ESTADÍSTICA ESPAÑOLA
Vol. 52, núm. 173, 2010, págs. 127 a 153

Descomposición de series de tiempo cortas: un enfoque bayesiano(1)

por
JOSÉ LUIS ROJO GARCÍA
y
JOSÉ ANTONIO SANZ GÓMEZ
Departamento de Economía Aplicada. Universidad de Valladolid

RESUMEN

La descomposición de series de tiempo es una actividad básica en análisis de la coyuntura económica. Los métodos más populares y recomendados por Eurostat son Tramo-Seats y X12-RegArima. La literatura constata el deterioro de la calidad de las descomposiciones para series cortas en su dimensión temporal.

Los autores desarrollan un modelo bayesiano jerárquico para la extracción de ciclo-tendencia y estacionalidad en una serie de tiempo. Dicho modelo resulta especialmente útil para series cortas, no presentando el deterioro de los métodos tradicionales.

El artículo se completa con una ilustración y con una validación del procedimiento mediante series simuladas que confirman su interés

(1) Trabajo OTP/08/05 realizado bajo subvención de la Consejería de Hacienda de la Junta de Castilla y León (Dirección General de Estadística), concedida por resolución de 31-7-08 (convocatoria de 7 de abril) de concurso público para la realización de trabajos y proyectos de investigación en materia de estadística aplicada, BOCYL nº 71 de 14-4-08.

Palabras clave: Modelos bayesianos, Series de tiempo cortas, Análisis de series de tiempo, Ajuste estacional

Clasificación AMS: 62C10 (Bayesian problems), 62F15 (Bayesian inference), 37M10 (Times Series analysis)

1. INTRODUCCIÓN

La descomposición de series temporales es una actividad fundamental además de frecuente en los análisis de coyuntura económica y, en general, en el estudio de la evolución económica y social de las sociedades. A veces, el objetivo es la obtención de la componente estacional, mientras que, en otras, se desea extraer la tendencia o la denominada ciclo-tendencia, componente que agrupa la tendencia secular y la componente cíclica. Como es conocido, ambas son componentes inidentificables de la serie temporal, en el sentido de que sin restricciones adicionales la descomposición es arbitraria.

Procedimientos clásicos de descomposición son, por ejemplo, el a veces denominado *Método clásico* (Martín-Guzmán, 1989, pp. 250 a 256) o los métodos de *alisado exponencial*, en especial el denominado de Holt-Winters (Winters, R., 1960)

Hoy día, no obstante, los dos programas más populares para realizar esta descomposición son Tramo-Seats (TS, Gómez y Maravall, 1997) y X12-RegArima (U.S. Census Bureau, 2002). Mientras este último es un procedimiento empiricista por la naturaleza de los filtros empleados, el primero es de los que se denominan basados en modelos. El calificativo *populares* hace referencia, entre otras cosas, al hecho de que son los procedimientos recomendados por Eurostat para los países de la Unión Europea. En la práctica ambos son las elecciones más frecuentes por parte de las oficinas estadísticas nacionales. Pese a la calidad contrastada de estos métodos para la descomposición de series de tiempo generales, se plantea un problema cuando la serie a estudiar alcanza únicamente unos pocos años (digamos, cuatro o cinco años).

La literatura estadística no es muy prolija en este tema, aunque algunas referencias pueden resultar de utilidad. Por ejemplo, Cholette (1979) apunta que, en general, debe evitarse la extracción de señal para series de cinco años o menos. Para ello, estudia las funciones de fase y de ganancia de los filtros X-11 para series de 3, 4, 5 y 7 años. Hood, Ashley y Findley (2000) utilizan 54 series simuladas de 12 años de longitud, y concluyen que, tanto X12 como TS deterioran sus estimaciones de la componente estacional cuando la muestra se reduce a 4 años (el deterioro se constata visualizando las desviaciones absolutas y la raíz cuadrática media relativa con respecto a los valores originales). El deterioro es superior para

TS. Matas Mir y Rondonetti (2003) hacen un estudio similar (con series simuladas) evaluando la calidad de X12. Concluyen que el deterioro de la calidad es superior en los dos primeros años de la serie que en los más recientes (y, sin duda, los más importantes en el análisis de la coyuntura). Findley y Martin (2003) realizan un estudio teórico en el dominio de las frecuencias, concluyendo que la calidad de los filtros TS y X12 viene determinada por los coeficientes del filtro de líneas aéreas estimados.

Mazzi y Savio (2005) analizan el deterioro resultante de acortar series provenientes de la base NewChronos de Eurostat, utilizando un amplio panel de indicadores de calidad para el ajuste estacional tomados de Ladiray y Mazzi (2003). Su conclusión es que en media, X12 funciona ligeramente mejor que TS cuando se reduce la muestra, un resultado que los autores achacan a la mayor inestabilidad de los enfoques basados en modelos cuando se pierde una parte importante de la información. Concluyen asimismo que el deterioro es superior cuando se pasa de series largas a series de longitud media (10 años) que al pasar de estas últimas a series cortas (5 años en su aplicación).

No son abundantes los procedimientos bayesianos orientados hacia la estimación de las componentes de una serie de tiempo. El análisis desde esta perspectiva se ha centrado más bien en cuestiones como la predicción y la regresión (véase, p. ej. Broemeling (1985)) y pocos trabajos, si bien notables, se han dedicado a la extracción de señales y/o componentes estacionales. Citemos en este sentido el clásico artículo de Akaike (1980) o el de Young (1996). Ninguno de estos autores se orienta a la descomposición de series de tiempo cortas. En relación con esta cuestión, sólo conocemos el trabajo de Mendoza y de Alba (2006), aunque no está orientado a la descomposición sino a la predicción.

Nuestro proyecto se plantea establecer un modelo bayesiano jerárquico normal-gamma para la descomposición de series de tiempo. Obviamente, el modelo puede utilizarse para la extracción de ciclo-tendencia y de la componente estacional de cualquier serie de tiempo, si bien su mayor habilidad se plasma en las series de tiempo cortas. El objetivo será la obtención explícita de las estimaciones de Bayes de la ciclo-tendencia y de las estacionalidades, así como las medidas de precisión asociadas a cada una de ellas.

Los modelos bayesianos normal-gamma han sido utilizados con éxito por los autores en varios contextos: para desagregar Contabilidades regionales, para pequeñas áreas (Rojo y Sanz, 2008), en procedimientos de trimestralización (Rojo y Sanz, 2005) o para armonizar desagregaciones espacio-temporales de series, con aplicación a la trimestralización de Contabilidades provinciales(2). Estos modelos permiten, generalmente, encontrar soluciones explícitas para los estimadores de Bayes, y manejar sin dificultades problemas con un número de parámetros especialmente altos.

La utilización de procedimientos bayesianos plantea el problema de la asignación de valores a los parámetros que intervienen en las distribuciones a priori. Una de las soluciones factibles pasa por considerar distribuciones a priori neutrales en el sentido de Jeffreys (véase Zellner (1971), pp. 41). Sin embargo, es posible introducir información a priori cuando se disponga de ella, para lo que es preciso obtener las distribuciones marginales a priori. Nuestra propuesta de trabajo consistiría en asignar valores razonables a las constantes que intervienen en las distribuciones a priori con lo que se obtiene, en general, una menor dispersión de las estimaciones y, por tanto, estimaciones más ajustadas.

Por lo que se refiere al interés del procedimiento, citemos un texto de Mazzi y Savio (Mazzi y Savio, 2005) refiriéndose a las estadísticas europeas: "*hoy día, una causa importante de preocupación en Eurostat reside en el ajuste estacional de series de tiempo cortas. Varias razones explican por qué; al nivel de los Estados Miembros algunas series están definidas en un periodo corto de tiempo; en concreto, cambios en la metodología y en las definiciones, actualización de los sistemas de clasificación estadística, uso de nuevas fuentes de información, etc. En el ámbito de Eurostat, las recientes y futuras incorporaciones a la Unión Europea incrementarán el número de estadísticas infra-anuales definidas para cortos periodos, especialmente las que resulten de agregar los datos nacionales disponibles*".

El objetivo de nuestro trabajo, en definitiva, consiste en la estimación de la tendencia y estacionalidad de una serie cronológica con un procedimiento bayesiano que sea de aplicación a series cortas. Para ello, en el siguiente apartado realizamos un planteamiento formal del problema, y en el tercero se establecen las distribuciones a priori así como la función de verosimilitud, concluyendo con las distribuciones a posteriori y, en consecuencia, obteniendo el estimador de Bayes óptimo. En el cuarto apartado, se obtienen las distribuciones marginales a priori, a partir de las cuales es posible asignar valores razonables a varios de los parámetros implicados en las mismas. El quinto apartado ejemplifica el procedimiento obtenido mediante su aplicación a la descomposición de la serie de matriculaciones de turismos en España, ilustrando las ventajas relativas del mismo especialmente cuando la longi-

(2) Dicho documento corresponde a un trabajo de los autores, no publicado, elaborado dentro de un proyecto competitivo de la Dirección General de Estadística de la Junta de Castilla y León (OTP/06/06) con título "Conciliación espacial-sectorial de trimestralización de series. Una aplicación a las Contabilidades provinciales".

tud de la serie es pequeña. En el mismo apartado se realiza un ejercicio de simulación, comprobando la calidad del método para series cortas simuladas (trimestrales durante tres años) en distintas condiciones de amplitud de la estacionalidad y de tamaño del ruido. En el sexto y último apartado se resumen las conclusiones más notables.

2. PLANTEAMIENTO DEL PROBLEMA

Se suponen conocidos los valores de una serie cronológica, $y_1,\ldots,y_N,y_{N+1},\ldots,y_{N+r}$ donde

• N es el número de periodos que corresponden con años completos conocidos, esto es $N = M \cdot p$ siendo p la estacionalidad y siendo M el número de años completos que disponemos.

• $r < p$, esto es, r es el número de períodos (meses, trimestres, etc.) de los que tenemos información dentro del año corriente.

Dicho de otra forma, disponemos de datos para los $p$meses(3) de cada uno de M años completos y, tal vez, para r meses del año corriente (si la información cubre un número entero de años, tomaremos $r = 0$).

El objeto de este artículo es la descomposición de la serie en una componente ciclo-tendencial, $T = (T_1,\ldots,T_{N+r})'$ y una estacionalidad fija, $S=(S_1,\ldots,S_p)'$, a la que se añade una componente irregular, $\varepsilon = (\varepsilon_1,\ldots,\varepsilon_{N+r})$ mediante un esquema aditivo(4).

$$y_t = T_t + S_j + \varepsilon_t, \quad t = 1,\ldots,N+r \qquad [1]$$

siendo $j = t - p \cdot [t/p]$ donde [.] es la función "parte entera", esto es, j es el ordinal del mes correspondiente al periodo t.

Como es habitual en los esquemas aditivos, la estacionalidad está sometida a la restricción

---

(3) En el trabajo hablaremos de meses, tanto si se trata de meses naturales como de cualquier periodo de frecuencia inferior a la anual.

(4) Como es conocido, un modelo aditivo puede utilizarse para esquemas multiplicativos a través de la transformación logarítmica. No obstante, en series cortas, muchas veces de tres o cuatro años, la distinción entre modelos aditivos y multiplicativos no viene avalada por los datos. Por otra parte, la tendencia puede considerarse ciclo-tendencia, y la distinción para series cortas no es obvia, ni empírica ni formalmente. Señalemos también que en esta versión de nuestro trabajo no introducimos efecto Semana Santa ni de número de días laborables. Es esta misma cortedad de las series la que hace poco verosímiles estacionalidades variables en el tiempo.

$$S_1+\cdots+S_p=0 \qquad [2]$$

La conveniencia de manejar modelos estocásticos que no incluyan restricciones en las variables nos plantea introducir la restricción en el modelo, eliminando para ello una componente estacional, que arbitrariamente será la última; así, llamando $S_{(p)}=(S_1,\ldots,S_{p-1})'$ podemos escribir

$$S=W\cdot S_{(p)} \qquad [3]$$

siendo $W=\begin{pmatrix} I_{p-1} \\ -\iota'_{p-1}\end{pmatrix}$ la matriz $p\times(p-1)$ anterior y denotando por ι al vector columna de la dimensión correspondiente con valores iguales a la unidad.

3. EL MODELO BAYESIANO JERÁRQUICO. ESTIMADORES ÓPTIMOS DE BAYES

Para la tendencia $T=\left(T_1,...,T_{N+r}\right)'$, consideraremos una distribución a priori, condicionada por la precisión, τ,

$$\pi(T\,|\,\tau,D)\propto\tau^{(N+r)/2}\exp\left\{-\frac{\tau}{2}\left[T'MT+\left(T-\bar{T}\right)'P_T\left(T-\bar{T}\right)\right]\right\} \qquad [4]$$

donde M es la matriz de la forma cuadrática cuyo objetivo será favorecer tendencias regulares. Para ello, la matriz se construye a partir de las matrices de primeras o segundas diferencias. Así, puede elegirse $M=D'_1D_1$, siendo $D_1=(d_{ij})$ la matriz $(N+r-1)\times(N+r)$ de rango $N+r-1$ que produce primeras diferencias, esto es, $d_{ij}=0$ excepto $d_{ii}=-1$, $d_{i,i+1}=1$ para $i=1,\ldots,N+r-1$, o bien, si se desean favorecer las tendencias lineales, $M=D'_2D_2$, siendo $D_2=(d_{ij})$ la matriz $(N+r-2)\times(N+r)$ de rango $N+r-2$ que produce las segundas diferencias, esto es, $d_{ij}=0$ excepto $d_{ii}=1$, $d_{i,i+1}=-2$, $d_{i,i+2}=1$ para $i=1,\ldots,N+r-2$,. La matriz P_T es una matriz cuadrada, no singular, de orden $N+r$. En definitiva, esta distribución a priori sigue un esquema gaussiano que hace más probables tendencias regulares, esto es, con primeras o segundas diferencias pequeñas.

Para τ suponemos una distribución marginal a priori gamma(5)

$$\pi(\tau | D) \propto \tau^{b-1} e^{-a\tau}, \quad \tau > 0 \qquad [5]$$

lo que garantiza que trabajaremos en un marco determinado por familias conjugadas de distribuciones.

A la componente estacional $S=(S_1,\ldots,S_p)'$ la asignamos una distribución a priori

$$\pi(S | \tau, D) \propto \tau^{p/2} \exp\left\{-\frac{\tau}{2}\left[\left(S-\bar{S}\right)' P_S \left(S-\bar{S}\right)\right]\right\} \qquad [6]$$

restringida por [2], y siendo P_S una matriz $p \times p$ no singular.

Introduciendo la restricción [3] en el modelo dispondremos de una priori sin restricciones; así, denotando por $W_S = W' P_S W$ y por $\mu_{S_{(p)}} = W_S^{-1} W' P_S \bar{S}$, se tiene que el exponente de [6] es,

$$\left(S-\bar{S}\right)' P_S \left(S-\bar{S}\right) = \left(S_{(p)} - \mu_{S_{(p)}}\right)' W_S \left(S_{(p)} - \mu_{S_{(p)}}\right) + \bar{S}' P_S \left[P_S^{-1} - W W_S^{-1} W'\right] P_S \bar{S}$$

con lo que, incorporando la restricción al modelo, y admitiendo independencia a priori entre T y $S_{(p)}$ condicionados por τ resulta la priori conjunta en el modelo restringido,

$$\pi(S_{(p)}, T, \tau | D) \propto \tau^{b-1+\frac{N+r}{2}+\frac{p}{2}} e^{-a\tau} \exp\left\{-\frac{\tau}{2}\left[T' D'_2 D_2 T + \left(T-\bar{T}\right)' P_T \left(T-\bar{T}\right) + \right.\right.$$

$$\left.\left. + \left(S_{(p)} - \mu_{S_{(p)}}\right)' W_S \left(S_{(p)} - \mu_{S_{(p)}}\right) + \bar{S}' P_S \left[P_S^{-1} - W W_S^{-1} W'\right] P_S \bar{S}\right]\right\} \qquad [7]$$

El modelo de verosimilitud es el modelo [1] de descomposición de la serie, con ciertas especificaciones aleatorias.

---

(5) Existen varias versiones de la distribución gamma de acuerdo con los parámetros que se utilicen. Nosotros utilizaremos la versión según la cual si $X \to \gamma(a,b)$, su función de densidad viene dada por $f(x) = \frac{a^b}{\Gamma(b)} x^{b-1} e^{-ax}$, para $x > 0$. Con esta parametrización, su esperanza es b/a y su varianza b/a^2.

Denotemos por E_S la matriz que expande la estacionalidad al conjunto de los $N+r$ meses, es decir, $E_S = \begin{pmatrix} \iota_M \otimes I_p \\ I_r \,|\, O_{r,p-r} \end{pmatrix}$, donde el bloque inferior es la matriz $r \times p$ que selecciona las estacionalidades de los r primeros meses del año, esto es, la matriz $O_{r,p-r}$ es una matriz nula de dimensión $r \times (p-r)$.

Recordando la expresión [3] podemos escribir el modelo de verosimilitud como $y = T + E_S S + \varepsilon = T + E_S W S_{(p)} + \varepsilon$, esto es

$$y = T + E_W S_{(p)} + \varepsilon \qquad [8]$$

siendo $y = (y_1, \ldots, y_{N+r})'$, y $E_W = E_S W = \begin{pmatrix} \iota_M \otimes W \\ I_r \,|\, O_{r,p-r-1} \end{pmatrix}$,donde el bloque inferior es la matriz $r \times p$ que selecciona las estacionalidades de los r primeros meses del año entre los $p-1$ primeros, esto es, la matriz $O_{r,p-r-1}$ es una matriz nula de dimensión $r \times (p-r-1)$.

Admitimos una distribución normal para el ruido, de manera que $(\varepsilon | \tau, D) \to N_{N+r}\left(O, \tau^{-1} P_\varepsilon^{-1}\right)$, es decir $\pi(\varepsilon | \tau, D) \propto \tau^{\frac{N+r}{2}} e^{-\frac{\tau}{2}\varepsilon' P_\varepsilon \varepsilon}$, donde P_ε es una matriz cuadrada de orden $N+r$ no singular. En definitiva, la verosimilitud del modelo resulta

$$\pi(\varepsilon | \tau, D) \propto \tau^{\frac{N+r}{2}} \exp\left\{-\frac{\tau}{2}\left(y - T - E_W S_{(p)}\right)' P_\varepsilon \left(y - T - E_W S_{(p)}\right)\right\} \qquad [9]$$

Multiplicando la priori conjunta [7] por la función de verosimilitud [9] obtendremos la conjunta a posteriori, detallada en el Anexo. Tras sencillas, aunque tediosas operaciones que en el mismo se detallan, resulta que la marginal a posteriori de T es una Student multivariante(6) de dimensión $N+r$ con $2b+N+r+1$

(6) Diremos que una variable z de dimensión d tiene una distribución t de Student multivariante si su función de densidad conjunta es de la forma $f(z) \propto \left[1 + \frac{1}{\nu}(z - \mu_z)' A (z - \mu_z)\right]^{-\left(\frac{d+\nu}{2}\right)}$. Sus grados de libertad son ν, su matriz de posición μ_Z y su matriz de escala A. Se demuestra que su matriz de varianzas y covarianzas resulta $(\nu/(\nu-2))A^{-1}$ para $\nu > 2$. Véase, por ejemplo, Zellner, (1971), pág. 383, donde se justifica la elección de esta versión de la distribución de Student multivariante en lugar de otras clásicas.

grados de libertad, matriz de posición(7) μ_T y matriz de varianzas-covarianzas

$$\frac{2\tilde{a}}{2b+N+r-1}P_{T\varepsilon}^{-1}.$$

Se obtiene también en dicho Anexo que la marginal a posteriori de $S_{(p)}$, es una multivariante de Student con $2b+N+r-1$ grados de libertad, matriz de posición $\tilde{\mu}_{S_{(p)}}$ y matriz de varianzas y covarianzas $\frac{2a^*}{2b+N+r-1}P_{S_{(p)}}^{-1}$.

En definitiva, la obtención de las distribuciones marginales a posteriori de T y $S_{(p)}$ proporciona el estimador de Bayes óptimo que, como es conocido, es la esperanza de la distribución marginal a posteriori si suponemos una función de pérdida cuadrática. Pueden observarse en el Anexo dichas esperanzas así como sus matrices de varianzas y covarianzas a posteriori.

4. DISTRIBUCIONES MARGINALES A PRIORI

El conocimiento de las distribuciones marginales a priori y, en especial, la relación de ciertas características de dichas marginales con los parámetros del modelo, permite asignar valores a dichos parámetros cuando se disponga de información a priori sobre la serie. Obtengamos, en consecuencia, dichas marginales. La marginal a priori de τ, [5], ya fue descrita anteriormente. Se trata de una gamma $\gamma(a,b)$ con lo que se conocen sus momentos, $E_\pi(\tau)=b/a$ y $Var_\pi(\tau)=b/a^2$. Igualmente, cálculos directos (véase Broemeling (1985) o Zellner (1971)) proporcionan los momentos de τ^{-1} que serán utilizados más adelante $E_\pi(\tau^{-1})=\frac{a}{b-1}$, $b>1$ y $Var_\pi(\tau^{-1})=\frac{a^2}{(b-1)^2(b-2)}$ si $b>2$. Señalemos que desde el punto de vista de la asignación de valores a los parámetros, es preferible el enfoque que utiliza los momentos de τ^{-1} a la utilización de los momentos de τ, ya que es más intuitivo trabajar en términos de varianza que de la precisión.

Teniendo en cuenta la distribución a priori de τ de [5], la distribución conjunta a priori de (ε,τ) es $\pi(\varepsilon,\tau|D)\propto\tau^{\frac{N+r}{2}+b-1}e^{-\frac{\tau}{2}[2a+\varepsilon'P_\varepsilon\varepsilon]}$. Integrando en $\tau\in(0,\infty)$ obtenemos la marginal a priori de ε

(7) En el Anexo se detallan las expresiones no definidas hasta el momento, tanto para esta distribución como para las restantes.

$$\pi\left(\varepsilon \middle| D\right) \propto \left[2a + \varepsilon' P_{\varepsilon}\varepsilon\right]^{-\left(\frac{N+r}{2}+b\right)} \qquad [10]$$

Se trata de una Student multivariante con $2b$ grados de libertad (obsérvese que $N+r$ es la dimensión de ε), matriz de posición igual a la matriz nula de dimensión $(N+r)\times 1$ y matriz de varianzas-covarianzas

$$V(\varepsilon) = \frac{a}{b-1} P_{\varepsilon}^{-1} \qquad [11]$$

Para obtener la distribución a priori de T, partimos de la priori conjunta de (T,τ),

$$\pi(T,\tau \mid D) \propto \tau^{b-1+(N+r)/2} \exp\left\{-\frac{\tau}{2}\left[2a + T'D'_2D_2T + \left(T-\bar{T}\right)'P_T\left(T-\bar{T}\right)\right]\right\} \qquad [12]$$

Integrando en τ dicha distribución resulta entonces

$$\pi(T \mid D) \propto \left[2\underline{a}_T + \left(T-\mu_{T\pi}\right)'P_{T\pi}\left(T-\mu_{T\pi}\right)\right]^{-\left(b+\frac{N+r}{2}\right)} \qquad [13]$$

donde denotamos por $P_{T\pi} = D'_2D_2 + P_T$, por $\mu_{T\pi} = P_{T\pi}^{-1}P_T\bar{T}$ y por $2\underline{a}_T = 2a + T'P_T\left(P_T^{-1} - P_{T\pi}^{-1}\right)P_T\bar{T}$.

La marginal a priori de T [13] es, por tanto, una distribución de Student multivariante con $2b$ grados de libertad, matriz de posición $\mu_{T\pi}$ y matriz de varianzas-covarianzas $\frac{\underline{a}_T}{b-1}P_{T\pi}^{-1}$.

Obtengamos ahora la marginal a priori de $S_{(p)}$, de forma análoga al trabajo realizado para obtener la marginal a priori de T. La distribución a priori conjunta $S_{(p)}$ y de τ es

$$\pi(S_{(p)},\tau \mid D) \propto \tau^{b-1+p/2}\exp\left\{-\frac{\tau}{2}\left[2a+\bar{S}'P_S\left(P_S^{-1}-WW_S^{-1}W'\right)P_S\bar{S}+\left(S_{(p)}-\mu_{S_{(p)}}\right)'W_S\left(S_{(p)}-\mu_{S_{(p)}}\right)\right]\right\} \qquad [14]$$

e integrando en τ resulta

$$\pi(S_{(p)} \mid D) \propto \left[2a_S + \left(S_{(p)}-\mu_{S_{(p)}}\right)'W_S\left(S_{(p)}-\mu_{S_{(p)}}\right)\right]^{-b-\frac{p}{2}} \qquad [15]$$

donde hemos llamado $2a_S = 2a + \bar{S}'P_S\left(P_S^{-1} - WW_S^{-1}W'\right)P_S\bar{S}$. En consecuencia, la distribución a priori marginal de $S_{(p)}$ es una Student multivariante con $2b+1$ grados de libertad, matriz de posición $\mu_{S_{(p)}}$ y matriz de varianzas-covarianzas $\frac{2a_S}{2b-1}W_S^{-1}$.

Decíamos anteriormente que estas distribuciones marginales a priori pueden utilizarse para asignar valores a los parámetros del modelo. Dichos parámetros son las matrices P_S, P_T y P_ε, y los escalares a y b que determinan la distribución a priori de τ.

La asignación de valores a los parámetros de las distribuciones a priori empleando información extramuestral es una técnica empleada por los autores en otras situaciones (véase, por ejemplo, Rojo y Sanz, 2002 y 2005). La idea que subyace a la utilidad de dicha asignación es que en muchos modelos económicos, la información muestral es escasa, por lo que la distribución a priori llega a dominar a la verosimilitud. En consecuencia, qué parámetros hayamos asignado a la distribución a priori resulta de extrema importancia.

Si disponemos de una serie *vicaria*, para la que se pueda suponer el mismo modelo de descomposición (puede ser una serie más larga correspondiente a otro ámbito geográfico una vez corregida de escala para que su orden de magnitud sea el de la serie a estudiar) pueden asignarse valores a los parámetros siguiendo las siguientes pautas:

1. Se obtiene una tendencia y una estacionalidad provisionales de la serie vicaria por procedimientos descriptivos, y en consecuencia, una asignación de valores a la serie de ruido.

2. Haciendo paquetes (*batchs*) a dicho ruido podemos obtener varios valores de τ^{-1} (véase, por ejemplo, Dhrymes (1984), págs. 14 y 102, dependiendo del tipo de modelo empleado) con los que obtener una estimación a priori de, $E_\pi\left[\tau^{-1}\right]$ y de $Var_\pi\left[\tau^{-1}\right]$ (la media y cuasivarianza muestral de los valores de τ^{-1} obtenidos) y, por tanto, de a y b.

3. Tomando ahora la serie a descomponer, podemos obtener un valor de $\bar{T}$ y de $\bar{S}$ por procedimientos descriptivos, y por tanto, una estimación del ruido de nuestra serie. A partir de esta estimación, y suponiendo un modelo para el ruido, estimamos a priori la matriz de varianzas y covarianzas del ruido utilizando la expresión [11] y los valores de a y b obtenidos en el paso anterior.

4. No existe información suficiente para asignar valores razonables a P_S y P_T por su alta dimensión. Parece prudente entonces asignarles matrices iguales a la

identidad del orden correspondiente. Ello no prejuzga que las matrices de varianzas y covarianzas a posteriori de T y de S sean distintas a la identidad.

Puede también optarse por emplear distribuciones a priori neutrales en el sentido de Jeffreys (ver Zellner, (1971), pp. 41). En concreto:

1. Obtenemos valores de $\bar{T}$ y de $\bar{S}$ por procedimientos descriptivos (véase, por ejemplo, Martín-Guzmán, 1989, pp. 250 a 256).

2. Tomamos matrices identidad del orden correspondiente para P_S, P_T y P_ε.

3. Tomamos valores de a y b muy pequeños (en concreto, los hemos tomado iguales a 0.001) de forma que la distribución a priori de τ sea aproximadamente neutral. Ello permite que la distribución a priori de τ siga siendo una gamma, garantizando que nos mantenemos dentro de las familias conjugadas de distribuciones (esta práctica es frecuente en los modelos bayesianos jerárquicos, véase, por ejemplo Spiegelhalter y otros, (1999))

5. APLICACIONES Y VALIDACIONES DEL PROCEDIMIENTO OBTENIDO

En este apartado ejemplificamos el funcionamiento del procedimiento diseñado en los epígrafes anteriores, a la vez que comparamos sus resultados con los que proporciona el método de alisado exponencial con estacionalidad (Holt-Winters aditivo), y con otros dos procedimientos clásicos(8) mencionados anteriormente: Tramo-Seats(9) y X12-Regarima(10), que denotaremos abreviadamente como TS y X12. Las versiones utilizadas para los tres son las que tiene implementadas Eviews(11). Se presenta una ilustración del procedimiento y una evaluación más sistemática mediante procedimientos de simulación.

El ejemplo propuesto como ilustración utiliza la serie mensual de Matriculaciones de turismos (Dirección General de Tráfico, www.dgt.es) para España. En este caso se dispone de una serie larga (en el momento de cerrar la información, se disponía de datos básicamente homogéneos desde enero de 1985 hasta septiembre de 2009). La longitud de la serie permite comparar los cuatro métodos (nuestro método bayesiano con el alisado exponencial, con X12 y con TS), así como evaluar el grado de deterioro de las estimaciones cuando la serie pierde longitud. Así, se

(8) Recomendados por Eurostat a las Oficinas estadísticas de los países miembros.

(9) Véase Gómez y Maravall, 1997 y 1998.

(10) Véase US Census Bureau, 2002.

(11) Desarrollado por Quantitative Micro Software. Utilizamos la versión 4.1 del programa.

descomponen las series para muestras que se inician en enero de 1985, de 1990, de 1995, de 2000 y de 2005 (el valor final es siempre septiembre de 2009).

Finalmente presentaremos una evaluación más detenida del método propuesto realizada con elementos de simulación. Se han realizado 500 simulaciones para una serie corta (trimestral de 3 años, esto es, 12 datos en total) con una tendencia constante y una estacionalidad fija, con composición aditiva. A estas dos componentes se les ha añadido un ruido blanco gaussiano. Las simulaciones se realizan en 3x10=30 situaciones diferentes, correspondiendo a tres dispersiones del ruido (poco disperso, medio disperso, muy disperso) y 10 situaciones en lo referente a la estacionalidad, desde una estacionalidad leve hasta una estacionalidad amplia.

A continuación describimos brevemente las herramientas de comparación empleadas. Utilizaremos medidas numéricas, de carácter absoluto o relativo. Son medidas clásicas para comparar dos series, donde habitualmente una de ellas es la "verdadera" serie y la otra es la estimada. Nosotros las utilizamos para comparar la serie estimada (tendencia más estacionalidad) con la serie original, o (en la ilustración simulada) para comparar la tendencia lineal con las tendencias estimadas por nuestro método o por X12.

Utilizaremos tres medidas absolutas de discrepancias, que son clásicas en la evaluación de series cronológicas. Si llamamos y_t a la serie a comparar e $\hat{y}_t$ a la serie estimada, utilizaremos dos medidas absolutas de error, MAD (*Mean Absolute Deviation*, desviación absoluta media), es decir, $\mathrm{MAD}(y,\hat{y}) = \frac{1}{m}\sum_{t=1}^{m}\left|\hat{y}_t - y_t\right|$. y RMSE (*Root of Mean Squared Error*, raíz del error cuadrático medio), esto es,

$$\mathrm{RMSE}(y,\hat{y}) = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(\hat{y}_t - y_t\right)^2}\ .$$

Como medidas relativas, además del coeficiente de correlación lineal de Pearson, usaremos MAPE, (*Mean Absolute Percentage Error*, error absoluto medio porcentual), $\mathrm{MAPE}(y,\hat{y}) = \frac{1}{m}\sum_{t=1}^{m}\left|\frac{\hat{y}_t - y_t}{y_t}\right| \cdot 100$ y RMSPE (*Root of Mean Squared Percentage Error*, raíz del error cuadrático medio porcentual), $\mathrm{RMSPE}(y,\hat{y}) = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(\frac{\hat{y}_t - y_t}{y}\right)^2} \cdot 100$

Una de las características que queremos evaluar es la regularidad de la tendencia estimada. Dagum (1979) utiliza la norma L_2 de las primeras diferencias (para una serie $A_1,\ldots,A_T$, la define como $B_1 = \sum_{t=2}^{T}\left(A_t - A_{t-1}\right)^2$). Utilizando esta idea, tomamos como medida de regularidad la que denominaremos $B1$, que para una serie $y_t,\ t=1,\ldots,T$ definiremos como $B1(y) = \frac{1}{T-1}\sum_{t=2}^{T}\left|y_t - y_{t-1}\right|$. Una modificación de $B1$ en términos relativos, que permite comparar la regularidad de series de diferentes tamaños es $RB1(y) = \frac{1}{\bar{y}}\frac{1}{T-1}\sum_{t=2}^{T}\left|y_t - y_{t-1}\right|$, donde $\bar{y}$ es la media aritmética de la serie, permite comparar la regularidad de series de diferentes escalas.

Otras dos medidas (una absoluta y otra relativa) permiten evaluar la linealidad de la serie (para una tendencia lineal, las segundas diferencias se anulan). Estas medidas son $B2(y)$ y $RB2(y)$, $B2(y) = \frac{1}{T-2}\sum_{t=3}^{T}\left|y_t - 2y_{t-1} + y_{t-2}\right|$ y

$$RB2(y) = \frac{1}{\bar{y}}\frac{1}{T-2}\sum_{t=3}^{T}\left|y_t - 2y_{t-1} + y_{t-2}\right|.$$

Asimismo, mediremos el tamaño de la estacionalidad estimada mediante la desviación típica de las componentes estacionales.

En cuanto a la descomposición de la serie de Matriculaciones de Turismos, la ventaja de esta ilustración se encuentra en la utilización de una serie larga para la que puede realizarse una estimación por TS. Adicionalmente su longitud nos va a permitir evaluar la evolución de la calidad del procedimiento cuando la serie se acorta. Señalemos también que, como ocurre con frecuencia para series económicas, el aspecto de la serie parece abonar más una modelización que incluya una composición multiplicativa. Por ello, se ha realizado una transformación logarítmica previa en la serie antes de proceder a la estimación de las componentes.

Las cinco primeras columnas de la Tabla 1 evalúan la semejanza entre la serie original (el neperiano de las matriculaciones) y su estimación como suma de tendencia y estacionalidad. Puede observarse que la peor calidad se presenta para TS. El resultado para este método viene condicionado por el hecho de que se ha hecho trabajar al programa con las opciones por defecto, y no ha encontrado significativos los parámetros estacionales, por lo que no los ha incluido en la estimación. Puede también notarse que, en general, la estimación bayesiana presenta mejores (menores) valores de los estadísticos, especialmente para series cortas, indicando una mayor proximidad de la serie estimada con la original.

Tabla 1

MEDIDAS DE DISCREPANCIA EN LA DESCOMPOSICIÓN DE LA SERIE DE MATRICULACIONES

	MAPE	RMSPE	RMSE	MAD	Pearson	σ_S	RB1
desde 1985 Bayesiano	0,385	0,545	0,060	0,043	0,987	0,182	0,002
desde 1985 X12	0,380	0,601	0,066	0,043	0,984	0,188	0,001
desde 1985 TS	1,328	1,656	0,187	0,151	0,860		0,002
desde 1985 Alisado	0,657	0,924	0,102	0,074	0,961	0,184	0,003
desde 1990 Bayesiano	0,359	0,478	0,054	0,041	0,987	0,184	0,002
desde 1990 X12	0,348	0,519	0,058	0,040	0,984	0,170	0,001
desde 1990 TS	1,336	1,636	0,187	0,153	0,821		0,002
desde 1990 Alisado	0,601	0,828	0,093	0,069	0,959	0,188	0,003
desde 1995 Bayesiano	0,316	0,404	0,046	0,036	0,988	0,188	0,002
desde 1995 X12	0,308	0,448	0,051	0,036	0,986	0,217	0,001
desde 1995 TS	1,329	1,637	0,188	0,153	0,788		0,001
desde 1995 Alisado	0,534	0,698	0,080	0,062	0,965	0,193	0,003
desde 2000 Bayesiano	0,295	0,372	0,043	0,034	0,987	0,179	0,002
desde 2000 X12	0,296	0,401	0,046	0,034	0,984	0,191	0,001
desde 2000 TS	1,251	1,528	0,178	0,146	0,734		0,001
desde 2000 Alisado	0,517	0,640	0,074	0,060	0,960	0,185	0,003
desde 2005 Bayesiano	0,331	0,417	0,048	0,038	0,988	0,168	0,002
desde 2005 X12	0,346	0,435	0,050	0,040	0,987	0,173	0,002
desde 2005 TS	1,217	1,468	0,170	0,142	0,836		0,002
desde 2005 Alisado	0,532	0,702	0,081	0,062	0,966	0,177	0,004

En cuanto al comportamiento cuando la serie se acorta, nótese que TS mejora muy ligeramente, mientras que X12 también lo hace según se acorta la serie, salvo cuando la serie es muy corta (desde 2005 hasta septiembre de 2009, una serie de cuatro años y nueve meses), en cuyo caso empeora. El procedimiento bayesiano mejora según la serie se acorta, lo que muestra que su mayor eficiencia del diseño se presenta en la descomposición de series cortas.

Las dos últimas columnas evalúan la regularidad de la estimación de la tendencia, mediante la medida RB1. Nótese que, aunque las diferencias son escasas, la estimación por el procedimiento X12 proporciona tendencias más regulares. Por otro lado, proporciona en general mayores componentes estacionales, como muestra la columna de desviaciones típicas de los coeficientes estacionales (columna σ_S).

El balance general es que el procedimiento bayesiano y X12 funcionan de manera muy parecida en este caso, incluso para series cortas; el ligero deterioro de ciertos estadísticos para el periodo 2005-2009 puede deberse al mayor peso del periodo más reciente, que hace que el cambio brusco en la serie de matriculaciones en el verano de 2008 debido a la crisis tenga mayor peso en el conjunto de la serie (13 meses en un total de 57) lo que aconsejaría, probablemente, un tratamiento diferenciado de ambos periodos.

Presenta peores resultados para todas las selecciones temporales el Alisado exponencial (Holt-Winters estacional aditivo), si bien supera en todos los casos a TS. Señalemos que hemos dejado que el programa estimase los parámetros de alisado por mínimos cuadrados.

Finalmente, describamos el procedimiento de simulación empleado en la validación de nuestro método. Como señalamos en la introducción de este apartado, la simulación se ha realizado para series trimestrales de tres años consecutivos. Se ha tomado una tendencia constante, $T_j = 10,\ j = 1,\dots,12$. Para la estacionalidad se han propuesto 10 escenarios, siempre con estacionalidad fija y composición (con la tendencia) aditiva. Más explícitamente, se generan los cuatro valores estacionales $S_j = u_j \bullet \frac{k}{2},\ j = 1,\dots,4$, donde $k = 1,\dots,10$ recorren los escenarios y u_j son realizaciones de una distribución uniforme $U(0,1)$. Se consigue así que la estacionalidad varíe en el intervalo $\left(0, \frac{k}{2}\right)$. Esta estacionalidad se corrige para que el efecto estacional medio sea nulo, es decir, se obtienen $S_j^{'} = S_j - \frac{1}{4}\sum_{i=1}^{4} S_i,\ j = 1,\dots,4$. Con ello se consiguen escenarios más estacionales (valores altos de k) o menos estacionales (valores pequeños de k). En las tablas, se describen estos 10 escenarios mediante R_S^{max} , recorrido máximo de la estacionalidad, que varía desde ½ hasta 5.

Otro de los parámetros de la simulación es el nivel de ruido. En concreto, el ruido se simula como un ruido blanco, con $\varepsilon_i \rightarrow \text{Niid}(0,\sigma)$, donde $i = 1,\dots,12$, y σ igual a 1, ½ ó ¼.

La comparación entre la tendencia original y las estimadas se presenta en la Tabla 2, y el valor presentado es un ratio entre el valor que se obtendría con el procedimiento que proponemos y la obtenida mediante X12. En cada situación se ha calculado el promedio de la medida para las 500 simulaciones realizadas en cada caso. El cociente de ambos es el valor que se recoge en la tabla, de modo que los valores menores que la unidad indican que el resultado obtenido con el procedimiento bayesiano propuesto es mejor (menor valor del índice correspondiente) que el obtenido mediante X12. Por ejemplo, para $R_S^{max} = 2,5$ y σ =½, MAPE vale 0,81, lo que indica que el valor de MAPE en el procedimiento bayesiano es un 81% del obtenido para X12. En el margen derecho e inferior de cada subtabla se presentan los ratios de los valores medios de las medidas para los distintos escenarios de simulación.

Puede verse que nuestro procedimiento replica mejor la tendencia de la serie, valorado con cualquiera de las medidas. Además, nótese que el procedimiento no es muy sensible al tamaño de la estacionalidad. Así, los resultados no son muy diferentes para los distintos valores de R_S^{max}. Por el contrario, la comparación es muy sensible al tamaño del ruido, mostrando una mayor similaridad de ambos procedimientos según disminuye el tamaño de la estacionalidad. Los resultados completos (que no se presentan) muestran adicionalmente que los indicadores mejoran para ambos procedimientos según disminuye la varianza del ruido.

Tabla 2

MEDIDAS DE DISCREPANCIA ENTRE TENDENCIA ORIGINAL Y ESTIMADA

MAPE	*σ=1*	*σ=1/2*	*σ=1/4*	***Media***	***MAD***	*σ=1*	*σ=1/2*	*σ=1/4*	***Media***
$R_S^{max}=0,5$	0,82	0,81	0,89	0,83	$R_S^{max}=0,5$	0,80	0,79	0,89	0,81
$R_S^{max}=1$	0,80	0,81	0,91	0,82	$R_S^{max}=1$	0,79	0,81	0,93	0,81
$R_S^{max}=1,5$	0,80	0,82	0,89	0,82	$R_S^{max}=1,5$	0,78	0,81	0,92	0,81
$R_S^{max}=2$	0,81	0,82	0,88	0,82	$R_S^{max}=2$	0,81	0,84	0,88	0,83
$R_S^{max}=2,5$	0,79	0,81	0,88	0,81	$R_S^{max}=2,5$	0,79	0,83	0,87	0,81
$R_S^{max}=3$	0,80	0,82	0,88	0,82	$R_S^{max}=3$	0,78	0,82	0,87	0,81
$R_S^{max}=3,5$	0,80	0,80	0,89	0,81	$R_S^{max}=3,5$	0,79	0,81	0,88	0,81
$R_S^{max}=4$	0,80	0,81	0,90	0,82	$R_S^{max}=4$	0,80	0,80	0,88	0,81
$R_S^{max}=4,5$	0,80	0,82	0,89	0,82	$R_S^{max}=4,5$	0,79	0,82	0,89	0,82
$R_S^{max}=5$	0,81	0,82	0,90	0,82	$R_S^{max}=5$	0,80	0,81	0,90	0,82
media	0,80	0,81	0,89		**media**	0,79	0,81	0,89	

RMSPE	*σ=1*	*σ=1/2*	*σ=1/4*	***Media***	***RMSE***	*σ=1*	*σ=1/2*	*σ=1/4*	***Media***
$R_S^{max}=0,5$	0,80	0,79	0,90	0,81	$R_S^{max}=0,5$	0,73	0,51	0,41	0,58
$R_S^{max}=1$	0,77	0,80	0,95	0,80	$R_S^{max}=1$	0,72	0,52	0,43	0,59
$R_S^{max}=1,5$	0,78	0,83	0,94	0,81	$R_S^{max}=1,5$	0,71	0,53	0,41	0,59
$R_S^{max}=2$	0,80	0,83	0,88	0,82	$R_S^{max}=2$	0,73	0,53	0,40	0,59
$R_S^{max}=2,5$	0,78	0,81	0,90	0,80	$R_S^{max}=2,5$	0,71	0,53	0,41	0,58
$R_S^{max}=3$	0,77	0,83	0,89	0,81	$R_S^{max}=3$	0,70	0,53	0,40	0,58
$R_S^{max}=3,5$	0,77	0,79	0,92	0,80	$R_S^{max}=3,5$	0,69	0,52	0,41	0,57
$R_S^{max}=4$	0,78	0,80	0,90	0,80	$R_S^{max}=4$	0,71	0,51	0,40	0,58
$R_S^{max}=4,5$	0,77	0,81	0,91	0,80	$R_S^{max}=4,5$	0,70	0,52	0,42	0,58
$R_S^{max}=5$	0,78	0,79	0,92	0,81	$R_S^{max}=5$	0,71	0,51	0,42	0,58
media	0,78	0,81	0,91		**media**	0,71	0,52	0,41	

Otros estadísticos revisados son los correspondientes a las primeras o segundas diferencias de la tendencia. El primero, B1, mide la regularidad de la tendencia, mientras que B2 mide el grado de linealidad. Al principio de este apartado se definieron estos indicadores en términos absolutos y relativos, si bien en la simulación actual pueden utilizarse los absolutos por ser las series del mismo orden de magnitud. Puede observarse en la Tabla 3 que nuestro procedimiento proporciona tendencias más regulares y más lineales que X12 (este último hecho es especialmente importante ya que la tendencia "verdadera" es lineal). Su evolución en los distintos escenarios es similar a la descrita para el resto de las medidas. Así, con los datos originales (que no se presentan) la regularidad y linealidad de la tendencia estimada es poco sensible al tamaño de la estacionalidad para ambos procedimientos, mientras que sí lo es a la dispersión del ruido, mejorando ambos métodos cuando el ruido pierde parte de su varianza. Sí muestra esta Tabla que nuestro procedimiento mejora su habilidad frente a X12 para escenarios de mayor varianza del ruido.

Tabla 3

PRIMERAS Y SEGUNDAS DIFERENCIAS

B1	*σ=1*	*σ=1/2*	*σ=1/4*	***Media***	***B2***	*σ=1*	*σ=1/2*	*σ=1/4*	***Media***
$R_S^{max} = 0,5$	0,86	0,92	0,93	0,90	$R_S^{max} = 0,5$	0,96	0,93	0,98	0,95
$R_S^{max} = 1$	0,83	0,91	0,92	0,89	$R_S^{max} = 1$	0,92	0,93	0,97	0,93
$R_S^{max} = 1,5$	0,84	0,92	0,93	0,89	$R_S^{max} = 1,5$	0,94	0,93	0,98	0,95
$R_S^{max} = 2$	0,84	0,92	0,92	0,89	$R_S^{max} = 2$	0,92	0,94	0,98	0,94
$R_S^{max} = 2,5$	0,84	0,91	0,93	0,89	$R_S^{max} = 2,5$	0,94	0,90	0,98	0,94
$R_S^{max} = 3$	0,84	0,91	0,92	0,89	$R_S^{max} = 3$	0,92	0,95	0,98	0,94
$R_S^{max} = 3,5$	0,85	0,91	0,93	0,89	$R_S^{max} = 3,5$	0,94	0,92	0,97	0,94
$R_S^{max} = 4$	0,84	0,91	0,93	0,89	$R_S^{max} = 4$	0,93	0,92	1,002	0,94
$R_S^{max} = 4,5$	0,85	0,92	0,93	0,89	$R_S^{max} = 4,5$	0,94	0,93	0,93	0,94
$R_S^{max} = 5$	0,84	0,92	0,93	0,89	$R_S^{max} = 5$	0,93	0,94	0,99	0,94
media	0,84	0,91	0,93		**media**	0,93	0,93	0,98	

CONCLUSIONES

Hemos desarrollado en este trabajo un procedimiento bayesiano jerárquico para descomponer una serie cronológica corta (con pocas observaciones) en una tendencia o ciclo-tendencia regular y una estacionalidad fija, para un esquema aditivo.

En los apartados 2 y 3 se desarrollan los aspectos distribucionales del problema Así, las distribuciones empleadas en la cadena jerárquica son gausianas, salvo para el parámetro de precisión, al que se asigna una distribución gamma. Ello garantiza que nos movemos en un marco de distribuciones conjugadas. La verosimilitud se construye a partir de un ruido gausiano para la componente irregular de la serie. Puede utilizarse un ruido blanco a priori, aunque es posible introducir otros esquemas, modificando de forma conveniente la matriz de varianzas y covarianzas.

El procedimiento proporciona explícitamente las soluciones de Bayes que constituyen las estimaciones de la tendencia de la serie y de su componente estacional. Se obtienen también medidas de precisión de las estimaciones. El hecho de que las soluciones sean explícitas facilita su uso en procedimientos automáticos de extracción de señal a grupos numerosos de series.

En el apartado 4 se obtienen las distribuciones marginales a priori para los parámetros implicados. Estas distribuciones son el vehículo adecuado para introducir información extramuestral a priori en el procedimiento de estimación, especialmente cuando se disponga de una serie vicaria (*proxy*) de la que se trata de descomponer.

El apartado 5 presenta ilustraciones y validaciones del procedimiento. En primer lugar, hemos descompuesto la serie de Matriculaciones de turismos en el ámbito nacional. Si bien por limitaciones de espacio no se presentan las series resultantes de la descomposición, se muestra la calidad del procedimiento comparada con el X12-Arima, con Tramo-Seats y con el Alisado exponencial (Holt-Winters). El balance general es que el procedimiento bayesiano y X12 funcionan de manera muy parecida para esta serie, incluso para series cortas; proporciona mejores resultados que el alisado exponencial, siendo TS el que peores resultados produce, si bien no se han modificado las opciones por defecto en este último. La ventaja que proporciona la longitud de la serie es que permite observar el comportamiento estable de nuestro procedimiento cuando la longitud se acorta.

La segunda ilustración es una evaluación del procedimiento realizada mediante simulaciones de una serie corta (trimestral con 12 observaciones) compuesta de una tendencia constante y con distintas amplitudes para la estacionalidad y para la dispersión del ruido. Los resultados permiten avalar la calidad de nuestro procedimiento, en cuanto a que es capaz de estimar tendencias regulares y a reconocer las estacionalidades de la serie.

Uno de los propósitos de los autores fue la combinación de la descomposición con la predicción coyuntural de series cortas, en especial para pequeñas áreas. En estos casos, la serie no contiene suficiente información relevante, y los métodos habituales de predicción prolongan hacia el futuro pautas erráticas. La escasa información de las series cortas dificulta integrar en un único procedimiento descomposición y predicción, por lo que nuestra propuesta en ese terreno es, por el momento, secuencial. En primer lugar se procedería a la descomposición óptima de la serie mediante el trabajo aquí descrito, para posteriormente realizar la predicción coyuntural de la tendencia mediante un método elaborado por Mendoza y de Alba (véase Mendoza y Alba, 2006) mejorado por los autores (Rojo y Sanz, 2008) para incorporar información extramuestral a priori.

Los resultados permiten, en definitiva, abordar la extracción de la señal y la predicción de series cortas, series que se presentan frecuentemente en el ámbito de las estadísticas económicas y sociales, en particular cuando se producen cambios metodológicos. Al tratarse de un procedimiento bayesiano, resulta especialmente de interés en ámbitos correspondientes a pequeñas áreas, como son las desagregaciones espaciales a nivel de NUTS3 o en ramas y subramas de actividad. Las frecuentes apariciones de volatilidades en las series pueden ser combatidas mediante la imposición de condiciones de regularidad en los resultados.

Los autores se proponen en un futuro próximo mejorar el modelo propuesto analizando modelos más complejos de comportamiento para los errores en el modelo de descomposición. Uno de ellos sería el denominado de líneas aéreas(12), frecuente en la modelización de los errores en series temporales. Señalemos, no obstante, que la escasa longitud de las series cortas hace a veces poco realista la asignación de modelos complejos, en los que intervienen un número importante de periodos. Es por ello por lo que debemos ser cautelosos en el planteamiento y motivación de estos trabajos.

(12) Un modelo $\text{ARIMA}(0,1,1)\times(0,1,1)_{12}$

ANEXO. OBTENCIÓN DE LAS DISTRIBUCIONES MARGINALES A POSTERIORI

En este Anexo se construyen de forma detallada las distribuciones a posteriori de la tendencia y la estacionalidad.

Multiplicando la expresión de la priori conjunta, [7], una vez incorporada la restricción [2], es decir, en el modelo restringido, por la expresión de la función de verosimilitud, [9], se obtiene la distribución conjunta a posteriori,

$$\pi\left(S_{(p)},T,\tau\middle|D\right)\propto \tau^{b-1+N+r+\frac{p}{2}}\, e^{-a\tau}\exp\left\{-\frac{\tau}{2}\left[T'D_2'D_2T+\left(T-\overline{T}\right)'P_T\left(T-\overline{T}\right)+\right.\right.$$

$$+\left(S_{(p)}-\mu_{S_{(p)}}\right)'W_S\left(S_{(p)}-\mu_{S_{(p)}}\right)+\overline{S}'P_S\left(P_S^{-1}-W\,W_S^{-1}\,W'\right)P_S\,\overline{S}+ \qquad \text{[A-1]}$$

$$\left.\left.+\left(y-T-E_W\,S_{(p)}\right)'P_\varepsilon\left(y-T-E_W\,S_{(p)}\right)\right]\right\}$$

Integrando en $S_{(p)}$ se obtiene la distribución conjunta a posteriori de T y τ

$$\pi\left(T,\tau\middle|D\right)\propto\tau^{\frac{p-1}{2}}\,\tau^{b-1+N+r+\frac{p}{2}}\,e^{-a\tau}\exp\left\{-\frac{\tau}{2}\left[T'D'_2D_2T+\left(T-\overline{T}\right)+\right.\right.$$

$$+\overline{S}'P_S\left(P_S^{-1}W\;W_S^{-1}W\right)P_S\,\overline{S}+\mu'_{S_{(p)}}\,W_S\,\mu_{S_{(p)}}+\left(y-T\right)'P_\varepsilon\left(y-T\right)+ \qquad \text{[A-2]}$$

$$\left.\left.-\left(\mu'_{S_{(p)}}\,W_S+\left(y-T\right)'P_\varepsilon E_W\right)P_{\varepsilon W}^{-1}\left[E_W'P_\varepsilon(y-T)+W_S\mu_{S_{(p)}}\right]\right]\right\}$$

donde $P_{\varepsilon W}=W_S+E'_W\,P_\varepsilon E_W$.

Integrando en T la distribución a posteriori conjunta [A-2], obtenemos la marginal a posteriori de τ ,

$$\pi\left(\tau\middle|D\right)\propto\tau^{b-1+N+r-\frac{N+r}{2}+\frac{p}{2}-\frac{p-1}{2}}\ \exp\left\{-\frac{\tau}{2}\left[2a+\bar{S}'P_S\left(P_S^{-1}-W\ W_S^{-1}W'\right)P_S\bar{S}+\right.\right.$$

$$+\mu'_{S_{(P)}}\ W_S\,\mu_{S_{(P)}}-\mu'_{S_{(p)}}\ W_S\,P_{\varepsilon W}^{-1}\ W_S\,\mu_{S_{(p)}}-\mu'_T\,P_{T\varepsilon}\,\mu_T+\bar{T}'P_T\bar{T}+ \qquad [A\text{-}3]$$

$$\left.\left.+\ y'P_\varepsilon\ P_{\varepsilon W\varepsilon}P_\varepsilon\,y-2\,\mu'_{S_{(p)}}\ W_S\,P_{\varepsilon W}^{-1}\ E'_W\,P_\varepsilon y\ \right]\right\}$$

donde denotamos por $P_{\varepsilon W\varepsilon}=P_\varepsilon^{-1}-E_W P_{\varepsilon W}^{-1}E'_W$, por $P_{T\varepsilon}=D'_2\,D_2+P_T+P_\varepsilon P_{\varepsilon W\varepsilon}P_\varepsilon$, y por

$\mu_T=P_{T\varepsilon}^{-1}\left[P_T\bar{T}+P_\varepsilon P_{\varepsilon W\varepsilon}P_\varepsilon y-P_\varepsilon E_W P_{\varepsilon W}^{-1}W_S\mu_{S_{(p)}}\right]$. Finalmente, llamando $\bar{y}=E'_W\,P_\varepsilon\,y$, la marginal a posteriori de τ , que figura en la expresión [A-3], queda reducida a

$$\pi\left(\tau\middle|D\right)\propto\tau^{b+\frac{N+r}{2}-\frac{1}{2}}\ e^{-\frac{\tau}{2}\left(2\hat{a}\right)} \qquad [A\text{-}4]$$

donde hemos denotado por

$$2\hat{a}=2a+\bar{T}'P_T\bar{T}+\bar{S}'P_S\bar{S}+y'P_\varepsilon y-\mu'_T\,P_{T\varepsilon}\mu_T-\left(\bar{y}+W'P_S\bar{S}\right)'P_{\varepsilon W}^{-1}\left(\bar{y}+W'P_S\bar{S}\right) \qquad [A\text{-}5]$$

Se trata de una $\gamma\left(\hat{a},b+\frac{N+r}{2}+\frac{1}{2}\right)$ (ver la Nota 5).

Para obtener la marginal a posteriori de T integramos la conjunta de T y τ que hemos obtenido en [A-2], con respecto a τ. Resulta

$$\begin{aligned}\pi\left(T\middle|D\right)\propto\Big\{&2a+\bar{T}'D'_2D_2T+(T-\bar{T})'P_T(T-\bar{T})+\\&\bar{S}'P_S\left[P_S^{-1}-WW_S^{-1}W'\right]P_S\bar{S}+\mu'_{S_{(p)}}\ W_S\mu_{S_{(p)}}+\left(y-T\right)'P_\varepsilon\left(y-T\right)+\\&\left[\mu'_{S_{(p)}}\ W+\left(y-T\right)'P_\varepsilon E_W\right]P_{\varepsilon W}^{-1}\left[E'_W P_\varepsilon(y-T)+W_S\mu_{S_{(p)}}\right]\Big\}^{-(b+N+r+1/2)}\end{aligned} \qquad [A\text{-}6]$$

Llamando

$$\begin{aligned}2\tilde{a}=2a-\mu'_T\,P_{T\varepsilon}\mu_T+\bar{T}'P_T\bar{T}+y'P_\varepsilon P_{\varepsilon W\varepsilon}\,P_\varepsilon y-2\mu'_{S_{(p)}}\ W_S P_{\varepsilon W}^{-1}E'_W\,P_\varepsilon y+\\+\bar{S}'P_S\left[P_S^{-1}-WW_S^{-1}W'\right]P_S\bar{S}+\mu'_{S_{(p)}}\ W_S\mu_{S_{(p)}}-\mu'_{S_{(p)}}\ W_S\,P_{\varepsilon W}^{-1}W_S\,\mu_{S_{(p)}}\end{aligned} \qquad [A\text{-}7]$$

podemos escribir la marginal a posteriori de T , [A-6], como

$$\pi\left(T\middle|D\right) \propto \left\{2\tilde{a} + \left(T-\mu_T\right)'P_{T\varepsilon}\left(T-\mu_T\right)\right\}^{-(b+N+r+1/2)} \qquad \text{[A-8]}$$

Se trata de una Student multivariante(13) de dimensión $N+r$ con $2b+N+r+1$ grados de libertad, matriz de posición μ_T y matriz de varianzas-covarianzas

$$\frac{2\tilde{a}}{2b+N+r-1}P_{T\varepsilon}^{-1}.$$

Finalmente, obtengamos la marginal a posteriori de $S_{(p)}$ a partir de [A-1]. Integrando primero en T y después en τ y escribiendo

$$P_{TD} = D_2^{'} D_2 + P_T + P_\varepsilon,$$

$$P_{S_{(p)}} = W_S + E'_W\left(P_\varepsilon - P_\varepsilon P_{TD}^{-1} P_\varepsilon\right)E_W,$$

$$\tilde{\mu}_{S_{(p)}} = P^{-1}{}_S\left[W_S\mu_{S_{(p)}} - E'_W P_\varepsilon P_{TD}^{-1} P_T \overline{T} + E'_W\left(P_\varepsilon - P_\varepsilon P_{TD}^{-1} P_\varepsilon\right)y\right] \text{ y} \qquad \text{[A-9]}$$

$$2a^* = 2a - \tilde{\mu}_{S_{(p)}}^{'} P_{S_{(p)}} \tilde{\mu}_{S_{(p)}} + \overline{S}'P_S\overline{S} + \overline{T}'P_T\overline{T} + y'P_\varepsilon y - \left(P_T\overline{T} + P_\varepsilon y\right)'P_{TD}^{-1}\left(P_T\overline{T} + P_\varepsilon y\right)$$

obtenemos la marginal a posteriori de $S_{(p)}$,

$$\pi\left(S_{(p)}\middle|D\right) \propto \left\{2a^* + \left(S_{(p)} - \tilde{\mu}_{S_{(p)}}\right)'P_{S_{(p)}}\left(S_{(p)} - \tilde{\mu}_{S_{(p)}}\right)\right\}^{-\left(b+\frac{N+r}{2}+\frac{p}{2}\right)} \qquad \text{[A-10]}$$

que es una multivariante de Student con $2b+N+r-1$ grados de libertad, matriz de posición $\tilde{\mu}_{S_{(p)}}$ y matriz de varianzas y covarianzas $\dfrac{2a^*}{2b+N+r-1}P_{S_{(p)}}^{-1}$.

La obtención de las distribuciones marginales de T y $S_{(p)}$, [A-8] y [A-10] proporciona el estimador de Bayes óptimo que, como es conocido, es la esperanza

(13) Véase la Nota 6 para la definición empleada para esta distribución.

(parámetro de posición) de la distribución marginal a posteriori correspondiente si suponemos una función de pérdida cuadrática.

REFERENCIAS

AKAIKE, H. (1980). «Seasonal adjustment by a Bayesian modelling». *Journal of Time Series Analysis*, Vol. 1, No. 1, pp. 1-13.

BROEMELING, L.D. (1985). «Bayesian Analysis of Linear Models». Marcel Dekker, Inc. New York.

CHOLETTE, P.A. (1979). «Spectral Diagnosis and Marginal Improvements of the X-11 Seasonal Adjustment Method for Short Series». *Technical Report of the Seasonal Adjustment and Time Series Staff*, Ottawa, Canada: Statistics Canada.

DAGUM, E. B. (1979). «On the Seasonal Adjustment of economic Time Series Aggregates: A Case Study of the Unemployment Rate», in Counting the Labor Force, National Common Employment and Unemployment Statistics, Appendix, 2, 317-344, Washington.

DHRYMES, PH.J. (1984). «Econometría». Ed. AC, Madrid.

FINDLEY, D.F. Y MARTIN, D.E.K. (2003). «Frequency Domain Analyses of SEATS and X-11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series». *Research Report Series*, Statistics, 2003-02, Washington D.C.: U.S. Bureau of the Census, pp. 1-32.

GÓMEZ, V. Y MARAVALL, A. (1997). «Programs Tramo and Seats»*. Instructions for the User*, guía incluida en el programa, Madrid.

GÓMEZ, V. Y MARAVALL, A. (1998). «Guide for Using the Programs Tramo and Seats». *Documento de trabajo 9805*, Servicio de Estudios del Banco de España, Madrid.

HOOD, C.C., ASHLEY, J.D. Y FINDLEY, D.F. (2000). «An empirical Evaluation of Tramo/Seats on Simulated Series». *Proceedings of the Business and Economic Statistics Section*, American Statistical Association.

LADIRAY, D. Y MAZZI, G.L. (2003). «Seasonal Adjustement of European Aggregates: Direct versus Indirect Approach» in M. Manna y R. Peronaci (eds.) *Seasonal Adjustment*, Frankfurt am Maim, germany: European Central Bank, pp. 37-65.

MARTÍN-GUZMÁN, M. P. Y MARTÍN PLIEGO, F. J. (1989). «Curso básico de estadística económica». AC: Madrid.

MATAS MIR, A. Y RONDONOTTI, V. (2003). «The Performance of X-12 in the Seasonal Adjustment of Short Time Series» in M. Manna y R. Peronaci (eds.) *Seasonal Adjustment*, Frankfurt am Maim, Germany: European Central Bank, pp. 149-159.

MAZZI, G.L. Y SAVIO, G. (2005). «The Seasonal Adjustment of Short Time Series» en *Working Papers and Studies*, European Commision.

MENDOZA, M. Y DE ALBA, E. (2006). «Forecasting an accumulated series based on partial accumulation II: A new Bayesian method for short series with stable seasonal patterns». *International Journal of Forecasting*, vol. 22, pp. 781-798.

ROJO, J.L. Y SANZ, J.A. (2002). «Estimaciones para pequeñas áreas: un enfoque bayesiano a un problema de distribución». *Estudios de Economía Aplicada*. 20-I, pp. 217-240.

ROJO, J.L. Y SANZ, J.A. (2005). «A Bayesian benchmarking method with applications to the Quarterly National Accounts». *Working Papers and Studies*. ISSN: 1725-4825, ISBN: 92-79-01310-6, Office for Official Publications of the European Communities, 2005. Luxembourg.

ROJO, J.L. Y SANZ, J.A. (2008). «Predicción de series cortas con estacionalidad estable usando distribuciones a priori informativas». *Anales de Economía Aplicada*, Nº XXII (Triclás y Moslares, editores), ISBN: 978-84-92453-25-2, pp. 265-276.

SPIEGELHALTER, D., THOMAS, A. Y BEST, N. (1999): *WinBUGS, V. 1.2: User Manual* MRC Biostatistics Unit, Cambridge, U.K.

U.S. CENSUS BUREAU (2002). «X-12-ARIMA Reference Manual, Final Version 0.2» Washington DC, USA: US Census Bureau.

WINTERS, P.R. (1960). «Forecasting sales by exponentially weighted moving averages». Management Science, 6, 324-342

YOUNG, M.R. (1996). «Robust Seasonal Adjustment by Bayesian Modelling». *Journal of Forecasting*, 15, pp. 355-367.

ZELLNER, A. (1971). «An Introduction to Bayesian Inference in Econometrics». J. Wiley & Sons, Inc. New York.

DECOMPOSITION FOR SHORT TIME SERIES: A BAYESIAN APPROACH

ABSTRACT

The decomposition of Time Series is an essential task for the specialists in Conjunctural Analysis. The most commonly used devices are Tramo-Seats and X12-RegArima. Both devices have been recommended by Eurostat. Nowadays the Statistics literature has stated the deterioration of the quality of the decomposition for short time series.

The authors have carried out a Hierarchical Bayes model for the trend-cycle extraction and for the seasonal component estimation in time series. Its usefulness is greatest for short ones, not showing any sign of deterioration, much unlike the classic methods.

The paper ends with two examples. Firstly, we present an illustration and, finally, we show a validation of the approach by using simulation for several scenarios. The conclusions confirm the interest of the procedure in decomposing a short time series

Key Words: Bayesian models, Short Time Series, Times Series Analysis, Seasonal Adjustment

AMS Classificación: 62C10 (Bayesian problems), 62F15 (Bayesian inference), 37M10 (Times Series Analysis)