



APLICACIÓN DEL ANÁLISIS DE SERIE DE TIEMPO AL REGISTRO DEL NÚMERO DE ANIMALES DEL PROGRAMA PARA LA ERRADICACIÓN Y PREVENCIÓN DEL GUSANO BARRENADOR (*Cochliomyia hominivorax*) DEL GANADO

⁽¹⁾**Ana María Concepción Sánchez,** ⁽²⁾**Gonzalo I. Carrasco O,** ⁽³⁾**Rafael Camarena A.**

⁽¹⁾Universidad de Panamá, Facultad de Ciencias Naturales Exactas y Tecnología, Escuela de Estadística, E-mail: azulplateado1100@yahoo.com

⁽²⁾ Universidad de Panamá, Facultad de Ciencias Naturales Exactas y Tecnología, Departamento de Estadística. Centro de Investigación y Consultoría Estadística. E-mail: gcarrasco27@hotmail.com

⁽³⁾ Universidad de Panamá, Facultad de Ciencias Naturales Exactas y Tecnología, Departamento de Estadística.

RESUMEN

Este trabajo presenta la aplicación del análisis de serie de tiempo al número de animales inspeccionados diariamente entre los meses de enero a octubre de 2004, en 6 (seis) de los 7 (siete) puestos del departamento Control de Movilización de Animales (CMA) del programa de erradicación y prevención del gusano barrenador del ganado de COPEG (Comisión Panamá-Estados Unidos para La Erradicación y Prevención del Gusano Barrenador del Ganado). En el estudio se utiliza la metodología desarrollada por Box-Jenkins para ajustar a una serie un tipo especial de modelos, denominados ARIMA (Autorregresive Integrated Moving Average). Los resultados mostraron que el número de animales inspeccionados diariamente presentan una media constante a través del tiempo, es decir que en los meses estudiados el registro promedio diario es aproximadamente el mismo, no ocurre un crecimiento o disminución considerable que indique algún cambio en la fluidez de animales inspeccionados en estos puestos de control; también esta serie resultó estar dominada por un componente estacional, con una periodicidad aproximada de 7 días, cuya ecuación de pronóstico se determinó.

PALABRAS CLAVES

Gusano Barrenador (*Cochliomyia hominivorax*), correlograma parcial y simple, media móvil, modelos ARIMA.

ABSTRACT

This work shows the application of the analysis of time series to the number of animals inspected daily between January and October of 2004, in 6 (six) out of 7(seven) posts of the department of Control of Animals Mobilization (CMA), which is part of the eradication program and prevention of the barrenador worm of the COPEG cattle (Panama- United States Commission for the Eradication and Prevention of the cattle worm (*Cochliomyia hominivorax*) program. The methodology developed by Box-Jenkins is used in order to fit to a series a special type of models, named ARIMA (Autorregresive Integrated Moving Average). The results showed that the number of animals inspected daily presents an average constant through time, meaning that in the studied months the daily registry average is approximately the same, does not happen a growth or considerable diminution that may indicate some change in the animal fluidity inspected in these control posts; also this series turned out to be dominated by a seasonal component, with an approximated regularity of 7 days.

KEYWORDS

Barrenador Word (*Cochliomyia hominivorax*), partial and simple correlogram, moving average ARIMA models.

INTRODUCCIÓN

Los animales que son trasladados dentro y fuera de la República de Panamá, como medida sanitaria, deben ser controlados mediante inspección, para certificar que tengan las autorizaciones reglamentarias para el movimiento como animales sanos y libres de cualquier patología. El departamento de Control de Movilización de Animales (CMA) es el encargado de las inspecciones. Es parte estructural de la comisión encargada de Erradicar y Prevenir el Gusano Barrenador del Ganado (COPEG). (COPEG 2007). También inspecciona los animales que se encuentran en subastas ganaderas fijas y temporales del país, y la verificación de la llegada y sacrificio de especies sensibles a la fiebre aftosa procedentes de la provincia de Darién a los mataderos oficiales. Este programa forma parte de la planificación, que a nivel regional de Norte a Centroamérica ha venido realizando el Gobierno de Estados Unidos con apoyo de cada uno de los países beneficiados. Inicia en 1994 cuando se firma el Acuerdo Cooperativo entre el Ministerio de

Desarrollo Agropecuario (MIDA) y el Departamento de Agricultura de los Estados Unidos (USDA) para la Erradicación y Prevención del Gusano Barrenador del Ganado (COPEG).

Teniendo como premisa la necesidad de llevar a cabo un análisis de serie de tiempo al número de animales inspeccionados en todos los puestos de control de movilización de animales y con el propósito de optimizar los escasos recursos del CMA, se llevó a cabo este estudio que permite pronosticar los animales a inspeccionar. Se utilizaron modelos ARIMA (Proceso de promedio móvil autorregresivo integrado) para determinar la ecuación de pronóstico según la metodología desarrollada por Box-Jenkins (Camarena, 1997 & Maddala, 1996) y se utilizaron pruebas estadísticas Kruskal-Wallis y Levene para la comprobación de la homogeneidad de varianza e igualdad de media. (Newbold, 1997 y Richard & Dennis, 1986).

MATERIALES Y METODOLOGÍA

La información utilizada en el estudio fue proporcionada por el departamento de Control de Movilización de Animales (CMA). La serie estudiada cuenta con 305 observaciones diarias de inspecciones hechas en los diez (10) primeros meses del año 2004. Cada observación es el resultado de la suma total del número de animales inspeccionados en los seis (6) puestos que se encuentran ubicados en la Provincia de Panamá, Colón y Darién.

$X(t)$ = Serie compuesta por la variable número de animales inspeccionados diariamente en los puestos de inspección.

t = dato diario ($t = 1, 2, \dots, 305$).

Formulación de la serie

$$X_t = \sum_{j=1}^6 X_{jt}$$

j = Puestos (1, 2, 3, 4, 5, 6).

t = Días (1, 2, 3, 4, , 305).

X_{1t} = Serie de animales inspeccionados en el puesto de La "Y" de Colón

X_{2t} = Serie de animales inspeccionados en el puesto de Capira .

X_{3t} = Serie de animales inspeccionados en el puesto de Gatún.

X_{4t} = Serie de animales inspeccionados en el puesto de Agua Fría.

X_{5t} = Serie de animales inspeccionados en el puesto de Platanilla.

X_{6t} = Serie de animales inspeccionados en el puesto de Palmas Bellas

En el análisis predictivo se utilizó la técnica estadística de Box-Jenkins que es una de las metodologías de uso más amplio para el análisis de las series de tiempo, debido a su generalidad ya que puede manejar cualquier serie estacionaria o no, con o sin elementos estacionales (G.S. Maddala, 1996). Este modelo se denota por la expresión: ARIMA (p, d, q) (en caso no estacional), y es una combinación de técnicas autorregresivas (AR (p)) y medias móviles (MA(q)) aplicadas sobre datos no diferenciados ($d = 0$) o diferenciados ($d > 0$) como el diferenciado de primer o segundo orden (∇^d ; $d = 1$ ó 2), $d = 1$ cuando la tendencia es lineal y $d = 2$ cuando se tiene tendencia por ejemplo cuadrática. O bien, diferenciado estacional (∇_s), para atenuar algún efecto estacional de período s . En el caso estacional los modelos Box-Jenkins como: ARIMA (p, d, q)* (P, D, Q) s , en donde el primer paréntesis se expresa la parte regular de la serie y el segundo paréntesis se refiere a la parte estacional (Saijas, 2003).

Los pasos a seguir en la metodología de Box-Jenkins son los siguientes (Cornejo, 2002):

1. Primeramente se estudia descriptivamente la serie para determinar los componentes predominantes tendencia, ciclo, estacional y aleatorio o irregular; en este punto también se verificará si la serie es estacionaria que no más que la condición de media y varianza constante. Para la comprobación de estacionalidad en media se utilizará la prueba Kruskal-Wallis $H = \frac{1}{\hat{S}^2} \left[\sum_{i=1}^I \frac{R_i^2}{n_i} - \frac{n(n+1)^2}{4} \right]$, alternativa no paramétrica al ANOVA, si la suposición de normalidad no es asumible, esta prueba contrasta la hipótesis nula de que las n muestras independientes proceden de la misma población y, en particular, todas ellas tienen la misma esperanza (medias). En caso de la varianza se hará uso del estudio gráfico de la dispersión del número de animales inspeccionados por día, la cual visualmente nos permitirá determinar su estacionaridad.
2. En el segundo paso se continúa con la identificación de la estructura del modelo al comprobar la no estacionaridad, en tal caso se aplicará el diferenciado que fuese necesario para estabilizar la varianza, la cual una vez diferenciada se verificará a través de la

$$SC_{intra} = \sum_{i=1}^n \sum_{j=1}^k (D_{ij} - \bar{D}_j)^2, \quad SC_{inter} = \sum_{j=1}^k n_j (\bar{D}_j - \bar{D}_t)^2,$$

prueba de Levene que es la que emplea SPSS y robusta al supuesto de normalidad, esta consiste en llevar a cabo un análisis de varianza de un factor utilizando como variable dependiente la diferencia en valor absoluto entre cada puntuación individual y la media (o la mediana, o la mediana recortada) de su grupo (Pérez L., 2001); en el caso que se necesite estabilizar la (media) se hará a través de las transformación Box-Cox. Para convertir el proceso subyacente en estacionario y determinar la estructura del proceso estacionario resultante, examinando los correlograma (simple y parcial) para decidir los órdenes apropiados de los componentes AR (proceso autorregresivo), MA (proceso de promedio móvil) y ARMA (proceso de promedio móvil autorregresivo). En base a esto, se llega a un modelo ARIMA tentativo a utilizar.

3. Luego del segundo paso se procede a estimar los parámetros del modelo ARIMA tentativo identificado.
4. Seguido de la verificación del diagnóstico del modelo, que consiste en comprobar si los errores del modelo son una realización de un ruido blanco, es decir, se comprobará que los errores tienen media cero, que son estables en varianza, que sus observaciones están incorrelacionadas y que proceden de una distribución normal.
5. Por último, se utilizará el modelo de ecuación que mejor se ajuste a la serie de datos para pronosticar.

La comprobación de la homogeneidad de varianza e igualdad de media se realiza mediante pruebas estadísticas de Kruskall-Wallis y Levene siguiendo una hipótesis con los siguientes planteamientos:

Primer Planteamiento:

H0: No existe diferencia significativa en el promedio de animales inspeccionados diariamente.

H1: Existe diferencia significativa en al menos dos promedios de animales inspeccionados diariamente.

Segundo Planteamiento:

H0: No existe diferencia significativa en la variabilidad del número de animales inspeccionados diariamente.

H1: Existe diferencia significativa en la variabilidad del número de animales inspeccionados diariamente.

Tercer Planteamiento:

H0: La varianza de los residuos del modelo es igual para los siete días.

H1: La varianza de los residuos del modelo es distinta para al menos uno de los siete días.

RESULTADOS Y DISCUSIÓN

Estudio Descriptivo

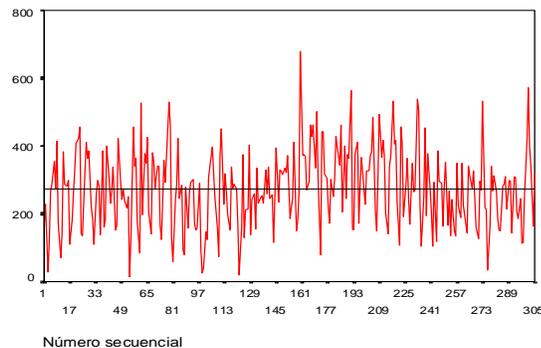


Fig. 1. Número de animales inspeccionados diariamente. Enero - octubre 2004.

Tendencia: Como podemos observar en la Figura 1 de la serie estudiada, su aspecto no manifiesta evolución sostenida ni hacia arriba ni hacia abajo o curva evidente que muestre alguna tendencia específica; en su lugar se observa un comportamiento oscilatorio casi uniforme acompañado de mucha irregularidad, de modo que los datos con respecto al tiempo se comportan casi igual a lo largo de la serie.

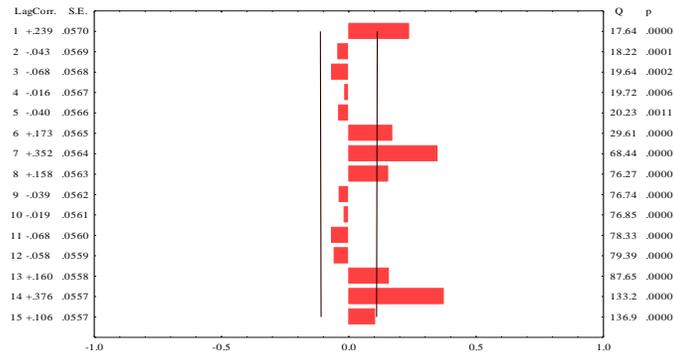


Fig. 2. Correlograma simple del número de animales inspeccionados diariamente. Enero-octubre 2004.

Estacional: En la Figura 2, indica valores elevados para $k = 1, 6, 7, 8, 13$ y 14 , esto significa que las correlaciones son significativas, pero específicamente los rezagos $7, 14, \dots, k*7$ registran una estacionalidad a 7 días.

Aleatoriedad o Irregularidad: La sucesión de barras se muestran en ambos lados del correlograma, es decir positivamente y negativamente con respecto al eje donde el valor de correlación es cero, lo cual nos demuestra un grado de aleatoriedad en los datos.

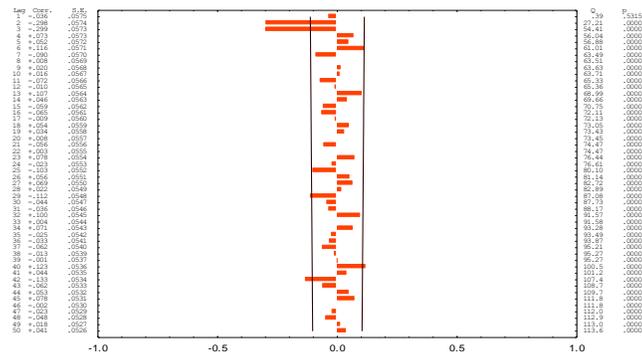


Fig. 3. Correlograma simple de la serie ajustada estacionalmente.

Cíclico: En la Figura 3, se observa que las barras 2 y 3 resultan significativas a diferencia del resto que se mantienen dentro de la banda de confianza, indicando así que la participación del efecto cíclico no es importante; sin embargo sabemos que la serie cuenta con

relativamente pocos datos (8 meses) y el componente cíclico se observa en periodos largos de dos o más años; por consiguiente no hay evidencia suficiente para sustentar la participación del componente cíclico, de modo que se descarta del modelo.

Tabla 1. Resultado de la prueba Kruskal Wallis del número de animales inspeccionados diariamente: enero-octubre 2004.

Estadísticos de contraste ^{a,b}

	NÚMERO
Chi-cuadrado	109.645
gl	6
Sig. asintót.	.000

a. Prueba de Kruskal-Wallis

b. Variable de agrupación: DAY, período

La Tabla 1 nos indica que la probabilidad del estadístico calculado es menor que el nivel de significación (P_c es un valor cercano a $0 < Pr .05$) por lo tanto se rechaza la hipótesis nula y se concluye que existe diferencia significativa en al menos uno de los promedios de animales inspeccionados diariamente, por lo tanto la serie no es constante en media.

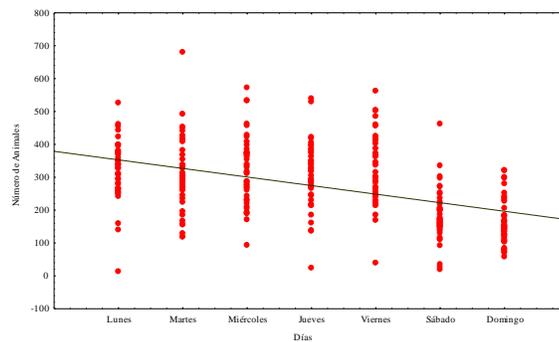


Fig. 4. Dispersión del número de animales inspeccionados diariamente por día. Enero - octubre 2004.

Observando la Figura 4, la nube de puntos de cada uno de los días no están uniformemente espaciados y comparativamente no están a la misma altura o igualmente distribuidos, específicamente los días

sábados y domingos los dos son mucho más bajos, lo que nos indica que la serie no es constante en varianza.

Por medio del análisis hecho a la serie en los puntos anteriores, ésta resulto ser no estacionaria debido a que no es constante en media ni en varianza, producto del efecto del componente estacional.

Identificación de la Estructura del Modelo

Como la serie no es estacionaria se hará uso de las transformaciones Box-Cox ($\ln X_t =$ Logaritmo natural, $X^2(t) =$ Potencia al cuadrado, $\sqrt{X_t} =$ Raíz cuadrada) para estabilizar la varianza; a continuación se le aplicará un diferenciado estacional a 7 períodos, para atenuar el componente estacional que se da cada 7 días y estabilizar la media.

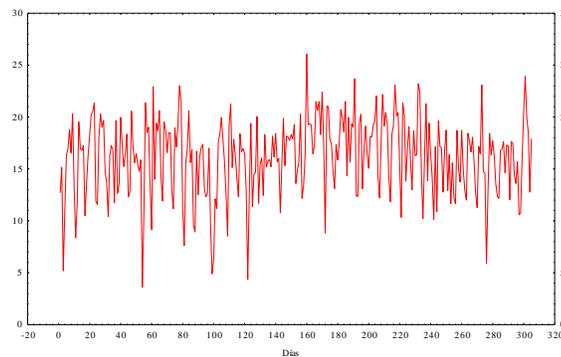


Fig. 5. Transformación del número de animales inspeccionados diariamente a raíz cuadrada $x^{.5}$: enero-octubre 2004.

Tabla 2. Prueba de Levene para la homogeneidad de varianza de la transformación Box-Cox del número de animales inspeccionados diariamente a raíz cuadrada: enero - octubre 2004.

Prueba de homogeneidad de varianzas

RAÍZ

Estadístico de Levene	gl1	gl2	Sig.
.291	6	298	.941

De las tres transformaciones realizadas, la raíz cuadrada fue la que mejor estabilizó la varianza respaldado por la prueba de Levene, la cual nos indica que la probabilidad del estadístico calculado es mayor que el nivel de significación ($P_c 0.0941 > Pr 0.05$) aceptándose la igualdad de varianza.

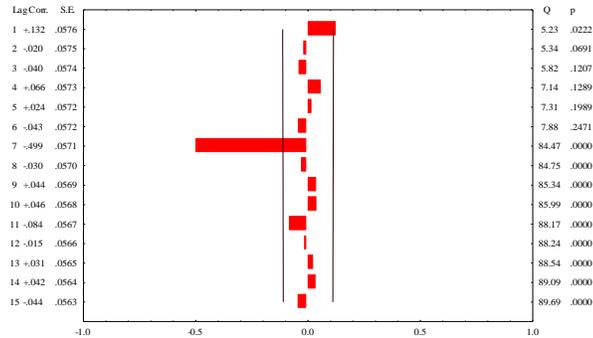


Fig. 6. Correlograma simple de la transformación del número de animales inspeccionados diariamente a raíz cuadrada con un diferenciado de orden 7 $x^{.5}$; D(-7): enero - octubre 2004.

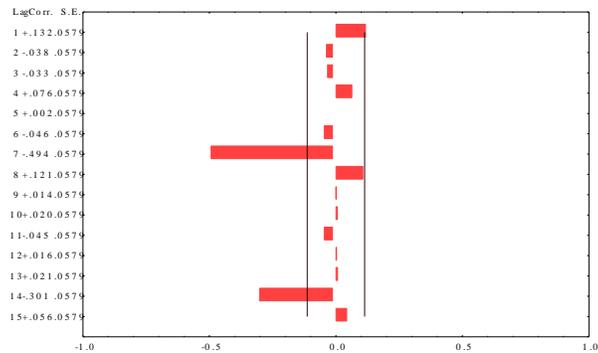


Fig. 7. Correlograma parcial del número de animales inspeccionados diariamente a raíz cuadrada con un diferenciado de orden 7 $x^{.5}$; D(-7): enero - octubre 2004.

Al comparar el correlograma simple y parcial de la serie transformada: $\sqrt{x_{(t)}}$; D(-7), con los correlogramas teóricos de ARIMA se identifican los modelos Media Móvil estacional puro: MA

(0, 1, 1)⁷ y MA (0, 1, 2)⁷. Luego de hacer la validación, resultó mas adecuado el modelo: MA (0,1,1)⁷.

Estimación de los Parámetros del Modelo

Tabla 3. Parámetros del modelo ARIMA del número de animales inspeccionados diariamente. Enero-octubre 2004.

Input:NÚMERO DE ANIMALES (new.sta)						
Transformations: x [^] .5,D(7)						
Model:(0,0,0)(0,1,1)		Seasonal lag: 7			MS Residual=10.592	
Parámetro	Param.	Asympt. Std.Err.	Asympt. t(296)	p	Lower 95% Conf	Upper 95% Conf
Constant	0.038473554	0.041075382	0.936657248	0.349698456	-0.042363239	0.119310346
Qs(1)	0.79896044	0.041888965	19.0732912	un valor cercano a 0	0.716522489	0.881398355

En la Tabla 3 se resume el análisis estadístico en la estimación del parámetro modelo.

- La Constante del Modelo \longrightarrow ($\mu = 0.03847354$). Como se observa el valor de la constante del modelo no es significativa, por consiguiente se descarta de la ecuación de pronósticos.
- El Valor del Parámetro Estacional \longrightarrow ($\theta = 0.79896044$). Como se puede observar el parámetro estimado es significativo, por consiguiente se incluye en la ecuación de pronóstico.

El modelo ARIMA de predicción para la serie del Número de Animales Inspeccionados está dado por (Camarena A., 1997):

$$x'_t - x'_{t-7} = e_t - \theta e_{t-7}$$

$$x'_t - x'_{t-7} = e_t - 0.79896044 e_{t-7}$$

donde;

$$x'_t = \sqrt{x_t}$$

e_t = se obtiene de un proceso ruido blanco con distribución normal de varianza σ^2

θ = parámetro estimado de Q (1)

t = tiempo

Validación del Modelo

En este punto se debe comprobar si los residuos del modelo son una realización de un proceso de ruido blanco.

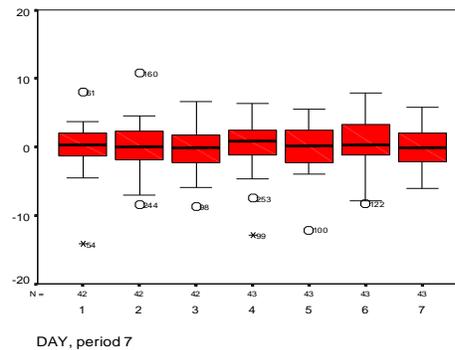


Fig. 8. Residuos del modelo ARIMA (0,0,0)(0,1,1). Enero - octubre 2004.

El gráfico de caja de los residuales (Fig.8), indica que las medias para cada uno de los días es aproximada igual a cero, esto sugiere medias residuales nulas.

Tabla 4. Prueba de Levene para la homogeneidad de varianza de los residuos del modelo: enero-octubre 2004.

Prueba de homogeneidad de varianzas

RESIDUOS			
Estadístico de Levene	gl1	gl2	Sig.
.276	6	291	.948

En la Tabla 4 se observa que los residuos del modelo, resultaron con una probabilidad mayor que el nivel de significación ($P_c 0.948 > P_r 0.05$) por consiguiente se acepta la hipótesis nula de que la varianza de los residuos son iguales para los siete días, por lo tanto los residuos son de varianza constante.

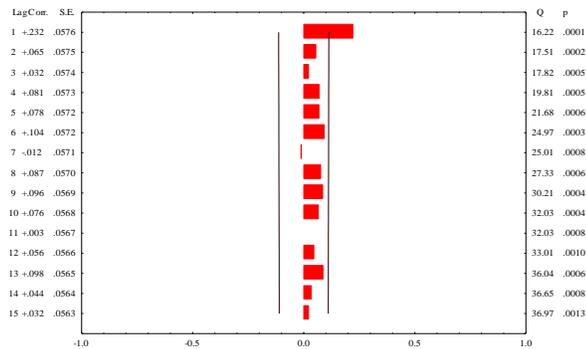


Fig. 9. Correlograma simple de los residuos del modelo ARIMA (0,0,0)(0,1,1): Enero - octubre 2004.

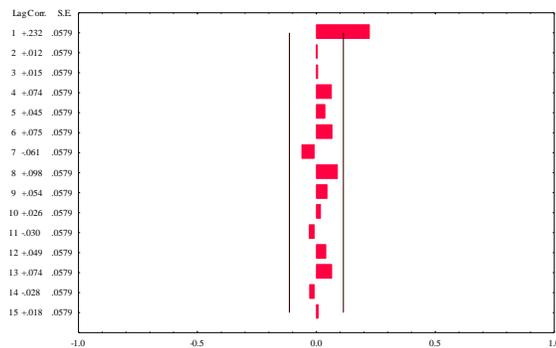


Fig. 10. Correlograma parcial de los residuos del modelo ARIMA (0,0,0)(0,1,1): Enero - octubre 2004.

En las Figuras 9 y 10 los residuos del modelo se observa que después del primer rezago se cae abruptamente a cero y los demás rezagos se mantienen dentro de la banda de confianza, esto demuestra que sus observaciones están incorrelacionadas, por lo tanto la serie de errores es aleatoria.

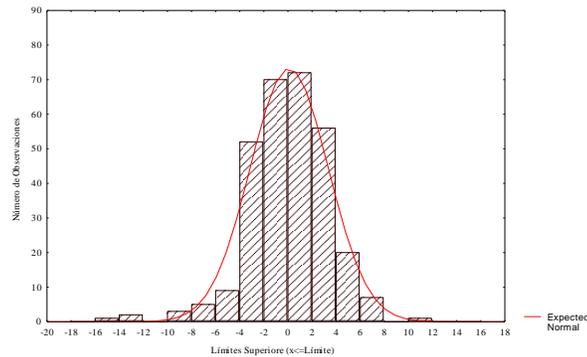


Fig. 11. Histograma de los residuos del modelo ARIMA (0,0,0)(0,1,1): Enero-octubre 2004.

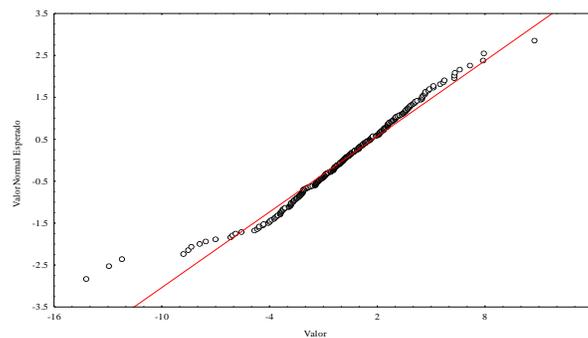


Fig. 12. Probabilidad normal de los residuos del modelo ARIMA (0,0,0)(0,1,1): Enero-octubre 2004.

En la Figura 11, muestra que los residuos del modelo parecen estar distribuidos normalmente. También la Figura 12 de la probabilidad normal, parece sugerir lo mismo exceptuando algunos datos atípicos que podrían resultar de errores de medición, por consiguiente los residuos provienen de una distribución normal.

El análisis hecho a los residuales, nos comprueban que es un proceso de ruido blanco y el modelo es adecuado para pronosticar.

Pronóstico

Para los datos diarios de la serie número de animales inspeccionados diariamente tenemos que:

$$x_t = \nabla_7 x'_t = (1 - B^7) x'_t = x_t - x_{t-7} = x'_t - x'_{t-7} = e_t - \theta e_{t-7} = x'_t = x'_{t-7} + e_t - \theta e_{t-7}$$

Ecuación (a) \longrightarrow $x'_t = x'_{t-7} + e_t - \theta e_{t-7}$

donde \longrightarrow x'_t es la transformación Box-Cox $\sqrt{x_t}$

\hat{x} = Pronóstico

$\hat{x}_{t+h} = \hat{x}_t(h)$

t = Primer subíndice proporciona el origen del pronóstico.

h = Segundo subíndice proporciona el horizonte del pronóstico

Pronósticos a corto plazo

Utilizando la ecuación (a) se pronosticará, para h = 1

$$\hat{x}'_{305+1} = \hat{x}'_{305}(1) = x'_{305-7} + e_{305} - \theta e_{305-7}$$

$$\hat{x}'_{306} = x'_{298} + (x'_{305} - \hat{x}'_{305}) - 0.79896044(x'_{298} - \hat{x}'_{298})$$

$$\sqrt{\hat{x}'_{306}} = \sqrt{116} + \sqrt{321} - \sqrt{148} - 0.79896044(\sqrt{116} - \sqrt{155})$$

$$\left(\sqrt{\hat{x}'_{306}}\right)^2 = (15.17774646)^2$$

$$\hat{x}_{306} \approx 230 \text{ (Newbold, P., 1997)}$$

CONCLUSIONES

El número de animales inspeccionados diariamente se presenta con un comportamiento esencialmente constante a través del tiempo, es decir, en el periodo de estudio el registro promedio diario es similar, no ocurre un crecimiento o disminución considerable que indique algún cambio en la cantidad de animales inspeccionados.

La serie resultó estar dominada por el componente estacional, el cual es un comportamiento periódico de 7 días. Por otro lado, el promedio de animales inspeccionados para los días de la semana (lunes, martes, miércoles, jueves y viernes) es relativamente más alto que los fines de semana (sábado y domingo).

Respecto a la validación del modelo este resulto un modelo con componente estacional ARIMA (0, 1, 1)₇ y cuya ecuación de pronostico a corto plazo es $x'_t - x'_{t-7} = e_t - 0.79896044e_{t-7}$

REFERENCIAS

Camarena A., R. 1997. Monografía: Algunos Modelos de Predicción para Series de Tiempo Univariadas. Universidad de Panamá, Facultad de Ciencias Naturales Exactas y Tecnología, Departamento de Estadística. 28pp.

COPEG. 2007. En <http://www.copeg.org/> visitada en diciembre de 2007.

Cornejo, G. 2003. Proyección de Pasajeros para Lan Chile en la Ruta Santiago Isla de Pascua a través de un Modelo Arima [Internet] Disponible en: <http://www.monografias.com/trabajos12/proyisla/proyisla.shtml>.

Maddala, G.S. 1996. Introducción a la Econometría. Segunda Edición. Editorial. 715 pp.

Newbold, P. 1997. Estadística para los Negocios y la Economía. Cuarta Edición. 752pp.

Pérez López, C. 2001. Técnicas Estadísticas con SPSS. 571pp.

Seijas, C. 2003. Modelo Estocástico de la Serie de Tiempo Económica “ Inflación en Venezuela (Junio 95 a Julio2000) [Internet] Disponible en: <http://servicio.cid.uc.edu.ve/ingenieria/revista/a9n1/9-1-6.pdf>.

William, M., L. Richard & D. Dennis. 1986: Estadística Matemática con Aplicaciones. 629 pp.

Recibido junio de 2006, aceptado diciembre de 2007.