.040	--	--	0.021	0.017	0.005	0.021	0.034	0.016	0.028

Año: 1980

AR1	0.791	0.899	0.803	0.831	0.692	0.783	--	--	0.651	0.753	0.801	0.793
MA1	0.843	0.964	0.926	0.878	0.786	0.888	--	--	0.885	0.962	0.889	0.940
res	0.028	0.023	0.026	0.022	0.042	0.014	--	--	0.010	0.026	0.030	0.024

Año: 1981

AR1	--	0.809	0.691	--	0.752	--	0.668	0.643	0.648	0.672	0.842	0.799
MA1	--	0.910	0.833	--	0.898	--	0.826	0.781	0.903	0.814	0.937	0.858
res	--	0.035	0.039	--	0.030	--	0.014	0.014	0.018	0.023	0.011	0.033

Año: 1982

AR1	--	0.790	0.843	0.794	--	0.734	0.734	0.746	0.781	--	0.796	--
MA1	--	0.952	0.845	0.740	--	0.692	0.772	0.663	0.935	--	0.864	--
res	--	0.037	0.014	0.029	--	0.021	0.014	0.013	0.023	--	0.024	--

## MADRID (...)

Año: 1983

AR1	0.754	0.732	0.626	--	0.735	0.714	--	0.679	0.558	--	--	0.844
MA1	0.731	0.773	0.914	--	0.773	0.862	--	0.956	0.690	--	--	0.868
res	0.007	0.025	0.015	--	0.025	0.014	--	0.020	0.01	--	--	0.022

Año: 1984

AR1	0.815	0.721	--	0.880	0.777	0.784	0.822	0.516	0.667	--	0.826	0.839
MA1	0.927	0.688	--	0.804	0.866	0.821	0.610	0.615	0.951	--	0.838	0.856
res	0.029	0.022	--	0.021	0.032	0.026	0.002	0.013	0.009	--	0.031	0.021

Año: 1985

AR1	0.873	0.736	0.768	0.832	0.690	--	0.634	0.644	0.474	0.713	0.825	0.803
MA1	0.965	0.912	0.760	0.743	0.936	--	0.909	0.695	0.938	0.770	0.952	0.876
res	0.027	0.032	0.025	0.023	0.036	--	0.01	0.004	0.006	0.01	0.024	0.023

Año: 1986

AR1	0.774	0.685	0.661	0.777	0.824	0.675	0.780	0.772	0.772	0.693	0.828	0.717
MA1	0.774	0.725	0.805	0.887	0.947	0.957	0.861	0.921	0.957	0.741	0.931	0.913
res	0.030	0.046	0.039	0.038	0.018	0.008	0.008	0.008	0.024	0.033	0.017	0.024



MALAGA

Año: 1977

AR1	0.845	0.744	0.751	0.650	0.679	0.656	0.829	0.609	0.625	0.768	0.867	0.787
MA1	0.918	0.880	0.910	0.667	0.742	0.694	0.709	0.848	0.836	0.926	0.891	0.829
res	0.034	0.031	0.016	0.027	0.018	0.018	0.012	0.010	0.012	0.027	0.017	0.037

Año: 1978

AR1	0.644	0.739	0.702	0.746	0.693	--	--	0.748	--	--	0.786	0.633
MA1	0.844	0.904	0.881	0.862	0.847	--	--	0.894	--	--	0.959	0.839
res	0.042	0.030	0.018	0.031	0.023	--	--	0.007	--	--	0.027	0.052

Año: 1979

AR1	0.843	0.745	0.760	0.688	0.800	0.681	0.818	0.616	0.730	0.767	0.584	0.666
MA1	0.906	0.830	0.813	0.796	0.901	0.789	0.879	0.710	0.762	0.851	0.624	0.627
res	0.032	0.039	0.028	0.015	0.016	0.012	0.008	0.005	0.017	0.034	0.029	0.027

Año: 1980

AR1	0.790	0.752	0.738	0.853	0.775	--	0.683	--	0.778	0.731	0.779	0.846
MA1	0.709	0.934	0.747	0.948	0.832	--	0.957	--	0.688	0.896	0.814	0.929
res	0.025	0.024	0.033	0.021	0.022	--	0.005	--	0.015	0.018	0.027	0.015

MALAGA (...)

Año: 1981

AR1	0.735	0.813	0.605	0.739	0.676	--	0.681	0.729	0.712	0.718	0.648	--
MA1	0.865	0.909	0.901	0.873	0.717	--	0.734	0.945	0.785	0.708	0.819	--
res	0.024	0.024	0.033	0.040	0.018	--	0.007	0.009	0.014	0.019	0.015	--

Año: 1982

AR1	0.883	0.818	0.792	0.775	0.834	0.553	0.819	--	0.699	0.606	--	0.704
MA1	0.968	0.980	0.863	0.778	0.943	0.588	0.700	--	0.720	0.690	--	0.932
res	0.017	0.034	0.018	0.028	0.014	0.008	0.011	--	0.016	0.017	--	0.018

Año: 1983

AR1	0.696	0.800	0.798	0.708	--	0.819	--	0.837	0.657	0.686	0.810	0.781
MA1	0.748	0.804	0.861	0.750	--	0.869	--	0.978	0.886	0.781	0.899	0.846
res	0.014	0.025	0.015	0.028	--	0.011	--	0.007	0.006	0.018	0.028	0.029

Año: 1984

AR1	0.629	0.660	--	0.725	0.640	--	0.693	0.636	0.598	0.617	0.780	0.636
MA1	0.961	0.796	--	0.524	0.813	--	0.520	0.861	0.675	0.670	0.826	0.746
res	0.023	0.036	--	0.037	0.045	--	0.012	0.015	0.01	0.018	0.040	0.030

MURCIA

Año: 1977

AR1	0.811	0.692	0.739	0.828	--	0.654	0.719	0.724	0.513	0.709	0.829	0.797
MA1	0.801	0.807	0.848	0.953	--	0.654	0.731	0.848	0.745	0.884	0.826	0.858
res	0.037	0.037	0.022	0.021	--	0.019	0.019	0.014	0.014	0.027	0.022	0.028

Año: 1978

AR1	0.724	0.751	0.796	0.614	--	0.723	--	0.724	0.521	0.666	0.811	--
MA1	0.842	0.810	0.842	0.830	--	0.910	--	0.733	0.614	0.678	0.881	--
res	0.032	0.025	0.022	0.036	--	0.019	--	0.013	0.012	0.023	0.023	--

Año: 1979

AR1	0.828	0.746	0.643	0.565	--	0.532	0.694	0.539	0.616	0.715	0.696	--
MA1	0.915	0.848	0.842	0.738	--	0.638	0.650	0.538	0.971	0.782	0.905	--
res	0.038	0.035	0.040	0.032	--	0.022	0.012	0.009	0.026	0.036	0.026	--

Año: 1980

AR1	0.803	0.832	0.742	0.804	0.688	--	0.556	0.795	0.647	0.719	0.807	0.876
MA1	0.653	0.799	0.847	0.896	0.885	--	0.723	0.766	0.674	0.925	0.711	0.916
res	0.027	0.019	0.034	0.028	0.038	--	0.014	0.015	0.019	0.018	0.029	0.025

## MURCIA (...)

Año: 1981

AR1	0.852	--	0.566	0.724	--	0.776	0.694	0.768	0.706	0.819	0.749	0.722
MA1	0.992	--	0.769	0.892	--	0.821	0.941	0.703	0.805	0.657	0.895	0.934
res	0.025	--	0.037	0.049	--	0.015	0.005	0.018	0.017	0.020	0.015	0.035

Año: 1982

AR1	0.788	0.805	0.795	0.719	0.752	0.676	0.698	0.674	0.763	0.739	0.832	0.783
MA1	0.898	0.946	0.929	0.716	0.737	0.909	0.839	0.840	0.694	0.780	0.842	0.942
res	0.033	0.037	0.029	0.035	0.028	0.014	0.010	0.018	0.029	0.027	0.026	0.025

Año: 1983

AR1	0.685	0.769	0.810	0.590	0.702	0.697	--	--	0.647	0.790	0.840	0.595
MA1	0.857	0.918	0.923	0.790	0.920	0.759	--	--	0.759	0.928	0.963	0.841
res	0.017	0.031	0.020	0.030	0.022	0.022	--	--	0.011	0.018	0.031	0.034

Año: 1984

AR1	0.750	0.820	0.731	0.708	0.763	0.807	--	0.611	0.690	--	0.759	0.600
MA1	0.756	0.920	0.748	0.663	0.787	0.950	--	0.887	0.919	--	0.810	0.850
res	0.023	0.020	0.033	0.029	0.033	0.014	--	0.012	0.014	--	0.024	0.036

OVIEDO

Año: 1977

AR1	0.738	0.635	0.742	0.756	--	0.882	0.791	--	0.867	0.804	0.729	0.702
MA1	0.872	0.833	0.846	0.843	--	0.935	0.833	--	0.920	0.819	0.906	0.870
res	0.049	0.057	0.030	0.037	--	0.029	0.034	--	0.023	0.029	0.035	0.048

Año: 1978

AR1	0.746	0.716	0.766	0.784	--	0.821	0.794	0.822	0.803	0.715	0.868	0.671
MA1	0.873	0.925	0.784	0.893	--	0.854	0.868	0.799	0.782	0.752	0.982	0.895
res	0.033	0.041	0.039	0.036	--	0.030	0.029	0.023	0.028	0.046	0.020	0.049

Año: 1979

AR1	0.849	0.763	0.708	0.699	0.808	0.850	--	0.823	0.803	0.716	0.846	0.728
MA1	0.883	0.829	0.793	0.861	0.766	0.825	--	0.913	0.891	0.881	0.890	0.885
res	0.027	0.050	0.040	0.039	0.030	0.025	--	0.027	0.024	0.031	0.021	0.031

Año: 1980

AR1	0.762	0.745	0.705	0.891	0.752	0.812	0.808	0.842	0.892	0.786	0.799	0.825
MA1	0.781	0.789	0.814	0.863	0.758	0.881	0.810	0.834	0.943	0.893	0.914	0.857
res	0.040	0.029	0.044	0.023	0.027	0.026	0.031	0.020	0.022	0.022	0.030	0.021

OVIEDO (...)

Año: 1981

AR1	--	0.770	0.717	0.722	0.634	0.830	0.819	0.895	0.664	0.656	0.718	0.645
MA1	--	0.960	0.786	0.763	0.874	0.673	0.861	0.740	0.881	0.710	0.895	0.821
res	--	0.033	0.038	0.045	0.041	0.025	0.028	0.019	0.031	0.035	0.018	0.041

Año: 1982

AR1	0.693	0.780	0.796	0.773	0.824	0.777	0.743	0.832	0.756	--	0.753	--
MA1	0.861	0.812	0.853	0.819	0.899	0.838	0.790	0.792	0.780	--	0.892	--
res	0.026	0.030	0.029	0.024	0.027	0.028	0.031	0.028	0.025	--	0.025	--

Año: 1983

AR1	0.825	0.680	--	--	0.711	0.866	0.821	0.779	0.744	0.884	0.817	--
MA1	0.794	0.803	--	--	0.778	0.702	0.834	0.764	0.805	0.939	0.885	--
res	0.013	0.033	--	--	0.032	0.017	0.026	0.026	0.030	0.021	0.022	--

Año: 1984

AR1	0.778	0.792	0.788	0.837	0.734	0.806	0.789	--	0.744	0.858	0.687	0.751
MA1	0.780	0.689	0.673	0.688	0.781	0.737	0.690	--	0.878	0.862	0.814	0.823
res	0.031	0.038	0.031	0.021	0.026	0.027	0.029	--	0.029	0.025	0.038	0.030

P. MALLORCA

Año: 1977

AR1	0.746	0.636	0.829	0.803	0.763	0.760	0.776	0.762	--	0.781	0.638	0.742
MA1	0.752	0.724	0.710	0.800	0.817	0.830	0.865	0.850	--	0.929	0.708	0.676
res	0.027	0.031	0.017	0.027	0.030	0.022	0.012	0.017	--	0.019	0.038	0.034

Año: 1978

AR1	0.804	0.693	0.707	0.776	--	0.729	0.693	0.672	--	--	--	0.754
MA1	0.934	0.855	0.838	0.768	--	0.951	0.656	0.753	--	--	--	0.937
res	0.031	0.035	0.025	0.028	--	0.021	0.01	0.014	--	--	--	0.029

Año: 1979

AR1	0.856	0.737	0.848	0.680	0.764	0.658	0.746	0.662	0.752	0.803	0.623	0.796
MA1	0.889	0.908	0.899	0.748	0.777	0.669	0.760	0.936	0.856	0.937	0.690	0.946
res	0.025	0.040	0.031	0.039	0.013	0.017	0.012	0.011	0.018	0.027	0.032	0.022

Año: 1980

AR1	0.787	0.790	0.783	--	0.732	--	0.770	--	0.790	--	0.676	0.851
MA1	0.782	0.860	0.888	--	0.918	--	0.980	--	0.800	--	0.952	0.781
res	0.026	0.030	0.023	--	0.025	--	0.009	--	0.012	--	0.030	0.025

P. MALLORCA (...)

Año: 1981

AR1	--	0.847	0.558	0.745	0.772	0.781	0.870	0.746	0.651	0.674	0.767	0.674
MA1	--	0.903	0.856	0.855	0.758	0.802	0.990	0.644	0.914	0.778	0.958	0.909
res	--	0.027	0.032	0.035	0.026	0.015	0.01	0.011	0.015	0.019	0.018	0.028

Año: 1982

AR1	0.638	0.790	--	0.770	0.769	0.682	--	--	0.744	--	0.748	0.735
MA1	0.893	0.922	--	0.782	0.699	0.637	--	--	0.785	--	0.832	0.846
res	0.027	0.032	--	0.025	0.027	0.010	--	--	0.025	--	0.030	0.025

Año: 1983

AR1	0.630	0.795	0.748	0.769	0.793	0.800	--	--	0.630	0.695	0.807	0.703
MA1	0.850	0.910	0.724	0.786	0.921	0.801	--	--	0.533	0.930	0.948	0.858
res	0.011	0.029	0.023	0.019	0.013	0.015	--	--	0.019	0.017	0.028	0.028

Año: 1984

AR1	0.744	0.667	0.774	0.814	0.767	--	0.839	--	0.843	0.720	0.845	0.754
MA1	0.940	0.717	0.777	0.715	0.848	--	0.760	--	0.957	0.844	0.980	0.812
res	0.022	0.035	0.033	0.030	0.033	--	0.005	--	0.024	0.020	0.023	0.024



SEVILLA

Año: 1977

AR1	0.672	0.819	0.517	0.755	--	--	--	--	--	--	--	--
MA1	0.698	0.860	0.828	0.928	--	--	--	--	--	--	--	--
res	0.037	0.036	0.022	0.020	--	--	--	--	--	--	--	--

Año: 1978

AR1	--	--	--	--	--	--	--	--	--	--	0.779	0.763
MA1	--	--	--	--	--	--	--	--	--	--	0.752	0.823
res	--	--	--	--	--	--	--	--	--	--	0.021	0.040

Año: 1979

AR1	--	0.850	0.812	0.764	--	0.624	--	0.388	--	0.705	--	0.795
MA1	--	0.926	0.924	0.699	--	0.740	--	0.622	--	0.899	--	0.874
res	--	0.039	0.030	0.023	--	0.014	--	0.008	--	0.048	--	0.027

Año: 1980

AR1	0.715	0.757	0.724	0.753	0.689	--	0.558	0.778	0.716	0.682	0.686	0.714
MA1	0.719	0.844	0.690	0.922	0.785	--	0.685	0.636	0.860	0.906	0.792	0.872
res	0.031	0.031	0.034	0.031	0.027	--	0.002	0.013	0.018	0.027	0.039	0.022

SEVILLA (...)

Año: 1981

AR1	0.646	0.749	0.664	0.712	0.685	0.598	--	0.573	0.662	--	--	0.828
MA1	0.905	0.870	0.836	0.910	0.615	0.710	--	0.698	0.919	--	--	0.923
res	0.017	0.033	0.039	0.039	0.019	0.015	--	0.008	0.015	--	--	0.033

Año: 1982

AR1	0.730	0.723	--	0.687	0.665	0.649	--	--	0.713	0.782	--	0.830
MA1	0.843	0.781	--	0.695	0.892	0.721	--	--	0.962	0.900	--	0.830
res	0.042	0.038	--	0.032	0.015	0.012	--	--	0.018	0.016	--	0.027

Año: 1983

AR1	0.754	0.806	--	0.656	0.697	--	0.555	--	0.791	0.755	0.685	--
MA1	0.695	0.909	--	0.860	0.788	--	0.707	--	0.882	0.892	0.766	--
res	0.012	0.024	--	0.030	0.025	--	0.007	--	0.007	0.015	0.042	--

Año: 1984

AR1	0.765	0.787	0.678	--	0.658	0.817	--	0.553	0.716	0.775	0.777	0.660
MA1	0.838	0.676	0.803	--	0.837	0.698	--	0.895	0.839	0.712	0.775	0.718
res	0.021	0.018	0.025	--	0.026	0.014	--	0.004	0.009	0.017	0.027	0.025

TORTOSA

Año: 1980

AR1	0.738	0.851	0.778	--	0.762	0.813	0.790	0.743	0.639	0.736	0.818	0.588
MA1	0.810	0.879	0.836	--	0.882	0.908	0.901	0.845	0.697	0.937	0.979	0.866
res	0.036	0.028	0.034	--	0.032	0.024	0.018	0.019	0.026	0.031	0.028	0.031

Año: 1981

AR1	0.793	0.848	0.726	0.754	0.748	0.611	--	0.725	--	0.687	0.876	--
MA1	0.840	0.949	0.723	0.739	0.854	0.795	--	0.697	--	0.795	0.956	--
res	0.027	0.032	0.038	0.042	0.040	0.021	--	0.021	--	0.028	0.015	--

Año: 1982

AR1	0.788	0.836	0.827	0.821	--	--	0.743	0.805	0.730	0.832	0.864	0.759
MA1	0.860	0.866	0.803	0.646	--	--	0.929	0.789	0.795	0.948	0.905	0.765
res	0.033	0.031	0.014	0.020	--	--	0.014	0.013	0.016	0.018	0.016	0.019

Año: 1983

AR1	0.814	0.877	0.903	0.725	0.661	0.822	0.733	0.779	0.610	0.728	0.792	0.749
MA1	0.865	0.951	0.755	0.744	0.762	0.826	0.776	0.886	0.662	0.954	0.835	0.709
res	0.005	0.015	0.015	0.032	0.033	0.022	0.013	0.024	0.025	0.020	0.037	0.035

TORTOSA (...)

Año: 1984

AR1	0.705	0.843	0.834	--	--	--	--	--	--	--	--	--
MA1	0.777	0.711	0.827	--	--	--	--	--	--	--	--	--
res	0.045	0.035	0.028	--	--	--	--	--	--	--	--	--

## **ANEXO III**

### III.1 CONTRASTES DE IGUALDAD DE VARIANZAS

Supondremos primero que queremos hacer un contraste sobre el valor de la varianza del término de error,  $\sigma^2$ , de un modelo ARMA. Dado un número real positivo  $c$ , el contraste:

$$\begin{aligned} H_0: & \sigma^2 = c \\ H_1: & \sigma^2 \neq c \end{aligned} \quad (1)$$

se puede efectuar teniendo en cuenta lo siguiente: llamando  $\hat{\sigma}^2$  al estimador de la varianza residual obtenido, se tiene que:

$$\frac{(T-K)}{\sigma^2} \hat{\sigma}^2 \approx \chi^2_{T-K} \quad (2)$$

donde  $T$  es el número de observaciones,  $K$  es el número de parámetros estimados y  $\sigma^2$  es el verdadero valor de la varianza del término de error. Así pues, bajo la hipótesis nula se tiene que la expresión (2) es cierta reemplazando  $\sigma^2$  por  $c$ . Por tanto, el contraste se realizará, con un nivel de significación  $\alpha$ , del modo siguiente:

(i) Calculamos:

$$\frac{T-K}{c} \hat{\sigma}^2 = \chi^2_{muestral} \quad (3)$$

(ii) Con las tablas de la distribución  $\chi^2$  obtenemos  $a$  y  $b$  tales que:

$$P(\chi^2_{T-K} \leq a) = \frac{\alpha}{2} \quad y \quad P(\chi^2_{T-K} \geq b) = \frac{\alpha}{2} \quad (4)$$

(iii) Si  $\chi^2_{muestral}$  pertenece al intervalo  $(a, b)$  aceptamos  $H_0$ ; en caso contrario rechazamos  $H_0$ .

Supongamos ahora que queremos contrastar si dos modelos estimados tienen igual varianza, es decir:

$$\begin{aligned} H_0: \sigma^2(1) &= \sigma^2(2) \\ H_1: \sigma^2(1) &\neq \sigma^2(2) \end{aligned} \quad (5)$$

Llamando:

$\hat{\sigma}^2(1)$ : *estimador de la varianza residual ( $\sigma^2(1)$ ) del modelo 1*

$\hat{\sigma}^2(2)$ : *estimador de la varianza residual ( $\sigma^2(2)$ ) del modelo 2*

entonces se tiene que, bajo  $H_0$ :

$$\frac{\frac{\hat{\sigma}^2(1)}{T-K}}{\frac{\hat{\sigma}^2(2)}{T^*-K^*}} \approx F_{T-K, T^*-K^*} \quad (8)$$

siendo  $T, T^*$  las observaciones de los modelos 1 y 2, respectivamente y  $K, K^*$  el número de parámetros estimados en cada uno de estos modelos. Por tanto, el contraste con nivel de significación  $\alpha$  se realizará del modo siguiente:

(i) Calculamos el valor de la expresión:

$$\frac{T^*-K^*}{T-K} \frac{\hat{\sigma}^2(1)}{\hat{\sigma}^2(2)} = F_{muestral} \quad (9)$$

(ii) Utilizando las tablas de la distribución F obtenemos valores  $a$  y  $b$  tales que:

$$P(F_{T-K, T^*-K^*} \leq a) = \frac{\alpha}{2} \quad \text{y} \quad P(F_{T-K, T^*-K^*} \geq b) = \frac{\alpha}{2} \quad (10)$$

(iii) Si  $F_{muestral} \in (a, b)$  aceptamos  $H_0$ , en el caso contrario, rechazamos  $H_0$ .

### III.2 CONTRASTE DE IGUALDAD DE LOS PARÁMETROS ESTIMADOS EN UN MODELO ARMA MULTIPLICATIVO $(1,0) \times (0,1)$ ,

Sean  $\phi$  y  $\theta$  dos parámetros estimados en un modelo ARMA. Dados dos números reales  $d$ ,  $e$  queremos contrastar:

$$\begin{aligned} H_0: \phi &= d, \theta = e \\ H_1: \phi &\neq d, \theta \neq e \end{aligned} \quad (11)$$

Llamando  $\hat{\phi}$ ,  $\hat{\theta}$  a los estimadores de los parámetros se tiene que, bajo  $H_0$ :

$$(\hat{\phi} - d, \hat{\theta} - e) \Sigma_1^{-1} \begin{pmatrix} \hat{\phi} - d \\ \hat{\theta} - e \end{pmatrix} \approx \chi^2_2 \quad (12)$$

donde  $\Sigma_1$  es la matriz de varianzas y covarianzas de los estimadores. Por tanto, el contraste puede realizarse con un nivel de significación  $\alpha$ , del modo siguiente:

(i) Calculamos:

$$(\hat{\phi} - d, \hat{\theta} - e) \Sigma_1^{-1} \begin{pmatrix} \hat{\phi} - d \\ \hat{\theta} - e \end{pmatrix} \equiv \chi^2_{muestral} \quad (13)$$

(ii) Utilizando las tablas de la distribución  $\chi^2$  con dos grados de libertad obtenemos un valor  $c$ , tal que:

$$P(\chi^2_2 \leq c) = 1 - \alpha \quad (14)$$

(iii) Si  $\chi^2_{muestral}$  pertenece al intervalo  $(0, c)$  aceptamos  $H_0$ ; en caso contrario rechazamos  $H_0$ .



### III.3 DERIVACIÓN DE LAS COVARIANZAS Y CORRELACIONES DE UN PROCESO ARMA(1,0)xARMA(0,1), ESTACIONARIO

Para la derivación de las covarianzas y correlaciones de un proceso ARMA multiplicativo (1,0)x(0,1), seguiremos los razonamientos detallados en la referencia /15/.

En un proceso ARMA(1,0)xARMA(0,1), el valor de la serie en un instante  $t$  viene dado por una expresión del tipo:

$$y_t = \phi_1 y_{t-1} - \tilde{\theta}_s a_{t-s} + a_t \quad (15)$$

Donde  $s$  es el periodo de la parte estacional. Utilizaremos en el desarrollo el valor de  $s=12$ . Suponemos que:

$$\begin{aligned} E[y_t] &= 0 \quad \forall t \\ \gamma_0 &= \text{Var}(y_t) \\ \gamma_k &= \text{Cov}(y_t, y_{t-k}) \end{aligned} \quad (16)$$

Llamamos primero:

$$\begin{aligned} \sigma_a^2 &= \text{Var}(a_t) \\ \gamma_{ya}(k) &= E[y_{t-k} a_t] \end{aligned} \quad (17)$$

y queremos obtener una expresión para  $\gamma_{ya}(k)$ . Debido a que la innovación es siempre independiente de las observaciones pasadas, podemos escribir:

$$\text{Para } k=1 \quad E[y_{t-1}a_t]=0$$

$$\text{Para } k=2 \quad E[y_{t-2}a_t]=0$$

$$\text{Y en general, si } k>0 \quad E[y_{t-k}a_t]=0$$

Para  $k=0$ , tenemos

$$E[y_t a_t] = \phi_1 E[y_{t-1} a_t] - \tilde{\theta}_s E[y_t a_{t-s}] + E[a_t^2]$$

luego:

$$E[y_t a_t] = \sigma_a^2$$

Para  $k=-1$ ,

$$\begin{aligned} E[y_{t+1} a_t] &= \phi_1 E[y_t a_t] - \tilde{\theta}_s E[a_{t-s-1} a_t] + \\ &+ E[a_{t+1} a_t] = \phi_1 \sigma_a^2 \end{aligned}$$

Para  $k=-10$

$$\begin{aligned} E[y_{t+10} a_t] &= \phi_1 E[y_{t+9} a_t] - \tilde{\theta}_s E[a_{t-2} a_t] + \\ &+ E[a_{t+10} a_t] = \phi_1^{10} \sigma_a^2 \end{aligned}$$

Para  $k=-11$ , análogamente

$$E[y_{t+11} a_t] = \phi_1^{11} \sigma_a^2$$

Para  $k=-12$

$$\begin{aligned} E[y_{t+12} a_t] &= \phi_1 (\phi_1^{12} \sigma_a^2 - \tilde{\theta}_{12} \sigma_a^2) - \phi_{12} E[a_{t+1} a_t] + \\ &+ E[a_{t+13} a_t] = \phi_1 \sigma_a^2 (\phi_1^{12} - \tilde{\theta}_{12}) \end{aligned}$$

En resumen, podemos escribir:

$$\gamma_{ya}(k) = \begin{cases} 0 & \text{si } k > 0 \\ \phi_1 \sigma_a^2 & \text{si } -11 \leq k \leq 0 \\ \phi_1^{-k-12} \sigma_a^2 (\phi_1^{12} - \tilde{\theta}_{12}) & \text{si } k \leq -12 \end{cases}$$

Con esta información pasamos a calcular las varianzas:

$$\begin{aligned} \gamma_0 = E[y_t^2] &= \phi_1 E[y_t y_{t-1}] - \phi_{12} E[y_t a_{t-12}] + E[y_t a_t] = \\ &= \phi_1 \gamma_1 - \tilde{\theta}_{12} \gamma_{ya}(-12) + \gamma_{ya}(0) \end{aligned}$$

es decir

$$\gamma_0 = \phi_1 \gamma_1 - \tilde{\theta}_{12} \sigma_a^2 (\phi_1^{12} - \tilde{\theta}_{12}) + \sigma_a^2 \quad (18)$$

Por otra parte:

$$\begin{aligned} \gamma_1 = E[y_t y_{t-1}] &= \phi_1 \gamma_0 - \tilde{\theta}_{12} E[y_{t-1} a_{t-12}] + E[y_{t-1} a_t] = \\ &= \phi_1 \gamma_0 - \tilde{\theta}_{12} \gamma_{ya}(-11) + \gamma_{ya}(1) \end{aligned}$$

es decir:

$$\gamma_1 = \phi_1 \gamma_0 - \tilde{\theta}_{12} \sigma_a^2 \phi_1^{11} \quad (19)$$

Con lo que tenemos un sistema con dos ecuaciones, (18) y (19), y dos incógnitas ( $\gamma_1, \gamma_0$ ). Podemos obtener el valor de  $\gamma_0$ :

$$\gamma_0 = \phi_1^2 \gamma_0 - \phi_1 \tilde{\theta}_{12} \sigma_a^2 \phi_1^{11} - \tilde{\theta}_{12} \sigma_a^2 (\phi_1^{12} - \tilde{\theta}_{12}) + \sigma_a^2$$

$$\gamma_0 = \frac{1 + \tilde{\theta}_{12}^2 - 2\tilde{\theta}_{12}\phi_1^{12}}{1 - \phi_1^2} \sigma_a^2$$

Análogamente se calcularían  $\gamma_1, \gamma_2, \dots$ , y de ahí  $\rho_1, \rho_2$ , etc.

## NOMENCLATURA

- $a_t$  : ruido blanco
- $\alpha$  : nivel de significación (típicamente 0.05)
- $\beta$  : inclinación superficie
- $d$  : día
- $D$  : radiación difusa
- $D_s$  : operador diferencias de orden  $s$
- $\delta$  : declinación
- $G$  : radiación global
- $G_{b,max}$  : valor de exposición horaria de radiación global máximo para la hora  $h$
- $\gamma$  : altura solar
- $h$  : hora
- $I$  : radiación directa
- $K$  : índice de transparencia atmosférico, cociente entre radiación global (horaria o diaria) y radiación extraterrestre (horaria o diaria)
- $m$  : valor medio
- $\Phi_i$  : coeficiente autorregresivo de orden  $i$  de la parte regular
- $\tilde{\Phi}_i$  : coeficiente autorregresivo de orden  $i$  de la parte estacional
- $\phi$  : latitud
- $Q(h)$  : estadístico Box-Ljung
- $r_s$  : coeficiente de correlación muestral
- $s$  : número de horas al día (12, 10 ó 8)
- $\Theta_i$  : coeficiente de medias móviles de orden  $i$  de la parte regular
- $\tilde{\Theta}_i$  : coeficiente de medias móviles de orden  $i$  de la parte estacional
- $X_b$  : índice horario definido como cociente entre el valor de exposición horaria de radiación global y el valor máximo esperable de exposición horaria de radiación global para esa misma hora

## **BIBLIOGRAFIA**

- /1/ R.J. Aguiar, M. Collares-Pereira, J.P. Conde, "Simple procedure for generating sequences of daily radiation values using a library of Markov Transition Matrices". *Solar Energy*, Vol.40, N.3, pp. 269-279. 1988.
- /2/ R. Aguiar, M. Collares-Pereira, "Statistical properties of hourly global radiation". *Solar Energy*, Vol. 48, N.3, pp. 157-167, 1992.
- /3/ R. Aguiar, M. Collares-Pereira, "Tag: A time-dependent, autoregressive, gaussian model for generating synthetic hourly radiation". *Solar Energy*, Vol. 49, N.3, pp. 167-174, 1992.
- /4/ U. Amato, A. Andreta, B. Bartoli, B. Coluzzi, "Stochastic modelling of solar radiation data". *Il Nuovo Cimento*, Vol. 8C, N.3. 1985.

- 
- /5/ U. Amato, A. Andretta, B. Bartoli, B. Coluzzi, V. Cuomo. "Markov processes and Fourier analysis as a tool to describe and simulate daily solar irradiance". *Solar Energy*, Vol.37, N.3, pp.179-194, 1986.
- /6/ U. Amato, V. Cuomo, F. Fontana, C. Serio, P. Silvestrini, "Behavior of hourly solar irradiance in the Italian climate", *Solar Energy*, Vol.40, N.1, pp.65-79, 1988.
- /7/ A. Asle Olseth, A. Skartveit, "A probability density function for daily insolation within the temperature storm belts". *Solar Energy*, Vol.33, N.6, pp.533-542, 1984.
- /8/ A. Asle Olseth, A. Skartveit, "A probability density model for hourly total and beam irradiance on arbitrarily orientated planes". *Solar Energy*, Vol.39, N.4, pp. 343-351. 1987.
- /9/ A.K. Athienitis, "A method and algorithm for estimation of transmitted solar radiation and its long-term averages using Fourier series". *Solar Energy*, Vol.45, N.5, pp. 257-263, 1990.
- /10/ A. Balouktsis, Ph. Tsalides, "Stochastic simulation model of hourly total solar radiation". *Solar Energy*, Vol.37, N.2, pp. 119-126. 1986.
- /11/ B. Bartoli, B. Coluzzi, V. Cuomo, M. Francesca, C. Serio, "Autocorrelation of daily global solar radiation". *Il nuovo cemento*, Vol. 4c, N.2.
- /12/ P. Bendt, M. Collares-Pereira, A. Rabl, "The frequency distribution of daily insolation values". *Solar Energy*, Vol. 27 pp 1-5. 1981.

- 
- /13/ A. J. Biga, R. Rosa, "Statistical behaviour of solar irradiation over consecutive days". *Solar Energy*, Vol. 27, pp 149-157. 1981.
- /14/ G. Boch, E. Boileau, C. Bernard, "Modelisation of the hourly solar global irradiation in Trappes (France) by ARMA processes". *LSS-EEE*, pp.2418-2423.
- /15/ E. Boileau, "Use of some simple statistical models in solar meteorology". *Solar Energy*, Vol. 30, N.4, pp. 333-339, 1983.
- /16/ J.E.P. Box, G.M. Jenkins, "Time series analysis, forecasting and control". Holden-Day, 575 pp. 1976.
- /17/ B.J. Brinkworth, "Autocorrelation and stochastic modelling of insolation sequences". *Solar Energy*, Vol. 19, pp. 343-347. 1977.
- /18/ L.L.Bucciarelli, Jr. "Estimating loss-of-power probabilities of stand-alone photovoltaic solar energy systems". *Solar Energy* 32, 205-209, 1984.
- /19/ J.M. Caridad y Ocerín, A. Gallego, J.G. Sepúlveda, "Análisis de series temporales en modelización del número de horas de sol e irradiación medida". Departamento de Estadística, Universidad de Córdoba, 1986.
- /20/ J. Dann Engels, S.M. Pollock, J.A. Clark, "Observations on the statistical nature of terrestrial irradiation". *Solar Energy*, Vol. 26, pp. 91-92. 1981.
- /21/ M.A. Egido, "El dimensionado de sistemas fotovoltaicos autónomos". Cap.17 del libro "Fundamentos, dimensionado y aplicaciones de la energía solar fotovoltaica". CIEMAT, 1992.

- 
- /22/ R. Festa, C.F. Ratto, DeGol, "A procedure to obtain average daily values of meteorological parameters from monthly averages". *Solar Energy*, Vol.40, N.4, pp. 309-313, 1988.
- /23/ W.A.Fuller, "Introduction to Statistical Time Series". Wiley & Sons, 470pp., 1976.
- /24/ K.K. Gopinathan, "Computing the monthly mean daily diffuse radiation from clearness index and percent possible sunshine". *Solar Energy*, Vol.41, pp.379-385, 1988.
- /25/ J.M. Gordon, M. Hochman, "On the random nature of solar radiation". *Solar Energy*, Vol. 32, N.1, 1984.
- /26/ J.M. Gordon. "Stand-alone photovoltaic systems: A stochastic analytic model for predicting loss-of-load". Proc. ASES Annual Meeting, Boulder, CO, 373-375, 1986.
- /27/ J.M. Gordon. "Optimal sizing of stand-alone photovoltaic solar power systems". *Solar Cells*, 20, pp.295-313, 1987.
- /28/ J.M. Gordon, T.A. Reddy, "Time series analysis of daily horizontal solar radiation". *Solar Energy*, Vol.41, N.3, pp.215-226, 1988.
- /29/ V.A. Graham, K.G.T. Hollands, T.E. Unny, "A time series model for  $K_t$  with application to global synthetic weather generation", *Solar Energy*, Vol.40, N.2, pp.83-92, 1988.
- /30/ V.A. Graham, K.G.T. Hollands, "A method to generate synthetic hourly solar radiation globally". *Solar Energy*, Vol.44, N.6, pp. 333-341, 1990.



- 
- /31/ D. Guinea, Ll. Mora, E. Palomo, "Análisis estadístico de series de exposición horaria de radiación solar", 4º Congreso Ibérico, 2º Iberoamericano de Energía -Solar, Oporto, Portugal, 1988.
- /32/ A.C. Harvey, "Time Series Models". Harvester Wheatsheaf, 1993.
- /33/ M.Iqbal, "An Introduction to solar radiation". Acad. Press Canada, 389 pp., 1983.
- /34/ M.K. Karabaev, I. Kh. Karimov, M.M. Kenisarin, N.P. Tkachenkova, "Model of hourly incidence of solar radiation on a horizontal surface for conditions of Tashkent". Geliotekhnika, Vol. 24, N.1, pp. 48-52. 1988.
- /35/ S.A. Klein and W.A. Beckman. "Loss-of-load probabilities for stand-alone photovoltaic systems". Solar Energy, Vol. 39, No. 6, pp. 499-512, 1987.
- /36/ K.M. Knight, S.A., S.A. Klein, J.A. Duffie, "A methodology for the synthesis of hourly weather data". Solar Energy, Vol. 46, N.2, pp.109-120, 1991.
- /37/ K.YA. Kondratyev, "Radiation in the atmosphere". Academic Press, 912 pp., 1969.
- /38/ B. Liu, R. Jordan, "The interrelationship and characteristic distribution of direct, diffuse and total solar radiation". Solar Energy 4, 1-19, 1960.
- /39/ Ll. Mora, "Generación de secuencias horarias de radiación global utilizando matrices de transición de Markov, para la localidad de Madrid", Trabajo de Tercer Ciclo. Universidad Complutense de Madrid, 1989.

- 
- /40/ Ll. Mora, M.A. Egido, M. Sidrach de Cardona. " Utilization of simulated hourly global solar radiation series for sizing stand-alone photovoltaic power systems. 9th European Photovoltaic Solar Energy Conference. Freiburg, RFA, Sep., 1989.
- /41/ Ll. Mora, M. Sidrach de Cardona, "Comparison between results for sizing stand-alone photovoltaic systems with variable load using real data, simulated data and TRY of radiation values". 12th European Photovoltaic Solar Energy Conference. Amsterdam, The Netherlands, 1994.
- /42/ C. Mustacchi, V. Cena, M. Rocchi, "Stochastic simulation of hourly global radiation sequences". Solar Energy, Vol. 23, pp. 47-51. 1979.
- /43/ J. A. Olseth, A. Skartveit, "A probability density model for hourly total and beam irradiance on arbitrarily orientated planes". Solar Energy, Vol. 39, N.4, pp. 343-351, 1987.
- /44/ J.A. Olseth, A. Skartveit, "Characteristic of hourly global irradiance modelled from cloud data", Solar Energy, Vol.51, N.3, pp.197-204, 1993.
- /45/ Orgills, Hollands, "Correlation equation for hourly difusse radiation on a horizontal surface". Solar Energy, Vol.19, 1977.
- /46/ E. Palomo, "Hourly solar radiation time series as first-order Markov chains", Proc. ISES Solar World Congress, Kobe, Japan, 1989.
- /47/ C.W.Paltridge, C.M.R. Platt, "Radiative processes in meteorology and climatology". Elsevier, 318 pp., 1976.

- 
- /48/ A. Papoulis, "Probabilidad, variables aleatorias y procesos estocásticos". EUNIBAR, Barcelona, 1980.
- /49/ M. Pardo González-Nandín, C. Gómez Camacho, "Determinación de áreas climáticas en la España Peninsular según exposición solar, precipitación y humedad relativa". VII Congreso Ibérico de Energía Solar, Vigo, Mayo, 1994.
- /50/ D. Peña Sánchez de Rivera, "Estadística. Modelos y métodos. 1. Fundamentos". Alianza Universidad Textos, Madrid. 571 pp. 1992.
- /51/ D. Peña Sánchez de Rivera, "Estadística. Modelos y métodos. 1. Modelos lineales y series temporales". Alianza Universidad Textos, Madrid. 745 pp. 1992.
- /52/ R. Perez, P. Ineichen, R. Seals, A. Zelenka, "Making full use of the clearness index for parameterizing hourly insolation conditions". Solar Energy, Vol.45, N.2, pp.111-114, 1990.
- /53/ Perrin de Brichambaut et al., "Meteorological aspects of the utilization of solar radiation as an energy source". World Meteorological organization, N.557, 299pp.
- /54/ R.T. Pinker, L.M. Militana, "The asymmetry of Global Solar Radiation Around Solar Noon". Journal of applied meteorology, Vol. 27, pp. 318-322, Marzo 1988.
- /55/ M.B.Priestley, "Spectral analysis and time series". Acad. Press Ltd. 890pp., 1992.
- /56/ G. Y. Saunier, T.A. Reddy, S. Kumar, "A monthly probability distribution

- 
- function of daily global irradiation values appropriate for both tropical and temperate locations". *Solar Energy*, Vol. 38, N.3, pp. 169-177. 1987.
- /57/ M. Shinozuka, C.M. Jan, "Digital simulation of random processes and its applications". *J. of Sound and Vibration*, 25(1), pp. 111-128, 1972.
- /58/ M. Sidrach de Cardona, Ll. Mora, "Optimization of hybrid photovoltaic/auxiliary generator systems for installations of rural electrification". 11th European Photovoltaic Solar Energy Conference. Montreux, Switzerland, 1992.
- /59/ M. Sidrach de Cardona, Ll. Mora, "Utilización de sistemas mixtos fotovoltaico-generador auxiliar para electrificación rural: análisis energético y económico". VII Congreso Ibérico de Energía Solar, Vigo, 1994.
- /60/ H. Suehrcke, P.G. McCormick, "The distribution of average instantaneous terrestrial solar radiation over the day". *Solar Energy*, Vol.42, N.4, pp.303-309, 1989
- /61/ E.Uriel, "Análisis de series temporales. Modelos ARIMA". Paraninfo, 270pp, 1992