

PREVISIÓN EN TIEMPO REAL CON MODELOS DE BOX-JENKINS UNIVARIANTES

José Parreño Fernández¹, Isabel Fernández Quesada², Alberto Gómez Gómez³, Nazario García Fernández⁴

¹ETSII e II Gijón, Campus de Viesques, parreno@etsiig.uniovi.es

²ETSII e II Gijón, Campus de Viesques, bel@etsiig.uniovi.es

³ETSII e II Gijón, Campus de Viesques, agomez@etsiig.uniovi.es

⁴ETSII e II Gijón, Campus de Viesques, ngarcia@econo.uniovi.es

RESUMEN

En este artículo se hace una aproximación a la previsión en tiempo real de series temporales. Para comprobar la eficacia de esta metodología se han analizado ocho series temporales, para las que se han obtenido modelos ARIMA univariantes, mediante la metodología Box-Jenkins. Estos modelos se han utilizado para hacer previsiones de dos maneras distintas: en primer lugar, se han obtenido previsiones de una manera convencional, esto es, para un cierto número N de periodos hacia delante; en segundo lugar, se han realizado previsiones un periodo hacia delante N veces, para lo que es necesario ir incorporando al modelo un dato real nuevo cada vez que se hace una previsión, con lo que se consigue simular una situación en la que se obtienen previsiones en tiempo real. Los resultados obtenidos al comparar estas dos formas de hacer previsiones, muestran la conveniencia de realizar previsiones en tiempo real siempre que sea posible, debido a la gran mejora en la calidad de las mismas.

1. Introducción.

En este artículo se hace una aproximación a la previsión en tiempo real de series temporales. Para comprobar la eficacia de esta metodología se han analizado ocho series temporales, para las que se han obtenido modelos ARIMA univariantes, mediante la metodología Box-Jenkins (1). Las ocho series temporales analizadas incluyen algunas estacionarias, otras no estacionarias, algunas estacionales y otras no estacionales, con la intención de que el estudio abarque a series temporales con características muy variadas, de forma que los resultados obtenidos sean generalizables para la mayoría de series temporales que se puedan encontrar en la práctica. El objetivo es intentar mejorar la calidad de las previsiones obtenidas con la metodología más tradicional, consistente en obtener de una sola vez previsiones para N periodos hacia delante.

La estructura de este artículo es la siguiente: en el apartado 2 se hace un análisis detallado de las características de las ocho series que van a ser utilizadas en el estudio; en el apartado 3 se muestran los modelos obtenidos para cada una de las series; en el apartado 4 se presentan los resultados obtenidos al comparar los dos métodos de previsión propuestos, y en el apartado 5 se muestran las conclusiones extraídas del estudio realizado.

2. Características de las series temporales utilizadas.

Como se indicó en el apartado anterior, se van a realizar previsiones con ocho series temporales cuyas características de estacionalidad y estacionariedad se indican en la Tabla 1.

SERIE	FUENTE	ESTACIONAL	ESTACIONARIA
UN04	Reilly (2)	NO	SI
BJ06	Box y Jenkins (1)	NO	SI
AL03	Abraham (3)	SI	SI
AL11	Abraham (3)	SI	SI
UN14	Reilly (2)	NO	NO
UN21	Reilly (2)	NO	NO
UN01	Reilly (2)	SI	NO
UN06	Reilly (2)	SI	NO

Tabla 1: Características de las series

Como se puede comprobar, se va a trabajar con dos series temporales que son no estacionales y estacionarias (UN04 y BJ06), otras dos que son estacionales y estacionarias (AL03 y AL11), otras dos que son no estacionales y no estacionarias (UN14 y UN21), y otras dos series que son estacionales y no estacionarias (UN01 y UN06).

3. Obtención de los modelos univariantes.

En la tabla 2 se muestran los modelos ARIMA que se han obtenido para las series temporales estudiadas.

SERIE	MODELO ARIMA
UN04	$(1-B^{12})_X = (1-\theta_1 B) a_t$
BJ06	$(1-\phi_1 B - \phi_2 B^2)_X = k + (1-\theta_1 B) a_t$
AL03	$(1-\phi_1 B)(1-\phi_{12} B^{12})(1-B^{12})_X = k + (1-\theta_{12} B^{12}) a_t$
AL11	$(1-\phi_1 B)(1-B^{12})_X = k + (1-\theta_{12} B^{12}) a_t$
UN14	$(1-\phi_1 B - \phi_2 B^2)(1-B)_X = k + a_t$
UN21	$(1-\phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1-B)_X = k + a_t$
UN01	$(1-\phi_{12} B^{12})(1-B)(1-B^{12})_X = (1-\theta_1 B) (1-\theta_{12} B^{12}) a_t$
UN06	$(1-B)(1-B^4)_X = (1-\theta_1 B) a_t$

Tabla 2: Modelos ARIMA

Como se puede apreciar a la vista de la tabla 2, estos modelos recogen las características de las series que se habían apuntado en el apartado 2.

4. Presentación de resultados.

El objetivo principal que persigue este trabajo es la obtención de previsiones. La manera tradicional de obtener estas previsiones es realizarlas en un solo paso, es decir, utilizar los datos disponibles hasta el período actual para obtener previsiones para los siguientes N períodos. Lo que en este trabajo se propone, siguiendo a Faraway y Chatfield (4), es ir incorporando nuevos datos reales al modelo según se van obteniendo, y realizar, por tanto, N veces previsiones para un solo período, simulando por tanto una previsión en tiempo real. Para comparar la calidad de las previsiones obtenidas con ambas técnicas (5), se utiliza el Error Absoluto Medio (MAD, Mean Absolute Deviation).

Para realizar las previsiones en tiempo real es necesario, cuando se incorpora un nuevo dato a la serie, comprobar la validez del modelo para la nueva serie, y en cualquier caso, reestimar el valor de sus parámetros.

Para obtener las previsiones de la manera convencional, se ha empleado la técnica ex-post, es decir, se dividen los datos de todas las series temporales en dos grupos (6,7): un primer subconjunto que servirá para la obtención del modelo ARIMA correspondiente, con el que se realizarán las previsiones para los siguientes N períodos; y un segundo subconjunto, formado por las restantes observaciones de cada serie, que servirá para comparar las previsiones obtenidas con los datos reales.

En las figuras 1 a 8 se comparan gráficamente los datos reales de las series con las previsiones obtenidas con ambos métodos.

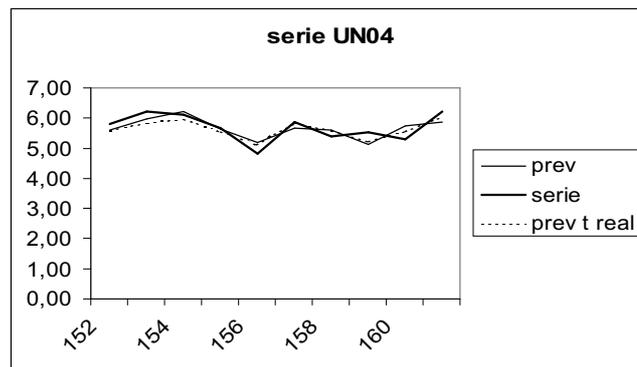


Figura 1. Previsiones serie UN04

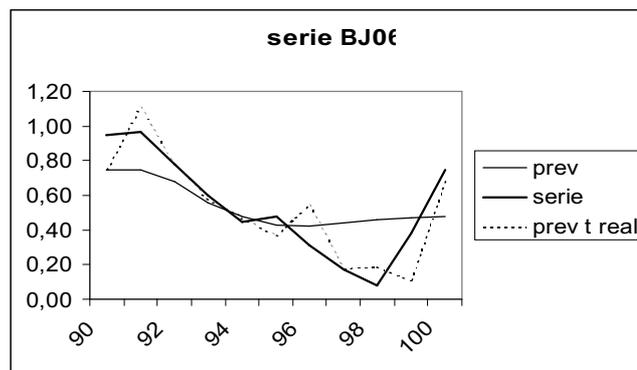


Figura 2. Previsiones serie BJ06

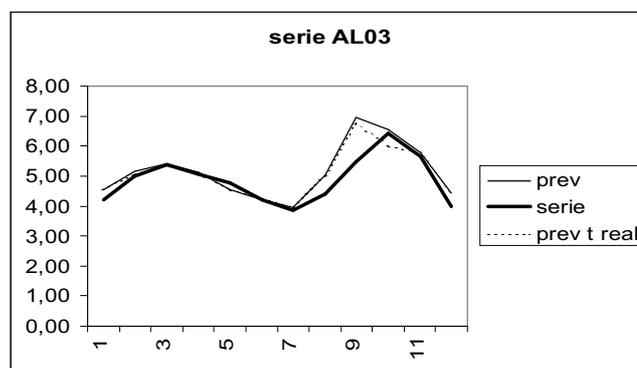


Figura 3. Previsiones serie AL03

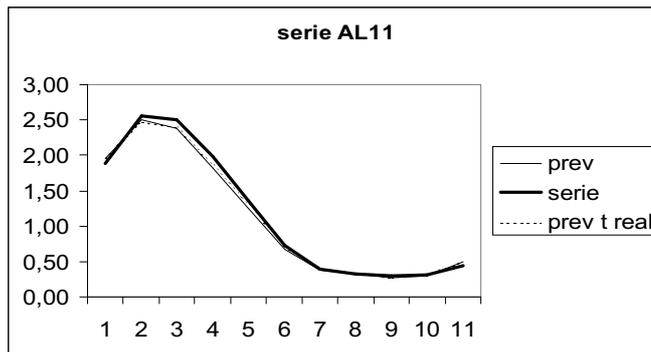


Figura 4. Previsiones serie AL11

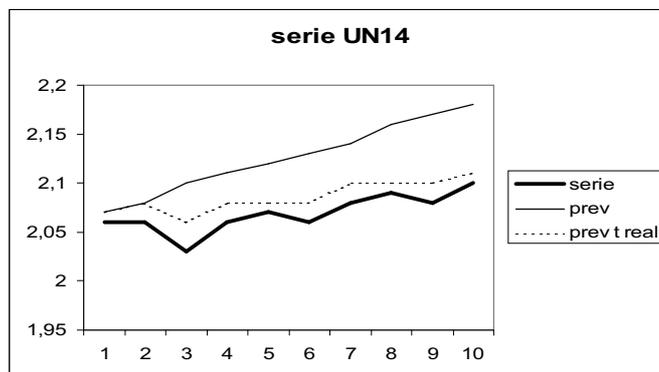


Figura 5. Previsiones serie UN14

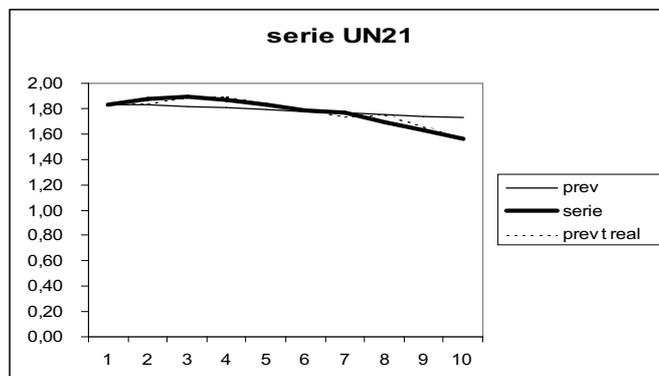


Figura 6. Previsiones serie UN21

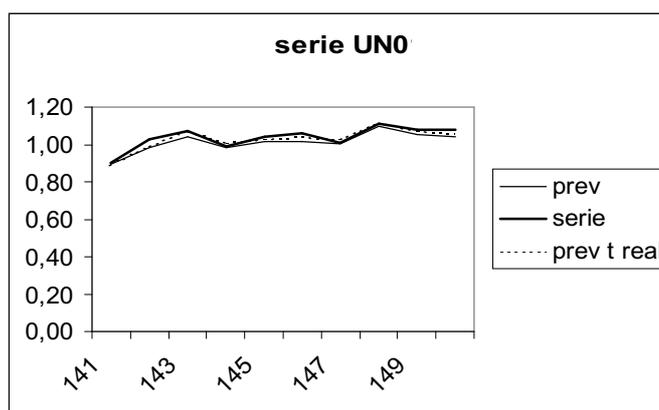


Figura 7. Previsiones serie UN01

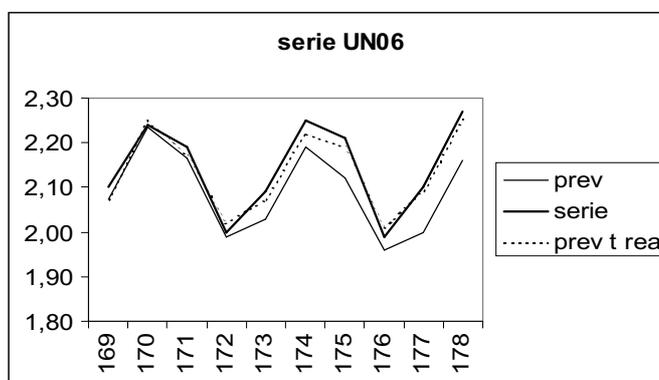


Figura 8. Previsiones serie UN06

A la vista de las figuras anteriores queda claro, incluso gráficamente, que las previsiones obtenidas con la técnica tiempo real superan en calidad a las obtenidas de forma convencional. De todas maneras, en la tabla 3 se muestran los MAD obtenidos por ambas técnicas al realizar previsiones con todas las series temporales. También se muestra, en la última columna de dicha tabla, el porcentaje de mejora que supone la realización de las previsiones en tiempo real con respecto a las convencionales.

SERIE	MAD PREV CONVENCIONAL	MAD PREV EN T REAL	MEJORA
UN04	0.258	0.209	19%
BJ06	0.160	0.107	33.1%
AL03	0.313	0.303	3.2%
AL11	0.062	0.047	24.2%
UN14	0.057	0.02	64.9%
UN21	0.057	0.021	63.2%
UN01	0.024	0.016	33.3%
UN06	0.052	0.020	61.5%

Tabla 3: Errores Medios Absolutos

Conviene reseñar que realizar previsiones en tiempo real implica mayor trabajo que realizarlas de la manera convencional, por lo menos a nivel de investigación, ya que hay que ir incorporando uno a uno los nuevos datos que se van obteniendo en cada serie, para a

continuación reestimar el modelo correspondiente. Este inconveniente dejaría de ser tal en la práctica, ya que la mayoría de las series temporales miden sus observaciones mensual, trimestral o anualmente, con lo que se dispondría de todo un mes, como mínimo, para incorporar el nuevo dato y obtener y estimar el nuevo modelo. Donde no sería aplicable esta técnica sería en casos de series temporales cuyas observaciones se midan a intervalos de un minuto o menos, en cuyo caso sería materialmente imposible la obtención de un nuevo modelo.

En la tabla 4 se muestra, a modo de ejemplo, la variación de los coeficientes y del RSE, para el modelo univariante de la serie UN14, a medida que se va añadiendo una observación nueva al modelo.

n° observaciones	ϕ_1	ϕ_2	k	RSE
137	-0.428	-0.263	0.021	0.0232
138	-0.427	-0.265	0.021	0.0232
139	-0.417	-0.253	0.020	0.0235
140	-0.418	-0.254	0.020	0.0234
141	-0.419	-0.250	0.020	0.0233
142	-0.419	-0.255	0.020	0.0233
143	-0.420	-0.256	0.020	0.0232
144	-0.421	-0.254	0.020	0.0231
145	-0.420	-0.257	0.020	0.0231
146	-0.420	-0.257	0.020	0.0231

Tabla 4. Variación de los coeficientes

5. CONCLUSIONES.

En este artículo se han comparado dos técnicas de previsión basándose siempre en los modelos Box-Jenkins univariantes. Estas dos técnicas son, en primer lugar, la obtención de previsiones de la manera convencional, esto es, de una sola vez y para un número N de períodos; y en segundo lugar, lo que denominamos previsión en tiempo real, consistente en realizar N veces previsiones para un solo período. Esta última técnica tiene el inconveniente de que hay que ir incorporando, uno a uno, nuevos datos a la serie temporal, lo que conlleva la reestimación del modelo cada vez.

Los resultados obtenidos muestran claramente que empleando la técnica de previsión en tiempo real se obtienen previsiones de mejor calidad, a la vista de los MAD obtenidos con ambas técnicas. La limitación de esta técnica se encuentra cuando las observaciones de la serie temporal se producen a intervalos muy cortos de tiempo, por ejemplo segundos o unos pocos minutos, ya que resultaría inviable la estimación de un modelo de Box-Jenkins en un período de tiempo tan corto.

Por otro lado, esta técnica tampoco sería de aplicación en aquellos casos en que haya que tomar decisiones con mucha antelación (8), ya que no quedaría más remedio que hacer previsiones con un número de períodos de antelación y actuar en consecuencia.

REFERENCIAS

(1) Box, G.E.P., and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*. 2nd. ed. San Francisco: Holden Day.

- (2) Reilly, D.P. (1980). "Experience with an Automatic Box-Jenkins Modeling Algorithm in Time Series Analysis". *O.D. Anderson (Amsterdam:North-Holland)*, pp. 493-508.
- (3) Abraham, B., and Ledolter, J. (1983). *Statistical Methods for Forecasting*. New York: Wiley.
- (4) Faraway, J. y Chatfield, C. (1998). "Time series forecasting with neural networks: a comparative study using the airline data". *Journal of Royal Statistical Society: Series C (Applied Statistics)* 47 (2), pp. 231-250.
- (5) Xie, M. and Ho, S.L. (1999). "Analysis of repairable system failure data using time series models". *Journal of Quality in Maintenance Engineering*. Vol. 5 No.1, pp. 50-61.
- (6) Atienza, O.O., Ang, B.W. y Tang, L.C. (1997). "Statistical process control and forecasting". *International Journal of Quality Science*, Vol. 2 No.1 págs. 37-51.
- (7) Tanwari, A.U. y Betts, J. (1999). "Impact of Forecasting on Demand Planning". *Production and Inventory Management Journal*. Third Quarter, 1999 pp.31-35.
- (8) García-Ferrer, A. y Queralt, R.A. (1997). "A note on forecasting international tourism demand in Spain". *International Journal of Forecasting* 13, págs. 539-549.