

# Herramientas gráficas para el análisis descriptivo de series temporales en la investigación médica



Aurelio Tobías<sup>a,b</sup>, Marc Sáez<sup>c</sup> e Iñaki Galán<sup>b</sup>

<sup>a</sup>Departamento de Estadística y Econometría. Universidad Carlos III de Madrid. Getafe. Madrid.

<sup>b</sup>Instituto de Salud Pública, Servicio de Epidemiología. Consejería de Sanidad. Comunidad de Madrid. Madrid. España.

<sup>c</sup>Grup de Recerca en Estadística, Economia Aplicada i Salut (GRECS), Departament d'Economia. Universitat de Girona. Girona. España.

Una serie temporal es una colección de observaciones ordenadas que se han recogido de manera periódica en el tiempo. Las aplicaciones del análisis de series temporales en el ámbito de la investigación médica han sido ampliamente descritas<sup>1</sup>. Aunque en la bibliografía nacional de la última década encontramos numerosos ejemplos del análisis de series temporales para la vigilancia y predicción de la infección nosocomial<sup>2</sup>, epidemias de asma<sup>3</sup>, fiebre tifoidea<sup>4</sup> y, recientemente, de la enfermedad meningocócica<sup>5</sup>, así como análisis de regresión con series temporales, principalmente en el estudio de los efectos a corto plazo de la contaminación atmosférica sobre indicadores de mortalidad<sup>6-9</sup> y morbilidad<sup>7,8</sup>, observamos que se ha prestado muy poca atención a sus análisis descriptivos. Asimismo, aunque existen revisiones sobre los diversos métodos y modelos estadísticos para el análisis de series temporales en ciencias de la salud<sup>10,11</sup>, ninguno de ellos hace hincapié en las herramientas gráficas disponibles para su análisis descriptivo.

Los gráficos descriptivos pueden utilizarse para presentar y describir tendencias y/o relaciones complejas de una manera simple e informal que permite al investigador entender de una forma sencilla la estructura subyacente de los datos<sup>12,13</sup>. El objetivo de este artículo es presentar diversos métodos gráficos que permitan mejorar el entendimiento de la variación subyacente dentro de los datos de una serie temporal. Para examinar la estructura subyacente dentro de una serie temporal, comentaremos los gráficos temporales y la aplicación de diversos tipos de suavizados, gráficos de cajas y de subseries estacionales y, por último, los gráficos de descomposición estacional. Tal como mostraremos, dichas herramientas gráficas resultan de gran importancia para estudiar tanto la magnitud de la tendencia como la estructura de la estacionalidad de la serie temporal.

Ninguno de los métodos que se presentan en este artículo es original. Se han venido utilizando desde ya hace años en diversos campos de investigación, principalmente en el de la economía. El presente artículo pretende mostrar que las técnicas gráficas resultan de gran utilidad en el análisis de series temporales en la investigación médica. Para ello, ilustramos diversos ejemplos a través del análisis descriptivo del estudio de los efectos a corto plazo de la contaminación atmosférica<sup>14</sup> y las concentraciones de polen<sup>15</sup> en las urgencias hospitalarias por asma en la ciudad de Madrid.

Este estudio ha sido promovido por la comisión asesora del Programa Regional de Prevención y Control del Asma de la Comunidad de Madrid.

Correspondencia: Prof. A. Tobías.  
Departamento de Estadística y Econometría.  
Universidad Carlos III de Madrid.  
Madrid, 126. 28903 Getafe. Madrid. España.  
Correo electrónico: atobias@est.econ.uc3m.es

Recibido el 9-7-2003; aceptado para su publicación el 12-2-2004.

## Material y método

### Estructura de una serie temporal

Un enfoque habitual en el análisis de series temporales es considerar la serie temporal como una mezcla de varias componentes. En la figura 1 se muestra la estructura de una serie temporal que asume una descomposición aditiva:  $Y_t = T_t + S_t + a_t$ , donde  $Y_t$  denota la serie temporal de interés,  $T_t$  denota la componente de la tendencia,  $S_t$  denota la componente estacional y  $a_t$  la componente residual o irregular<sup>16</sup>.

Es importante identificar la tendencia y las componentes estacionales, y eliminarlas de la serie temporal cuando queremos relacionar series temporales. Cuando esto no se hace, las series con tendencia y/o estacionales pueden parecer relacionadas, pero esto es puramente debido a su distribución temporal, no a la existencia de una relación causal. Un ejemplo clásico de confusión temporal por omisión es la relación entre la tasa de nacimientos y el número de cigüeñas en Oldenburg (Alemania) entre 1930 y 1936<sup>17</sup>. Ambas variables presentaban una alta correlación debida a que las dos aumentaban simultáneamente a lo largo del tiempo y dicha tendencia se omitía en el análisis de correlación. Existen numerosos ejemplos de confusión temporal por omisión de tendencia (número de bibliotecas y consumo de drogas) y/o estacionalidad (consumo de helados y fallecimientos por ahogo)<sup>18</sup>. Por consiguiente, la estimación de la tendencia y las condiciones estacionales, y el cálculo de series ajustadas que eliminen su efecto, son problemas importantes que deben contemplarse en cualquier análisis de series temporales<sup>16</sup>.

## Herramientas gráficas

### Gráficos temporales y suavizados

El diagrama de la serie temporal es la herramienta gráfica utilizada habitualmente. Resulta de utilidad para examinar el modelo subyacente a largo plazo de una serie temporal. El tiempo  $t$  (días, meses, años, etc.) se representa en el eje de ordenadas ( $X_t$ ), y la serie temporal se representa en el eje de abscisas ( $Y_t$ ). Aunque estos diagramas proporcionan una perspectiva de la tendencia a largo plazo, no siempre permiten una comprensión visual clara de las fluctuaciones de cada observación desde un intervalo de tiempo al siguiente. La tendencia puede acentuarse, y se pueden examinar visualmente intervalos de tiempo específicos superponiendo al gráfico una línea de suavizado.

**Suavizado paramétrico.** La mayoría de las series temporales en ciencias de la salud presentan distribuciones no estacionarias, es decir, con tendencia y/o con componentes estacionales de periodicidad constante o casi constante. El suavizado paramétrico resulta de un modelo que parte de

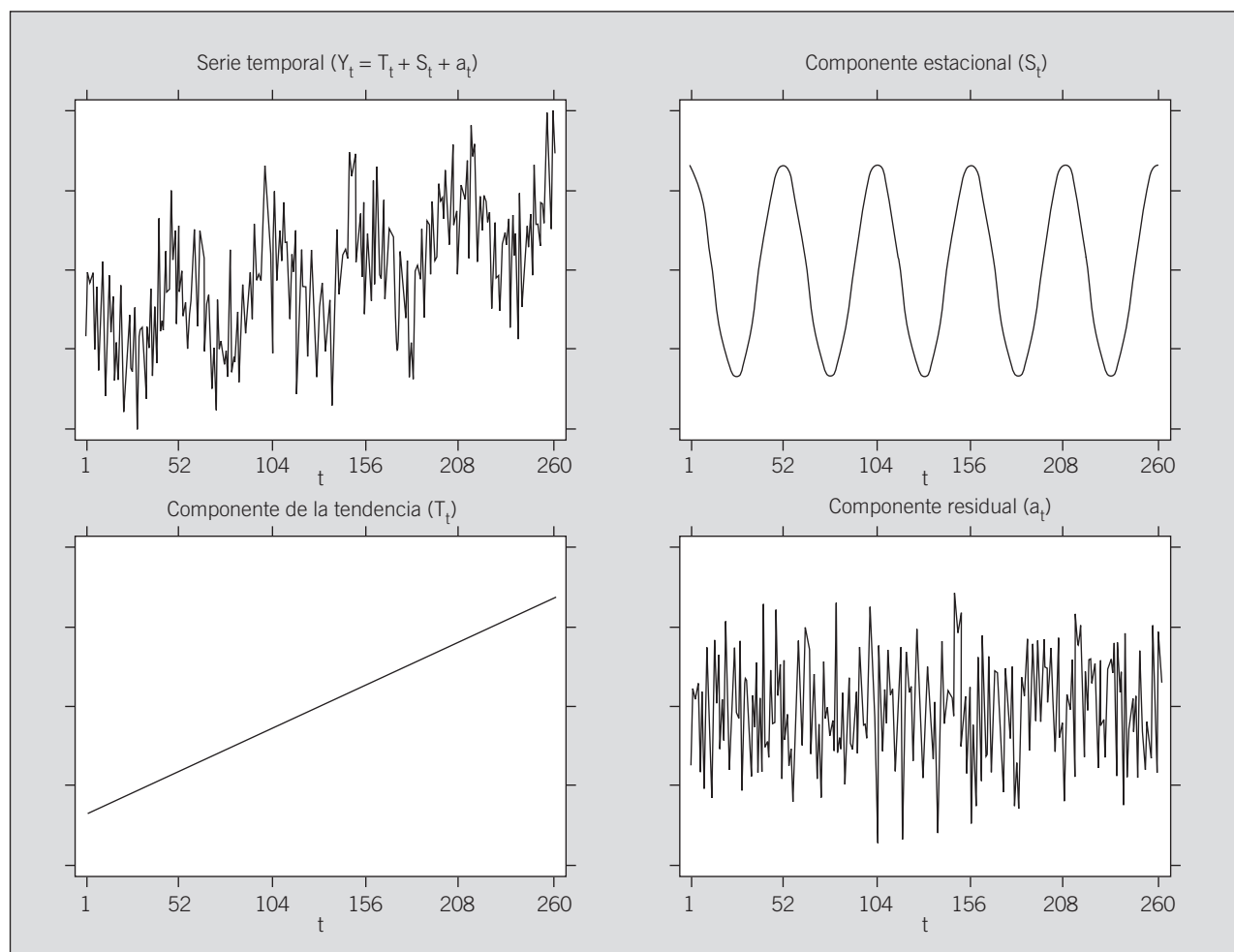


Fig. 1. Descomposición aditiva de una serie temporal.  $Y_t$ : serie temporal de interés;  $T_t$ : componente de la tendencia;  $S_t$ : componente estacional;  $a_t$ : componente residual o irregular.

suponer que la serie temporal sigue una función que combina un término lineal, que describe la tendencia ( $T_t$ ), y términos no lineales de senos y cosenos, que describen la variación estacional ( $S_t$ )<sup>19-21</sup>:

$$E[Y_t | t] = \mu + \beta T_t + \sum_{i=1}^k [\gamma_i \sin(\theta(t)) + \delta_i \cos(\theta(t))] + a_t$$

donde  $\beta$  es el parámetro de tendencia lineal;  $\gamma_i$  y  $\delta_i$  los parámetros de las componentes estacionales, y  $\theta(t) = 2\pi t/T$ .

**Suavizado no paramétrico.** Las curvas de suavizado no paramétricas no establecen *a priori* ninguna restricción sobre la estructura de la serie temporal y así permiten que los datos nos indiquen su verdadera forma funcional. Una curva de suavizado puede estimarse a partir de diversas técnicas; las más frecuentes son los alisadores (o suavizados): *kernel*, *cubic splines* y *lowess*<sup>22,23</sup>. En general, cualquiera de estos alisadores proporciona una estimación no paramétrica del valor medio de la serie temporal para un instante dado de tiempo. Por lo tanto, la curva de suavizado no paramétrica  $s(t)$  es una función que permite modelar simultáneamente tendencia y estacionalidad con el mismo intervalo cuando los valores en un instante de tiempo  $t$  cumplen que:

$$E[Y_t | t] = s(t) + a_t$$

### Subseries estacionales

Tendencia y estacionalidad son los principales elementos de confusión en el análisis de datos diarios de mortalidad y/o morbilidad<sup>10,11,21</sup>. Por consiguiente, es importante identificar la fuerza y magnitud tanto de la tendencia como de la componente estacional en una serie temporal. Para ello, se pueden construir gráficos de subseries estacionales mediante la estratificación de la serie temporal agregando, y representado gráficamente, sus valores por diversas unidades de tiempo (años, meses, etc.) a fin de evaluar la conducta de cada subserie.

**Diagramas de cajas.** En el análisis mediante gráficos de subseries estacionales, cabe la posibilidad de representar las subseries mediante diagramas de cajas. El diagrama de cajas<sup>12</sup> permite mostrar tanto el comportamiento del modelo global como el de las distintas subseries mensuales incluidas dentro de la serie temporal. El gráfico de cajas estacional se construye situando el tiempo sobre el eje horizontal, y cada caja representa la distribución de los datos para una determinada unidad de tiempo (años, meses, etc.) dentro de la serie temporal. Observando cada una de las cajas después de agregar por la mayor unidad de tiempo (habitualmente años), se puede determinar visualmente la mediana de la tendencia de las series, mientras que a partir de la representación gráfica de la agregación por unidades menores de

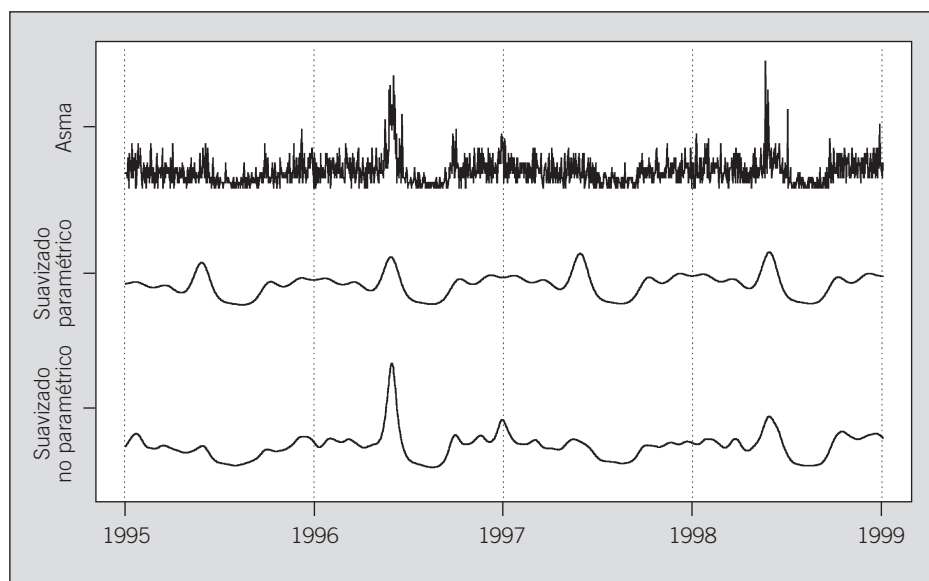


Fig. 2. Gráfico de la serie temporal, y de sus suavizados paramétrico y no paramétrico (lowess), de las urgencias hospitalarias diarias por asma en Madrid entre 1995 y 1998.

tiempo (p. ej., meses) puede determinarse la estructura de la componente estacional (p. ej., a partir de la variabilidad mensual). En cualquier caso, una caja muy amplia implicará que los datos poseen mucha variabilidad, mientras que una caja muy estrecha querrá decir lo contrario, esto es que, por ejemplo, un determinado mes se comporta de forma muy similar en los sucesivos años que componen la serie temporal.

**Subseries estacionales.** El gráfico de subseries estacionales<sup>24</sup> se construye representando, para cada unidad de tiempo menor (p. ej., los meses), la media de los valores de la serie temporal mediante una línea horizontal, mientras que las líneas verticales que parten de la horizontal son los valores por cada subserie. Así, de nuevo se puede evaluar la estacionalidad global de las series mediante las medias

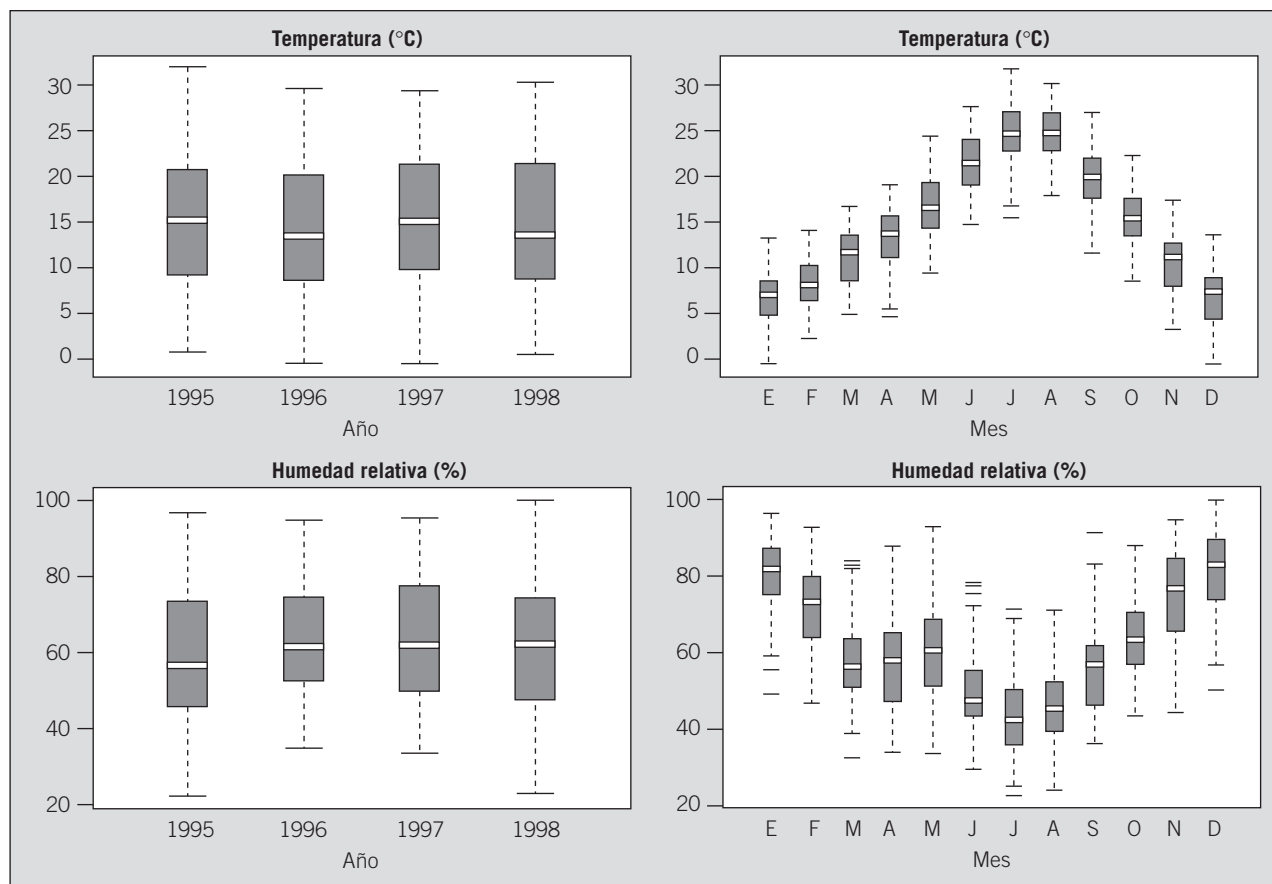


Fig. 3. Diagramas de cajas estacionales para variables meteorológicas (temperatura y humedad relativa) en Madrid entre 1995 y 1998.

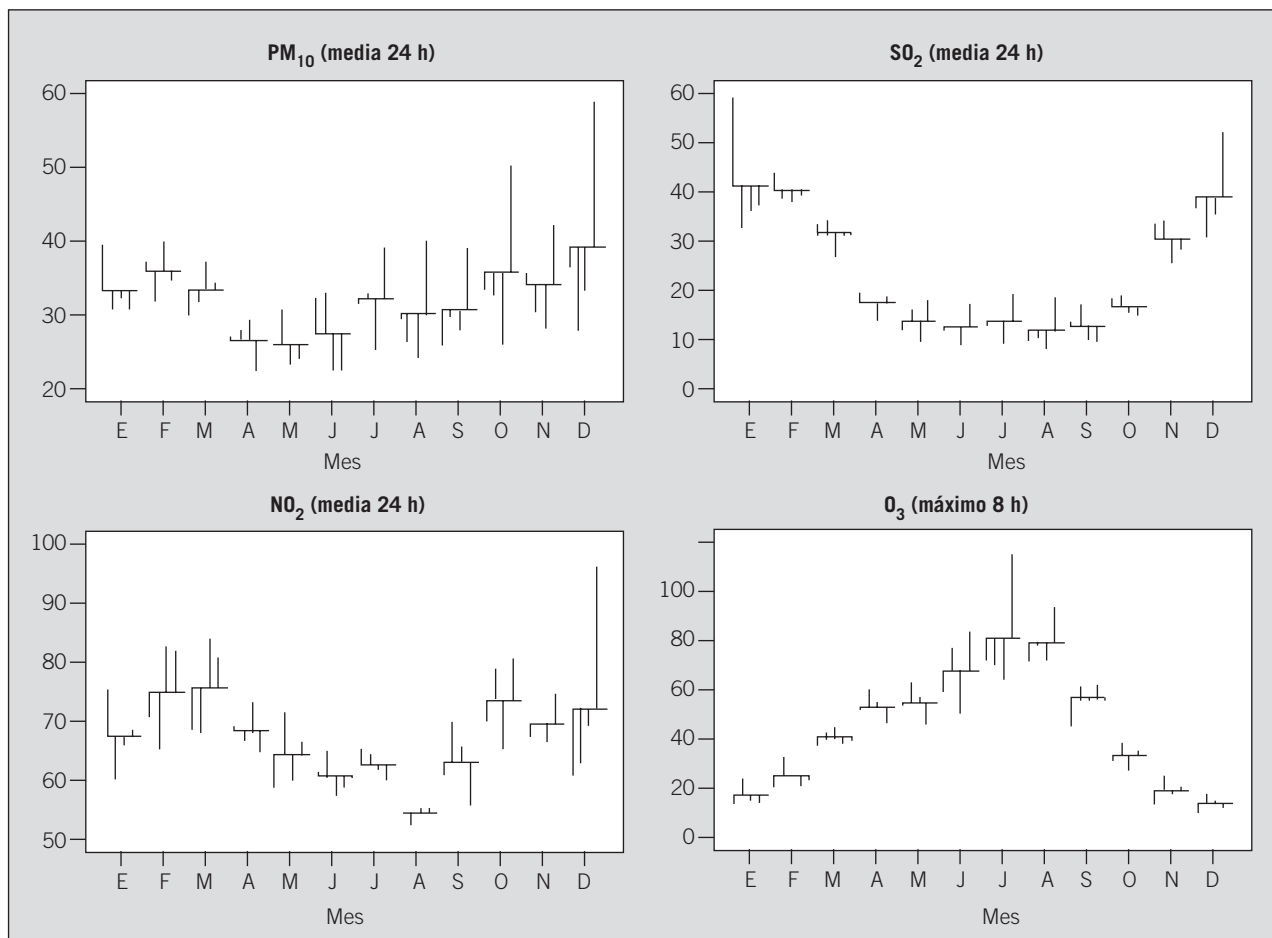


Fig. 4. Gráficos de subseries estacionales para contaminantes atmosféricos ( $PM_{10}$ ,  $SO_2$ ,  $NO_2$  y  $O_3$ ) en Madrid entre 1995 y 1998.

de cada subserie y también se puede evaluar la variación dentro de cada subserie mediante las líneas verticales.

#### Gráficos de descomposición

Los métodos de descomposición para series temporales proporcionan una valoración de la fuerza de la tendencia y de la componente estacional. Tras dicha identificación, se suprimen ambas componentes; queda entonces la serie desestacionalizada y proporcionan un cuadro más claro de otras características de la serie temporal.

Existen varios métodos de descomposición de series temporales. El método de descomposición clásico basado en el modelo paramétrico descrito previamente es relativamente sencillo de implementar utilizando modelos de regresión estándar, pero no permite mostrar una componente estacional que varíe suavemente<sup>16</sup>. Para superar esta limitación, se pueden adoptar procedimientos de descomposición de tendencia estacional basados en el método *lowess* (STL)<sup>25</sup>. Las series desestacionalizadas (o estacionales ajustadas)  $Y_t^*$  se calculan simplemente eliminando la tendencia y la componente estacional estimada de la serie original:

$$Y_t^* = Y_t - (\hat{T}_t + \hat{S}_t)$$

El método STL consiste en una sucesión de aplicaciones del suavizado del *lowess*, que proporciona estimaciones robustas de las componentes  $T_t$ ,  $S_t$  y  $a_t$ . Dicho método involucra un algoritmo iterativo, que progresivamente refina y mejora las estimaciones de la tendencia y la componente estacional. Después del ajuste estacional apropiado de la

serie temporal se puede construir los gráficos de descomposición. Así, los gráficos de descomposición se componen de 4 gráficos: a) el gráfico de la serie original; b) el gráfico de la componente de la tendencia; c) el gráfico de la componente estacional, y d) el gráfico de la componente residual.

#### Datos

Para ilustrar la aplicación de las herramientas gráficas descritas previamente, se ha recogido la demanda diaria por asma del Servicio de Urgencias del Hospital Gregorio Marañón durante el período 1995-1998. La temperatura media y la humedad relativa media se obtuvieron del observatorio meteorológico de Barajas, situado a 8 km al noroeste de la ciudad. Los datos de contaminación se recabaron a través de la red automática del Sistema Integral de Vigilancia, Predicción e Información de la Contaminación Atmosférica de la Ciudad de Madrid. Los contaminantes utilizados fueron: promedios diarios de  $PM_{10}$ ,  $SO_2$  y  $NO_2$ , y promedio de los valores máximos de cada 8 h de  $O_3$ . Los datos de polen provienen de la Red Polinológica de la Consejería de Sanidad de la Comunidad de Madrid, compuesta por 10 estaciones captadoras. Se recogieron los promedios diarios de los principales tipos de polen con capacidad alergénica en Madrid (*Olea*, *Plantago*, *Poaceae* y *Urticaceae*). Tanto los métodos de recogida como la descripción de los datos se han descrito con mayor detalle en trabajos previos<sup>14,15</sup>. Los datos se analizaron con el paquete estadístico S-Plus (Insightful Corporation, Seattle, WA, EE.UU.).

## Resultados

La figura 2 corresponde al gráfico de la serie temporal perteneciente al número de ingresos hospitalarios causados por asma. Aunque se observan las diversas fluctuaciones en la tendencia a lo largo de la serie, es difícil determinar visualmente el modelo de variación entre los intervalos de tiempo. En la misma figura se observa que tanto el suavizado paramétrico como el no paramétrico utilizando el alisador *lowess* proporcionan una interpretación visual más clara del modelo a largo plazo en los ingresos por asma. El suavizado paramétrico, dada las restricciones impuestas por los términos lineales y sinusoidales, mantiene una estructura fija de estacionalidad que se repite en los diferentes años que componen la serie temporal. Por el contrario, el suavizado no paramétrico no establece *a priori* ninguna restricción sobre la estructura de la serie temporal y se adapta mejor a la verdadera forma funcional de los datos.

Los gráficos superiores de la figura 3 representan los diagramas de cajas estacionales para la serie de temperatura. El de la izquierda es un gráfico de cajas anual y el de la derecha, mensual. Los gráficos inferiores representan la humedad relativa. Los gráficos de la izquierda nos indican que ambas series son estables tanto en mediana como en variancia. Esto se determina observando tanto la amplitud de las cajas como la ubicación de las medianas. En los gráficos de la derecha se observa el claro comportamiento estacional de ambas series. La temperatura presenta un aumento paulatino hasta el mes de julio, cuando alcanza su valor máximo, que se mantiene en agosto, y un decrecimiento a partir de septiembre que culmina en enero, que encontramos el mínimo. Por tanto, la estacionalidad presenta una forma de V invertida. Para la humedad relativa se observa una forma de W: decrecimiento de

enero a marzo, a continuación un ligero aumento hasta mayo y posteriormente un decrecimiento hasta julio, mes en el cual se alcanza el mínimo; a partir de agosto se observa un nuevo aumento de la humedad que culmina en enero.

En la figura 4 se describen las concentraciones de diversos contaminantes atmosféricos. Para las de  $PM_{10}$  se observa cierta estabilidad que se ve influida por un aumento de la variabilidad conforme pasa el tiempo, mientras que las de  $SO_2$  presentan una clara forma de U, con valores estables entre mayo y septiembre, acompañados de un crecimiento de noviembre a enero y un decrecimiento de febrero a mayo. En cuanto a los contaminantes fotoquímicos, el  $NO_2$  presenta un comportamiento irregular, con concentraciones mínimas durante el mes de agosto. En cambio, las cifras de  $O_3$  tienen un comportamiento opuesto a las de  $SO_2$ , con forma de V invertida, con un crecimiento de enero a julio y un decrecimiento de agosto a noviembre.

Por último, en la figura 5 se muestran las concentraciones de distintos tipos de polen donde se ha aplicado el método STL para descomponer tendencia y estacionalidad. La descomposición de la tendencia nos indica, para todas las series, que en el año 1995 los valores de polen eran muy bajos, para experimentar una brusca explosión en 1996. En los años sucesivos las concentraciones de polen parecen estabilizarse, a excepción de las de *Poaceae*, que experimentan una nueva explosión en el año 1998. Al observar la componente estacional se puede determinar que en todas las series las concentraciones de polen alcanzan sus máximos durante los meses de primavera. Las de *Urticaceae* presentan una distribución bimodal; aumentan durante los primeros meses de primavera, a continuación disminuyen y vuelven a aumentar durante los meses de verano. Finalmente, el gráfico de las componen-

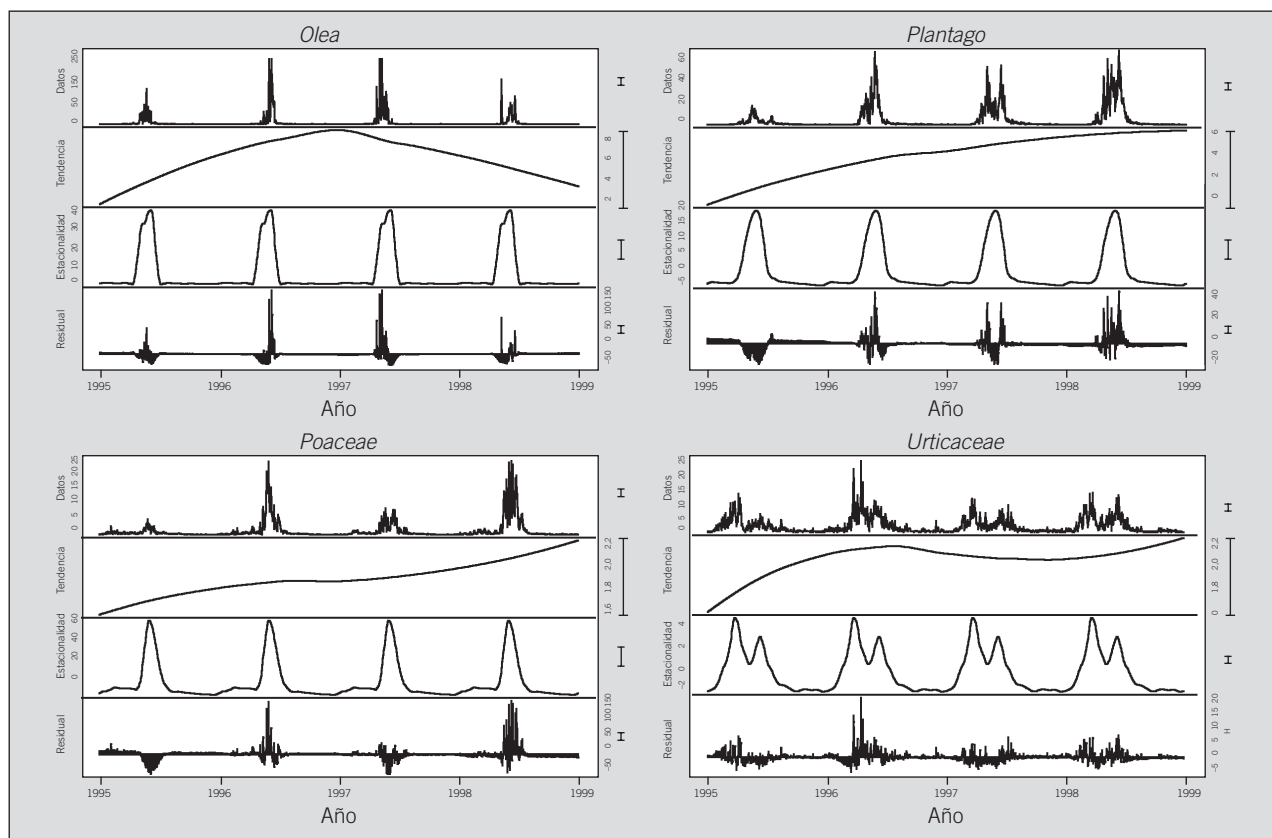


Fig. 5. Gráficos de descomposición (método lowess; STL) para concentraciones de polen con capacidad alérgica (*Olea*, *Plantago*, *Poaceae* y *Urticaceae*) en Madrid entre 1995 y 1998.

tes residuales muestra cómo quedaría cada una de las series temporales tras extraer las componentes de tendencia y estacionalidad. Las barras situadas en el margen derecho de cada uno de los gráficos muestran la variación de los datos y de las componentes; de esta manera, el gráfico de descomposición puede usarse para evaluar la fuerza y la magnitud de la tendencia y de la componente estacional dentro de una serie temporal. Todas las barras deben tener la misma longitud, pero están dibujadas a diferente escala. Así pues, observando, por ejemplo, la barra situada a la derecha del tercer gráfico de las concentraciones de *Poaceae*, su longitud indica que una parte de la variación en la serie original está causada por la estacionalidad. El mismo comportamiento se observa para los otros tipos de polen.

## Discusión

En este artículo hemos comentado algunos métodos simples de visualización de datos que permiten un examen gráfico eficaz de la información cuantitativa de una serie temporal: gráfico de la serie temporal y suavizados, gráficos de subseries y gráficos de descomposición.

Hemos visto que los diversos tipos de suavizado, paramétrico y no paramétrico, del gráfico temporal de los valores de la serie original proporcionan una representación visual mejorada del modelo a largo plazo. Además, muestran de una manera más clara la variación a corto plazo que se da en intervalos de tiempo dentro de la serie temporal. Aunque el suavizado paramétrico resulta más restrictivo en cuanto a las asunciones del modelo, a partir de la estimación de los términos lineales y sinusoidales, su estructura fija de estacionalidad, que se repite a través del tiempo, permite que se emplee también para la predicción más allá del rango de observación<sup>19</sup>. Por el contrario, el suavizado no paramétrico, dado que no establece *a priori* ninguna restricción sobre la estructura de la serie temporal, permite una mejor adaptación a la verdadera forma funcional de los datos, pero por el mismo motivo no resulta útil en términos de predicción.

Los gráficos de subseries se utilizan para evaluar la variación dentro de cada año o de cada mes, y entre los meses sucesivos en una serie temporal, utilizando diagramas de cajas y subseries estacionales. Permiten el estudio visual de la tendencia y especialmente la estructura de estacionalidad de la serie temporal<sup>24</sup>. Finalmente, los gráficos de descomposición, utilizando el método STL, proporcionan una mejor comprensión del comportamiento de la serie temporal, a partir de la descomposición de sus componentes de tendencia subyacente, estacionalidad, y residuales, además de mostrar la fuerza relativa de dichas componentes<sup>25</sup>. Es necesario remarcar que el método STL descrito se basa en la descomposición aditiva de la serie estacional. Así, si la serie temporal presenta una estructura más compleja que, por ejemplo, necesitase de una descomposición multiplicativa tal que  $Y_t = T_t \times S_t \times a_t$ , el *software* estándar que se ha utilizado en nuestro análisis debería modificarse de acuerdo con el modelo supuesto.

La visualización de los datos es una parte integrada para los modelos estadísticos<sup>12,13</sup>. Algunas de las herramientas de visualización que hemos presentado no son específicas del análisis de series temporales, ya que pueden usarse en la mayoría de las aplicaciones de métodos estadísticos. En este sentido, aunque en los últimos años se han producido grandes avances metodológicos relacionados con los análisis de regresión de series temporales en la investigación médica<sup>26,27</sup>, la inclusión de los gráficos descriptivos como herramienta fundamental para explorar el comportamiento de diversos factores de riesgo no se ha tenido debidamente en cuenta hasta el momento.

## Agradecimiento

A Francisco Javier García Agudo, por su inestimable ayuda en la edición del manuscrito y en la realización de los gráficos con S-Plus, y a José Ramón Banegas, por sus comentarios a versiones previas de este trabajo.

Expresamos nuestra gratitud por los datos suministrados a: Red Polinológica de la Consejería de Sanidad, Departamento de Control de Contaminación Atmosférica del Ayuntamiento de Madrid, Subdirección General de Calidad Ambiental del Ministerio de Medio Ambiente, Programa Regional de Prevención y Control del Asma, y al Servicio de Urgencias del Hospital Gregorio Marañón.

## REFERENCIAS BIBLIOGRÁFICAS

1. Antó JM, Murillo C. Aplicaciones sanitarias del análisis de series temporales. *Gac Sanit* 1991;5:3-4.
2. López del Val JA, Calvete Fernández HI, Carreter Oróñez CA, Abaurrea León J, Muniesa Cuenca MP, García Mata JR, et al. Análisis de series temporales aplicado a infección nosocomial. *Med Clin (Barc)* 1992;99:52-6.
3. Murillo C, Sáez M. Seguimiento y control de las predicciones: el caso de las epidemias de asma en Barcelona. *Gac Sanit* 1993;7:116-22.
4. Pérez D, Ríos M, García JM, Cubedo M. Análisis de series temporales en la epidemiología de la fiebre tifoidea en España. *Med Clin (Barc)* 1996;106:686-9.
5. Cano L, Ríos M, Sánchez JA. La estacionalidad y los recientes cambios de la enfermedad meningocócica en España. *Gac Sanit* 2001;15:336-40.
6. Díaz J, Alberdi JC, Montero JC, Mirón IJ. Asociación entre la contaminación atmosférica por dióxido de azufre y partículas totales en suspensión y la mortalidad diaria en la ciudad de Madrid (1986-1992). *Gac Sanit* 1998;12:207-15.
7. Tobias A, Sunyer J, Castellsagué J, Sáez M, Antó JM. Impacto de la contaminación atmosférica sobre la mortalidad y las urgencias por enfermedad pulmonar obstructiva crónica y asma en Barcelona. *Gac Sanit* 1998;12:223-30.
8. Ballester F, Iñiguez C, Pérez-Hoyos S, Tenías JM. Contaminación atmosférica por partículas y salud en Valencia, 1994-1996. *Gac Sanit* 2002;16:464-79.
9. Ballester F, Iñiguez C, Sáez M, Pérez-Hoyos S, Daponte A, Ordóñez JM, et al, en nombre del grupo EMECAM-EMECAS. Relación a corto plazo de la contaminación atmosférica y la mortalidad en 13 ciudades españolas. *Med Clin (Barc)* 2003;121:684-9.
10. Sáez M, Pérez-Hoyos S, Tobias A, Saurina C, Barceló MA, Ballester F. Métodos de series temporales en los estudios epidemiológicos sobre contaminación atmosférica. *Rev Esp Salud Publica* 1999;73:119-30.
11. Tobias A. Métodos para el análisis del impacto de la contaminación atmosférica sobre la salud. *Gac Sanit* 1999;13:167-8.
12. Tuckey JW. *Exploratory data analysis*. Massachusetts: Addison-Wesley, 1997.
13. Cleveland WS. *The elements of graphing data*. New Jersey: Hobart Press, 1994.
14. Galán I, Tobias A, Banegas JR, Aránguez E. Short-term effects of air pollution on daily asthma emergency room admissions in Madrid, Spain. *Eur Respir J* 2003;22:802-8.
15. Tobias A, Galán I, Banegas JR, Aránguez E. Short-term effects of airborne pollen concentrations on asthma epidemic. *Thorax* 2003;58:708-10.
16. Greene WH. *Econometric analysis*. 5th ed. Prentice Hall, 2003.
17. Box GEP, Hunter WG, Hunter JS. *Statistics for experimenters: an introduction to design, data analysis, and model building*. New York: Wiley, 1989.
18. Johnson B, Christensen L. *Educational research: quantitative, qualitative, and mixed approaches*. 2nd ed. Allyn & Bacon, 2003.
19. Serfling ER. Methods for current statistical analysis of excess of pneumonia and influenza deaths. *Public Health Rep* 1963;78:494-506.
20. Kuhn L, Davidson LL, Durkin MS. Use of Poisson regression and time series analysis for detecting changes over time in rates of child injury following a prevention program. *Am J Epidemiol* 1994;140:943-55.
21. Stolwijk AM, Straatman H, Zielhuis GA. Studying seasonality by using sine and cosine functions in regression analysis. *J Epidemiol Community Health* 1999;53:235-8.
22. Hastie T, Tibshirani RJ. *Generalised additive models*. London: Chapman and Hall, 1990.
23. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* 1979;74:829-36.
24. Cleveland WS, Terpenning IJ. Graphical methods for seasonal adjustment. *J Am Stat Assoc* 1982;77:52-62.
25. Cleveland RB, Cleveland WS, McRae JE, Terpenning I. STL: a seasonal-trend decomposition procedure based on lowess. *J Official Stat* 1990;6:3-73.
26. Schwartz J, Spix C, Touloumi G, Bacharova L, Barumamdzadeh T, Le Tertre A, et al. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *J Epidemiol Community Health* 1996;50(Suppl):3-11.
27. Dominici F, McDermott A, Zeger SL, Samet JM. On the use of generalised additive models in time-series of air pollution and health. *Am J Epidemiol* 2002;156:193-203.