

**TRAMO Y SEATS: UN MARCO COMPLETO PARA EL
ANÁLISIS UNIVARIANTE Y EXTRACCIÓN DE SEÑALES
DE SERIES TEMPORALES¹**

Fernando Lorenzo y J. Manuel Revuelta²

Resumen

Los programas **TRAMO** y **SEATS** constituyen dos herramientas independientes, pero complementarias, para el análisis univariante y la extracción de señales en series temporales. Es precisamente esa complementariedad, reforzada por la posibilidad de una fácil integración de ambos entornos, la que permite referirse a ellos en conjunto como un marco completo de análisis de series temporales, de fácil manejo y de gran eficiencia. Con esta presentación se pretende motivar y orientar a posibles nuevos usuarios de estos programas, realizando una valoración de sus principales prestaciones en relación a otros programas disponibles en el mercado.

Palabras Clave:

TRAMO; SEATS; Modelización automática; Modelos ARIMA; Extracción de señales.

¹ Los autores desean agradecer los comentarios recibidos de Antoni Espasa, Regina Kaiser, J.M. Martínez, Esther Ruiz y Eva Senra, así como las aclaraciones y precisiones de Víctor Gómez y Agustín Maravall. Se agradece, asimismo, la ayuda recibida de la DGICYT, proyecto PB93-0236. Este trabajo se publicará próximamente en la *Revista de Economía Aplicada*.

² Departamento de Estadística y Econometría, Universidad Carlos III de Madrid; E-mail: lorenzo@est-econ.uc3m.es y revuelta@est-econ.uc3m.es.



INTRODUCCIÓN

Los programas **TRAMO** (*Time Series Regression with ARIMA Noise, Missing Observations and Outliers*) y **SEATS** (*Signal Extraction in ARIMA Time Series*) constituyen dos herramientas independientes, pero complementarias, para el análisis univariante y la extracción de señales en series temporales. Es precisamente esa complementariedad, reforzada por la posibilidad de una fácil integración de ambos entornos, la que permite referirse a ellos en conjunto como un marco completo de análisis de series temporales, de fácil manejo y de gran eficiencia.

En 1986 Agustín Maravall³ y Víctor Gómez⁴ comienzan a trabajar sobre una metodología alternativa a los métodos empiricistas de desestacionalización, tomando como punto de partida estudios previos de J.P. Burman (Banco de Inglaterra). Como resultado surgen las primeras versiones de SEATS. Este programa de extracción de señales, en contraposición al empiricismo en que se basan los métodos anteriores, tiene tras de sí el marco de los llamados **métodos de desestacionalización basados en modelos**. Esta metodología plantea la necesidad de una modelización univariante previa de las series objeto de estudio, lo que hace deseable una sistematización de este proceso, especialmente en el caso de que se pretenda un tratamiento masivo de series. Es esta necesidad la que motiva a los autores a desarrollar un entorno de modelización de series temporales, con un especial hincapié puesto en la automatización del proceso: el programa TRAMO. Ambos programas se desarrollan definitivamente durante la estancia de Agustín Maravall en el Instituto Universitario Europeo de Florencia.

En 1992 la agencia EUROSTAT encarga a un equipo internacional de expertos la actualización y homogeneización de sus actividades relacionadas con la extracción de señales de series temporales. TRAMO y SEATS son recomendados y comienzan a utilizarse en el tratamiento periódico que realiza esta agencia de miles de series. A partir de ese momento, otras agencias y organismos los han ido incorporando en sus actividades de forma generalizada. También han despertado el interés del mundo académico por su efectividad y lo innovador de muchas de sus ideas. Todo parece indicar que en un futuro próximo se podrá hablar de ellos como de un estándar en este tipo de tratamientos.

Las consideraciones anteriores pueden llevar a confusión respecto al tipo de usuario al que estos programas van dirigidos. En realidad el perfil puede ser variado. El entorno de ambos programas está estructurado mediante un conjunto de parámetros, de forma que cada usuario pueda

³ Suboficina de Información y Métodos de Cálculo. Servicio de Estudios del Banco de España.

⁴ Dirección General de Análisis y Programación Presupuestaria. Ministerio de Economía y Hacienda.

definir las alternativas más convenientes en función de sus necesidades y formación. Incluso existe la posibilidad de una automatización prácticamente absoluta del estudio de una serie particular. Por tanto, no es necesario ser un experto en series temporales para su uso, aunque evidentemente, será éste el que saque mayor partido a un entorno que incluye utilidades de gran potencial para el tratamiento de la mayoría de los aspectos involucrados en el análisis univariante de series temporales. No obstante, el usuario experto que requiera de estudios específicos fuera de los habituales, que surgen sobre todo en la modelización, se encontrará limitado en este marco. En este sentido, estos programas no pueden sustituir, ni lo pretenden, a entornos como SCA, TSP, etc., que permiten una verdadera programación personalizada. Estos últimos son colecciones de programas estadísticos ligados por un lenguaje, mientras que TRAMO y SEATS son dos programas orientados a objetivos concretos no cubiertos, en general, por los anteriores.

Cabe destacar asimismo, la transparencia en lo que se refiere a la metodología estadística utilizada, que es en gran medida original, y que queda perfectamente ilustrada en las referencias de los autores que aparecen en la bibliografía incluida al final del presente trabajo.

El *hardware* necesario para la última versión es, básicamente, un PC con procesador Intel 80486 o *PENTIUM*, 2 MB libres de capacidad en el disco duro y al menos 4 MB de memoria RAM. Respecto al *software*, es suficiente disponer de la versión de MS-DOS 3.3 o superior. Cabe resaltar que estos programas no están sujetos a ningún coste comercial y son de libre acceso.⁵ Las versiones actuales no son todavía definitivas (son versiones Beta) aunque son plenamente operativas. La última revisión es de marzo de 1996 y es sobre la que se basa esta presentación.

La exposición comienza con una descripción independiente tanto del programa TRAMO (sección 1) como del programa SEATS (sección 2), comparándolos con otros entornos existentes. En la sección 3 se analizan las posibilidades de integración de ambos entornos, planteando distintas alternativas de uso. El artículo se cierra con un apartado final de conclusiones. Evidentemente, no es el objetivo del presente trabajo hacer un estudio completo sino, fundamentalmente, motivar y orientar a posibles nuevos usuarios, tratando de hacer una valoración de sus principales prestaciones.

⁵ Es posible encontrar versiones de estos programas a través de INTERNET. Para ello sólo hace falta conectarse a "FTP 149.139.6.101", con *login* "anonymous" y *password* la dirección de correo electrónico del usuario, siguiendo a continuación las instrucciones del fichero "readme" en el directorio "seats-tramo". Los manuales de ambos programas aparecen como ficheros *postscript* (*.ps) dentro de este entorno.

1. EL PROGRAMA TRAMO

1.1. EL MODELO DE PARTIDA

El modelo que constituye la base del programa TRAMO responde a la expresión (respetamos la notación del manual):

$$z_t = y_t' \beta + v_t,$$

donde z_t son las observaciones de una serie temporal $\{t=1, \dots, T\}$, y_t un vector de n variables de regresión, entre las que se puede incluir una constante y cualquier conjunto de variables cualitativas que recojan efectos deterministas sobre z_t , β el vector de coeficientes de regresión y v_t el componente estocástico de la serie considerada. Para modelizar esta última se considera un modelo ARIMA estacional multiplicativo general:

$$\phi(L)\delta(L)v_t = \theta(L)a_t,$$

donde L es el operador de retardos ($Lx_t = x_{t-1}$). Los polinomios tienen la forma:

$$\delta(L) = (1-L)^d(1-L^s)^D$$

$$\phi(L) = (1 + \phi_1 L + \dots + \phi_p L^p)(1 + \Phi_1 L^s + \dots + \Phi_r L^{s \times r})$$

$$\theta(L) = (1 + \theta_1 L + \dots + \theta_q L^q)(1 + \Theta_1 L^s + \dots + \Theta_Q L^{s \times Q})$$

siendo s el número de observaciones por año. Para modelos directamente introducidos por el usuario no existe limitación en los órdenes de los polinomios. Esta limitación sí se da en el caso de activar la opción de modelización automática o si se desea utilizar el modelo en procesamientos posteriores con SEATS.

Por lo tanto, el usuario debe comprobar la adecuación de los datos objeto de estudio a este modelo de partida. Esto no suele ser un problema, pues se trata de una parametrización muy general que suele resultar apropiada para series temporales mensuales, trimestrales, etc., enmarcándose en la **metodología Box-Jenkins**. Es, por tanto, para este tipo de datos para los que el programa está especialmente pensado. En series de menor nivel de agregación temporal, como las diarias, en las cuales puede aparecer más de una estacionalidad estocástica, el esquema anterior puede no ser el más idóneo. Otro aspecto a tener en cuenta es que en todo momento estamos tratando con modelos lineales.

El modelo que el programa asume por defecto para la parte estocástica responde a la expresión:

$$(1-L)(1-L^s)v_t = (1 + \theta_1 L)(1 + \Theta_1 L^s)a_t,$$

que se suele denominar "modelo de las líneas aéreas". Este modelo es de uso frecuente en la literatura, ya que ofrece una representación univariante razonable de muchas series temporales, y en

particular de las series de actividad económica. En la opción por defecto se realiza, también, un diagnóstico sobre la media de los residuos. En caso de ser significativamente distinta de cero se incorpora una constante al modelo.

1.2. EL ENTORNO DEL PROGRAMA

El programa se estructura en torno a tres módulos principales que, si se usa la rutina de instalación por defecto de TRAMO, se corresponden a tres subdirectorios: **SERIES**, **GRAPH** y **OUTPUT**. No obstante, dos variables de entorno (**SLIB** y **SERIES**) permiten una configuración a la medida del usuario. En esta presentación se suponen siempre activas las opciones por defecto. A continuación se describe cada uno de esos bloques. Las cuestiones de formato de ambos programas es conveniente verlas más detalladamente en los manuales.

1.2.1 La Introducción de Datos y Parámetros de Control

La base del proceso es un fichero de entrada (por defecto "c:\TRAMO\SERIE") que contiene las series de datos y el valor de los distintos parámetros de control que indican al programa qué hacer y cómo hacerlo. La rutina **INPUT** permite la creación de forma simple de este fichero, aunque lo más habitual si se trabaja sobre una misma serie es modificar alguno de los ficheros ya existentes mediante su edición en cualquier editor de textos, dado que las modificaciones necesarias sobre estudios previos suelen ser pequeñas.

La versión actual de TRAMO, al igual que la de SEATS, se encuentra compilada para un máximo de 600 observaciones y un mínimo dependiente de las operaciones que exijamos del programa.

El fichero SERIE se estructura en tres partes: (1) la propia serie temporal ("n_serie") y sus valores de configuración; (2) parámetros de control (siguen el formato `&INPUT PAR1=v1, PAR2=v2, ..., /`); (3) parámetros referentes a las variables de regresión del modelo (`® ... /`). En general, el número de parámetros que el usuario debe fijar suele ser reducido, ya que en buena parte de ellos los valores por defecto resultan adecuados. No obstante, el programa no dispone de ayuda en línea, lo cual facilitaría en gran medida esta operación. Parte de los parámetros de control están relacionados con la introducción de un modelo distinto del que TRAMO considera por defecto. Aunque esta tarea resulta simple, sería más cómoda si se pudiera introducir el modelo de forma explícita, como ocurre con otros programas de series temporales.

Por último, de cara a agilizar el tiempo de interacción entre programa y usuario, sobre todo en el tratamiento masivo de series, el programa puede analizar simultáneamente varios modelos sobre

una misma serie temporal ($ITER=1$)⁶, un mismo modelo sobre varias series ($ITER=2$) o distintas series con distintos modelos ($ITER=3$).

1.2.2 Resultados del Programa

Los resultados del programa aparecen en "c:\TRAMO\OUTPUT\n_serie.out". La extensión de la información puede ser modificada a través del parámetro **OUT**. Además de los resultados finales de la identificación, estimación, diagnóstico y predicción, el fichero en su forma más extensa ($OUT=0$ (D))⁷) contiene información completa sobre el proceso de decisión del programa, con lo cual se consigue transparencia en el proceso, fundamentalmente si éste incluye alguna opción automática.

Un aspecto que conviene subrayar es el apartado de diagnóstico de los resultados. El programa presenta información sobre los t-valores de los parámetros estimados, sus correlaciones, estadísticos y contrastes de rachas sobre los residuos y los estadísticos de Ljung-Box, entre otros. Aunque su utilidad es un tema de amplia discusión en el mundo académico, se echa en falta algún tipo de contraste sobre raíces unitarias en la frecuencia cero y en las distintas frecuencias estacionales, que faciliten el juicio de las transformaciones estacionarias adoptadas.

En términos generales, puede concluirse que la salida del programa es completa y se presenta de forma clara.

1.2.3 Gráficos y Vectores Auxiliares

La rutina **GRAPH** permite la visualización gráfica de un conjunto de resultados que el programa almacena automáticamente en ficheros situados en el subdirectorio del mismo nombre. Las opciones que se presentan son:

1. Menú *Series*: Serie original y transformada, serie extrapolada de predicciones y series corregidas de atípicos, de efectos especiales (efecto calendario y Semana Santa) y de efectos de regresión en general.
2. Menú *ACF*: Correlograma y correlograma parcial de los residuos y correlograma de la serie diferenciada.
3. Menú *Regoutse*: Efecto separado y conjunto de los atípicos, variables de regresión y efectos especiales.

En definitiva, un amplio glosario de vectores auxiliares que pueden ser recuperados y

⁶ La primera vez que se cite un parámetro, aparecerá en negrilla.

⁷ (D) indica la opción por defecto del programa.

exportados para tratamientos posteriores con otros programas (hojas de cálculo, Matlab, etc.) y de gran ayuda, sobre todo, para el usuario experto. Su visualización gráfica se realiza a través de un cómodo sistema de menús. La facilidad *overlay* permite la superposición de varios gráficos, facilitando las tareas de comparación.

El programa permite la posibilidad, además, de almacenar automáticamente todo el análisis realizado sobre una serie temporal y compararlo con sucesivos estudios que se vayan haciendo.

Las prestaciones gráficas que acabamos de describir son muy notables, sobre todo si se compara con otros programas de análisis de series temporales que carecen de entornos gráficos adecuados.

1.3. PRINCIPALES PRESTACIONES

En esta sección, junto con la descripción de las principales facilidades que ofrece TRAMO, se describen los parámetros fundamentales requeridos para la especificación de una configuración distinta a la que el programa considera por defecto. En este sentido, se consideran recomendaciones de configuración dependiendo del tipo de usuario (experto o no experto) y del tipo de actividad a desarrollar (tratamiento específico o masivo de series con características heterogéneas). TRAMO ofrece otros parámetros, aparte de los descritos, que controlan aspectos que se han considerado como secundarios a efectos de esta presentación (gestión de memoria, características del entorno, etc.) o en los que la opción por defecto resulta apropiada para la mayoría de las aplicaciones que se plantean en la práctica. Por su especial relevancia, se ha reservado un apartado especial a la opción de modelización automática univariante que ofrece el programa (sección 1.4). Por otro lado, todos los aspectos relacionados con la interconexión con el programa SEATS se tratan en la sección 3.

1.3.1 Transformaciones Previas

Existe la posibilidad de un diagnóstico sobre la posible idoneidad de una transformación logarítmica de los datos (LAM). El usuario experto puede decidir sobre este aspecto con información *a priori* (LAM=0 (D), logaritmo). Sin embargo, para el usuario no experto o para el tratamiento masivo de series heterogéneas es conveniente la diagnosis automática (LAM=-1).

1.3.2 Estimación

La estimación de los parámetros del modelo se realiza mediante algoritmos de máxima verosimilitud, bien exacta (INCON=0 (D)) o aplicando mínimos cuadrados no condicionales (INCON=1). Esta última posibilidad se suele utilizar fundamentalmente cuando el modelo que se

estima es puramente autorregresivo. TRAMO permite elegir entre distintos métodos (IFILT) como el algoritmo de Morf, Sidhu y Kailath (D), el filtro de Kalman ordinario o en su versión raíz cuadrada y mínimos cuadrados condicionales. Asimismo, es posible fijar varios parámetros relacionados con los algoritmos como son valores iniciales, criterios de convergencia, etc. Con este menú de opciones se cubren adecuadamente las necesidades a este respecto.

1.3.3 Tratamiento de Observaciones Atípicas.

Uno de los aspectos destacables del programa se relaciona con la detección y corrección de distintos tipos de observaciones atípicas (*outliers*). El procedimiento es una versión mejorada del método propuesto por Chen y Liu (1993). El análisis más general se realiza sobre los cuatro tipos de atípicos considerados en la literatura: (1) innovativos; (2) aditivos; (3) cambios de nivel y (4) cambios transitorios de nivel. El parámetro que define la sensibilidad de la búsqueda (VA) puede ser fijado por el usuario. Su valor por defecto depende de la longitud de la serie.

Como en otros programas de series temporales, la detección y corrección de efectos atípicos puede hacerse de forma automática (IATIP=1) o no hacerse (IATIP=0 (D)). Si se opta por la opción automática, es posible elegir el tipo de atípicos a identificar (AIO) y el rango de la muestra en que se realiza la búsqueda (INT1 e INT2), extendiéndose por defecto a toda la muestra. Nuestra recomendación es activar la búsqueda automática exclusivamente sobre atípicos del tipo aditivo, cambio de nivel y cambio transitorio de nivel (AIO=2 (D)). Esta recomendación se basa en estudios que muestran que puede haber problemas en la identificación de atípicos innovativos en presencia de otros atípicos (por ejemplo, Balke, 1993). Por otro lado, es conveniente realizar la búsqueda sobre toda la muestra, sobre todo teniendo en cuenta que en la última versión del programa se han incorporado mejoras importantes en la identificación de atípicos en los extremos de la serie. Todas estas flexibilidades no están, en general, incorporadas en otros programas de series temporales.

El procedimiento de detección automática exige la reestimación sucesiva del modelo. Esto puede suponer un coste computacional considerable, especialmente en series largas que presenten gran número de atípicos. Por ello, el programa ofrece la posibilidad de elegir, para esta etapa, entre distintos métodos de estimación. Esto lo controla el parámetro IMVX. El manual recomienda IMVX=2 para situaciones en las que se prevee mucho cálculo. Esto puede ser especialmente importante en el tratamiento masivo de un gran número de series.

1.3.4 Predicción

Las utilidades que TRAMO ofrece para el análisis de predicción tienen dos vertientes. Por un lado, el programa permite reservar un cierto número de observaciones de la parte final de la muestra para evaluar las predicciones que se obtienen del modelo univariante que ha sido estimado

con la primera parte de ésta. Esto aporta información sobre la estabilidad postmuestral del modelo. En TRAMO este aspecto lo cubre el parámetro **NBACK**. Si $NBACK < 0$, el programa genera los errores de predicción un período hacia adelante de las últimas $|NBACK|$ observaciones y realiza un contraste de estabilidad postmuestral basado en un estadístico F .

Por otro lado, TRAMO permite generar predicciones a distintos horizontes temporales a partir de una fecha que puede ser fijada por el usuario ($NBACK \geq 0$). El número de observaciones futuras a predecir con el modelo obtenido se indica en TRAMO mediante el parámetro **NPRED**.

Si se ha aplicado a la variable de partida una transformación logarítmica, las predicciones sobre el nivel original pueden calcularse por medio de la distribución lognormal ($LOGN=1$) o aplicando la transformación exponencial sobre la predicción en logaritmos ($LOGN=0$ (D)).

La información que brindan estas utilidades es, en general, suficiente. No obstante, el analista interesado en el desempeño del modelo a horizontes de predicción más largos que un período hacia adelante puede verse limitado, al no permitir TRAMO la programación de rutinas de cálculo de errores cuadráticos medios de predicción a esos horizontes.

1.3.5 Observaciones Ausentes

Conviene destacar la importancia de contar con un conjunto de opciones automáticas de estimación de observaciones ausentes como paso previo a cualquier estudio. Este aspecto se trata en TRAMO de forma más completa que en el resto de los entornos. En primer lugar, habrá que indicar al programa los lugares que se corresponden con observaciones ausentes en la serie, señalándolos mediante el número "-99999.". Posteriormente, el problema se puede enfocar de dos formas, según el valor del parámetro **INTERP**: (a) mediante su predicción óptima dado el resto de la muestra, aplicando el filtro de Kalman (*skipping approach*; D), o (b) mediante su consideración como atípico aditivo, corrigiendo la función de verosimilitud. Esta última opción es la que se ajusta cuando la modelización se realiza de forma automática.

1.3.6 Variables de Regresión

Aparte de las facilidades de que dispone TRAMO para la introducción de variables de regresión por parte del usuario y para la creación de diversos tipos de variables de intervención, se ofrece la capacidad de diagnóstico y generación automática de determinadas variables relacionadas con la corrección del efecto Semana Santa (**IEAST**) y de la heterogeneidad del número de laborables de cada mes (**ITRAD**). Por defecto, este apartado está desactivado, por lo que en series en las que el usuario considere que estos efectos puedan tener importancia, debe modificarse esta opción. Este será el caso, por ejemplo, si se trabaja con series mensuales de actividad (ventas, producción, series de transporte, etc.).

1.4. IDENTIFICACIÓN AUTOMÁTICA DE MODELOS

El procedimiento de identificación automática de un modelo univariante constituye una de las utilidades más sobresalientes del programa, sobre todo para el usuario no experto. El proceso de identificación se hace en dos etapas:

- a) **Transformación estacionaria.** Se determinan los valores de d y D que transforman a la serie original, a través de operaciones de diferenciación, en una serie estacionaria. Para ello, en una primera aproximación se utilizan modelos autorregresivos $AR(2) \times (1)$, junto con los resultados asintóticos de Tiao y Tsay (1983) y Tsay (1984). En una segunda fase se ajusta un modelo $ARMA(1,1) \times (1,1)$, con media sobre los residuos de la transformación elegida en la aproximación anterior, con el objetivo de diagnosticar la existencia de otras posibles raíces unitarias en los polinomios autorregresivos, que hubieran podido quedar enmascaradas por la presencia de términos de medias móviles con módulos cercanos a la unidad. Por último, en la fase de estimación aún se podrán incorporar raíces si quedan suficientemente cercanas al círculo unidad. Son tres los parámetros importantes a la hora de fijar la sensibilidad de este proceso de búsqueda: (1) **UB1** en la primera fase; (2) **UB2** en la segunda y (3) **UBP** en la estimación final del modelo. Aunque los valores por defecto de estos parámetros son adecuados, un usuario experto podrá optar por manipularlos, dada la importancia de esta parte en la especificación del modelo.

Con el parámetro **IDIF** puede restringirse el espacio paramétrico de búsqueda de raíces unitarias ($IDIF=1,3,4,5$), o incluso pueden ser directamente introducidas por el usuario ($IDIF=2$). Algunas de estas opciones incluyen la posibilidad de implementar una búsqueda de raíces unitarias complejas, aparte de las estacionales y regulares.

- b) **Estructura estacionaria.** Mediante el método BIC se ajustan los valores de p , P , q y Q . El parámetro **INIC** acota la búsqueda de los valores anteriores a unos márgenes que racionalicen el proceso de identificación. Para el usuario no experto se recomienda la búsqueda en el espacio más amplio ($INIC=4$), aunque puede verse limitado por el número de observaciones de la serie. Si se pretende una interacción posterior con **SEATS**, por cuestiones de compatibilidad, es conveniente tomar $INIC=3$.

A este proceso de identificación automática se le puede incorporar cualquiera de las prestaciones que se han tratado en puntos anteriores, como detección de atípicos, reconstrucción de observaciones ausentes, tratamiento del efecto calendario, etc., de forma que es posible afirmar que **TRAMO** ofrece un proceso completo de modelización automática, con una metodología original.

Los resultados que obtiene este proceso pueden considerarse satisfactorios y con costes computacionales muy reducidos. A nivel ilustrativo, una ejecución completa del programa con todas las utilidades activadas, sobre una serie real (IPI de bienes de consumo) de 160 observaciones, utilizando un PC con procesador *PENTIUM 100*, ha tardado 1,8 segundos.

Tan sólo en algunos casos se ha detectado una cierta tendencia a la sobrediferenciación, aunque siempre en situaciones en las que no tiene consecuencias graves para el estudio. Como se sabe, en general, los efectos de una sobrediferenciación son menos graves que los que tendría una infradiferenciación.

1.5. CONCLUSIONES SOBRE TRAMO

En función de todo lo expuesto anteriormente, podemos concluir que el posible potencial que para un usuario puede tener el programa TRAMO, es desglosable en tres vertientes:

- a) Como **rutina de modelización automática**. El programa resuelve de manera muy satisfactoria este problema. Enfrentado con series tanto reales como simuladas, TRAMO ha mostrado solidez y fiabilidad, con costes computacionales muy reducidos. Este es, sin duda, un aspecto de gran utilidad, no sólo para el usuario no experto, sino que los resultados de este procedimiento automático pueden considerarse como modelo de referencia válido por parte de usuarios cualificados. Por ello, se puede sugerir la elección de TRAMO como programa idóneo para el tratamiento masivo de series, sin la necesidad de una constante supervisión de expertos.
- b) Como **herramienta útil** en muchos de los aspectos relacionados con el análisis univariante de series temporales. TRAMO ofrece un amplio conjunto de opciones e instrumentos que cubren las operaciones más frecuentemente involucradas en este tipo de estudios y suficiente para la mayor parte de los posibles usuarios. Además, el coste de aprendizaje es pequeño, la velocidad de procesamiento es muy alta, las salidas son completas, dispone de un buen entorno gráfico y permite la exportación fácil de un gran número de vectores de datos de interés para posteriores tratamientos en otros entornos. No obstante, el programa no permite la programación de rutinas o macros específicas, lo que en determinadas circunstancias puede ser una limitación.
- c) **En interconexión con el programa SEATS**. Este aspecto será tratado en la sección 3. Se puede adelantar aquí que su potencial es evidente, sobre todo cuando se pretende una automatización completa del proceso de modelización y extracción de señales de una serie temporal.

2. EL PROGRAMA SEATS

El objetivo fundamental del programa SEATS es la extracción de señales a partir de una serie temporal. Sobre este tema existen cuestiones metodológicas que es importante conocer para una mejor valoración del programa. A continuación se expone una breve introducción.

2.1. LA EXTRACCIÓN DE SEÑALES

Tradicionalmente, la extracción de señales sobre una serie temporal (z_t) se realiza sobre cuatro tipos de componentes: (a) **tendencia** (t_t), que recoge la evolución subyacente de la serie; (b) **estacionalidad** (s_t), que condensa las oscilaciones sistemáticas cuasiregulares de la serie dentro del año; (c) **ciclo** (c_t), que aglomera desviaciones sistemáticas respecto de la tendencia de la serie ajustada de estacionalidad; (d) **irregular** (i_t) o ruido residual. De este modo, la serie original podría desagregarse de la siguiente forma:

$$z_t = t_t + s_t + c_t + i_t,$$

tomando un modelo aditivo (si fuese multiplicativo llegaríamos a él mediante logaritmos). En la práctica es todavía usual considerar otras descomposiciones. La más habitual es:

$$z_t = z_{at} + s_t,$$

siendo z_{at} la **serie ajustada de estacionalidad**. Este es un dato que las agencias estadísticas a menudo proporcionan junto o en sustitución de la serie original.

De forma sintética, se pueden distinguir dos formas de implementar este tipo de descomposiciones:⁸

1) **Métodos empiricistas**

Se caracterizan por hacer una descomposición basada en unos filtros cuya estructura y parámetros no dependen de la naturaleza de los datos sino que tienen valores prefijados. Los procedimientos más representativos en economía son el filtro de Hodrick-Prescott, los métodos de alisado y, fundamentalmente, el método *X-11* o su posterior extensión el *X-11-ARIMA*,⁹ cuyo desarrollo se basó en la experiencia de J. Shiskin con miles de series en la *Bureau of Census* de Estados Unidos. Estos son aún los métodos de uso más generalizado, por cuestiones históricas y por ofrecer, en general, resultados satisfactorios. No obstante, no tienen la flexibilidad de los métodos basados en modelos que veremos a continuación y, además, carecen de una teoría estadística que los respalde. Esto hace que su uso sea cada vez más restringido.

⁸ En Espasa y Cancelo (1993) se puede encontrar un desarrollo amplio de estos procedimientos.

⁹ Actualmente existe en el mercado una nueva beta-versión: *X-12-ARIMA-REGARIMA*.

2) Métodos basados en modelos

En ellos se plantea de forma explícita un modelo para los distintos componentes, a partir de los cuales se implementa la extracción de señales. Estos procedimientos tienen en cuenta en su forma de operar las características particulares de cada serie temporal, existiendo toda una literatura que les da soporte teórico. De partida, el planteamiento de estos métodos es más fiable y satisfactorio que el de los empiricistas, y la tendencia en su uso es creciente.

Existen dos enfoques dentro de este tipo de procedimientos. Por un lado, están los **métodos basados en modelos de forma reducida** (ver, por ejemplo, Maravall, 1994). Estos métodos definen la estructura de los filtros a partir de un modelo univariante específico estimado para la serie objeto de estudio. Como principal limitación, tienen un problema de identificabilidad y existe una cierta arbitrariedad a la hora de definir los componentes. La calidad del método depende de lo adecuado que sea el modelo ajustado y de lo razonables que sean los criterios posteriores de descomposición.

Por otro lado están los **métodos basados en modelos estructurales**, ampliamente descritos en Harvey (1989). En ellos se formulan *a priori* los modelos teóricos de los componentes. Para ello se recurre a determinadas características que se supone éstos deben tener. Esta metodología está implementada en el programa STAMP. El tipo de modelo univariante que este programa puede tratar es más limitado que el que permite SEATS, a la vez que presenta problemas a la hora de aplicarlo de forma automática a gran número de series. Su uso está restringido actualmente al campo académico.

2.2. LA METODOLOGÍA SEATS

La metodología SEATS se encuadra dentro de los métodos basados en modelos de forma reducida. Para afrontar el problema de identificabilidad anteriormente mencionado recurre a la llamada **descomposición canónica**, imponiendo la restricción de que el componente irregular absorba el máximo de variabilidad (sea de máxima varianza), de forma que el resto de los componentes sea lo más estable posible (Maravall, 1985).

Por otro lado, la propia evolución del programa SEATS ha puesto de manifiesto el problema de indefinición de los componentes, fundamentalmente en lo que se refiere a la parte estocástica del modelo. Esta descomposición se basa en las raíces de la parte autorregresiva del modelo, fundamentalmente en las de módulo unidad. Para profundizar en este punto, se recomienda consultar el artículo de Maravall (1994). En las últimas versiones se ha optado por la siguiente definición de los componentes estocásticos:

1. **Tendencia** En ella se incluyen las raíces autorregresivas reales positivas de módulo mayor al valor del parámetro **RMOD** (la opción por defecto de 0.5 puede ser excesivamente baja).

- 2. Estacional** Recoge aquellas raíces autorregresivas que en el dominio frecuencial quedan dentro de un entorno (de amplitud **EPSPHI**) de las frecuencias estacionales.
- 3. Ciclo** Aglomera al resto de la estructura regular del modelo que no puede ser incluida en los componentes anteriores.
- 4. Irregular** Ruido blanco residual.

En versiones anteriores de SEATS se optaba por una definición cerrada del componente cíclico, en sintonía con su interpretación como ciclo periódico de largo plazo, asignándole las raíces autorregresivas de período superior a un año. Con esta definición determinados modelos carecían de una descomposición admisible lo cual creaba problemas computacionales y conceptuales. Esto era lo que ocurría con los denominados modelos no equilibrados o no balanceados (mayor orden en el polinomio de medias móviles que en el autorregresivo). El nuevo enfoque adoptado merece un comentario particular por su sensatez, al basar la descomposición en la definición de los dos componentes más importantes, tendencia y estacionalidad, definiendo el ciclo de forma complementaria. Este planteamiento ha aportado al método una mayor eficacia. Por ejemplo, permite una fácil descomposición de los modelos no equilibrados.

2.3. DESCRIPCIÓN DEL PROGRAMA SEATS

La forma de actuar de SEATS, en lo que se refiere a la extracción de señales, requiere los siguientes pasos (en torno a los cuales girará el significado de los distintos parámetros):

- 1.** Especificación de un modelo para la serie observada de entre la familia de modelos definida en el apartado 1.1 para TRAMO.
- 2.** Descomposición de la parte ARIMA en modelos ortogonales para cada uno de los componentes.
- 3.** Asignación de las variables determinísticas a los componentes en función de la naturaleza de los efectos modelizados.
- 4.** Estimación de los distintos componentes, basada en la metodología de extracción de señales.
- 5.** Diagnóstico de los resultados obtenidos y cálculo de variables de interés relacionadas. SEATS calcula, por ejemplo, los errores de las señales, así como las predicciones óptimas de éstas (con sus errores estándar). Muchas de estas aportaciones son originales y no se encuentran en ningún otro programa de extracción de señales, siendo de gran interés para el análisis de coyuntura y de política económica.

Por tanto, el primer paso es ajustar un modelo a los datos. Existen tres posibilidades: (1) introducir un modelo especificado por el usuario, (2) utilizar el modelo automático ajustado por TRAMO ó (3) estimar el modelo por defecto del programa SEATS, el "modelo de las líneas aéreas". En muchas de las aplicaciones este último planteamiento es suficientemente aceptable. De hecho,

tomando como modelo de partida un "modelo de líneas aéreas" con unos determinados valores para los coeficientes de medias móviles se consigue una buena aproximación a los filtros del método X-11. En general, la descomposición que implementa SEATS es bastante robusta respecto al modelo de partida.

Cuando las series tengan una modelización compleja por la existencia de observaciones ausentes, atípicos, etc., lo óptimo, especialmente para usuarios no expertos, será recurrir al programa TRAMO para obtener un modelo de partida para SEATS. Personas con experiencia en el análisis de series temporales en otros entornos también pueden introducir fácilmente sus modelos. No obstante, nuestra recomendación es recurrir siempre a TRAMO para obtener un modelo de referencia y control.

Otros aspectos relacionados con el modelo ajustado y su análisis, como la posibilidad de tomar o no logaritmos, estimación, predicción, etc., también se contemplan de forma independiente en SEATS.

Una característica sobresaliente del programa es la posibilidad de asignar las distintas variables de regresión al componente al que el usuario considere que correspondan, realizando el programa automáticamente los ajustes que la teoría impone sobre las variables para su adaptación. El parámetro que se debe definir para cada elemento a incluir es **REGEFF**, que por defecto no incluye la variable en la descomposición. Esta facilidad se muestra aún más atractiva si el fichero de entrada proviene de TRAMO, pues en este caso todas las variables de regresión introducidas automáticamente en el modelo, tales como el efecto calendario, atípicos, etc., tendrán asignados los componentes a los que corresponden sin que el usuario deba preocuparse por ello.

Aspectos tales como los referentes a la introducción del fichero de entrada, la estructuración del entorno y muchas de las facilidades, o bien son equivalentes a las ya analizadas cuando se habló de TRAMO o las consideramos secundarias para nuestros objetivos, por lo que remitimos al lector al manual.

2.4. SALIDAS DEL PROGRAMA

Sobre este aspecto de nuevo cabe decir que la salida es completa y satisface los requerimientos de cualquier usuario.

2.4.1 Gráficos y Vectores Auxiliares

La presentación y la forma de operar del programa en este apartado es en todo equivalente a la ya descrita cuando hemos tratado TRAMO. De nuevo la rutina **GRAPH** permite la cómoda visualización de los principales resultados de interés. Estos se subdividen en cinco menús que agrupan

las principales prestaciones que cubre el programa:

1. Menú *Series*: Serie original y transformada, residuos, tendencia¹⁰, estacionalidad, serie ajustada de estacionalidad, pseudo-innovaciones de los modelos parciales de los componentes, tasas de crecimiento, etc.
2. Menú *ACF*: Correlogramas de la serie estacionaria y de los residuos. Correlogramas teóricos y estimados de los componentes.
3. Menú *Espectro*: Espectro teórico de los modelos de la serie original y de los componentes y espectro estimado de la serie original y de las señales obtenidas.
4. Menú *Filtro*: Todos los filtros y pesos utilizados en la descomposición.
5. Menú *Predicción*: Predicciones de la serie original con sus bandas de confianza, serie transformada, tendencia y estacionalidad.

De nuevo, todos los resultados obtenidos, y en particular todos los correspondientes a los gráficos anteriores, se almacenan de forma ordenada, siendo posible su recuperación de forma simple para tratamientos posteriores en otros entornos. En el manual aparece descrito cada uno de estos ficheros.

2.4.2 Resultados del Programa

Aparte de la información anterior, al igual que en TRAMO, en el fichero "c:\SEATS\OUTPUT\n_serie.out" aparece toda la información fundamental relativa a la aplicación que SEATS ha ejecutado. En este caso, dada la enorme cantidad de información que SEATS puede generar y su compleja interpretación, sí que puede resultar conveniente la manipulación del parámetro **OUT**. La salida muy reducida (OUT=2) permite una rápida y fácil inspección para comprobar que tanto el modelo estimado como los componentes obtenidos son adecuados. La opción reducida (OUT=1) proporciona, básicamente, un listado de todas las series de interés, un resumen de la estimación del modelo, un estudio detallado de los residuos, los modelos de los componentes, señales extraídas y errores de revisión. Por otro lado, la opción detallada (OUT=0 (D)) amplía la anterior permitiendo, además, el seguimiento de todo el proceso que haya tenido lugar. Para el usuario no experto esta opción puede resultar excesiva.

2.5. CONCLUSIONES SOBRE SEATS

El programa SEATS se puede considerar una potente, rápida, sólida y metodológicamente novedosa herramienta que cubre perfectamente el objetivo para el que fue creado: la extracción de

¹⁰ El programa presenta tanto las señales logarítmicas que componen el modelo aditivo ("componentes"), como las señales sin transformar del modelo multiplicativo ("factores").

señales. Una idea de su rapidez la da el hecho de que la opción más completa del programa, ejecutada sobre la misma serie y ordenador que se utilizó para TRAMO, no superó los dos segundos.

Es habitual utilizar la extracción de señales como herramienta para otros estudios, por ejemplo de coyuntura. En este caso el programa permite una fácil exportación de vectores de datos para tratamientos personalizados en otros entornos.

3. INTERCONEXIÓN ENTRE TRAMO Y SEATS; TRATAMIENTO AUTOMÁTICO

Hasta ahora se ha descrito el funcionamiento independiente de TRAMO y SEATS. Sin embargo, es en su funcionamiento conjunto donde la potencialidad de los programas se revela más claramente. La interconexión entre ambos entornos es un aporte en sí mismo.

El parámetro **SEATS** en el programa TRAMO es el que se encarga de crear un fichero de salida con los resultados obtenidos por éste en el formato de entrada de SEATS, de modo que el usuario no deba interactuar en un proceso que podría resultar tedioso. Este proceso incluye cualquier variable de regresión creada o no automáticamente por TRAMO. Como ya se ha señalado, en el caso de las variables no generadas automáticamente, un determinado valor del parámetro REGEFF habrá indicado a TRAMO cómo hacer la asignación. Este fichero puente se llama "seats.itr" y deberá ser transferido al entorno adecuado para que SEATS lo pueda ejecutar. Para automatizar el proceso se puede utilizar el siguiente fichero "*.bat":

```
TRAMO
copy c:\tramo\seats.itr c:\seats\serie
cd c:\seats
SEATS
```

Una de las posibilidades más atractivas que aporta el uso conjunto de estos programas es su habilidad para llevar a cabo de forma automática el tratamiento completo y masivo de series temporales. En la implementación de esta opción interviene, fundamentalmente, el parámetro **RSA** (0(D) a 4). Según su valor se fijan automáticamente determinados valores de los parámetros de TRAMO y SEATS (en el manual aparece una descripción detallada), de forma que el usuario pueda desentenderse de todo este proceso. Con ellos, TRAMO genera automáticamente un modelo y el fichero de entrada para SEATS con los parámetros adecuados. Problemas de incompatibilidad, como por ejemplo modelos generados por TRAMO con descomposición no admisible en SEATS, son solucionados automáticamente.

Las cuatro opciones de RSA que activan el proceso automático (RSA = 1 a 4) comparten lo siguiente:

- Contraste sobre la transformación logarítmica.

- Identificación y corrección de observaciones atípicas (aditivos, transitorios y cambios de nivel).
- Identificación y estimación por máxima verosimilitud exacta del modelo ARIMA.
- Interpolación de observaciones ausentes.
- Predicción de las series.
- Descomposición del modelo; si éste no admite una descomposición admisible, se utiliza una aproximación.
- Estimación y predicción de las señales.
- Diagnóstico e inferencia asociados con la descomposición.

Dependiendo del valor del parámetro, existe la opción de contrastar el efecto Semana Santa y el efecto calendario y la posibilidad de que el proceso de identificación del modelo sea rápida y sencilla o más cuidadosa y completa.

4. COMENTARIOS Y VALORACIONES FINALES

TRAMO y SEATS son dos programas especialmente diseñados para el análisis univariante y la extracción de señales en series temporales. Las versiones actuales, aún siendo totalmente operativas, son beta-versiones, a las que puede accederse sin coste comercial alguno. En general, las utilidades que ofrecen resultan de interés para usuarios con diferentes niveles de formación en el análisis estadístico-econométrico de series temporales y con un coste de aprendizaje pequeño. No obstante, la comprensión de algunos aspectos de los manuales puede resultar difícil para un usuario no cualificado. Cabe destacar la facilidad con que se realiza la introducción de datos, la sencillez en la ejecución y la rapidez con que se efectúan los procesamientos. La disponibilidad de ayudas en línea facilitarían aún más estos procesos. Las salidas de los programas incluyen un conjunto de información suficiente sobre los resultados obtenidos. Por otra parte, las utilidades gráficas son de calidad y constituyen un valioso instrumento para el análisis y valoración de los resultados.

TRAMO ofrece un amplio abanico de instrumentos que cubren las etapas de identificación, estimación, diagnóstico y predicción con modelos ARIMA. No obstante, la imposibilidad de programación de rutinas y macros específicas hace que TRAMO no pueda considerarse como un sustituto de otros programas de series temporales disponibles en el mercado.

Entre las aportaciones fundamentales de TRAMO merece una mención especial la rutina de identificación automática de modelos. Otras herramientas, como las relacionadas con la detección y modelización de atípicos y con el tratamiento de observaciones ausentes son también eficientes y novedosas desde el punto de vista metodológico.

El programa SEATS implementa de manera eficiente la metodología de extracción de señales basada en modelos univariantes de forma reducida y debe considerarse como un valioso instrumento para el análisis de coyuntura y de política económica. Pensamos que en el futuro la tendencia en su uso será creciente en detrimento de otros procedimientos de desestacionalización como X-11 o sus variantes.

Es sin duda en la utilización conjunta de TRAMO y SEATS donde se encuentran las mayores potencialidades de estos programas, ya que su fácil interconexión permite una automatización prácticamente completa del análisis de una serie temporal. El procedimiento está pensado para aportar comodidad, eficacia y fiabilidad al proceso, por lo que no requiere la constante atención de expertos. Esta opción es ideal para agencias u organizaciones que trabajen intensivamente con series temporales, aunque cualquier usuario puede aprovecharse de sus prestaciones. El aval que respalda los buenos resultados de ambos programas es el uso que de ellos hacen agencias de la importancia de EUROSTAT.

BIBLIOGRAFÍA

- Balke, N.S. (1993), "Detecting Level Shifts in Time Series", *Journal of Business & Economic Statistics*, 11, 81-92.
- Chen, C. and Liu, L.M. (1993), "Joint Estimation of Model Parameters and Outlier Effects in Time Series", *Journal of the American Statistical Association*, 88, 284-297.
- Espasa, A. y Cancelo, J.R. (1993), *Métodos Cuantitativos para el Análisis de la Coyuntura Económica*, Alianza Economía.
- Gómez, V. and Maravall, A. (1994a), "Estimation, Prediction, and Interpolation for Nonstationary Series With the Kalman Filter", *Journal of the American Statistical Association*, 89, 611-624.
- Gómez, V. and Maravall, A. (1994b), "Program TRAMO. Time Series Regression with ARIMA Noise, Missing Observations and Outliers. Instructions for the User", *European University Institute Working Paper ECO No. 94/31*.
- Gómez, V. and Maravall, A. (1995), "Programs TRAMO and SEATS. Update: December 1995", *European University Institute Working Paper ECO No. 95/46*.
- Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Maravall, A. (1985), "On Structural Time Series Models and the Characterization of Components", *Journal of Business & Economic Statistics*, 3, 350-355.
- Maravall, A. (1994), "Unobserved Components in Economic Time Series", en Pesaran, H. H. Schmidt, P. and Wickens, M. (eds.), *The Handbook of Applied Econometrics*, vol. 1, Oxford: Basil Blackwell.
- Maravall, A. and Gómez, V. (1992), "Signal Extraction in ARIMA Time Series: Program SEATS", *European University Institute Working Paper ECO No. 92/65*.
- Tiao, G.C. and Tsay, R.S. (1983), "Consistency Properties of Least Squares Estimates of Autorregressive Parameters in Arima Models", *The Annals of Statistics*, 11, 856-871.
- Tsay, R.S. (1984), "Regression Models with Time Series Errors", *Journal of The American Statistical Association*, 79, 118-124.

WORKING PAPERS 1996

Business Economics Series

- 96-10 (01) David Camino
"The role of information and trading volume on intradaily and weekly returns patterns in the Spanish stock market"
- 96-11 (02) David Camino
"A transaction cost approach to strategic alliances in telecommunications"
- 96-13 (03) Clara Cardone
"A single european union deposit insurance scheme? An overview"
- 96-20 (04) Jaime Rivera Camino
"Reexamining the adoption of the marketing concept"
- 96-32 (05) M^a José Alvarez and Juan Ceron
"How well is Telefonica de España performing? An international comparison of telecom operators' productivity"
- 96-33 (06) Nora Lado Cousté and Jaime Rivera Camino
"Forms of market strategies in the european insurance sector"
- 96-34 (07) Nora Lado, Albert Maydeu-Olivares and Jaime Rivera Camino
"Modelling market orientation: a structural equation approach"
- 96-35 (08) Jaime Rivera Camino and Gérard Verna
"The spanish management style: an exploratory comparison with the french managers"

Economics Series

- 96-01 (01) Praveen Kujal and Roland Michelitsch
"Market power, inelastic elasticity of demand, and terms of trade"
- 96-02 (02) Emmanuel Petrakis and Minas Vlassis
"Endogenous wage-bargaining institutions in oligopolistic industries"
- 96-03 (03) Coral del Río and Javier Ruiz-Castillo
"Intermediate inequality and welfare. The case of Spain, 1980-81 to 1990-91"
- 96-04 (04) Javier Ruiz-Castillo
"A simplified model for social welfare analysis. An application to Spain, 1973-74 to 1980-81"
- 96-08 (05) Juan José Ganuza
"Optimal procurement mechanism with observable quality"