

## CARTOGRAFÍA PREDICTIVA DE VARIABLES CLIMÁTICAS: COMPARACIÓN DE DISTINTOS MODELOS DE INTERPOLACIÓN DE LA TEMPERATURA EN ESPAÑA PENINSULAR

J. Bustamante\*

### RESUMEN

Se evalúa la capacidad de modelos de interpolación de tres familias: superficies de tendencias, regresiones múltiples con predictores derivados de un modelo digital de elevaciones (MDE) y kriging, para interpolar la temperatura en España peninsular. Los modelos se validaron con un conjunto independiente de estaciones control seleccionadas al azar. Los resultados indican que es posible estimar la temperatura en puntos donde no existan estaciones de registro con un error medio en torno al 9%.

Las superficies de tendencias lineales, superficies de tendencias polinómicas y el kriging (siempre corrigiendo la temperatura por la altitud) interpolaron significativamente peor que las superficies de interpolación mediante splines de suavizado, regresiones locales, y los modelos de regresión múltiple con predictores del MDE. Sin embargo, estos últimos no difirieron entre sí. Los modelos que mejor funcionan para interpolar una media de 30 años son también los que mejor funcionan para interpolar la media de un año. Así mismo, tienden a ser el mismo subconjunto de predictores topográficos derivados del MDE los mejores estimadores de la temperatura media independientemente del mes o la duración del periodo.

Del conjunto de los mejores modelos se eligió uno de los modelos de regresión del MDE para elaborar mapas de temperatura real mensual para el periodo 1965-1995 en España peninsular.

**Palabras clave:** mapas temperatura real, interpolación, kriging, regresión múltiple, loess, MDE, splines.

### ABSTRACT

#### Predictive cartography of climatic variables: Comparing interpolation models of temperature in peninsular Spain

Models of three families of interpolation techniques: trend surfaces, multiple linear regression with predictors derived from a digital elevation model (DEM) and kriging, were tested to interpolate temperatures in peninsular Spain. The models were validated with an independent random sample of meteorological stations. Results indicate that it is possible to estimate temperatures in points where there are no recording stations with an error around 9%.

Linear or polynomial trend surfaces, and kriging (always correcting temperature for altitude) performed significantly poorer as interpolators than smoothing spline surfaces, local regression surfaces, or multiple regression models with predictors derived from a DEM. The better models did not differ significantly among themselves. Models that perform better interpolating 30 years mean temperatures are also the best for a 1 year mean. Also, the same subset of topographic predictors derived from the DEM tend to be the better estimators of mean temperature independently of the month or the duration of the period interpolated.

One of the best multiple regression models with predictors from the DEM was used to produce maps of actual monthly temperatures for the period 1965-1995 in peninsular Spain.

**Key words:** actual temperature maps, interpolation, kriging, multiple regression, loess, DEM, splines.

---

\* Grupo de Ecología Espacial, Departamento de Biología Aplicada, Estación Biológica de Doñana, CSIC. Avda María Luisa s/n, 41013 – Sevilla, España.

## Introducción

El clima en la tierra se ha calentado aproximadamente 0,6°C a lo largo de los últimos 100 años (IPCC, 2002; Walther *et al.*, 2002) y los modelos climáticos predicen que las temperaturas continuarán incrementándose como resultado del aumento de la concentración de gases como CO<sub>2</sub> y metano que provocan un efecto invernadero (IPCC 2002). Esto ha provocado un renovado interés en el clima, en su funcionamiento y su distribución espacial, y hace especialmente deseable el ser capaz de estimar la distribución en el espacio de variables climáticas. El funcionamiento de los ecosistemas y la distribución de los organismos son algunos de los muchos aspectos que se sabe están afectados por el clima en la tierra. Ante un escenario de inminente cambio climático tanto para la agricultura como para la ecología resulta de gran interés la estimación con precisión de variables climáticas en lugares donde no existen estaciones que las registren. En este artículo se trata el problema de estimar la temperatura media mensual, o media anual, en un lugar donde no exista una estación meteorológica. Dado que inevitablemente existe un límite al número de estaciones meteorológicas en un territorio, en el momento que estemos interesados en los valores de temperatura con resoluciones iguales o inferiores a unos 10.000 km<sup>2</sup> nos encontraremos con que no existen estaciones suficientes para cubrir el territorio con una malla regular de datos registrados.

Existe una cierta controversia a cerca de cual es el mejor método para producir una superficie climática interpolada partiendo de un conjunto limitado de estaciones meteorológicas. Hay métodos muy diversos de interpolación, con casi infinitas variantes (ver por ejemplo: Malczewski, 1997), y cabe esperar además que varíen en su comportamiento dependiendo de la variable climática que se desee interpolar, la distribución de las estaciones, las características del área de estudio y la resolución espacial y temporal deseada. En el caso de Gran Bretaña, White y Smith (White, 1979; White *et al.*, 1982) usaron modelos de regresión múltiple para estimar un conjunto de variables climáticas a partir de predictores topográficos y de posición espacial. Sin embargo, esta metodología fue inicialmente muy criticada por una autoridad en el tema (Gregory, 1982). En Australia y Nueva Zelanda ha habido una preferencia por un método de interpolación no paramétrico denominado "thin plate splines" propuesto inicialmente por Wahba y Wendelberger (1980) y usado para interpolar la temperatura, precipitación y otras variables climá-

ticas a escala continental. Las superficies así generadas han sido muy utilizadas para la predicción del hábitat potencial de diversas especies de animales y plantas (Braithwaite *et al.*, 1989; Austin *et al.*, 1990; Lindenmayer *et al.*, 1990; Walker, 1990; Lindenmayer *et al.*, 1991; Thomas *et al.*, 1991). En Gran Bretaña, un estudio reciente (Lennon *et al.*, 1995) en el que se comparaban hasta 20 métodos de interpolación para la temperatura proponía como mejor método un método mixto en el que se usaba para interpolar un "thin plate spline" para las coordenadas espaciales (latitud, longitud) después de corregir los valores de temperatura mediante regresión lineal con un subconjunto de predictores derivados de un modelo digital de elevaciones (MDE). El siguiente mejor modelo era uno de regresión múltiple usando variables derivadas del MDE. Los autores concluían que los modelos de regresión o de splines usando exclusivamente latitud, longitud y altitud eran, en general, comparativamente insatisfactorios.

En el caso de España, no existen, hasta donde yo conozco, estudios comparativos similares publicados que permitan sugerir metodologías óptimas para interpolar superficies de temperatura. Se han publicado superficies interpoladas de variables climáticas para la Península Ibérica, entre otras, las temperaturas medias mensuales. El más conocido es el Atlas climático de España (Font Tullot, 1983a) que presenta mapas de temperaturas máximas, mínimas y medias mensuales a una escala 1:6.000.000 usando datos del periodo 1931-1960. El más reciente es la sección de climatología del Atlas Nacional de España (Arévalo Barroso, [1993]) que presenta superficies interpoladas de temperatura media anual y medias mensuales con datos del periodo 1956-1985 a una escala de 1:4.500.000. Sin embargo, en ninguno de estos atlas se detallan las metodologías que se han seguido para la interpolación. Según Lennon y Turner (1995) los mapas climáticos oficiales publicados por los institutos meteorológicos muestran las temperaturas mensuales mediante isolíneas interpoladas a mano. Las estimaciones para los puntos intermedios son pura conjetura. Además, los mapas de temperatura se corrigen al nivel del mar, lo que resulta conveniente en meteorología para eliminar el efecto dominante de la altitud en la temperatura. Desafortunadamente esto genera mapas de temperatura que carecen de sentido desde un punto de vista ecológico. Sin embargo, numerosos ecólogos no han sido conscientes de esto y han utilizado estas isolíneas ficticias para ajustarlas a la distribución de especies (v.g. Root, 1988) o como predicto-

res potenciales en modelos de su distribución (v.g. Donázar *et al.*, 1993; Bustamante, 1997) Desconozco con certeza si este es el caso concreto de los atlas climáticos de España publicados por el Instituto Nacional de Meteorología (Font Tullot, 1983a; Arévalo Barroso, [1993]), ya que no se indica expresamente, aunque es de sospechar que se utilice esta metodología. En el Atlas Nacional de España (Arévalo Barroso, [1993]) se señala "...se han seguido las recomendaciones de la Organización Meteorológica Mundial, con el fin de que esta información se pueda comparar con la contenida en atlas de países vecinos..."

Los mapas de temperatura real son difíciles de dibujar debido a la necesidad de superponer el mapa de altitudes al mapa de isotermas calculado a 0 m de altitud. Además, sería poco recomendable interpolar los valores brutos de las estaciones sin corregirlos a nivel del mar, ya que estaríamos intentando interpolar la superficie mucho más irregular de las altitudes con una red de puntos inadecuada. Los únicos mapas publicados que conozco de temperatura real interpolada de España peninsular son los de Sánchez Palomares *et al.* (1999) que utiliza para cada cuenca hidrográfica modelos de regresión múltiple por pasos usando como predictores la altitud, latitud y longitud y permitiendo el uso de polinomios de hasta grado 2. Existe una tesis doctoral (Ninyerola, 2000) que aborda la comparación de un grupo reducido de técnicas de interpolación para las temperaturas mínimas, medias y máximas en el caso de de la Península Ibérica. Los resultados, para el caso concreto de Cataluña, han sido publicados (Ninyerola *et al.*, 2000). Ninyerola emplea tanto para el caso de la Península (Ninyerola, 2000) como en el caso de Cataluña (Ninyerola *et al.*, 2000) modelos de regresión múltiple con un conjunto reducido de variables derivadas de un MDE y compara modelos globales frente a modelos realizados por cuencas hidrográficas.

Por todo esto, parece interesante abordar la comparación de un conjunto diverso de técnicas de interpolación, de las propuestas por distintos autores, de una manera sistemática, sobre el mismo conjunto de estaciones meteorológicas y para el mismo periodo temporal. Utilizando la técnica de interpolación más adecuada para el caso concreto de España peninsular, es de interés elaborar mapas de temperatura real que puedan ser de utilidad para ecólogos y agrónomos. Por ello he comparado distintos modelos de interpolación agrupados en tres familias: (1) superficies de tendencias con distintos grados de flexibilidad, usando como predictores latitud y longitud y corrigiendo para altitud, (2)

metodos de regresión usando como predictores un conjunto extenso de variables derivadas de un MDE, (3) kriging de los residuos de temperatura tras corregir el efecto de la altitud mediante una regresión lineal. Las preguntas que me planteo son: (1) Si las distintas familias de modelos, o modelos individuales difieren en su capacidad de interpolar la temperatura, (2) Si el comportamiento espacial de la temperatura en los distintos meses hace que unos modelos funcionen mejor en unos meses que en otros, o unos meses resulten más difíciles de interpolar que otros. (3) Si la integración temporal a la que se refieren los datos de temperatura media afectan al resultado de la interpolación (v.g. que tenga menos error la interpolación de la temperatura mensual de una media de 30 años que la del valor de un año, o que determinados modelos funcionen mejor con una serie larga que con una serie corta.

## Material y métodos

### GENERALES

El Instituto Nacional de Meteorología (INM) proporcionó en formato digital los datos de base consistentes en nombre e identificador de la estación, coordenadas geográficas en grados, minutos y segundos, altitud en metros, y temperatura media, en décimas de grado, para cada mes del periodo 1965-1995. Se seleccionaron 584 estaciones meteorológicas en España peninsular con la condición de un mínimo de 20 años de datos de temperatura media para cada uno de los meses del año en dicho periodo. Los valores de temperatura media mensual fueron calculados por el INM de la manera estándar, como la media de las temperaturas medias diarias, que a su vez se calculan como  $(TMAX+TMIN)/2$ , siendo TMAX la temperatura máxima diaria y TMIN la temperatura mínima diaria. La temperatura media en el periodo 1965-1995 para cada mes calendario (TM\_Enero a TM\_Diciembre) fue calculada como la media de los valores de todos los años con datos. La temperatura media anual (TMA) fue calculada como  $TMA = (TM\_Enero + TM\_Febrero + \dots + TM\_Diciembre) / 12$ .

Las estaciones meteorológicas se dividieron en dos grupos seleccionados al azar. El primero, denominado grupo de ajuste, con un 80% de las estaciones ( $n= 467$ ) se utilizó para la selección y el ajuste de los modelos y el segundo, denominado grupo control, con el 20% restante ( $n= 117$ ), se utilizó como para validar la capacidad predictiva de cada modelo (Fig. 1).

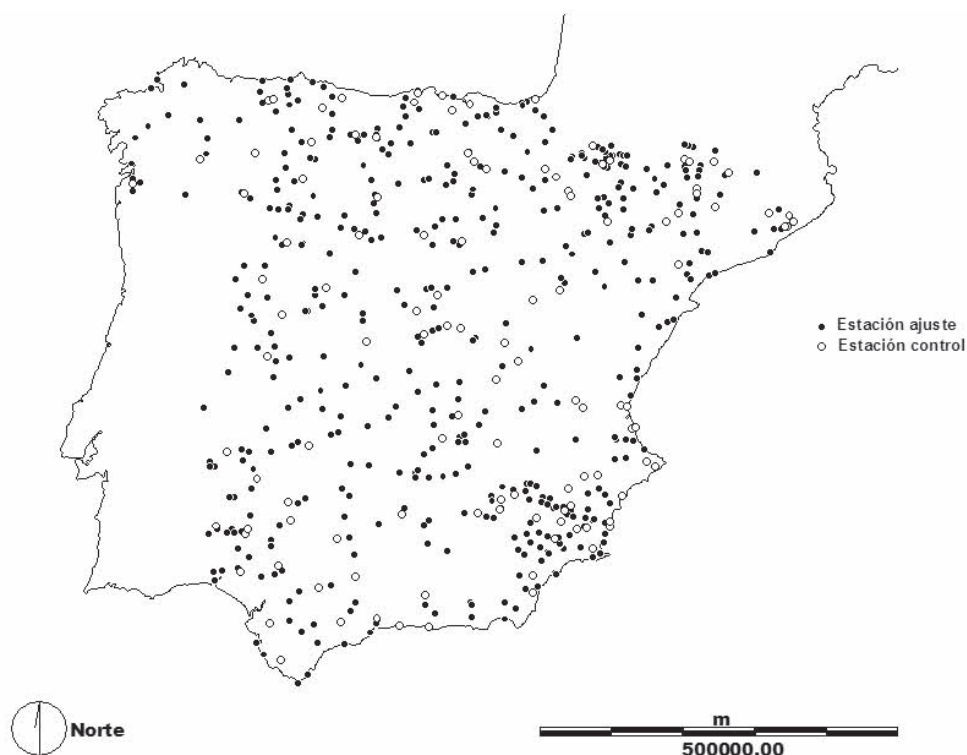


Fig. 1.— Localización de las estaciones meteorológicas empleadas en los modelos.

Fig. 1.— Location of weather stations employed in the models.

#### VARIABLES RESPUESTA

Se elaboraron modelos para la temperatura media anual (TMA) y temperaturas medias mensuales (TM\_Enero a TM\_Diciembre) del periodo completo 1965-1995. Siendo un total de 13 variables respuesta denominadas genericamente *temperaturas medias mensuales (30 años)*. También se seleccionaron al azar 30 meses entre Enero de 1965 y Diciembre de 1995, con la condición de tener al menos 340 estaciones con datos completos en el grupo de ajuste. Se ajustaron modelos para cada una de estas variables respuesta denominadas genericamente *temperaturas medias mensuales (1 año)*.

#### PREDICTORES TOPOGRÁFICOS

Para construir los modelos de regresión se generaron un total de 13 variables topográficas, geográficas y de insolación derivadas de un MDE (Tabla 1) que se utilizaron como predictores potenciales en los modelos. Se utilizó un MDE de la Península Ibérica en proyección UTM huso 30, Datum

Europeo 1950, en una retícula de 1x1 km que fue extraído del GTOPO30 del U.S.Geological Survey (<http://webgis.wr.usgs.gov/globalgis/gtopo30/gtopo30.htm>) que se encuentra en una proyección geográfica en grados decimales, con una resolución horizontal de 30 segundos de arco (aproximadamente 1 km), y usa como datum el WGS84.

Los predictores topográficos derivados del MDE son variables que desde un punto de vista teórico o empírico pueden tener una influencia en la temperatura media registrada en cada punto. Están, por una parte, la altitud nominal de la estación meteorológica y la altitud a distintas resoluciones espaciales derivadas del MDE. Estas variables pueden reflejar el efecto del enfriamiento adiabático del aire con la altitud (0.5°C por cada 100 m de altitud) y además del efecto de la orografía y topografía a distintas escalas espaciales (v.g. el enfriamiento del aire con la altura no es igual en un pico aislado, en una valle a gran altitud o en una meseta elevada). Las variables distancia al mar y porcentaje de terreno alrededor de la esta-

Tabla 1.— Variables empleadas como predictores potenciales en los modelos de regresión

Table 1.— Variables employed as potential predictors in the regression models.

| Nombre                | Descripción  |
|-----------------------|--|
| LATITUD               | Latitud en grados decimales  |
| LONGITUD              | Longitud en grados decimales   |
| ALTITUD               | Altitud nominal de la estación meteorológica según datos del INM, en m sobre el nivel del mar.   |
| ALT1KM                | Altitud media en un cuadrado de 1x1 km en el que se localiza la estación, obtenida del MDE, en m (s.n.m.).   |
| ALT7KM <sup>(1)</sup> | Altitud media en un cuadrado de 7x7 km centrado en la estación, obtenida del MDE, en m (s.n.m.).   |
| ALT11KM               | Altitud media en un cuadrado de 11x11 km centrado en la estación, obtenida del MDE, en m (s.n.m.).   |
| ALT21KM               | Altitud media en un cuadrado de 21x21 km centrado en la estación, obtenida del MDE, en m (s.n.m.).   |
| ALT41KM               | Altitud media en un cuadrado de 41x41 km centrado en la estación, obtenida del MDE, en m (s.n.m.).   |
| DIFALT                | Diferencia entre la altitud nominal de la estación estación y la del cuadrado de 7 x 7 km en el que está centrada (ALTITUD-ALT7KM)   |
| DISTMAR               | Distancia al mar en metros medida desde el el cuadrado de 1x1 km en el que se localiza la estación.  |
| DISTMONT              | Distancia a la cadena montañosa más próxima (distancia lineal a la zona más próxima con más de 1000 m de altitud)  |
| MAR41KM               | Superficie alrededor del punto cubierta por mar en un cuadrado de 41x41 km   |
| INSANU                | Insolación anual calculada en K <sub>j</sub> /cm <sup>2</sup> para un pixel de 1x1km con la orientación y la pendiente derivada del MDE. Se utiliza el programa Cagnard <sup>(2)</sup> para integrar cada 15 minutos la radiación teórica recibida por cada pixel del MDE (Durand <i>et al.</i> , 1981; Legros <i>et al.</i> , 1997). No tiene en cuenta las sombras proyectadas por el relieve ni la nubosidad. |

(1) Las estimas de altitud a distintas resoluciones se obtuvieron usando un filtro de media de forma cuadrada de la dimensión indicada sobre el DEM a 1 km de resolución en IDRISI32 (Eastman, 1999b)

(2) El programa Cagnard ha sido realizado por Jean-Paul Legros, INRA, Montpellier, Francia (legros@ensam.inra.fr)

ción que es mar intentan capturar el efecto amortiguador del mar en las temperaturas. La variable distancia a montaña pretende capturar el posible efecto barrera de estas así como el efecto de vientos fríos de ladera. La variable insolación anual podría de reflejar el efecto de las orientaciones de solana o de umbría en las zonas de montaña sobre la temperatura registrada. La latitud permite incorporar en los modelos el efecto del enfriamiento de las temperaturas hacia el norte debido a la distinta altura del sol sobre el horizonte.

Para los cambios de proyección y resolución del MDE así como para el cálculo de los predictores topográficos se utilizó IDRISI for Windows e IDRISI32 (Eastman, 1999b).

#### SUPERFICIES DE INTERPOLACIÓN

Las superficies de interpolación consisten en interpolar mediante una superficie más o menos flexible los valores de temperatura usando como predictores las coordenadas espaciales (LATITUD y LONGITUD) y corrigiendo simultáneamente el efecto de la altitud. En todos los casos, el efecto de la altitud se corrigió mediante una regresión lineal usando como predictor la altitud nominal de la estación (ALTITUD, Tabla 1). Como superficies de interpolación se evaluaron 5 modelos ditintos:

(M1) modelo lineal en el que se ajustó un término lineal para LATITUD, otro para LONGITUD y otro para su interacción.

$$T = a + b_1*(ALTITUD) + b_2*(LATITUD) + b_3*(LONGITUD) + b_4*(LATITUD*LONGITUD) \quad \text{Equ. 1}$$

(M2) modelo polinómico, en el que se consiguió una mayor flexibilidad de la superficie espacial mediante un polinomio de tercer grado.

$$T = a + b_1*(ALTITUD) + b_2*(LATITUD) + b_3*(LATITUD)^2 + b_4*(LATITUD)^3 + b_5*(LONGITUD) + b_6*(LONGITUD)^2 + b_7*(LONGITUD)^3 + b_8*(LATITUD*LONGITUD) + b_9*(LATITUD*LONGITUD)^2 + b_{10}*(LATITUD*LONGITUD)^3 \quad \text{Equ. 2}$$

(M3) modelo de splines en el que la mayor flexibilidad de la superficie espacial se consigue mediante el ajuste de splines de suavizado, con 3 grados de libertad (gl), a la LATITUD, LONGITUD y su interacción. s(X,3) equivale a un spline de suavizado con 3 gl.

$$T = a + b_1(ALTITUD) + s(LATITUD, 3) + s(LONGITUD, 3) + s(LATITUD*LONGITUD, 3) \quad \text{Equ. 3}$$



(M4) modelo de regresión local en el que se ajusta una superficie de suavizado no paramétrica tipo loess (Hastie *et al.*, 1990) a LATITUD y LONGITUD con entorno o “span”=0,8 y un polinomio de grado 2. En este tipo de regresión cada observación se estima mediante una superficie de tendencias polinómica de grado 2, utilizando un entorno del 80% de las observaciones en el que el peso de cada caso en el modelo de regresión es inversamente proporcional a la separación espacial respecto a la observación que se quiere estimar.

$$T = a + b_1(\text{ALTITUD}) + \text{lo}(\text{LATITUD}, \text{LONGITUD}, \text{span}=0.8, \text{gl}=2) \quad \text{Equ. 4}$$

(M5) Partiendo del modelo 4 se selecciona mediante un procedimiento por pasos el valor óptimo de entorno o “span” para la superficie entre 1 y 0,05

#### REGRESIÓN MÚLTIPLE CON PREDICTORES DERIVADOS DE UN MDE

Se basan en un proceso de selección de predictores de entre un conjunto de variables topográficas y espaciales, y el posterior ajuste de los coeficientes en un modelo de regresión lineal o aditivo (no paramétrico). Se evaluaron tres variantes:

(M6) Para la variable respuesta TMA se seleccionaron por pasos los predictores y la flexibilidad de suavizado óptimo utilizando para ello polinomios hasta un máximo de 3 gl. Se prefirieron polinomios por la facilidad de trasladar luego las ecuaciones a un sistema de información geográfica (SIG) para generar los mapas. El modelo final obtenido incluía los predictores: DIFALT, ALT7KM, MAR41KM, INSANU, LONGITUD y un polinomio de grado 3 de LATITUD. La exploración de este modelo mediante splines, testando la inclusión u omisión de otros predictores manualmente indicó que el polinomio de grado 3 representaba bien la respuesta de TMA con la latitud y que no existían comportamientos curvilíneos significativos para los restantes predictores. Partiendo de esta ecuación fija se estimaron para cada una de las variables respuesta de *temperaturas medias mensuales (30 años)* y *temperaturas medias mensuales (1 año)* los coeficientes de regresión óptimos de cada predictor.

(M7) Se seleccionaron para cada una de las variables respuesta los predictores óptimos en un

procedimiento de selección por pasos, permitiendo sólo el ajuste de modelos lineales (1 gl) para las variables del MDE y splines entre 1 y 3 gl para las coordenadas espaciales (LATITUD, LONGITUD y su interacción).

(M8) Se seleccionaron por pasos los predictores óptimos para cada una de las variables respuesta, ajustando en todos los casos relaciones de tipo curvilíneo mediante un spline de suavizado con 3 gl.

#### KRIGING

Kriging es un método de interpolación espacial propio de la geoestadística. Se se basa en el principio de que, en aquellas variables que cambian de manera continua en el espacio, los puntos próximos espacialmente tienden a tener valores más similares que los que están más alejados. Dado que el mayor efecto en la temperatura del aire medida en un lugar es el atribuible a la altitud se corrigieron previamente los valores de temperatura mediante regresión lineal, usando como predictor la altitud nominal de la estación. Los residuos de esta regresión se sometieron a kriging ordinario. Inicialmente con los residuos de TMA se analizaron los semivariogramas direccionales y omnidireccionales hasta conocer el comportamiento espacial de los residuos de temperatura. Con un semivariograma direccional representativo se ajustó visualmente un modelo matemático isotrópico (equ. 5) que explicara el patrón observado. Este modelo de dependencia espacial se usó para derivar los pesos de las estaciones cercanas en la estimación de la temperatura de cada punto usando un máximo de 30 estaciones vecinas.

$$\text{Variograma} = 30.000000 \text{Nug}(0) + 35 \text{Exp}(30000) + 75 \text{Pow}(1.8); \quad \text{Equ. 5}$$

Se asumió el mismo modelo de dependencia espacial para los residuos de todas las variables respuesta. El modelo generado mediante kriging (M9) consiste en estimar la temperatura en el punto en base a su altitud y estimar el residuo no explicado mediante kriging ordinario de las 30 estaciones más próximas.

#### AJUSTE DE MODELOS

Los modelos de superficies de interpolación y de regresión múltiple fueron ajustados en S-Plus 2000 (Mathsoft, 2000) usando los procedimientos glm para modelos lineales y gam para modelos aditivos (splines de suavizado y superficies de regresión local). Los modelos por pasos fueron ajustados

con el procedimiento *step.gam* que realiza una selección bidireccional de predictores tomando el AIC (Criterio de información de Akaike) como criterio de optimización (Sakamoto *et al.*, 1986).

Los modelos de kriging se ajustaron mediante una regresión lineal usando altitud como predictor en S-Plus y posteriormente interpolando los residuos de la regresión mediante kriging con la interfaz gráfica de IDRISI32 (Eastman, 1999a) para Gstat. Gstat ha sido desarrollado por Edzer J. Pebesma (Departamento de Geografía Física, Universidad de Utrech) y fue utilizado tanto para analizar los semivariogramas como para realizar el kriging y predecir para las estaciones control. Gstat está disponible de manera libre a través de una licencia GNU en <http://www.geog.uu.nl/gstat/>.

#### VALIDACIÓN CRUZADA

El porcentaje de varianza explicado por un modelo estimado a partir de los mismos datos utilizados para ajustarlo tiende a sobreestimar el valor real. Por ello, se utilizó un procedimiento de validación cruzada en el que los coeficientes de cada variable, y la selección de los predictores, si es el caso, se estiman con un conjunto de datos (en nuestro caso el 80% de las estaciones) y la estima de varianza explicada se hace con un conjunto independiente (el 20% restante) seleccionado al azar.

#### COMPARACIÓN ESTADÍSTICA DE LOS MODELOS

Se utilizó el porcentaje de varianza explicada sobre el conjunto de estaciones control ( $r^2 \cdot 100$ ) como medida de la capacidad explicativa de cada modelo. Dado que el porcentaje de varianza explicada está acotado entre 0 y 100, y no cabe esperar que se distribuya normalmente, se usaron test estadísticos no paramétricos para comparar modelos: test de Kruskal-Wallis para ANOVA de una vía y test de Wilcoxon-Mann-Whitney para comparación de pares de modelos (Siegel *et al.*, 1988).

#### DELINEACIÓN DE MAPAS

Las ecuaciones de los modelos finales seleccionados para cada temperatura media mensual se implementaron en IDRISI 32 generando superficies de temperatura media estimada para un retículo de 1x1 km de la Península Ibérica, usando como base el DEM anteriormente mencionado en proyección UTM 30. Aunque los modelos utilizados permiten estimar la temperatura también para Portugal la estima de varianza explicada por el modelo sólo es exacta para las zonas cubiertas por las estaciones meteorológicas, por lo tanto para la porción española de la Península Ibérica.

Los mapas originales, en formato digital con una resolución de 1 km<sup>2</sup> y temperaturas en décimas, pueden obtenerse directamente del autor o descargarse de: [ftp://ebd.csic.es/pub/busta/mapas\\_temperatura/](ftp://ebd.csic.es/pub/busta/mapas_temperatura/)

## Resultados

#### TEMPERATURA MEDIA ANUAL

La varianza explicada por cada modelo de interpolación para la temperatura media anual (TMA) en el periodo 1965-1995 sobre las estaciones control fue en orden creciente: M2 < M9 < M1 < M4 < M5 < M7 < M8 < M3 < M6 (Tabla 2). El mejor modelo, M6, incluía las variables: DIFALT, ALT7KM, MAR41KM, INSANU y LONGITUD como términos lineales y un polinomio de grado 3 de LATITUD y tuvo un error medio de 0,76 °C. El procedimiento de inclusión por pasos automático de predictores del MDE permitiendo sólo términos lineales (M7) incluyó además las variables: ALT41KM, DISTMAR y DISTMONT. Cuando se usaron splines de suavizado (M8) la ecuación de M7 perdió la variable MAR41KM. Sin embargo, ni M7 ni M8 fueron superiores a M6 evaluados con las observaciones de las estaciones control.

#### TEMPERATURA MEDIA MENSUAL (30 AÑOS)

Cuando se analizaron los modelos de interpolación para los datos de temperatura media de cada mes para el periodo de 30 años considerado, la mediana de la varianza explicada sobre los datos de las estaciones control osciló entre el 48,92% del modelo polinómico (M2) y el 90,76% de uno de los modelos de regresión múltiple con el MDE (M8) (Tabla 2). El porcentaje de varianza explicada por los distintos modelos de interpolación difirió significativamente entre modelos (Kruskal-Wallis  $c^2 = 61,70$ ;  $gl = 8$ ;  $p < 0.0001$ ). Como cabía esperar que la temperatura de los distintos meses difiriera en su comportamiento espacial (Fig. 2) y esto afectara de manera similar al porcentaje de varianza que pueden explicar los distintos modelos, se decidió tomar como referencia el modelo lineal (M1) y calcular el incremento de varianza respecto a éste para los demás modelos. Las diferencias entre modelos fueron aún mayores cuando se testaron sobre el incremento de la varianza explicada, (K-W  $c^2 = 63,52$ ;  $gl = 8$ ;  $p < 0.0001$ ). Sin embargo, no todos los modelos difirieron significativamente entre sí. Los modelos polinómico (M2) y de kriging (M9) fueron significativamente inferiores al modelo lineal (M1) con un incremento de varianza expli-

Tabla 2.— Validación estadística y error estandar medio de de los distintos modelos de interpolación empleados en este artículo.

Table 2.— Statistical validation and mean standar error of the different interpolation models used in this paper.

| Modelo | Tipo <sup>(1)</sup>  | T. media anual (TMA)  |                       | T. media mensual (30 años) <sup>(2)</sup> |        | T. media mensual (1 año) <sup>(2)</sup> |        |
|--------|--|-----------------------|-----------------------|---|--------|---|--------|
|        |  | % v.e. <sup>(3)</sup> | d.s.r. <sup>(4)</sup> | % v.e.                                    | d.s.r. | % v.e.                                  | d.s.r. |
| M1     | Superficie de tendencias lineal  | 90,81                 | 0,80                  | 85,51                                     | 1,05   | 80,73                                   | 1,29   |
| M2     | Superficie de tendencias polinómica  | 56,72                 | 1,74                  | 48,92                                     | 1,96   | 43,65                                   | 2,28   |
| M3     | Superficie de tendencias con splines   | 91,65                 | 0,76                  | 88,30                                     | 0,95   | 82,96                                   | 1,25   |
| M4     | Superficie de tendencias local (LOESS) entorno fijo                              | 90,84                 | 0,76                  | 89,03                                     | 0,86   | 82,18                                   | 1,20   |
| M5     | Superficie de tendencias local (LOESS) entorno variable                          | 90,85                 | 0,76                  | 89,20                                     | 0,87   | 82,44                                   | 1,19   |
| M6     | Regresión múltiple lineal del MDE con predictores fijos                          | 91,68                 | 0,76                  | 90,11                                     | 0,88   | 84,31                                   | 1,21   |
| M7     | Regresión múltiple lineal del MDE con selección por pasos de predictores         | 91,24                 | 0,78                  | 90,25                                     | 0,84   | 84,43                                   | 1,22   |
| M8     | Regresión múltiple no paramétrica del MDE con selección por pasos de predictores | 91,57                 | 0,77                  | 90,76                                     | 0,85   | 84,94                                   | 1,20   |
| M9     | Kriging  | 80,40                 | 1,16                  | 78,83                                     | 1,24   | 65,90                                   | 1,72   |

(1) Ver en métodos las diferencias entre cada tipo de modelo

(2) Mediana de los modelos para los distintos meses

(3) % v.e. = % varianza explicada

(4) d.s.r. = desviación estandar de los residuos en °C, calculado como  $\sqrt{\frac{(T_{obs} - T_{pred})^2}{n}}$ 

(5) Todas las estimas de validación calculadas sobre el conjunto independiente de 117 estaciones test

cada negativo de -32.4% y -5.8%, respectivamente (test de Wilcoxon-Mann-Whitney;  $Z = 4,41$ ;  $p < 0,0001$ ; y  $Z = 3,67$ ;  $p = 0,0002$ ; respectivamente) (Fig. 3). Los modelos de splines (M3), de regresión local (M4 y M5) y de regresión múltiple a partir del MDE (M6, M7 y M8) no difirieron significativamente entre si (K-W  $c^2 = 3,16$ ;  $gl = 5$ ;  $p = 0,68$ ); pero todos incrementaron la varianza explicada respecto al modelo lineal entre un 0,3% y un 5,2%, aunque el incremento no fue significativo para las regresiones locales (M4 y M5), (W-M-W; M3:  $Z = -4,41$ ;  $p < 0,0001$ ; M4:  $Z = -1,45$ ;  $p = 0,147$ ; M5:  $Z = -1,45$ ;  $p = 0,147$ ; M6:  $Z = -3,67$ ;  $p = 0,0002$ ; M7:  $Z = -3,67$ ;  $p = 0,0002$ ; M8:  $Z = -3,67$ ;  $p = 0,0002$ ).

Los modelos de selección por pasos de las variables topográficas y espaciales del MDE incluyeron siempre las variables DIFALT, ALT7KM y LATITUD, y casi siempre LONGITUD, INSANU y DISTMAR (Tabla 3). Los modelos lineales (M7) incluyeron por termino medio una variable más que los que usaron splines de suavizado (M8), tendiendo a incluir con más frecuencia MAR41KM, y los términos de altitud a distintas resoluciones, sobre todo ALT41KM.

#### TEMPERATURAS MEDIA MENSUAL (1 AÑO)

Cuando se analizaron los modelos de interpolación para la temperatura de un mes elegido al azar ( $n = 30$ ), la mediana de la varianza explicada sobre los datos de las estaciones control osciló entre el 43,65% del modelo polinómico (M2) y el 84,94% de uno de los modelos de regresión del MDE (M8) (Tabla 2). El porcentaje de varianza explicada por los distintos modelos de interpolación difirió significativamente entre modelos (K-W  $c^2 = 129,36$ ;  $gl = 8$ ;  $p < 0,0001$ ). Las diferencias entre modelos fueron aún mayores cuando se testaron sobre el incremento de la varianza explicada respecto al modelo lineal (M1), (K-W  $c^2 = 155,36$ ;  $gl = 8$ ;  $p < 0,0001$ ). Sin embargo, no todos los modelos difirieron significativamente entre si. Los modelos polinómico (M2) y de kriging (M9) fueron significativamente inferiores al modelo lineal (M1) con un incremento de varianza explicada negativo de -34,7% y -15.1%, respectivamente (W-M-W;  $Z = 6,65$ ;  $p < 0,0001$ ; y  $Z = 6,65$ ;  $p < 0,0001$ ; respectivamente) (Fig. 4). Los modelos de splines (M3), de regresión local (M4 y M5) y de regresión múltiple a partir del MDE (M6, M7 y M8) no difirieron significativamente entre si (K-W  $c^2 = 6,13$ ;  $gl = 5$ ;  $p = 0,29$ ); pero todos incre-



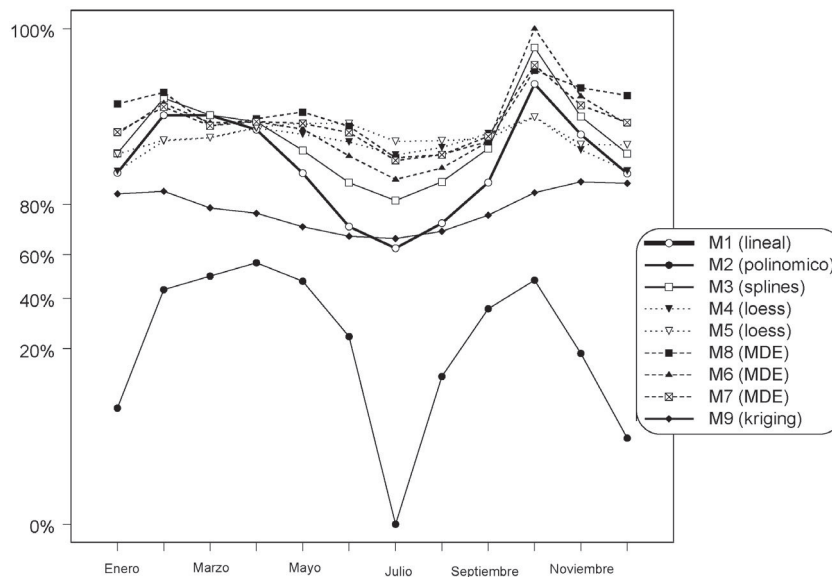


Fig. 2.— Porcentaje de varianza explicada de la temperatura media mensual por cada uno de los modelos de interpolación.

Fig. 2.— Percentage of explained variance of mean monthly temperatures for each interpolation model.

mentaron la varianza explicada respecto al modelo lineal entre un 0,8% y un 3,5%, aunque el incremento no fue significativo para las regresiones

Tabla 3.— Porcentaje de modelos de regresión múltiple por pasos que incluyen cada variable topográfica o espacial.

Table 3.— Percentage of stepwise multiple regression models that include each topographic or spatial variable.

|                                     | <i>medias mensuales<br/>30 años</i> |            | <i>medias mensuales<br/>1 año</i> |            |
|-------------------------------------|-------------------------------------|------------|-----------------------------------|------------|
|                                     | M7 lineal                           | M8 splines | M7 lineal                         | M8 splines |
| LATITUD                             | 100                                 | 100        | 100                               | 100        |
| LONGITUD                            | 92                                  | 92         | 87                                | 87         |
| ALT1KM                              | 17                                  | 0          | 20                                | 0          |
| ALT7KM                              | 100                                 | 100        | 100                               | 97         |
| ALT11KM                             | 25                                  | 0          | 17                                | 3          |
| ALT21KM                             | 0                                   | 0          | 3                                 | 0          |
| ALT41KM                             | 50                                  | 33         | 23                                | 0          |
| DIFALT                              | 100                                 | 100        | 100                               | 100        |
| DISTMAR                             | 83                                  | 83         | 93                                | 90         |
| DISTMONT                            | 33                                  | 25         | 23                                | 30         |
| MAR41KM                             | 83                                  | 63         | 80                                | 33         |
| INSANU                              | 92                                  | 83         | 46                                | 43         |
| Nº medio de variables<br>por modelo | 7,75                                | 6,75       | 6,90                              | 5,86       |

locales (M4 y M5), (W-M-W; M3:  $Z = -5,69$ ;  $p < 0,0001$ ; M4:  $Z = -0,47$ ;  $p = 0,64$ ; M5:  $Z = -0,94$ ;  $p = 0,34$ ; M6:  $Z = -4,74$ ;  $p < 0,0001$ ; M7:  $Z = -4,27$ ;  $p < 0,0001$ ; M8:  $Z = -4,27$ ;  $p < 0,0001$ ). Cuando se testó si había algún efecto atribuible al mes calendario que se pretendía interpolar no se encontraron diferencias significativas en la mayoría de los modelos, especialmente en los mejores, (K-W con 9 gl; M2:  $c^2 = 13,0$ ;  $p = 0,16$ ; M3:  $c^2 = 14,6$ ;  $p = 0,10$ ; M4:  $c^2 = 9,0$ ;  $p = 0,44$ ; M5:  $c^2 = 9,0$ ;  $p = 0,44$ ; M6:  $c^2 = 13,0$ ;  $p = 0,16$ ; M7:  $c^2 = 9,0$ ;  $p = 0,44$ ; M8:  $c^2 = 11,7$ ;  $p = 0,23$ ) aunque si se encontraron en dos de los modelos más ineficientes (K-W con 9 gl; M1:  $c^2 = 16,0$ ;  $p = 0,07$ ; y M9:  $c^2 = 18,0$ ;  $p = 0,04$ ).

Los modelos de selección por pasos de las variables topográficas y espaciales del MDE incluyeron siempre las variables DIFALT y LATITUD, y casi siempre ALT7KM, LONGITUD y DISTMAR (Tabla 3). Los modelos lineales (M7) incluyeron por término medio una variable más que los que usaron splines de suavizado (M8). Los modelos lineales tendieron a incluir con más frecuencia MAR41KM, y términos de altitud a distintas resoluciones: frecuentemente ALT41KM, ALT1KM, ALT11KM y solo en un caso ALT21KM.

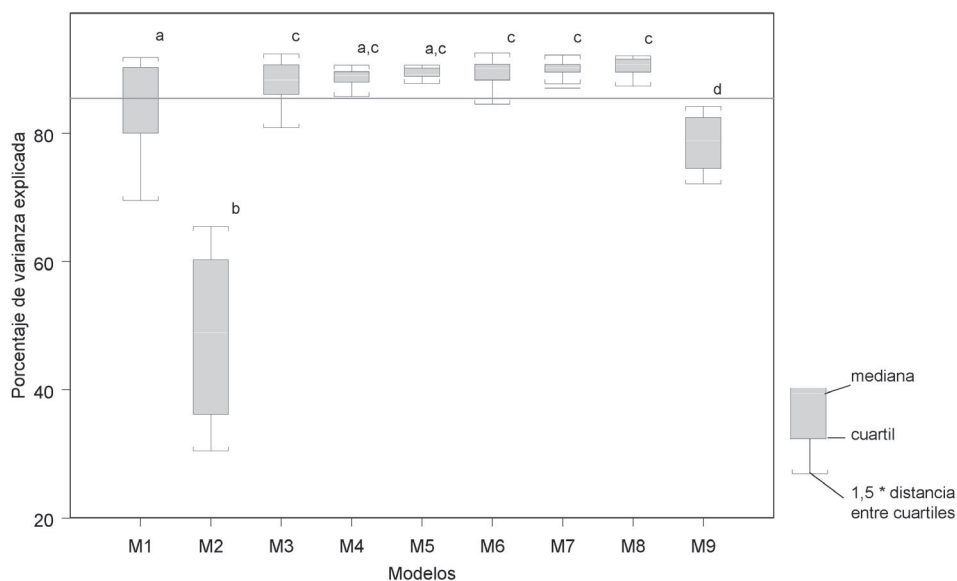


Fig. 3.— Distribución de los valores de porcentaje de varianza explicada para cada modelo para la temperatura media mensual (30 años). La línea central indica la mediana y las cajas el cuartil superior e inferior. Modelos con la misma letra no difieren significativamente entre si en base al incremento de varianza respecto al modelo lineal (M1).

Fig. 3.— Distribution of percentage of explained variance for each interpolation model for mean monthly temperature (30 years). Central line indicates the median and boxes the upper and lower quartile. Models with the same letter do not differ in their increase in variance in relation to the linear model (M1).

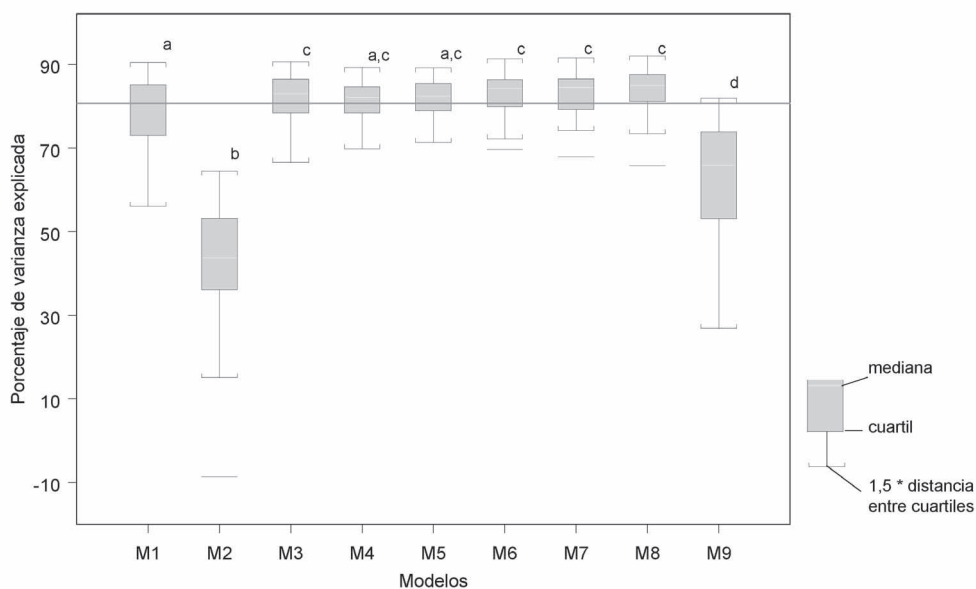


Fig. 4.— Distribución de los valores de porcentaje de varianza explicada para cada modelo para la temperatura media mensual (1 año). La línea central indica la mediana y las cajas el cuartil superior e inferior. Modelos con la misma letra no difieren significativamente entre si en base al incremento de varianza respecto al modelo lineal (M1).

Fig. 4.— Distribution of percentage of explained variance for each interpolation model for mean monthly temperature (1 year). Central line indicates the median and boxes the upper and lower quartile. Models with the same letter do not differ in their increase in variance in relation to the linear model (M1).

## COMPARACIÓN TEMPERATURAS MEDIAS MENSUALES PARA 1 Ó 30 AÑOS

Los modelos de interpolación predijeron una porción significativamente mayor de la varianza cuando se interpoló la media de 30 años que cuando se interpoló el valor de un año. El efecto fue significativo para todos los modelos usados (test W-M-W; M1:  $Z = 2,27$ ;  $p = 0,02$ ; M3:  $Z = 3,41$ ;  $p = 0,0006$ ; M4:  $Z = 4,64$ ,  $p < 0,0001$ ; M5:  $Z = 4,83$ ,  $p < 0,0001$ ; M6:  $Z = 3,99$ ,  $p = 0,0001$ ; M7:  $Z = 4,44$ ;  $p < 0,0001$ ; M8:  $Z = 4,30$ ,  $p < 0,0001$ ; M9:  $Z = 3,97$ ,  $p = 0,0001$ ) salvo para el polinómico (test W-M-W, M2:  $Z = 1,07$ ,  $p = 0,28$ ) que es, por otra parte, el más ineficiente. Por término medio, los mejores modelos (M3 a M8) explican entre el 90,8 y el 91,6% de la varianza de la *temperatura media anual*, entre el 88,3 y el 90,8% de las *temperaturas medias mensuales (30 años)*, y entre el 82,1 y el 84,9% de las *temperaturas medias mensuales (1 año)*. Sin embargo, los distintos modelos manifiestan el mismo comportamiento relativo cuando se trata de interpolar al *temperatura media anual*, *temperaturas medias mensuales de 1 ó 30 años*. Es decir, los mejores modelos M3 a M8 son los mismos para todas las variables respuesta y no difieren significativamente entre sí; pero si difieren del modelo de superficie de tendencias lineal (M1), del modelo de kriging (M9) y especialmente del modelo desuperficie de tendencias polinómica (M2), que resulta siempre el peor. Igualmente dentro de los modelos de regresión múltiple que utilizan predictores derivados del MDE, los mejores predictores tienden a ser siempre los mismos independientemente de que se interpole la media de un año o la de treinta.

### Discusión

Este estudio demuestra que es posible estimar los valores de temperatura observada con bastante precisión en lugares donde no existen estaciones de registro, y que estos modelos se pueden utilizar para generar mapas de temperatura real sobre un territorio como la Península Ibérica mediante un SIG.

No todos los modelos de interpolación se comportan igual. Un modelo relativamente simple como la interpolación lineal de altitud, latitud y longitud (M1) produce resultados bastante satisfactorios (entre un 80-90% de la varianza explicada); pero existen modelos significativamente mejores que incrementan la varianza explicada entre 1-5%. Curiosamente, hay modelos bastante dispares que

producen buenos resultados. Desde un punto de vista estadístico, no difieren los resultados obtenidos utilizando superficies de interpolación mediante splines de suavizado (M3), superficies de regresión local con entorno fijo (M4) o variable (M5), modelos de regresión con predictores espaciales y topográficos derivados de un MDE, tanto si se parte de un grupo fijo de predictores (M6), estos se seleccionan para cada modelo (M7) o incluso si además se utilizan modelos aditivos que permiten establecer relaciones curvilíneas no paramétricas (M8). De los modelos testados sólo las superficies polinómicas de tercer grado con coordenadas espaciales (M2) se comportaron mal. También la interpolación mediante kriging ordinario (M9) fue claramente inferior a otros modelos.

### ¿POR QUÉ FUNCIONAN MAL ALGUNOS MODELOS?

Los modelos polinómicos son aparentemente equivalentes en capacidad explicativa a otros modelos si se evalúan con los mismos datos de construcción, el grupo de estaciones de ajuste, sin embargo con un conjunto control se demuestra que esta varianza explicada está sesgada. Este resultado resalta la importancia de usar datos independientes para evaluar los modelos. Las regresiones polinómicas tienden a tener comportamientos erráticos en los extremos del intervalo de interpolación (Hastie *et al.*, 1990) y posiblemente por esta razón estiman valores alejados de la realidad para algunas estaciones control. Es por esto que los splines, que son funciones polinómicas con nudos fijos por los que se fuerza a pasar la curva, son preferidos a las regresiones polinómicas para el suavizado de datos (Hastie *et al.*, 1990).

El kriging al ser un interpolador exacto (la superficie de interpolación pasa por los puntos utilizados para construir el modelo) siempre explica el 100% de la varianza de los datos de las estaciones de ajuste y sólo puede ser evaluado mediante validación cruzada o con un conjunto independiente de observaciones control. Esta característica hace que errores en los datos (valores de temperatura, coordenadas o altitud de la estación) tengan más influencia potencial que en los modelos de regresión o en las superficies de tendencias. El modelo de kriging utilizado posiblemente funcione relativamente mal también por que al asumir la estacionaridad de la media no es capaz de considerar las tendencias espaciales a gran escala. Quizá otras técnicas de kriging como modelos anisotrópicos, el cokriging o el kriging con deriva fueran capaces de reflejar mejor estos efectos a gran escala en las temperaturas.

Tabla 4.— Coeficientes de los modelos usados para producir los mapas de temperatura media anual y temperatura media por meses (M6).

Table 4.— Coefficients of the models used to produce mean anual temperature and mean monthly temperature maps (M6)

| Modelo       | intercepto | DIFALT      | ALT7KM      | MAR41KM     | INSANU     | LONGITUD    | LATITUD   | LATITUD <sup>2</sup> | LATITUD <sup>3</sup> |
|--------------|------------|-------------|-------------|-------------|------------|-------------|-----------|----------------------|----------------------|
| TMA          | -10271,56  | -0,04669076 | -0,05148373 | -0,06559098 | 0,01319894 | 0,4466114   | 767,6764  | -18,63852            | 0,1489724            |
| T_Enero      | -27982,47  | -0,03924727 | -0,04786648 | 0,3666425   | 0,01548718 | -0,184876   | 2157,633  | -55,05766            | 0,4659942            |
| T_Febrero    | -14432,94  | -0,04775568 | -0,05485916 | 0,184218    | 0,01086873 | -0,07110641 | 1121,194  | -28,64409            | 0,2421529            |
| T_Marzo      | -6679,84   | -0,05092926 | -0,05534218 | -0,00323785 | 0,01097581 | 0,01001121  | 515,6513  | -12,87856            | 0,1057258            |
| T_Abril      | -1490,51   | -0,05501203 | -0,05786732 | -0,1343424  | 0,01196458 | 0,545128    | 110,4167  | -2,314463            | 0,01417406           |
| T_Mayo       | -746,0347  | -0,05189608 | -0,0556718  | -0,2997702  | 0,01247426 | 1,234512    | 35,48096  | 0,1047122            | -0,0104974           |
| T_Junio      | -594,4389  | -0,04886731 | -0,05107859 | -0,4257906  | 0,01330027 | 0,9662187   | -19,22028 | 2,653914             | -0,04171835          |
| T_Julio      | 1680,741   | -0,04462529 | -0,04458208 | -0,5434487  | 0,01418017 | 1,002498    | -207,7669 | 7,906954             | -0,09020951          |
| T_Agosto     | -3977,48   | -0,04281842 | -0,04561899 | -0,4305509  | 0,01529472 | 1,302746    | 235,2712  | -3,613969            | 0,0092539            |
| T_Septiembre | -9470,151  | -0,04910236 | -0,05163527 | -0,2170746  | 0,01550075 | 0,4133366   | 675,0639  | -15,38106            | -0,01138182          |
| T_Octubre    | -12297,78  | -0,04987019 | -0,05499399 | 0,03071714  | 0,01131059 | 0,5319385   | 933,0347  | -23,08484            | 0,1886361            |
| T_Noviembre  | -21513,27  | -0,04356873 | -0,05083692 | 0,2743191   | 0,01209765 | -0,04658796 | 1659,88   | -42,22385            | 0,3559117            |
| T_Diciembre  | -25754,51  | -0,03659654 | -0,047452   | 0,4112267   | 0,01493253 | -0,3444817  | 1995,477  | -51,12935            | 0,4344266            |

#### ¿QUÉ INDICAN LOS PREDICTORES TOPOGRÁFICOS EN LOS MODELOS DE REGRESIÓN?

Las variables derivadas del MDE que entran en los modelos de regresión indican la importancia de la altitud de la estación (DIFALT); pero también la altitud del entorno cercano, un área de 49 km<sup>2</sup> (ALT7KM). Las altitudes a otras escalas entran a veces en los modelos, especialmente si no se permiten relaciones curvilíneas con los predictores. Los coeficientes de ALT7KM y DIFALT en el modelo de predictores fijos (M6, Tabla 4) tienen un valor prácticamente constante y coinciden con la tasa de enfriamiento adiabático del aire en atmósfera libre (0,5°C cada 100 m). Debe tenerse en cuenta que las unidades de temperatura en los modelos están en décimas de grado.

La entrada siempre de dos o más altitudes en los modelos indica que además del enfriamiento adiabático del aire la topografía del entorno de la estación influye en la temperatura media registrada. Las coordenadas espaciales entran siempre en los modelos; pero mientras que la LONGITUD manifiesta casi siempre un comportamiento lineal la LATITUD tiene un comportamiento curvilíneo que requiere de un polinomio o un spline con al menos 3 gl. Este comportamiento de la LATITUD puede ser debido a la orientación este-oeste de gran parte de las cadenas montañosas que genera zonas latitudinales con distinto comportamiento térmico. El efecto de la LONGITUD cambia según el mes indicando un incremento de las temperaturas hacia el Este en los meses cálidos (Marzo a Octubre) y una disminución de ésta en los meses fríos (Noviembre a Febrero, Tabla 4). Este puede deberse a un incremento de la

continentalidad en las dos submesetas y en general de la Península en dirección Este (Font Tullot, 1983b). Las variables que reflejan el efecto amortiguador del mar, sobretodo DISTMAR, pero también MAR41KM entran con frecuencia en los modelos (Tabla 3). Los valores del coeficiente de MAR41KM indican este efecto amortiguador (efecto positivo en los meses fríos y negativo en los meses cálidos) Los modelos obtenidos también apoyan la hipótesis de que a esta escala las orientaciones de solana y umbría de las cadenas montañosas tienen un efecto en las temperaturas registradas. El efecto de la insolación es siempre positivo y relativamente constante (Tabla 4) aunque más marcado en verano e invierno que en primavera y otoño.

#### ¿CÓMO SE COMPARAN ESTOS MODELOS A OTROS EXISTENTES?

Resulta difícil comparar los modelos de este artículo con los de los atlas climáticos publicados (Font Tullot, 1983a; Arévalo Barroso, [1993]). La escasa información en estos últimos sobre la metodología empleada y la ausencia de cualquier tipo de validación estadística impide una comparación. Los modelos de Sánchez Palomares *et al.* (1999) son los únicos que conozco que estiman la temperatura real para España peninsular. Su metodología no coincide exactamente con ninguno de los modelos aquí empleados; pues divide España peninsular en cuencas hidrográficas y utiliza un procedimiento por pasos para el ajuste de polinomios de grado 2. Aunque cabe esperar que estos modelos se comporten de una manera intermedia entre el modelo lineal (M1) y el polinómico (M2); que son, en ambos casos, significativamente



peores que algunos otros modelos. Los valores de varianza explicada por los modelos de Sánchez Palomares *et al.* (1999) están entre 84-96% para la temperatura media anual y el 71-96% para la temperatura de cada mes. Estos valores son equivalentes a los de los modelos aquí desarrollados (90,8 - 91,6% para la temperatura media anual, y 88,3 - 90,8% para la temperatura de un mes). Sin embargo, las estimas de varianza explicada de los modelos de Sánchez Palomares *et al.* (1999) están sobreestimadas al calcularse sobre los mismos datos de construcción del modelo. Los modelos de Lenon y Turner (1995) para Gran Bretaña, basados en modelos de regresión con predictores derivados de un MDE y una superficie de splines de latitud-longitud proporcionaron resultados equivalentes a los de este artículo: 92,4% de varianza explicada para la temperatura media anual y entre 88,1 y 94,0% para la temperatura media de cada mes. La tesis de Ninyerola (2000) utiliza, para la Península Ibérica, un modelo de regresión múltiple por pasos con predictores derivados de un MDE intermedio entre mis modelos M6 y M7 y lo valida con estaciones meteorológicas independientes. Los predictores del MDE que él utiliza no son los mismos que en este trabajo. El tratamiento de la altitud y el relieve es menos exhaustivo, aunque se usa un modelo más sofisticado para calcular la insolación anual, que no suele entrar en los modelos. En su tesis, Ninyerola sólo proporciona valores de varianza explicada para los meses equinociales (marzo, junio, septiembre y diciembre) que resultan ser ligeramente inferiores a los de mi modelo M6 (Fig. 5), y es idéntico en el caso de la temperatura media anual ( $R^2 = 0,92$ ). Ninyerola concluye en su trabajo que un modelo por cuencas no mejora el modelo global para la Península Ibérica e incluye una sofisticación más consistente en interpolar los residuos del modelo de regresión mediante la inversa de la distancia al cuadrado, aunque no demuestra que ésto mejore significativamente los modelos.

#### ¿POR QUÉ SE INTERPOLAN MEJOR PERIODOS LARGOS QUE CORTOS?

Posiblemente porque las medias de periodos más largos están menos afectadas por errores de medida y son menos susceptibles a valores extremos difíciles de estimar. Entre la media de un año y la de treinta años hay un incremento significativo de un 6% de varianza explicada a favor de esta última y entre la media de un mes y la de todo el año para un periodo de 30 años hay una mejora de un 1% a favor de la media anual.

Aunque la capacidad predictiva de los modelos de interpolación es mejor cuanto más largo es el

periodo temporal que pretendemos interpolar (v.g. la media de treinta años frente a la media de un mes), la capacidad predictiva relativa de los distintos tipos de modelos testados no está afectada por la duración del intervalo. Más aún, las variables derivadas del MDE con mayor capacidad predictiva tampoco dependen de la duración del intervalo temporal y tienden a ser siempre las mismas. Por todo esto, un modelo que parta de predictores derivados de un MDE fijos, bien elegidos, funciona igual de bien que la selección de predictores caso a caso.

#### ¿QUÉ MODELO ELEGIR?

Ante la situación de distintos tipos de modelos que interpolan la temperatura de una manera equivalente hay otras razones que pueden tener peso a la hora de elegir uno u otro para generar mapas de temperaturas.

Las superficies de tendencias requieren sólo de tres variables (altitud, latitud y longitud) para generar una superficie interpolada; sin embargo los modelos más eficientes como son los splines y las regresiones locales al ser métodos no paramétricos carecen de una ecuación que se pueda implementar fácilmente en un SIG. El problema es solventable estimando el valor del término aditivo de cada variable, para cada valor distinto de esta, en un programa que pueda ajustar modelos tipo gam, como S-Plus y luego reclasificando los valores de esta variable a los del término aditivo en el SIG. El procedimiento es, cuando menos, trabajoso. Las regresiones locales tienen además el inconveniente de que no pueden realizar estimas fuera del polígono definido por las estaciones meteorológicas usadas, así que no se pueden hacer estimas para algunas zonas costeras de la Península Ibérica.

Los métodos de regresión múltiple con variables derivadas del MDE generan ecuaciones más complejas y con más predictores; pero son fácilmente implementables en un SIG si no requieren de términos no paramétricos. Dado que éstos no mejoran significativamente la capacidad predictiva del modelo, podemos prescindir de ellos. También resulta más simple ajustar modelos partiendo de un conjunto fijo de predictores que seleccionarlos paso a paso en cada caso. Computacionalmente el segundo procedimiento requiere tres veces más tiempo. Por estas razones he optado por usar el modelo M6, el más simple de calcular e implementar en un entorno SIG, y que no difiere significativamente de los otros, para producir mapas de temperatura media real en el periodo 1965-1995 para cada uno de los meses del año en España peninsular (Fig. 5). En la figura 6 se representa también el mapa de la temperatura media anual para el

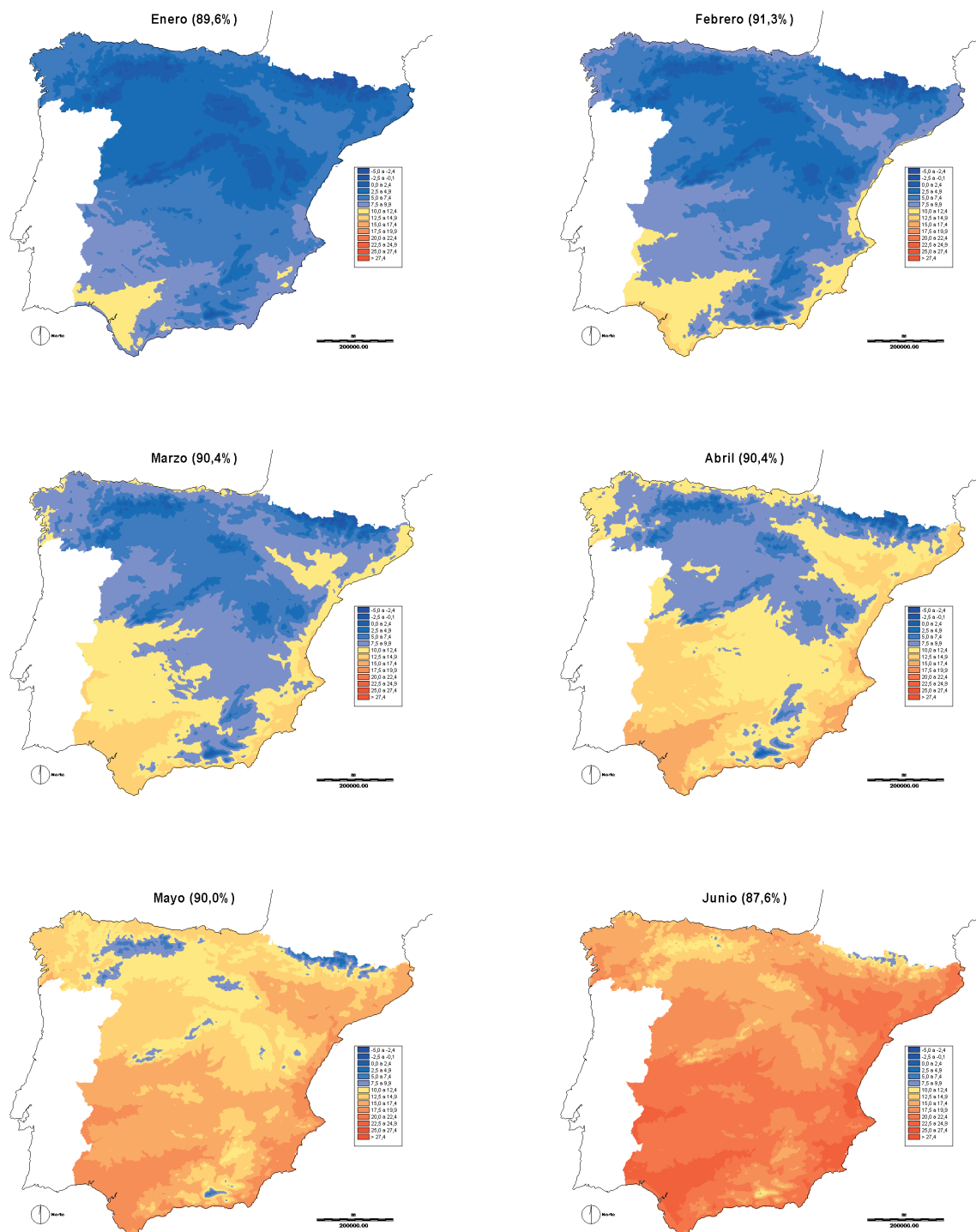


Fig. 5.— Mapas de la temperatura media mensual (1965-1995) reclasificados en intervalos de 2,5 °C, para facilitar la comparación con el atlas climático de España (Arévalo Barroso, [1993]). Se indica el porcentaje de varianza explicada por el modelo.

Fig. 5.— Mean monthly temperature maps (1965-1995). Percentage of explained deviance is indicated for each model.

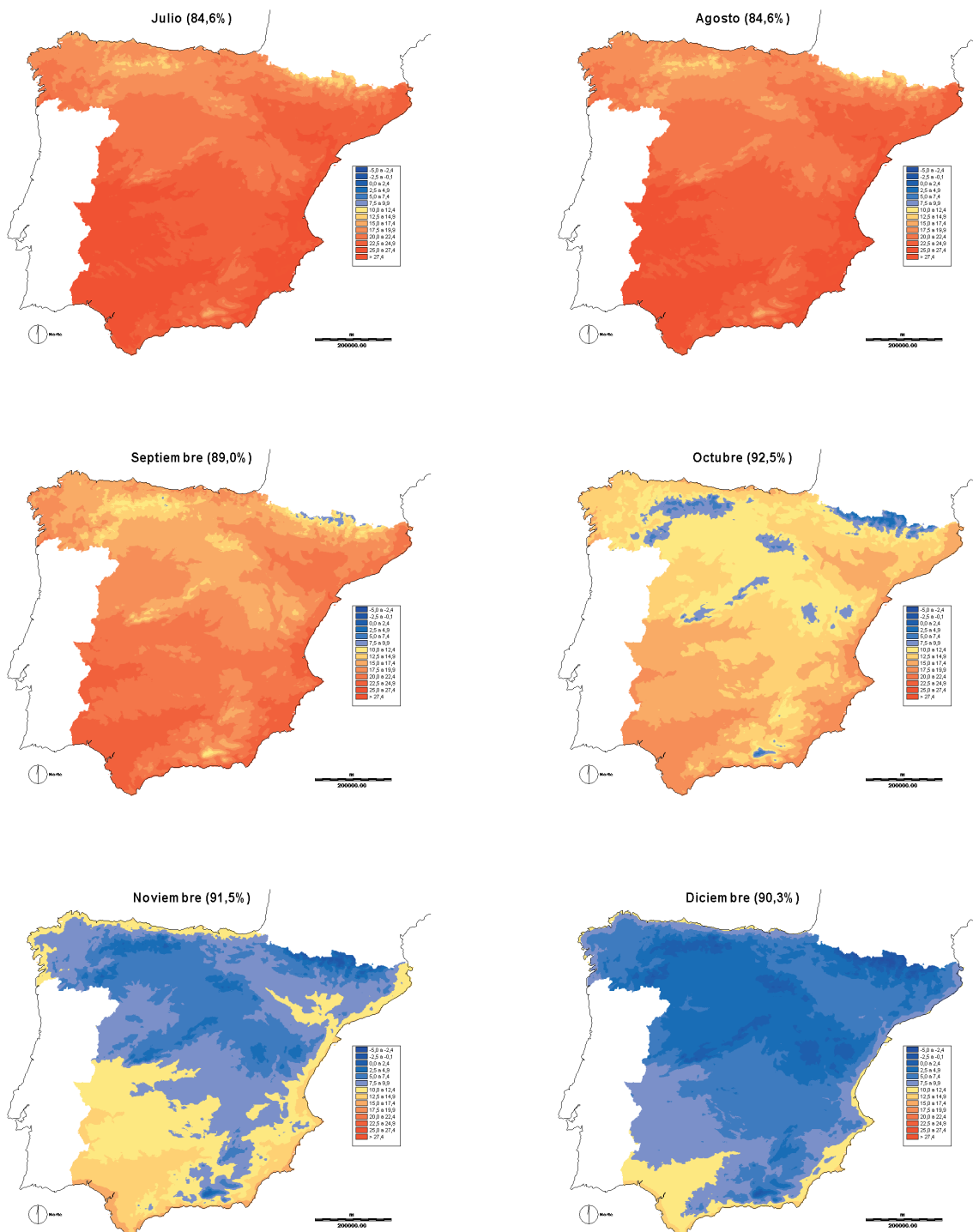


Fig. 5.— Continuación.

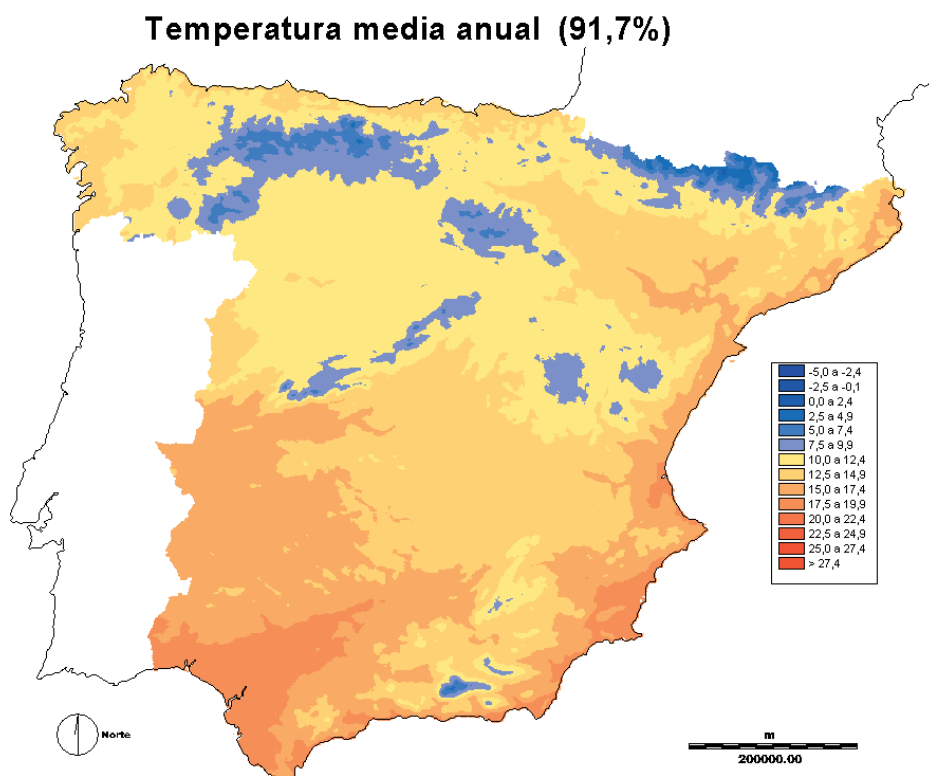


Fig. 6.— Mapa de la temperatura media anual 1965-1995 reclasificado en intervalos de 2,5 °C, para facilitar la comparación con el atlas climático (Arévalo Barroso, [1993]). Se indica el porcentaje de varianza explicada por el modelo.

Fig. 6.— Mean annual temperature map 1965-1995 reclassified in 2.5°C intervals to facilitate the comparison with climatic atlas (Arévalo Barroso, [1993]). The percentage of explained variance of the model is indicated.

mismo periodo. En todos los casos, para facilitar la interpretación y la comparación con los mapas publicados en atlas climáticos, los mapas se representan con los valores de temperaturas reclasificados a intervalos de 2,5 °C. Los mapas originales, con valores en décimas de grado y una resolución de 1 km, pueden conseguirse del autor o descargarse de [ftp://ebd.csic.es/pub/busta/mapas\\_temperatura/](ftp://ebd.csic.es/pub/busta/mapas_temperatura/)

#### ¿SON MEJORABLES ESTOS MODELOS?

Los modelos de kriging y los de regresión local parecen comportarse de forma distinta a los de regresión múltiple utilizando predictores del MDE y las superficies de interpolación lineales o mediante splines (Fig. 2); posiblemente porque los dos primeros reflejan mejor fenómenos locales y los últimos patrones más generales y a mayor escala. Una vía que

cabría explorar serían modelos que intentaran incorporar ambos tipos de patrones; por ejemplo modelos de regresión múltiple con variables derivadas del MDE que incorporaran un kriging o una regresión local de entorno variable de los residuos. Esta aproximación la emplea Ninyerola (2000), sin embargo no testa si mejora significativamente los modelos. En cualquier caso, no cabe esperar grandes mejoras ya que la varianza residual remanente 8-15% debe estar próxima a la varianza atribuible al error en los datos.

#### Conclusiones

Los resultados demuestran que es posible generar mapas de temperatura real mediante distintas técnicas de interpolación y que varias de estas téc-



nicas pueden llegar a generar mapas con una capacidad predictiva similar. Estos modelos de interpolación son muy superiores a los procedimientos de interpolación a mano propios de los atlas climáticos tradicionales, además permiten conocer el error de las estimas. Las mismas técnicas funcionan bien para interpolar datos de temperatura correspondientes a distintos meses del año y con datos con distinta resolución temporal. Todo ello hace aconsejable aplicar estas técnicas de modo operativo para generar mapas de temperaturas o de distintos índices climáticos de gran interés en un escenario de cambio climático donde serán de utilidad, entre otras cosas, para generar modelos de distribución potencial de especies

#### AGRADECIMIENTOS

A Jorge Lobo por invitarme a participar en este volumen homenaje a Fermín Martín Piera. Al Instituto Nacional de Meteorología que proporcionó los datos de temperatura y localización de las estaciones meteorológicas. A Manuel de la Puente, Javier Seoane y Ricardo Díaz Delgado por dicitir distintos aspectos de los modelos utilizados y por su ayuda para implementar el ajuste y validación de modelos en S-Plus y la producción de mapas en IDRISI. El análisis de los datos en los que se basa este artículo se realizó gracias a los proyectos REN2001-2134/GLO y 1DF97-0648 financiados con fondos FEDER.

#### Referencias

- ARÉVALO BARROSO, A., [1993]. *Atlas Nacional de España. Sección II. Grupo 9. Climatología*. Dirección General del Instituto Geográfico Nacional. Madrid. [31] pp.
- AUSTIN, M. P., NICHOLLS, A. O. & MARGULES, C. R., 1990. Measurement of the realized qualitative niche: environmental niches of five Eucalyptus species. *Ecological Monographs*, 60: 161-177.
- BRAITHWAITE, L. W., AUSTIN, M. P., CLAYTON, M., TURNER, J. & NICHOLLS, A. O., 1989. On predicting the presence of birds in eucalyptus forest types. *Biological Conservation*, 50: 33-50.
- BUSTAMANTE, J., 1997. Predictive models for lesser kestrel *Falco naumanni* distribution, abundance and extinction in southern Spain. *Biological Conservation*, 80: 153-160.
- DONÁZAR, J. A., HIRALDO, F. & BUSTAMANTE, J., 1993. Factors influencing nest site selection, breeding density and breeding success in the bearded vulture (*Gypaetus barbatus*). *Journal of Applied Ecology*, 30: 504-514.
- DURAND, R. & LEGROS, J. P., 1981. Cartographie automatique de l'énergie solaire en fonction du relief. *Annales Agronomiques*, 1(1): 31-39.
- EASTMAN, J. R., 1999a. *IDRISI32, version 1 32.04*. Clark Labs, Clark University. Worcester.
- EASTMAN, J. R., 1999b. *IDRISI32: Guide to GIS and Image Processing*, v. 1. Clark Labs, Clark University. Worcester. 193 pp.
- FONT TULLOT, I., 1983a. *Atlas climático de España*. Instituto Nacional de Meteorología. Madrid. [48] pp.
- FONT TULLOT, I., 1983b. *Climatología de España y Portugal*. Instituto Nacional de Meteorología. Madrid. 296 pp.
- GREGORY, S., 1982. [Review of White & Smith 1982]. *Weather*, 38: 284.
- HASTIE, T. J. & TIBSHIRANI, R. J., 1990. *Generalized Additive Models: Monographs on Statistics and Applied Probability*, v. 43. Chapman & Hall. London. 335 pp.
- IPCC (INTERGOVERNMENTAL PANEL OF CLIMATE CHANGE), 2002. *Climate Change and Biodiversity. IPCC Technical Paper V*. WMO-UNEP. 77 pp.
- LEGROS, J. P., ANTONIOLETTI, R., DURAND, R., FALIPOU, P. & ROBBEZ-MASSON, J. M., 1997. *FLASH III, Estimation des ressources en énergie solaire, Programmes Fortran et notice d'emploi*. INRA-EA. Montpellier. 23 pp.
- LENNON, J. J. & TURNER, J. R. G., 1995. Predicting the spatial distribution of climate: temperature in Great Britain. *Journal of Animal Ecology*, 64: 370-392.
- LINDENMAYER, D. B., NIX, H. A., MCMAHON, J. P., HUTCHINSON, M. F. & TANTON, M. T., 1991. The conservation of Leadbeater's possum *Gymnobelideus leadbeateri* (McCoy): a case study of the use of bioclimatic modelling. *Journal of Biogeography*, 18(4): 371-383.
- LINDENMAYER, D. B., NIX, H. A., MCMAHON, J. P. & HUTCHINSON, M. F., 1990. Bioclimatic modelling and wildlife conservation and management - a case study on Leadbeater's possum, *Gymnobelideus leadbeateri*. In: W. Clark & J. H. Seebeck (eds.). *Management and conservation of small populations*. Chicago Zoological Society. Chicago: 253-273.
- MALCZEWSKI, J. (ed.), 1997. *Special Issue: Spatial Interpolation Comparison 97*. Journal of Geographic Information and Decision Analysis, 2, 1-2. GIDA. Ontario. 251 pp.
- MATHSOFT, I., 2000. *S-Plus 2000, version 2000*. Seattle.
- NINYEROLA, M., 2000. *Modelització climàtica mitjançant tècniques SIG i la seva aplicació a l'anàlisi quantitativa de la distribució d'espècies vegetals a l'Espanya peninsular*. Tesis Doctoral. Universitat Autònoma de Barcelona. Barcelona. 202 pp.
- NINYEROLA, M., PONS, X. & ROURE, J. M., 2000. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, 20: 1823-1841.

- ROOT, T., 1988. Environmental factors associated with avian distributional boundaries. *Journal of Biogeography*, 15: 489-505.
- SAKAMOTO, Y., ISHIGURO, M. & KITIGAWA, G., 1986. *Akaike Information Criterion Statistics: Mathematics and its Applications*. D. Reidel Publishing Company. Tokyo. 288 pp.
- SÁNCHEZ PALOMARES, O., SÁNCHEZ SERRANO, F. & CARRETERO CARRERO, M. P., 1999. *Modelos y cartografía de estimaciones climáticas termopluviométricas para la España peninsular*. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria. Madrid. 192 pp.
- SIEGEL, S. & CASTELLAN, N. J., JR., 1988. *Nonparametric statistics for the behavioral sciences*. McGraw-Hill. New York. 399 pp.
- THOMAS, Y. & NEIL, M., 1991. Generalized additive models in plant ecology. *Journal of Vegetation Science*, 2: 587-602.
- WAHBA, G. & WENDELBERGER, J., 1980. Some new mathematical methods for variational objective analysis using splines and cross validation. *Monthly Weather Review*, 108: 1122-1143.
- WALKER, P. A., 1990. Modelling wildlife distribution using a geographic information system - kangaroos in relation to climate. *Journal of Biogeography*, 17: 279-289.
- WALTHER, G., POST, E., CONVEY, P., MENZEL, A., PARMESAN, C., BEEBEE, T. J. C., FROMENTIN, J., HOEGH-GULDBERG, O. & BAIRLEIN, F., 2002. Ecological responses to recent climate change. *Nature*, 416: 389-395.
- WHITE, E. J., 1979. The prediction and selection climatological data for ecological purposes in Great Britain. *Journal of Applied Ecology*, 16: 141-160.
- WHITE, E. J. & SMITH, R. I., 1982. *Climatological Maps of Great Britain*. Institute of Terrestrial Ecology. Midlothian.