

# Estado del arte en predicción de series temporales

J.J. Asensio, P.A. Castillo

Departamento de Arquitectura y Tecnología de Computadores  
ETSIIT, CITIC-UGR  
Universidad de Granada, España  
{asensio, pacv}@ugr.es

**Resumen** bonito abstract test

## 1. Introducción

Una serie temporal es una secuencia de observaciones dispuestas a intervalos regulares de aparición. Existen multitud de ejemplos donde se registran datos de esta forma, por citar algunos: precios de acciones en bolsa, temperatura mínima y máxima de cada día, accidentes domésticos... etc.

Aquí se destacan los trabajos de investigación publicados en revistas patrocinadas por el Instituto Internacional de Predictores (IIF), aunque también se cubren publicaciones clave de otras revistas. Se ofrece así una guía selectiva para la literatura en predicción de series temporales, cubriendo el periodo 1982-2005 y resumiendo cerca de 940 artículos que incluyen aproximadamente 340 artículos del IIF. También se revisan artículos y libros clave publicados en otros sitios que han sido altamente influyentes para varios desarrollos en este campo. Los trabajos referidos comprenden 380 artículos de revista y 20 libros y monografías.

Los artículos se han clasificado primero de acuerdo a los modelos introducidos en la literatura de series temporales, en lugar de hacerlo respecto a los métodos. Por ejemplo, los métodos bayesianos en general pueden aplicarse a todos los modelos. Los artículos que no conciernen con un modelo particular fueron entonces clasificados atendiendo a los diferentes problemas que abordan (medidas de precisión, combinación).

Puede haber artículos que no aparezcan o que sean citados por otros autores, aunque esto no significa que no sean importantes. Esta revisión es un breve recorrido histórico de los principales desarrollos.

## 2. Alisado exponencial

### 2.1. Preámbulo

En 1981, los métodos de alisado exponencial a menudo eran considerados como un conjunto de técnicas especiales para extrapolar varios tipos de series

temporales univariantes. Si bien los métodos de alisado exponencial eran ampliamente utilizados en la industria y los negocios, habían recibido poca atención por parte de los estadísticos y no tenían un fundamento estadístico bien desarrollado. Estos métodos se originaron en las décadas 50 y 60 con el trabajo de Brown (1959, 1963) [1] [2], Holt (1957, reimpreso en 2004)[3], y Winters (1960) [4]. Pegels (1969) [5] proporcionó una clasificación sencilla pero útil de la tendencia y los patrones estacionales dependiendo de si eran aditivos (lineales) o multiplicativos (no lineales).

Muth (1960) [6] fue el primero en sugerir una base estadística para el alisado exponencial simple (SES) demostrando que proporcionaba la predicción óptima para el camino aleatorio con ruido. Los siguientes pasos para poner en un marco estadístico el alisado exponencial fueron proporcionados por Box y Jenkins (1970), Roberts (1982)[7], y Abraham y Ledolter (1983, 1986) [8][9], quienes mostraron que algunas predicciones de alisado exponencial lineales resultan como casos especiales de los modelos ARIMA. Sin embargo, estos resultados no se extienden a cualquier método de alisado exponencial no lineal.

Los métodos de alisado exponencial recibieron un impulso debido a dos artículos publicados en 1985, los cuales establecieron la base para gran parte del trabajo posterior en el área. En primer lugar, Gardner (1985) [10] realizó una revisión exhaustiva y una síntesis del trabajo en alisado exponencial hasta la fecha y extendió la clasificación de Pegels para incluir tendencia amortiguada. Este artículo reunió gran cantidad del trabajo existente lo que estimuló el uso de estos métodos y promovió más trabajos de investigación al respecto. Posteriormente ese mismo año, Snyder (1985) [11] demostró que SES podría ser considerado como una forma resultante de una innovación del modelo de espacio de estados (modelo con una única fuente de error). Aunque esta revelación fue en gran parte inadvertida entonces, en años recientes ha proporcionado la base para una gran cantidad de trabajos en modelos de espacio de estados en los que subyacen los métodos de alisado exponencial.

La mayor parte del trabajo desde 1980 ha incluido el estudio de las propiedades empíricas de los métodos (Bartolomei & Sweet, 1989 [12]; Makridakis & Hibon, 1991 [13]), propuestas para nuevos métodos de estimación o inicialización (McClain, 1988 [14]; Sweet & Wilson, 1988 [15]), o ha estado relacionado con modelos estadísticos en los que se considera que subyacen los métodos (McKenzie, 1984 [16]).

Los métodos de alisado multiplicativo de Taylor (2003)[17] proporcionan los únicos métodos de alisado exponencial genuinamente nuevos. durante este periodo. Por supuesto, hay habido numerosos estudios que aplican métodos de alisado exponencial en varios contextos incluyendo componentes de computadores (Gardner, 1993 [18]) , pasajeros de vuelo (Grubb & Masa, 2001 [19]), y planificación de producción (Miller & Liberatore, 1993 [20]).

La taxonomía de Hyndman, Koehler, Snyder, y Grose (2002 [21] extendida por Taylor, 2003 [17]) proporciona una categorización útil para describir la variedad de métodos. Cada método se componen de una de cinco tipos de tendencia (ninguna, aditiva, aditiva alisada, multiplicativa y multiplicativa alisada) y uno

de tres tipos de estacionalidad (ninguna, aditiva y multiplicativa). Por tanto, hay 15 métodos diferentes, donde entre los más conocidos está SES (sin tendencia ni estacionalidad), el método lineal de Holt (tendencia aditiva, sin estacionalidad), el método aditivo de Holt-Winters (tendencia aditiva y estacionalidad aditiva), y el método no lineal de Holt-Winters (tendencia aditiva y estacionalidad multiplicativa).

## 2.2. Variantes

Se han propuesto numerosas variaciones de los métodos originales. Por ejemplo, Carreno y Madinaveitia (1990) [22] y Williams y Miller (1999) [23] propusieron modificaciones para tratar discontinuidades, y Rosas y Guerrero (1994) [24] consideraron predicciones con alisado exponencial sujetas a una o más restricciones. También hay variantes por cómo y cuándo las componentes estacionales deberían ser normalizadas. Lawton (1998) [25] abogó por una renormalización de los índices estacionales para cada periodo de tiempo, puesto que elimina sesgo en las estimaciones de los componentes estacionales y de nivel. Roberts (1982) [7] y McKenzie (1986) [26] dieron esquemas ligeramente diferentes de normalización. Archibald y Koehler (2003) [27] desarrollaron nuevas ecuaciones de normalización que eran más sencillas de usar dando el mismo grado de predicción que los métodos originales. Una variante útil a mitad de camino entre SES y el método de Holt, es SES con desviación. Es equivalente al método de Holt con el parámetro de tendencia a cero. Hyndman y Billah (2003) [28] mostraron que este método también era equivalente al método Theta de Assimakopoulos y Nikolopoulos (2000) [29], cuando el parámetro de desviación es igual a la mitad de la pendiente de una tendencia lineal ajustada a los datos. El método Theta dio un rendimiento extremadamente bueno en la competición M3, aunque el por qué esta particular elección de modelo y parámetros era buena todavía no se ha determinado. Ha habido extraordinariamente poco trabajo en desarrollar versiones multivariadas de los métodos de alisado exponencial para predicción. Una excepción notable es Pfeiffermann y Allon (1989) [30] quienes consideraron datos de turismo israelí. El SES Multivariado se usa para gráficos de control de procesos (Pan, 2005) [31], donde es llamado “medias móviles multivariadas ponderadas exponencialmente”, pero aquí la cuestión no es predicción.

## 2.3. Modelos de espacio de estados

Ord, Koehler y Snyder (1997) [32] a partir del trabajo de Snyder (1985)[11] propusieron una clase de modelos de espacio de estado de innovaciones en la que se puede considerar que subyacen algunos métodos de alisado exponencial. Hyndman et al (2002) [21] y Taylor (2003) [17] amplían esto para incluir los 15 métodos de alisado exponencial. De hecho, Hyndman et al (2002) [21] propusieron dos modelos de espacio de estado para cada método, correspondientes a los casos de error aditivo y multiplicativo. Estos modelos no son los únicos, y otros modelos relacionados para los métodos de alisado exponencial se presentaron en Koehler, Snyder y Ord (2001) [33] y Chatfield, Koehler, Ord y Snyder (2001)

[34]. Desde hace tiempo se conoce que algunos modelos ARIMA dan predicciones equivalentes a los métodos lineales de alisado exponencial. La importancia del trabajo reciente en modelos de espacio de estado de innovaciones es que los métodos de alisado exponencial no lineales también pueden derivarse de modelos estadísticos.

#### **2.4. Selección del modelo**

Gardner y McKenzie (1988) [35] proporcionaron algunas reglas sencillas basadas en las varianzas de series temporales diferenciadas para elegir un método de alisado exponencial apropiado. Tashman y Kruk (1996) [36] compararon estos resultados con otros propuestos por Collopy y Armstrong (1992) [37] y un enfoque basado en el criterio de información bayesiano (BIC). Hyndman et al. (2002) [21] también propuso un enfoque basado en criterio de información, pero usando los modelos de espacio de estado subyacentes.

#### **2.5. Robustez**

El buen rendimiento de la predicción de los métodos de alisado exponencial ha sido estudiado por varios autores. Satchell y Timmermann (1995) [38] y Chatfield et al. (2001) [34] mostraron que SES es óptimo para un rango amplio de procesos generadores de datos. En un pequeño estudio de simulación, Hyndman (2001) mostró que SES rendía mejor los modelos ARIMA de primer orden porque no está tan sujeto a problemas de selección de modelo, especialmente cuando los datos no siguen una distribución normal.

#### **2.6. Intervalos de predicción**

Una de las críticas a los métodos de alisado exponencial a principios de los 80 era que no había forma de producir intervalos de predicción para las predicciones. La primera aproximación analítica a este problema fue asumir que las series eran generadas mediante funciones deterministas del tiempo más ruido blanco (Brown, 1963 [2]; Gardner, 1985 [10]; McKenzie, 1986 [26]; Sweet, 1985 [39]). En tal caso, sería mejor usar un modelo de regresión en lugar de los métodos de alisado exponencial. Newbold y Bos (1989) [40] criticaron fuertemente cualquier enfoque basado en esta hipótesis. Otros autores buscaron intervalos de predicción mediante la equivalencia entre los métodos de alisado exponencial y los modelos estadísticos. Johnston y Harrison (1986) [41] establecieron variantes de predicción para los métodos SES, y los métodos de Holt de alisado exponencial para modelos de espacio de estado con múltiples fuentes de error. Yar y Chatfield (1990) [42] obtuvieron intervalos de predicción para el método aditivo de Holt-Winters derivando el equivalente método ARIMA subyacente. Chatfield y Yar (1991) [43] discutieron intervalos aproximados de predicción para el método multiplicativo de Holt-Winters, asumiendo que los errores en las predicciones a un paso fueran independientes. Koehler et al. (2001) [33] también

derivó una fórmula aproximada para la previsión de varianza para el método multiplicativo de Holt-Winters, distinguiéndose de Chatfield y Yar (1991) [43] solamente en cómo la desviación estándar del error en la predicción a un paso era estimada. Ord et al (1997) [32] y Hyndman et al. (2002) [21] usaron el modelo de espacio de innovación subyacente para simular rutas de muestras futuras, y de este modo obtuvieron intervalos de predicción para todos los métodos de alisado exponencial. Hyndman, Koehler, Ord y Snyder (2005) [44] usaron modelos de espacio de estado para derivar intervalos de predicción analíticos para 15 de los 30 métodos, incluyendo todos los métodos usados habitualmente. Proporcionaron el enfoque algebraico más comprensible hasta la fecha para manejar el problema de la distribución de predicción para la mayoría de los métodos de alisado exponencial.

## 2.7. Espacio de parámetros y propiedades del modelo

Una práctica habitual es restringir los parámetros de alisado al rango de 0 a 1. Sin embargo, ahora que los modelos estadísticos subyacentes están disponibles, el espacio natural (invertible) de parámetros se puede usar en su lugar. Archibald (1990) [45] mostró que es posible para parámetros de alisado dentro de los intervalos usuales producir modelos no invertibles. Consecuentemente, al predecir, el impacto del cambio en los valores pasados de la serie no es despreciable. Intuitivamente, tales parámetros producen predicciones pobres y deteriora el rendimiento de la predicción. Lawton (1998) [25] también discutió este problema.

## 3. Modelos ARIMA

### 3.1. Preámbulo

Los primeros intentos de estudiar series temporales, particularmente en el siglo 19, estaban generalmente caracterizados por la idea de un mundo determinista. Fue la mayor contribución de Yule (1927) la que impulsó la noción de aleatoriedad en series temporales postulando que cada serie temporal puede ser considerada como la realización de un proceso estocástico. Basado en esa sencilla idea, se desarrolló desde entonces un gran número de métodos de series temporales. Slutsky, Walker, Yaglom y Yule formularon el concepto de los modelos autorregresivos (AR) y media móvil (MA). El teorema de la descomposición de Wold condujo a la formulación y solución del problema de predicción lineal de Kolmogorov (1941). Desde entonces, una gran cantidad de literatura ha aparecido en el área de series temporales, tratando estimación de parámetros, identificación, comprobación de modelos y predicción, véase Newbold (1983) para un estudio anterior.

La publicación *Time Series Analysis: Forecasting and Control* de Box y Jenkins (1970) integró el conocimiento existente. Además, estos autores desarrollaron una ciclo coherente, versátil e iterativo en tres etapas para la identificación, estimación y verificación de series temporales (conocido como el enfoque Box-Jenkins). El libro ha tenido un enorme impacto en la teoría y práctica moderna

de análisis y predicción de series temporales. Con la llegada de los ordenadores, se popularizó el uso de los modelos ARIMA y se extendió a muchas áreas de la ciencia. De hecho, la predicción de series temporales discretas mediante modelos ARIMA univariantes, modelos de función de transferencia (regresión dinámica), y modelos ARIMA multivariados han generado bastantes artículos en IJF. A menudo estudios de naturaleza empírica, que usan uno o varios métodos/modelos como punto de referencia para comparar.

(incluir tabla)

### 3.2. Univariado

El éxito de la metodología de Box-Jenkins se basa en el hecho de que los modelos pueden imitar el comportamiento de los diversos tipos de series y sin requerir muchos parámetros a estimar en la elección final del modelo. Sin embargo, a mediados de los años sesenta, la selección de un modelo era en gran medida una cuestión subjetiva del investigador, no había ningún algoritmo para especificar un modelo único. Desde entonces, muchas técnicas y métodos han surgido para añadir rigor matemático en el proceso de búsqueda de un modelo ARMA, incluyendo el criterio de información Akaike (AIC), el error de predicción final de Akaike (FPE) y el criterio de información bayesiano (BIC). A menudo, estos criterios se reducen a la minimización de errores de predicción a un paso (en la muestra), con un término de penalización para el sobreajuste. FPE también ha sido generalizado para predicción a varios pasos (véase, por ejemplo, Bhansali, 1996 [46] y Bhansali, 1999 [47]), pero esta generalización no ha sido utilizada en el área aplicada. Este también parece ser el caso de criterios basados en principios de validación cruzada y validación por división muestral (véase, por ejemplo, West, 1996), haciendo uso de errores de predicción fuera-de-muestra; véase Peña y Sánchez (2005) para un enfoque relacionado a considerar.

Existen diferentes métodos para estimar los parámetros de un modelo ARMA (véase Box et al., 1994 [48]). Aunque estos métodos son asintóticamente equivalentes, en el sentido de que las estimaciones tienden a la misma distribución normal, hay grandes diferencias en las propiedades con muestra finitas. En un estudio comparativo de los paquetes de software, Newbold, Agiakloglou y Miller (1994) [?] mostraron que esta diferencia puede ser muy sustancial y, como consecuencia, puede influir en las predicciones. En consecuencia, recomendaron el uso de máxima verosimilitud completa. El efecto de que los errores de estimación para los parámetros limitan las predicciones también fue algo observado por Zellner (1971). Él utilizó un análisis bayesiano y derivó la distribución predicha de observaciones futuras tratando los parámetros del modelo ARIMA como variables aleatorias. Más recientemente, Kim (2003) [?] consideró una predicción y estimación de los parámetros de modelos AR con pequeñas muestras. Encontró que las estimaciones de parámetros con sesgo corregido producían predicciones más precisas que las de mínimos cuadrados. Landsman y Damodaran (1989) [?] presentaron la evidencia de que la estimación de parámetros de ARIMA James-Stein mejoraba la precisión de predicción con respecto a otros métodos, bajo el criterio de MSE.

Si una serie temporal se sabe que sigue un modelo ARIMA univariado, las predicciones que usan observaciones desagregadas son, en términos de MSE, al menos tan buenas como aquellas que usan observaciones agregadas. Sin embargo, en aplicaciones prácticas, existen otros factores a considerar, como valores perdidos en series desagregadas. Ledolter (1989) [?] y Hotta (1993) analizaron el efecto producido de un outlier aditivo en el intervalo de predicción cuando se estiman los parámetros del modelo ARIMA. Cuando el modelo es estacionario, Hotta y Cardoso Neto (1993) mostraron que la pérdida de eficiencia usando datos desagregados no era alta, incluso si el modelo no es conocido. Por tanto, la predicción podría hacerse tanto con modelos desagregados como agregados.

El problema de añadir información externa (prior) en las predicciones de ARIMA univariado fue considerado por Cholette (1982) [49], Guerrero (1991) [50] y de Alba (1993) [51]).

Como alternativa a la metodología de ARIMA univariado, Parzen (1982) [?] propuso la metodología ARARMA. La idea clave es que una serie temporal es transformada de un filtro AR de memoria larga, a un filtro de memoria corta, evitando así un operador de diferenciación más duro. Además, se usa una fase de identificación diferente al convencional de Box-Jenkins. En la competición M (Makridakis et al., 1982 [?]), los modelos ARARMA consiguieron el MAPE más bajo para horizontes de predicción mayores. Por lo tanto, es sorprendente encontrar que, a parte de artículo de Meade y Smith (1985) [?], la metodología ARARMA todavía no haya despegado en el área aplicada. Su valor definitivo, quizá pueda ser juzgado mediante el estudio de Meade (2000), quien comparó la capacidad de predicción de un método ARARMA automático y no automático.

El modelado automático de ARIMA univariado, ha mostrado una capacidad de predicción a un paso tan precisa como la de otros métodos competentes (Hill y Fildes, 1984, Liber, 1984, Poulos et al., 1987 y Texter y Ord, 1989). Varios proveedores de software han implementado métodos de predicción automáticos de series temporales (incluso métodos multivariados); véase Geriner y Ord (1991) [52], Tashman y Leach (1991) [?], y Tashman (2000) [?]. A menudo estos métodos actúan como caja negra. La tecnología de sistemas expertos (Mélard & Pasteels, 2000) se puede usar para evitar este problema. Algunas directrices en la elección de un método de predicción automático se da en Chatfield (1988) [53].

Mejor que adoptar un modelo AR para todos los horizontes de predicción, Kang (2003) investigó empíricamente el caso de seleccionar un modelo AR de forma diferenciada para cada horizonte en predicciones multi-paso. El rendimiento de predicción del procedimiento multi-paso parece depender entre otras cosas del criterio de selección de orden óptimo, los periodos de predicción, los horizontes de predicción y la serie a predecir.

### 3.3. Función de transferencia

La identificación de modelos de función de transferencia puede ser difícil cuando hay más de una variable de entrada. Edlund (1984) [54] presentó un método en dos pasos para identificación de una función respuesta impulso cuando

varias variables de entrada están correladas. Koreisha (1983) [?] estableció varias relaciones entre funciones de transferencia, implicaciones causales y especificación de modelo econométrico. Gupta (1987) [55] identificó los principales problemas en las comprobaciones de causalidad. Usando análisis de componentes principales, del Moral y Valderrama (1997) [?] sugirieron una representación parsimoniosa de modelo de función de transferencia. Krishnamur, Narayan y Raj (1989) [?] mostraron cómo estimaciones más precisas del impacto de las intervenciones en modelos de función de transferencia se pueden obtener usando una variable de control.

### 3.4. Multivariado

El modelo de vector ARIMA (VARIMA) es una generalización multivariada del modelo univariado ARIMA. Las características de población de procesos VARMA parecen haber sido derivados primero por Quenouille (1957) [?], aunque el software para su implementación llegó a estar disponible en los 80 y 90. Puesto que los modelos VARIMA pueden acomodar supuestos de exogeneidad y de relaciones de contemporaneidad, ofrecen nuevos retos a los encargados de predecir y elaborar de políticas. Riise y Tjostheim (1984) [?] mostraron como los filtros de suavizado pueden construirse dentro de los modelos VARMA. El alisado previene fluctuaciones irregulares en series temporales explicativas al migrar las predicciones de la serie dependiente. Para determinar el máximo horizonte de predicción para procesos VARMA, DeGooijer y Klein (1991) [?] establecieron las propiedades teóricas de las predicciones y el error de predicción acumulado a varios pasos. Lütkepohl (1986) [?] estudió los efectos de la agregación temporal y el muestreo sistemático en predicción, suponiendo que la variable desagregada (estacionaria) sigue un proceso VARMA de orden desconocido. Posteriormente, Bidarkota (1998) [?] consideró el mismo problema pero con las variables observadas integradas en lugar de estacionarias.

Los vectores de auto-regresión (VAR) constituyen un caso especial de una clase más general de modelos VARMA. En esencia, un modelo VAR es una aproximación flexible a la forma reducida de una variedad amplia de modelos econométricos dinámicos. Los modelos VAR se pueden especificar de varias formas. Funke (1990) [?] presentó 5 especificaciones diferentes de VAR y comparó sus rendimientos de predicción usando series mensuales de producción industrial. Dhrymes y Thomakos (1989) [?] disertó cuestiones relacionadas con la identificación de VARs estructurales. Hafer y Sheedhan (1989) [?] mostraron el efecto de cambios en la estructura del modelo en las predicciones de VAR. Ariño y Franses (2000) dieron expresiones explícitas para predicciones VAR por niveles; véase también Wieringa y Horváth (2005). Hansson, Jansson y Löf (2005) usaron un modelo de factor dinámico como punto de partida para obtener predicciones a partir de VARs parametrizados parsimoniosamente.

En general, los modelos VAR tienden a sufrir sobreajuste con demasiados parámetros libres y no significativos. Como resultado, estos modelos pueden dar pobres predicciones fuera de muestra, incluso aunque el ajuste en las muestras



sea bueno; véase Liu, Gerlow y Irwin (1994) [?] y Simkins (1995). En lugar de restringir algunos parámetros como habitualmente se hace, Litterman (1986) y otros impusieron una distribución preferente en los parámetros, expresando la idea de que muchas variables económicas se comportan como un camino aleatorio. Los modelos BVAR se han usado principalmente para predicción macroeconómica (Artis y Zhang, 1990 [?], Ashley, 1988 [?], Holden & Broomhead, 1990 and Kunst y Neusser, 1986), para predicción de cuotas de mercado (Ribeiro Ramos, 2003), predicción del mercado laboral (LeSage y Magura, 1991 [?]), para proyección de negocios (Spencer, 1993 [?]) o para predicción de economía local (LeSage, 1989 [?]). Kling y Bessler (1985) compararon predicciones fuera de muestra de varios métodos para series temporales multivariadas bien conocidos de la época, incluyendo el modelo BVAR de Litterman.

El concepto de cointegración de Engle y Granger (1987) [?] ha generado varias cuestiones interesantes en la capacidad de predicción de los modelos de corrección de error (ECMs) frente a VARs y BVARs no restringidos. Shoesmith (1992 [?], 1995 [?]), Tegene y Kuchler (1994) [?] y Wang y Bessler (2004) [?] dieron evidencias empíricas para sugerir que los ECMs mejoran el rendimiento de VARs por niveles, particularmente en horizontes de predicción mayores. Shoesmith (1995) [?], y posteriormente Villani (2001) [?], también mostraron cómo el enfoque bayesiano de Litterman (1986) [?] podía mejorar la predicción con VARs cointegrados. Reimers (1997) [?] estudió el rendimiento de predicción de procesos de series temporales vector cointegradas estacionalmente usando un ECM. Poskitt (2003) disertó sobre la especificación de sistemas VARMA cointegrados. Chevillon y Hendry (2005) analizó la relación entre estimaciones multi-paso directas de VARs estacionarios frente a no estacionarios y respectivas precisiones de predicción.

## 4. Estacionalidad

El enfoque más antiguo para controlar la estacionalidad en series temporales es extraerla usando un procedimiento de descomposición como el método X-11. En los últimos 25 años este método y sus variantes (incluyendo versiones más reciente, X-12-ARIMA, Findley, Monsell, Bell, Otto y Chen, 1998 [?], se han estudiado de forma extensa.

Una línea de investigación considera el efecto de usar predicción como parte del método de descomposición estacional. Por ejemplo Dagum (1982) [?] y Huot, Chiu y Higginson (1986) [?] observaron el uso de predicción en X-11-ARIMA para reducir el tamaño de las revisiones en el ajuste estacional de los datos, y Pfeiffermann, Morry y Wong (1995) exploraron el efecto de las predicciones en la varianza de los valores ajustados de tendencia y estacionalidad.

Quenneville, Ladiray y Lefrançois (2003) con otro enfoque observaron las predicciones implícitas mediante los filtros de medias móviles asimétricos del método x-11 y sus variantes.

Un tercer enfoque fue ver la efectividad de predicción usando datos ajustados estacionalmente a partir de un método de descomposición estacional. Milller y

Williams, 2003 [?] y 2004 [?] mostraron que se obtenía la mayor precisión de predicción cuando se encogía la componente estacional hacia cero. Los comentarios sobre este último artículo dieron sugerencias relacionadas con la implementación de esta idea.

Además de trabajar en el método X-11 y sus variantes, también se han desarrollado varios métodos nuevos para el ajuste estacional, entre los que destaca el enfoque basado en modelo de TRAMO-SEATS (Gómez y Maravall, 2001 [?] y Kaiser y Maravall 2005 [?]) y el método no paramétrico STL (Cleveland, Cleveland, McRae y Terpenning, 1990 [?]). Otra propuesta ha sido el uso de modelos sinusoidales (Simmons, 1990).

En la predicción de varias series similares, Withycombe (1989) [?] mostró que puede ser más eficiente estimar una componente estacional combinada para el grupo de series, en lugar de usar sus patrones individuales. Bunn y Vassilopoulos (1993) [?] demostró cómo usar agrupamiento para formar grupos apropiados para esta situación, y Bunn y Vassilopoulos (1999)[?] introdujo algunos estimadores mejorados para los índices de grupo estacional.

A principio de los 80, las pruebas de raíz unitarias acababan de ser inventadas y las pruebas de raíz unitaria estacional estaban por llegar en breve. Posteriormente hubo un trabajo considerable en el uso e implementación de las pruebas de raíz unitaria estacional incluyendo Hylleberg y Pagan (1997) [?], Taylor (1997)[?] y Franses y Koehler (1998) [?]. Paap, Franses y Hoek (1997) [?] y Clements y Hendry (1997) [?] estudió el rendimiento de predicción de modelos con pruebas de raíz especialmente en el contexto de los cambios de nivel.

Algunos autores han advertido contra el uso extendido de modelos de raíz unitaria estacional estándar para series temporales económicas. Osborn (1990) [?] argumentó que las componentes estacionales deterministas son más habituales en series económicas que en estacionalidad estadística. Franses y Romijn (1993) [?] sugirió que las raíces estacionales en modelos periódicos resultan en mejores predicciones. Los modelos de series temporales periódicas fueron también exploradas por Wells (1997) [?], Herwartz (1997) [?] y Novales y de Fruto (1997) [?], los cuales encontraron que los modelos periódicos pueden conducir a una mejora en el rendimiento de la predicción comparados con modelos no periódicos bajo algunas condiciones. La predicción de procesos ARMA periódicos multivariados fue considerada por Ullah (1993) [?].

Varios artículos han comparado empíricamente los modelos estacionales. Chen (1997) exploró la robustez de un modelo estructural, un modelo de regresión con estacionalidad, un modelo ARIMA, y el método de Holt-Winters, y encontró que los dos últimos daban predicciones relativamente robustas cuando el modelo no estaba completamente especificado. Noakes, McLeod, y Hipel (1985) [?], Albertson y Aylen (1996) [?], Kulendran y King (1997) [?] y Franses y van Dijk (2005) compararon todos el rendimiento de predicción de varios modelos estacionales aplicados a datos reales. El modelo con mejor predicción variaba a lo largo del estudio, dependiendo de qué modelo fuera usado y la naturaleza de los datos. Parece no haber consenso todavía en qué condiciones se puede preferir un modelo u otro.

## 5. Espacio de estado y modelos estructurales y el filtro de Kalman

A principios de 1980, los modelos de espacio de estado acababan de empezar a usarse por estadísticos para predecir series temporales, aunque las ideas ya habían estado presentes en la literatura de ingeniería desde el trabajo innovador de Kalman (1960) [?]. Los modelos de espacio de estado proporcionan un marco unificado en el que se pueda escribir cualquier modelo lineal de serie temporal. La contribución clave de predicción de Kalman (1960) fue dar un algoritmo recursivo (conocido como el filtro de Kalman) para computar las predicciones. Los estadísticos empezaron interesarse por los modelos de espacio de estado cuando Schweppe (1965) [?] mostró que el filtro de Kalman proporciona un método eficiente para computar los errores de la predicción a un paso y su varianza asociada necesaria para producir la función de probabilidad. Shumway y Stoffer (1982) combinaron el algoritmo EM con el filtro de Kalman para dar un enfoque general para predecir series temporales usando modelos de espacio de estado, incluyendo la posibilidad de observaciones inexistentes.

Una clase particular de modelos de espacio de estado, conocidos como *modelos lineales dinámicos* (DLM) fue introducida por Harrison y Stevens (1976) [?], quienes también propusieron un enfoque bayesiano para la estimación. Fildes (1983) [?] comparó las predicciones obtenidas usando el método de Harrison y Stevens con métodos más sencillos como el alisado exponencial concluyendo que la complejidad adicional no conducía a una mejora del rendimiento de predicción. El enfoque de modelado y estimación de Harrison y Stevens fue mejorado por (West, Harrison y Migon (1985) [?] y West y Harrison (1989) [?]. Harvey, 1984 [?] y 1989 [?] amplió esta clase de modelos y siguió un enfoque no bayesiano para la estimación. También renombró los modelos como *modelos estructurales*, aunque posteriormente usaría el término *modelos de componentes no observadas*. Harvey (2006) proporciona una revisión comprensible y una introducción para esta clase de modelos incluyendo variaciones no gaussianas y en tiempo continuo.

Estos modelos poseen muchas similitudes con los métodos de alisado exponencial, pero tienen múltiples fuentes de error aleatorio. En particular, el *modelo estructural básico* (BSM) es similar al método de Holt-Winters para datos estacionales e incluye componentes de nivel, tendencia y estacionalidad.

Ray (1989) [?] discutió la tasa de convergencia para el crecimiento lineal del modelo estructural y mostró que los estados iniciales (elegidos de forma subjetiva normalmente) tenían un impacto no despreciable en las predicciones. Harvey y Snyder (1990) [?] propusieron algunos modelos estructurales de tiempo continuo para predecir el tiempo de espera demandado para control de inventario. Proietti (2000)[?] discutió variaciones del BSM, comparando sus propiedades y evaluando las predicciones resultantes.

Los modelos estructurales no gaussianos, han sido objeto de numerosos artículos, empezando por el modelo de potencia estable de Smith (1979) [?] con posterior desarrollo por West et al. (1985) [?]. Por ejemplo, estos modelos fueron aplicados a la predicción de series temporales de dimensiones continuas por

Grundwald, Raftery y guttorp (1993) [?] y para recuentos por Harvey y Fernandes (1989)[?]. Sin embargo, Grundwald, Hamza y Hyndman (1997) mostraron que la mayoría de los modelos habitualmente usados tienen una sustancial imperfección en la trayectoria de muestras convergiendo a una constante cuando el espacio de muestras es menor que la línea real completa, haciéndolas inadecuadas para predicción de intervalos.

Otra clase de modelos de espacio de estado, conocida como *modelos balanceados de espacio de estado*, se ha usado principalmente para predicción de series temporales macro-económicas. Mittnik (1990) proporcionó un estudio de esta clase de modelos, y Vinod y Basu (1995) obtuvo predicciones para el consumo, ingresos y tasas de interés usando estos modelos. Estos sólo tienen una única fuente de error aleatorio y generalizan otros modelos como ARMAX, ARMA y modelos con retardo racional distribuido. Una clase de espacio de estados relacionada son los modelos de *fuentes únicas de error* en los que subyacen los métodos de alisado exponencial.

A parte de estos desarrollos metodológicos, han habido varios artículos proponiendo modelos innovadores de espacio de estados para resolver problemas prácticos de predicción. Entre ellos, Coomes (1992) quien usó un modelo para predecir empleos según la industria para regiones locales y Patterson (1995) quien usó un enfoque de espacio de estado para predicción de ingresos personales disponibles.

Los libros de Harvey (1989), West y Harrison (1989) y Durbin y Koopman (2001) han tenido un impacto sustancial en la literatura de series temporales en cuanto a investigación de modelos de espacio de estado, filtrado de Kalman y modelos estructurales discretos / continuos en tiempo.

## 6. Modelos no lineales

### 6.1. Preámbulo

Comparado con el estudio de series temporales lineales, el desarrollo de análisis y predicción no lineal de series temporales todavía está en su infancia. El comienzo se atribuye a Volterra (1930). Mostró que cualquier función continua no lineal en  $t$  podría ser aproximada por una serie finita de Volterra. Wiener (1958) se interesó en las ideas de representación funcional de series y desarrolló el material existente. Aunque las propiedades probabilistas de estos modelos se habían estudiado ampliamente, los problemas de estimación de parámetros, ajuste del modelo y predicción fueron descuidados durante un largo tiempo. Esto puede atribuirse a la complejidad del modelo propuesto por Wiener y sus formas simplificadas como el modelo bilineal (Poskitt y Tremayne, 1986). En su día, el ajuste de estos modelos conllevaba dificultades de cómputo insuperables.

Aunque la linealidad es un supuesto útil y una herramienta potente en muchas áreas, a finales de los 70 y principios de los 80, los modelos lineales eran insuficientes en muchas aplicaciones reales. Por ejemplo, los ciclos prolongados de tamaño de población animal, (los famosos datos del lince Canadiense), ciclos solares (número anual de manchas solares), flujo de energía y relaciones de

amplitud-frecuencia, son ejemplos donde no se adecuan los modelos lineales. Para satisfacer esta demanda, se propusieron modelos no lineales de series temporales útiles durante este mismo periodo. De Gooijer y Kumar (1992) ofrecieron un resumen de los desarrollos en el área al principio de los 90. Estos autores argumentaban que el rendimiento superior en predicción de estos modelos no lineales era irregular.

Un factor que probablemente ha retrasado difusión de publicaciones de predicciones no lineales es que hasta entonces no era posible obtener expresiones analíticas cerradas para predicciones multi-paso. Sin embargo, usando la ley de Chapman-Kolmogorov en principio se pueden obtener mediante complejas integraciones numéricas predicciones de mínimos cuadrados exactas multi-paso. Pemberton (1987) y AL-Qassem y Lane (1989) dieron algunos de los primeros ejemplos de este enfoque. Hoy día, las predicciones no lineales se obtienen mediante simulación Monte Carlo o por *bootstrapping* (remuestreo). El último caso se prefiere al no necesitar suposiciones sobre la distribución del proceso de error.

La monografía de Granger y Teräsvirta (1993) ha impulsado nuevos desarrollos en la estimación, evaluación y selección entre los diferentes modelos de predicción no lineales para series temporales de economía y finanzas. Un buen resumen del estado del arte en se encuentra en IJF Special Issue 20:2 (2004). En su artículo de introducción, Clements, Franses y Swanson (2004) subrayan varios temas para investigación futura. Concluyen que todavía queda mucho por hacer para que los procedimientos no lineales para predicción, estimación y especificación de modelos estén fácilmente disponibles.

## 6.2. Modelos de cambio de régimen

La clase de modelos no lineales autorregresivos con umbral auto-excitado (SETAR) fue promocionada de forma prominente mediante los libros de Tong, 1983 y 1990. Estos modelos, los cuales son modelos lineales a trozos, en su forma más básica, han atraído algo de atención en IJF. Clements y Smith (1997) compararon varios métodos para obtener predicciones multi-paso para modelos SETAR univariados discretos. Concluyeron que las predicciones mediante simulación Monte Carlo eran satisfactorias en casos donde se sabe que las perturbaciones en el modelo SETAR vienen de una distribución simétrica. En otro caso se prefiere el método de *bootstrapping*. De Gooijer y Vidiella-i-Anguera (2004) dieron resultados similares para modelos con umbral VAR. Brockwell y Hyndman (1992) obtuvieron predicciones de un paso para modelos con umbral continuos (CTAR). Puesto que el cálculo de predicciones multi-paso a partir de los modelos CTAR involucra integraciones complicadas de mayor dimensionalidad, el uso práctico de CTAR está limitado. El rendimiento de predicción fuera de muestra de varias variantes de modelos SETAR con respecto a los modelos lineales ha sido objeto de varios artículos en IJF, incluyendo Astatkiw, Watts y Watt (1997), Boero y Marrocu (2004) y Enders y Falk (1998).

Una desventaja del modelo SETAR es que la dinámica cambia de forma discontinua de un régimen a otro. En cambio, un modelo de transición suave (STAR) permite una transición más gradual entre los distintos regímenes. Sarantis (2001)

encontró evidencia de que los modelos tipo STAR pueden ser mejores que los modelos lineales autorregresivos y de camino aleatorio para predecir precios de acciones tanto a corto plazo como a medio plazo. El estudio de Bradley y Jansen (2004) parece refutar la conclusión de Sarantis.

Fok, van Dijk y Franses (2005) examinaron la siguiente cuestión clave: ¿Puede un modelo STAR multinivel de datos de panel para series desagregadas mejorar las predicciones de agregados macroeconómicos como ingresos totales o desempleo total?. El modelo STAR propuesto parece merecer la pena investigarse en más detalle puesto que permite que los parámetros que gobiernan el cambio de régimen sean distintos según los estados. Basándose en simulaciones y hallazgos empíricos, los autores afirman que de hecho se pueden conseguir mejoras en las predicciones a un paso.

Franses, Paap, y Vroomen (2004) propusieron un modelo con umbral AR(1) que permite una inferencia plausible sobre valores específicos de los parámetros. La idea clave es que los valores de los parámetros de AR dependen de una variable indicadora dominante. El modelo resultante mejora el rendimiento respecto a otros modelos no lineales cambiantes en tiempo, incluyendo el modelo de cambio de régimen de Markov, en términos de predicción.

### 6.3. Modelos con coeficiente funcional

Un modelo AR con coeficiente funcional (FCAR o FAR) es un modelo AR en el que los coeficientes AR pueden variar como una función suave de otra variable, como un valor retrasado de la propia serie o una variable exógena. El modelo FCAR incluye los modelos TAR y STAR como casos especiales y es análogo al modelo aditivo generalizado de Hastie y Tibshirani (1991). Chen y Tsay (1993) propusieron un procedimiento de modelado usando ideas de estadística paramétrica y no paramétrica. El enfoque supone poca información a priori sobre la estructura del modelo sin sufrir la *maldición de la dimensión*; véase también Cai, Fan y Yao (2000). Harvill y Ray (2005) presentaron resultados de predicción usando modelos univariados y multivariados con coeficiente funcional (V)FCAR. Estos autores limitaron su comparación a tres métodos, el complemento predictivo ingenuo, el predictor bootstrap y el predictor multi-etapa. Tanto los resultados empíricos como las simulaciones indicaban que el método bootstrap daba predicciones ligeramente más precisas. Un área potencialmente útil de investigación futura es saber si la potencia de los modelos VFCAR puede mejorarse usando variables exógenas.

### 6.4. Redes neuronales

Una red neuronal artificial (ANN) puede ser útil para procesos no lineales que tengan relaciones funcionales desconocidas y por tanto difíciles de ajustar (Darbellay y Slama, 2000). La idea principal con ANN es que las entradas, o variables dependientes, son filtradas mediante una o más capas ocultas cada una de las cuales consta de nodos antes de alcanzar la variable de salida. La salida

intermedia se enlaza a la salida final. Otros modelos son versiones específicas de ANN donde se impone cierta estructura (véase JoF Special Issue 17:5/6 (1998)).

Una de las mayores áreas de aplicación de ANN es predicción; véase Zhang, Patuwo y Hu (1998) y Hippert, Pedreira y Souza (2001) para revisiones de la literatura. Existen numerosos estudios que documentan el éxito de ANN en predicción de datos financieros. Sin embargo, en dos publicaciones de IJF, Chatfield, 1993 y 1995 se cuestionó si las ANNs se habían vendido demasiado como técnicas de predicción milagrosas. Seguidamente algunos artículos documentaron que modelos ingenuos como el de camino aleatorio podían mejorar el rendimiento de ANNs (véase Callen et al., 1996, Church y Curram, 1996, Conejo et al., 2005, Gorr et al., 1994 y Tkacz, 2001). Estas observaciones son consistentes con los resultados de Adya y Collopy (1998) donde se evalúa la efectividad de predicción basada en ANN para 48 estudios realizados entre 1988 y 1994.

Gorr (1994) y Hill, Marquez, OConnor y Remus (1994) sugirieron que la investigación a seguir debería tratar de definir mejor los límites en que ANN mejora las técnicas tradicionales y viceversa. Varios autores exploran este tema. Hill et al. (1994) advirtió que las ANNs probablemente funcionen mejor para datos financieros con alta frecuencia y Balkin and Ord (2000) también subrayaron la importancia del tamaño de la serie para asegurar resultados óptimos del entrenamiento de la ANN. Qi (2001) señaló que las ANNs probablemente mejoren otros métodos cuando los datos de entrada se mantienen lo más actualizados posible usando modelado recursivo (Olson y Mossman, 2003).

Un problema general con los modelos no lineales se encuentra en la complejidad de los modelos y su parametrización excesiva. Si se considera realmente importante el principio de parsimonia, es interesante comparar el rendimiento de predicción fuera de muestra de los modelos lineales frente a los no lineales, usando una amplia variedad de criterios de selección de modelo. Esta cuestión fue considerada en bastante profundidad por Swanson y White (1997). Sus resultados sugirieron que una simple ANN con una única capa oculta *feed-forward*, la cual era muy popular en series temporales econométricas, ofrece una alternativa útil y flexible a modelos lineales de especificación fijada, particularmente para horizontes de predicción mayores de un paso. En contraste con Swanson y White, Heravi, Osborn y Birchenhall (2004) encontraron que los modelos lineales producen predicciones más precisas de producción industrial mensual europea sin ajuste estacional que los modelos de ANN. Ghiassi, Saidane, y Zimbra (2005) presentaron una ANN dinámica y compararon su rendimiento de predicción frente a la ANN tradicional y los modelos ARIMA.

Con el tiempo, la importancia del riesgo de la parametrización excesiva y el sobreajuste ha sido reconocida por varios autores; véase Hippert, Bunn y Souza (2005) que usaron una gran ANN (50 entradas, 15 neuronas ocultas y 24 salidas) para predecir perfiles de carga de electricidad diaria. Sin embargo, la cuestión de si la ANN está sobre-parametrizada o no, sigue sin resolverse. Algunas ideas con valor potencial para construir ANNs minimizando el número de parámetros, usando inferencia estadística se sugieren en Teräsvirta, van Dijk y Medeiros (2005).

### 6.5. Dinámicas estadísticas frente a deterministas

La posibilidad de que las no linealidades de alta frecuencia en datos financieros (como rendimientos por hora) sean producidos por procesos caóticos deterministas de baja dimension, ha sido objeto de algunos estudios en IJF. Cecen y Erkal (1996) mostraron que no es posible explotar la dependencia determinista no lineal en los tipos de cambio diario para mejorar la predicción a corto plazo. Lisi y Medio (1997) reconstruyeron el espacio de estado para los tipos de cambio mensual y usando un método lineal local, aproximaron las dinámicas del sistema en ese espacio. La predicción fuera de muestra a un paso mostró que su método mejoraba un modelo de camino aleatorio. Un estudio similar fue realizado por Cao y Soofi (1999).

### 6.6. Otros

Otros modelos no lineales ni tan conocidos se han usado para predicción. Por ejemplo, Ludlow y Enders (2000) adoptaron coeficientes de Fourier para aproximar diferentes no linealidades presentes en series temporales. Herwartz (2001) amplió el vector lineal ECM para permitir asimetrías. Dahl y Hylleberg (2004) compararon el modelo de regresión no lineal flexible de Hamilton (2001), ANNs y dos versiones del modelo de regresión

## Referencias

1. Brown, R. Statistical Forecasting for Inventory Control (1959) cited By (since 1996)127.
2. Brown, R. Smoothing, Forecasting and Prediction of Discrete Time Series (1963) cited By (since 1996)192.
3. Holt, C.: Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting **20**(1) (2004) 5–10 cited By (since 1996)78.
4. Winters, P.: Forecasting sales by exponentially weighted moving averages. Management Science **6**(3) (1960) 324–342 cited By (since 1996)264.
5. Pegels, C.: Exponential forecasting: Some new variations. Management Science **12**(5) (1969) 311–315 cited By (since 1996)35.
6. Muth, J.: Optimal properties of exponentially weighted forecasts. Journal of the American Statistical Association **55**(290) (1960) 299–306 cited By (since 1996)184.
7. Roberts, S.: A general class of holt-winters type forecasting models. Management Science **28**(7) (1982) 808–820 cited By (since 1996)0.
8. Abraham, B., Ledolter, J. Statistical Methods for Forecasting (1983) cited By (since 1996)207.
9. Abraham, B., Ledolter, J.: Forecast functions implied by autoregressive integrated moving average models and other related forecast procedures. International Statistical Review **54** (1986) 51–66 cited By (since 1996)14.
10. Gardner E.S., J.: Exponential smoothing: The state of the art. Journal of Forecasting **4**(1) (1985) 1–28 cited By (since 1996)206.
11. Snyder, R.: Recursive estimation of dynamic linear models. Journal of the Royal Statistical Society, Series B **47**(2) (1985) 272–276 cited By (since 1996)40.



12. Bartolomei, S., Sweet, A.: A note on a comparison of exponential smoothing methods for forecasting seasonal series. *International Journal of Forecasting* **5**(1) (1989) 111–116 cited By (since 1996)8.
13. Makridakis, S., Hibon, M.: Exponential smoothing: The effect of initial values and loss functions on post-sample forecasting accuracy. *International Journal of Forecasting* **7**(3) (1991) 317–330 cited By (since 1996)8.
14. McClain, J.: Dominant tracking signals. *International Journal of Forecasting* **4**(4) (1988) 563–572 cited By (since 1996)7.
15. Sweet, A., Wilson, J.: Pitfalls in simulation-based evaluation of forecast monitoring schemes. *International Journal of Forecasting* **4**(4) (1988) 573–579 cited By (since 1996)1.
16. McKenzie, E.: General exponential smoothing and the equivalent arma process. *Journal of Forecasting* **3**(3) (1984) 333–344 cited By (since 1996)31.
17. Taylor, J.: Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting* **19**(4) (2003) 715–725 cited By (since 1996)42.
18. Gardner Jr., E.: Forecasting the failure of component parts in computer systems: A case study. *International Journal of Forecasting* **9**(2) (1993) 245–253 cited By (since 1996)5.
19. Grubb, H., Mason, A.: Long lead-time forecasting of uk air passengers by holt-winters methods with damped trend. *International Journal of Forecasting* **17**(1) (2001) 71–82 cited By (since 1996)27.
20. Miller, T., Liberatore, M.: Seasonal exponential smoothing with damped trends. an application for production planning. *International Journal of Forecasting* **9**(4) (1993) 509–515 cited By (since 1996)10.
21. Hyndman, R., Koehler, A., Snyder, R., Grose, S.: A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting* **18**(3) (2002) 439–454 cited By (since 1996)108.
22. Carreno, J., Madinaveitia, J.: A modification of time series forecasting methods for handling announced price increases. *International Journal of Forecasting* **6**(4) (1990) 479–484 cited By (since 1996)3.
23. Williams, D., Miller, D.: Level-adjusted exponential smoothing for modeling planned discontinuities. *International Journal of Forecasting* **15**(3) (1999) 273–289 cited By (since 1996)8.
24. Rosas, A., Guerrero, V.: Restricted forecasts using exponential smoothing techniques. *International Journal of Forecasting* **10**(4) (1994) 515–527 cited By (since 1996)5.
25. Lawton, R.: How should additive holt-winters estimates be corrected? *International Journal of Forecasting* **14**(3) (1998) 393–403 cited By (since 1996)10.
26. Mckenzie, E.: Error analysis for winters' additive seasonal forecasting system. *International Journal of Forecasting* **2**(3) (1986) 373–382 cited By (since 1996)4.
27. Archibald, B., Koehler, A.: Normalization of seasonal factors in winter's methods. *International Journal of Forecasting* **19**(1) (2003) 143–148 cited By (since 1996)6.
28. Hyndman, R., Billah, B.: Unmasking the theta method. *International Journal of Forecasting* **19**(2) (2003) 287–290 cited By (since 1996)23.
29. Assimakopoulos, V., Nikolopoulos, K.: The theta model: A decomposition approach to forecasting. *International Journal of Forecasting* **16**(4) (2000) 521–530 cited By (since 1996)48.
30. Pfeffermann, D., Allon, J.: Multivariate exponential smoothing: Method and practice. *International Journal of Forecasting* **5**(1) (1989) 83–98 cited By (since 1996)13.

31. Pan, X.: An alternative approach to multivariate ewma control chart. *Journal of Applied Statistics* **32**(7) (2005) 695–705 cited By (since 1996)4.
32. Ord, J., Koehler, A., Snyder, R.: Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association* **92**(440) (1997) 1621–1629 cited By (since 1996)80.
33. Koehler, A., Snyder, R., Ord, J.: Forecasting models and prediction intervals for the multiplicative holt-winters method. *International Journal of Forecasting* **17**(2) (2001) 269–286 cited By (since 1996)29.
34. Chatfield, C., Koehler, A., Ord, J., Snyder, R.: A new look at models for exponential smoothing. *Journal of the Royal Statistical Society Series D: The Statistician* **50**(2) (2001) 147–159 cited By (since 1996)36.
35. Gardner Jr., E., McKenzie, E.: Model identification in exponential smoothing. *Journal of the Operational Research Society* **39** (1988) 863–867 cited By (since 1996)17.
36. Tashman, L., Kruk, J.: The use of protocols to select exponential smoothing procedures: A reconsideration of forecasting competitions. *International Journal of Forecasting* **12**(2) (1996) 235–253 cited By (since 1996)12.
37. Collopy, F., Armstrong, J.: Rule-based forecasting: Development and validation of an expert systems approach to combining time series extrapolations. *Management Science* **38**(10) (1992) 1394–1414 cited By (since 1996)90.
38. Satchell, S., Timmermann, A.: On the optimality of adaptive expectations: Muth revisited. *International Journal of Forecasting* **11**(3) (1995) 407–416 cited By (since 1996)9.
39. Sweet, A.: Computing the variance of the forecast error for the holt-winters seasonal models. *Journal of Forecasting* **4** (1985) 235–243 cited By (since 1996)6.
40. Newbold, P., Bos, T.: On exponential smoothing and the assumption of deterministic trend plus white noise data-generating models. *International Journal of Forecasting* **5**(4) (1989) 523–527 cited By (since 1996)7.
41. Johnston, F., Harrison, P.: Variance of lead-time demand. *Journal of the Operational Research Society* **37**(3) (1986) 303–308 cited By (since 1996)14.
42. Yar, M., Chatfield, C.: Prediction intervals for the holt-winters forecasting procedure. *International Journal of Forecasting* **6**(1) (1990) 127–137 cited By (since 1996)31.
43. Chatfield, C., Yar, M.: Prediction intervals for multiplicative holt-winters. *International Journal of Forecasting* **7**(1) (1991) 31–37 cited By (since 1996)24.
44. Hyndman, R., Koehler, A., Ord, J., Snyder, R.: Prediction intervals for exponential smoothing using two new classes of state space models. *Journal of Forecasting* **24**(1) (2005) 17–37 cited By (since 1996)28.
45. Archibald, B.: Parameter space of the holt-winters' model. *International Journal of Forecasting* **6**(2) (1990) 199–209 cited By (since 1996)8.
46. Bhansali, R.: Asymptotically efficient autoregressive model selection for multistep prediction. *Annals of the Institute of Statistical Mathematics* **48**(3) (1996) 577–602 cited By (since 1996)37.
47. Bhansali, R.: Autoregressive model selection for multistep prediction. *Journal of Statistical Planning and Inference* **78**(1-2) (1999) 295–305 cited By (since 1996)5.
48. Box, G., Jenkins, G., Reinsel, G. *Time Series Analysis: Forecasting and Control* (1994) cited By (since 1996)3225.
49. Cholette, P.: Prior information and arima forecasting. *Journal of Forecasting* **1**(4) (1982) 375–383 cited By (since 1996)0.
50. Guerrero, V.: Arima forecasts with restrictions derived from a structural change. *International Journal of Forecasting* **7**(3) (1991) 339–347 cited By (since 1996)5.

51. de Alba, E.: Constrained forecasting in autoregressive time series models: A bayesian analysis. *International Journal of Forecasting* **9**(1) (1993) 95–108 cited By (since 1996)6.
52. Geriner, P., Ord, J.: Automatic forecasting using explanatory variables: A comparative study. *International Journal of Forecasting* **7**(2) (1991) 127–140 cited By (since 1996)7.
53. Chatfield, C.: What is the 'best' method of forecasting? *Journal of Applied Statistics* **15**(1) (1988) 19–38 cited By (since 1996)12.
54. Edlund, P.E.: Identification of the multi-input box-jenkins transfer functions model. *Journal of Forecasting* **3** (1984) 297–308 cited By (since 1996)5.
55. Gupta, S.: Testing causality. some caveats and a suggestion. *International Journal of Forecasting* **3**(2) (1987) 195–209 cited By (since 1996)1.