

Revista Cubana de Ciencia Agrícola
ISSN: 0034-7485
rcca@ica.co.cu
Instituto de Ciencia Animal
Cuba

Sánchez, Lyhen; Cabanas, Gladis; Abad, Yoandra; Torres, Verena
Utilización de modelos ARIMA para la predicción de la producción de leche. Estudio de caso en la UBPC “Maniabo”, Las Tunas
Revista Cubana de Ciencia Agrícola, vol. 48, núm. 3, 2014, pp. 213-218
Instituto de Ciencia Animal
La Habana, Cuba

Disponible en: http://www.redalyc.org/articulo.oa?id=193032133002

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

Sistema de Información Científica
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Utilización de modelos ARIMA para la predicción de la producción de leche. Estudio de caso en la UBPC "Maniabo", Las Tunas

Lyhen Sánchez[1], Gladis Cabanas[2],Yoandra Abad[3] y Verena Torres[1]
[1] *Instituto de Ciencia Animal, Apartado Postal 24, San José de las Lajas, Mayabeque, Cuba*
[2] *Facultad de Matemática y Computación, Universidad de la Habana, Ciudad de La Habana, Cuba*
[3] *Instituto de Medicina Tropical "Pedro Kourí", La Habana, Cuba*
Correo electrónico: lyhen@ipk.sld.cu

Se pronosticó, a corto plazo, la producción de leche por modelos ARIMA, en un estudio de caso en la Unidad Básica de Producción Cooperativa "Maniabo", de la provincia Las Tunas, en Cuba. Se utilizaron los datos de producción de leche mensual del período 2000-2010. Se utilizaron gráficas de ploteo horizontal y Box-Plot, para el análisis descriptivo de la serie y para encontrar la tendencia al suavizamiento polinomial. Con la aplicación de los períodogramas, se confirmaron los picos notables en los movimientos periódicos de la serie, la significación de una frecuencia y además, se detectó la estacionalidad, descrita con los índices estacionales. Se determinaron las correlaciones con la serie diferenciada y se estimaron los parámetros p y q del modelo. El mejor modelo ajustado fue un ARIMA (1,0,3) x (0,1,0) 12, con constante. Se pronosticó y se validó la producción de leche para el 2011.

Palabras clave: *series temporales, modelos ARIMA, producción de leche*

En la investigación estadística aplicada a la esfera agropecuaria se abordan problemas en los que no se cumplen las condiciones para la utilización del modelo lineal general, pues se realizan controles de la variable de forma histórica y, generalmente, las observaciones de una misma variable están correlacionadas. Además, el pronóstico es una especie de extrapolación (Aguirre 1994). Con la aplicación de los modelos de series temporales, se puede describir y predecir el comportamiento de un fenómeno que cambia en el tiempo, y que muestra dependencia entre las observaciones sucesivas (Arellano 2001).

Según Maté (1995), la producción de leche cubana ha experimentado variaciones durante años, lo que ha traído implicaciones negativas para el crecimiento del sector. La mayoría de los modelos para series temporales presuponen que los factores que provocan los cambios en la serie se mantengan constantes, por lo que los primeros análisis que se realizan parten de aislar estos factores para estudiar su comportamiento (Blumberg 1984).

Entre las series temporales, la modelación ARIMA (Autorregresive Integrated Moving Average) fue desarrollada por George Box y Gwilym Jenkins en la década de los setenta del siglo pasado y constituyó una revolución en el análisis de series de tiempo.

Se conoce que a partir del 2000, la Unidad Básica de Producción Cooperativa (UBPC) "Maniabo" no experimentó cambios que se consideren importantes en la producción de leche, por lo que se puede tomar la serie litros producidos a partir de este año para el objetivo del trabajo.

En Cuba se dispone de pocos trabajos que utilicen las series temporales y, en particular, los modelos ARIMA en el campo agropecuario. Es por ello que en este trabajo se pretendió pronosticar, a corto plazo, la producción de leche en la UBPC "Maniabo", de Las Tunas, Cuba, con la aplicación de este tipo de modelo.

Materiales y Métodos

El estudio se realizó en la UBPC "Maniabo", ubicada en Las Tunas. Los datos utilizados para el análisis fueron la producción de leche mensual de esta UBPC en el período 2000-2010.

Se utilizaron gráficas de ploteo horizontal y de Box-Plot para el análisis descriptivo de la serie. Se empleó el suavizamiento polinomial para detectar la tendencia de la misma (Coutín 2003).

Con los períodogramas (simple e integrado) se encontró la significación de la componente estacional y se hallaron los índices estacionales para estudiar el comportamiento de la estacionalidad de forma descriptiva (Maté 2011).

Con la serie con diferencia estacional se hallaron los coeficientes de correlación simple y parcial, mediante los que se estimaron los parámetros p, d y q y P, D y Q del modelo ARIMA a ajustar, cuya notación es:

ARIMA (p,d,q) (P,D,Q) k, donde k es la longitud del período estacional

La bondad de ajuste del modelo probado se realizó mediante los criterios estadísticos de errores clásicos: error medio (ME), error medio absoluto (MAE), error medio porcentual (MPE), error medio absoluto porcentual (MAPE) y error cuadrático medio (SME) (Torres *et al.* 2012).

Se validó el modelo seleccionado mediante la comprobación de la aleatoriedad de los residuos con la prueba de rachas y la normalidad residual del mismo (Maté 1995). Los resultados se procesaron mediante los paquetes estadísticos, R versión 2.15.0, el STATGRAPHICS, Centurión XV.II y el Microsoft Excel 2007.

Resultados y Discusión

Según criterios de Einsteinnet (2001), la serie analizada se consideró consistente, estable, y no se presentaron valores aberrantes.

El análisis descriptivo inicial de la serie (figura 1) evidenció notable estacionalidad. El período de mayor rendimiento en la producción de leche fue de mayo a octubre. Noviembre, a pesar de no estar contemplado en el período lluvioso en Cuba, es un mes en el que se logran buenas producciones de leche porque aún existe humedad en los suelos, lo que ayuda, en gran medida, al crecimiento del pasto. La media de producción en este mes está muy próxima a la media general.

La variabilidad mayor se da en los meses de mayor media, pero no tanto como para hacer visible esto en los coeficientes de variación. Se conoce que, en la época de lluvia, la variabilidad es mayor en los meses centrales de la temporada húmeda, debido a la relación de la lluvia con los pastos, y de estos con la producción de leche. La variabilidad de la producción de leche evidenciada en estos meses se podría relacionar causalmente con la de la lluvia. Estos resultados coinciden con los de un estudio realizado por Grubb (1992).

La figura 2, que muestra la serie temporal de producción de leche en el período analizado (2000-2010), sugiere tendencia decreciente en los últimos años. Esta no se evidencia muy marcada ni mantenida durante un período considerable, por lo que se aplicó un tipo de alisamiento que permite visualizar mejor el movimiento de la serie, y así comprobar si existe tendencia o no. Para esto se utilizó el suavizamiento polinomial (figura 3) para n=1 y n=3 y la comparación de los ajustes. Con ello se reafirmó el criterio de que la serie no presenta tendencia.

Según Colás *et al*. (2011), los períodogramas son un elemento que valida la significación de la componente estacional y la existencia de otros fenómenos periódicos en el comportamiento de la serie (componente cíclica). En el períodograma se evidencia que el movimiento estacional, cada doce meses (frecuencia 0,0833), es altamente significativo y presenta un gran pico en esta frecuencia (figura 4).

El períodograma integrado de la figura 5 muestra que el segmento que representa el incremento en esta misma frecuencia es mucho más largo que el ancho de la franja más externa, que marca 95 % de confianza. Además, no se ven otros ciclos que el períodograma integrado encuentre como significativos.

Teniendo en cuenta el análisis previo, que encontró presencia de estacionalidad y ausencia de tendencia, se puede modelar entonces la serie por un modelo ARIMA, con diferencia estacional y ninguna no estacional. Este resultado coincide con estudios realizados por Funes

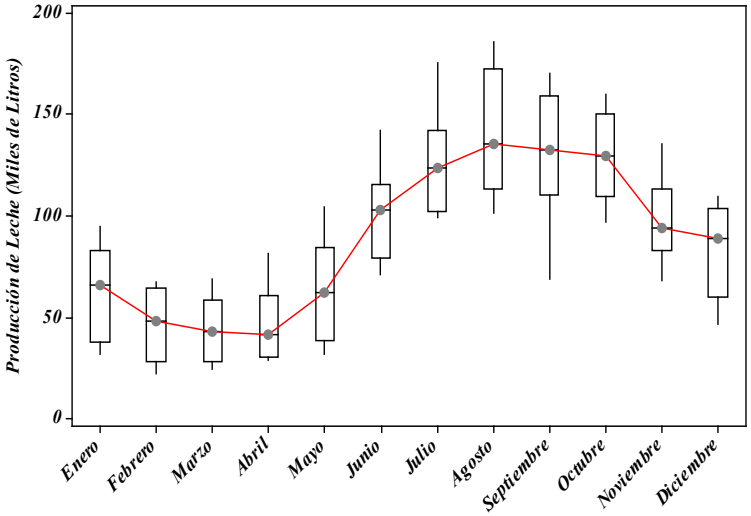


Figura 1. Análisis descriptivo de la serie de producción de leche por mes

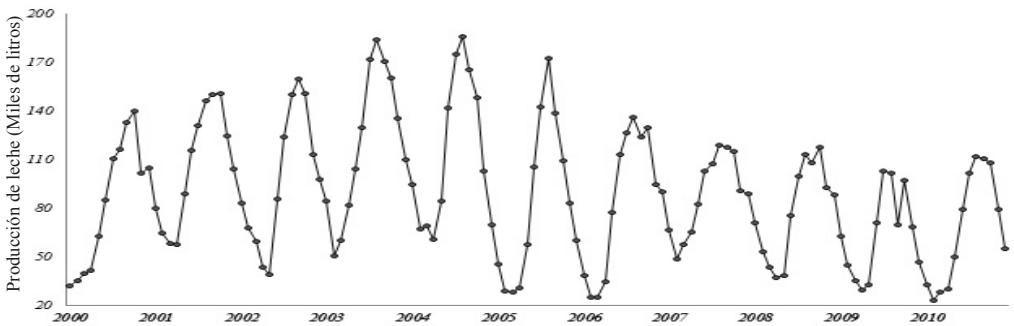


Figura 2. Comportamiento de la producción de leche en la UBPC "Maniabo" durante el período 2000-2010.

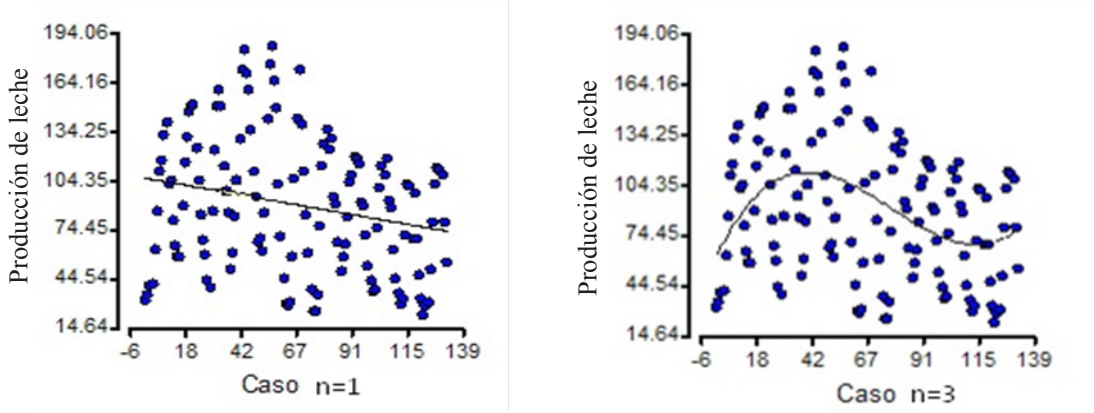


Figura 3. Ajuste de la tenencia de la serie de producción de leche, mediante suavizamiento polinomial, para n=1 y n=3

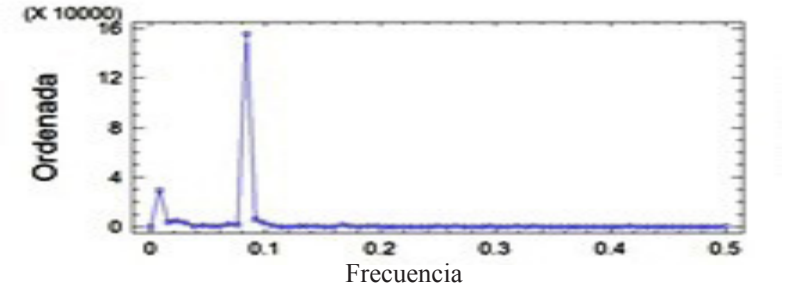


Figura 4. Períodograma de la serie de producción de leche en la UBPC "Maniabo" en el período 2000-2010

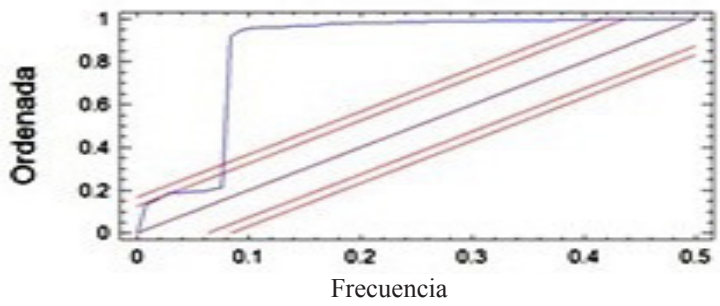


Figura 5. Períodograma integrado de la serie de producción de leche en la UBPC "Maniabo" durante el período 2000-2010.

et al. (2009).

En las funciones de autocorrelación y autocorrelación parcial de la serie diferenciada (figura 6), se estimó que los órdenes del modelo pueden ser $p \leq 1$ y $q \leq 3$. Sobre esta base se prueban diferentes modelos ARIMA.

Se probaron los modelos 1) ARIMA (1,0,0) x (0,1,0) 12, el 2) ARIMA (1,0,1) x (0,1,0)12, el 3) ARIMA (1,0,2) x (0,1,0) 12, 4) ARIMA (1,0,3) x (0,1,0) 12 y 5) ARIMA (0,0,3) x (0,1,0) 12, todos con constante.

Se realizó la comparación de los modelos probados mediante la utilización de la tabla de errores (tabla 1).

Como se muestra en la tabla 1, el modelo 4 resultó ser mejor en la mayoría de los parámetros evaluados, por lo que se considera el modelo de mejor ajuste. En cuanto al análisis del sesgo del modelo, se puede decir que tendió a dar valores algo menores del real, sobre todo en los picos de máxima producción. Esto, en las condiciones de la producción de leche en Cuba, unido a la poca diferencia que se tiene entre el pronóstico y los valores reales, se considera de muy poca importancia.

En la tabla 2 se muestran los parámetros del modelo elegido. Los valores P de los términos MA (1), MA (2) y MA (3) son estadísticamente significativos al 95 % (menores que 0.05), y el del AR (1) al 90 %. La desviación típica estimada del ruido blanco (error) es igual a 13.88.

Para validar el modelo se realizaron las pruebas de aleatoriedad (prueba de rachas) y normalidad (Shapiro-Wilks modificado) de los residuos (figura 7). Estas mostraron que los residuos de la serie de producción

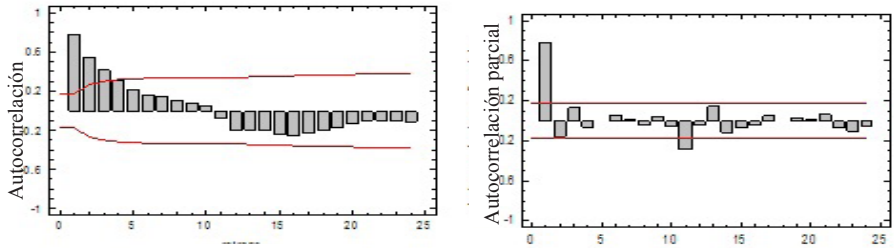


Figura 6. Funciones de autocorrelación y autocorrelación parcial para la serie de producción de leche diferenciada.

Tabla 1. Comparación de modelos probados para la serie de producción de leche

Modelo	Raíz del Error Cuadrático Medio (RMSE)	Error Medio Absoluto (MAE)	Error Medio Absoluto Porcentual (MAPE)	Error Medio (ME)	Error Medio Porcentual (MPE)
(1)	22.42	17.56	25.56	-0.00	-6.14
(2)	15.51	11.68	16.66	-0.03	-3.46
(3)	14.08	10.79	14.84	-0.08	-2.64
(4)	13.82	10.55	14.12	-0.14	-2.33
(5)	19.42	15.02	22.47	-0.00	-3.64

Tabla 2. Parámetros del modelo ARIMA (1,0,3) x (0,1,0) 12 con constante

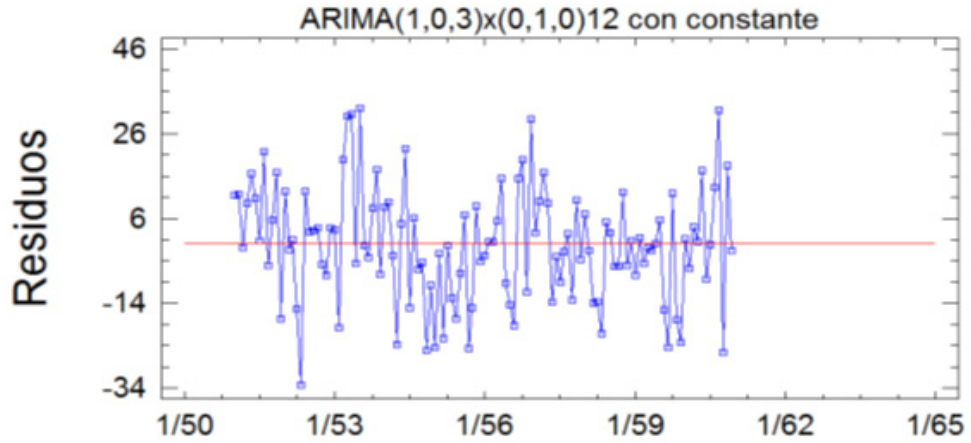
Parámetros	Estimado	Error Estd.	t	Valor-P
AR(1)	0.00	0.00	-1.73	0.09
MA(1)	-0.99	0.09	-10.99	0.00
MA(2)	-0.53	0.12	-4.48	0.00
MA(3)	-0.20	0.09	-2.14	0.03
Media	-1.59	3.44	-0.46	0.65
Constante	-1.59			

de leche por el modelo elegido fueron normales e incorrelacionados. Además, no se observaron cambios de media o varianza mantenidos en el tiempo. Resultados similares a estos encontró Castañeda *et al.* (2006) en un estudio del comportamiento productivo de ganado cruzado europeo por cebú en clima tropical, en Baja California, México. Resultado muy similar obtuvo Colás *et al.* (2011) en una evaluación epidemiológica de procesos respiratorios bacterianos en gallinas ponedoras, en Cuba.

Se pronosticó entonces la producción de leche para el año 2011 por el modelo elegido, con banda de confianza para el pronóstico, con nivel de confianza de 95 % (figura 8).

En la tabla 3 se muestran los valores reales y pronosticados de la producción de leche para el año 2011, los que evidencian que no existieron variaciones significativas entre el pronóstico del 2011.

La producción de leche en la UBPC se comportó con



Shapiro-Wilks modificado (residuales)
P-valor = 0.34
Prueba de Rachas (residuales)
P- valor = 0.53

Figura 7. Análisis de cumplimiento de supuestos en los residuales del modelo ARIMA ajustado

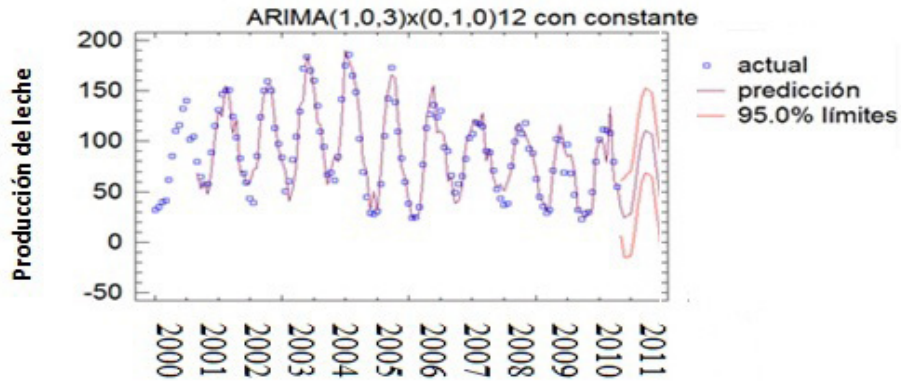


Figura 8. Pronóstico de producción de leche para el 2011-2013

Tabla 3. Valores reales y pronóstico de la producción de leche para el año 2011

Año 2011	Real	Pronóstico
Enero	36530	37208
Febrero	25500	25678
Marzo	29650	28225
Abril	30340	31370
Mayo	49870	48974
Junio	79170	85769
Julio	101190	110039
Agosto	111090	118844
Septiembre	106960	110453
Octubre	106960	110468
Noviembre	78390	77282
Diciembre	53940	60116

estacionalidad marcada, relacionada con los períodos húmedo y poco lluvioso del año, característica que se debe repetir en cualquier otra UBPC de Cuba, por sus peculiaridades y por el clima de la Isla. Con la aplicación de los modelos ARIMA, se obtuvo el pronóstico de la producción de leche en esta UBPC para el año 2011, con errores absolutos porcentuales por debajo del 15 %, lo que se considera una buena aproximación. La comparación con los reales del 2011, obtenidos con posterioridad, confirmaron la bondad del pronóstico y la posibilidad de prever el comportamiento de la serie en años próximos. Con la modelación ARIMA, se lograron realizar pronósticos de producción de leche en las UBPC, y se demostró que con su aplicación se contribuye a tomar decisiones en el sector agropecuario en Cuba.

Referencias

Aguirre, A.J. 1994. Introducción al tratamiento de series temporales. Aplicación a las Ciencias de la Salud. Departamento de Epidemiología y Bioestadística. Escuela Nacional de Sanidad. Ed. Díaz de Santos, Madrid. España. 585 p.

Arellano, M. 2001 Introducción al análisis clásico de series de tiempo. Disponible: http://ciberconta.unizar.es/LECCION/seriest/100.html. [Consultado:12/01/2012]

Blumberg, J.C.1984. Commnets on a simplified time-series analysis for evaluating treatment intervention. J. Applied Behavior Analysis. .17: 539-542

Box, G. E. P. & Jenkins, G. M. 1970. Time Series Analysis, Forecasting and Control. Ed. Holden-Day. San Francisco

Castañeda, M. O.G., Ávila, D. A. & Vinay, V. J. C. 2006. Comportamiento productivo de ganado cruzado europeo por cebú en clima tropical. XVI Reunión Internacional sobre Producción de Carne y Leche en Clima Cálido. Baja California, México. p. 195-199

Colás, M., Chávez, M. & Lamazares, D. 2011. Evaluación epidemiológica de procesos respiratorios bacterianos en gallinas ponedoras. Revista de Salud Animal. 33:69

Coutín, M.G. 2003. Las series temporales. Trabajo de Diploma. Facultad de Contabilidad y Finanzas. Universidad de La Habana. Cuba. 133 pp.

Einsteinnet, E. 2001. Análisis clásico de series temporales. Disponible: http://www.einsteinnet.com/econometria/seriestemp/acseriestemp. [Consultado: 25 /06/2010]

Funes, M., Lantiga, M. & van Keulen, H. 2009. Agro-Ecological Indicators (AEIs) for dairy and mixed farming systems classification: identifying alternatives for the Cuban livestock sector. J. Sustainable Agriculture. 33:435

Grubb, H. 1992. Multivariate Time Series Analysis of Some Flour Price Data. Appl. Statistics. 41:95

Maté, C.1995. Curso General Sobre STATGRAPHICS.

Procedimientos. Métodos Estadísticos. Aplicaciones. Ejercicios Resueltos. Tomo II. Editorial Universidad Pontificia Comillas. Madrid, España.

Maté, C. 2011. A multivariate analysis approach to forecasts combination. Application to Foreign Exchange (FX) markets. Revista Colombiana de Estadística. 34:347

Torres, V., Barbosa I., Meyer, R., Noda, A. & Sarduy, L. 2012. Crieteria of goodness of fit test in the selection of non-linear models for the description of biological performances. Cuban J. Agric. Sci. 46:345

Recibido: 15 de febrero de 2013