

5 Aplicación de la modelación ARIMA a un puente instrumentado

5.1 El caso de un puente en curva

La Secretaría de Comunicaciones y Transportes (SCT) construyó la carretera de cuota Tuxtla Gutiérrez – San Cristóbal de las Casas, la cual beneficia en forma directa a 275 mil habitantes de los municipios de Tuxtla Gutiérrez, Chiapa de Corzo, Ixtapa, Zinacantán y San Cristóbal de las Casas, este último considerado como el principal polo de desarrollo de la zona de Los Altos de Chiapas, beneficiando además al turismo nacional e internacional que visita los atractivos del Mundo Maya y a un flujo vehicular estimado del orden de 5,000 vehículos/día.

El proyecto carretero, tiene una longitud de 59.2 km., lo que significa un ahorro en tiempo de 60 minutos, al bajar el recorrido de 1 hora 45 minutos a 45 minutos, con mejores niveles de seguridad (SEMARNAT).

Dentro del sistema carretero que comunica la ciudad de Tuxtla Gutiérrez con San Cristóbal de las Casas, se encuentra el Puente San Cristóbal (*Figura 5.1*), el cual da continuidad a la obra al cruzar la barranca Chentic. Debido a la alineación geométrica de la carretera, profundidad de la barranca y tipo de superestructura, el puente San Cristóbal es la estructura más importante de toda la carretera.

El Puente San Cristóbal se ubica entre los Km. 24+766 al 25+090 de la carretera Tuxtla Gutiérrez –San Cristóbal de las Casas a una altura de 1600 m.s.n.m, sobre la barranca del Arroyo Chentic (*Figura 5.2*) (SEMARNAT).



Figura 5.1 Puente San Cristóbal



Figura 5.1 Ubicación Puente San Cristóbal

El puente cuenta con una longitud de 324.00 metros y un ancho total de 13.64 metros, con una pendiente longitudinal del 5% y pendiente transversal del 10% con un grado de curvatura de $2^{\circ}45''$ para una velocidad de 110 km/hrs. (Figura 5.3 y Figura 5.4).

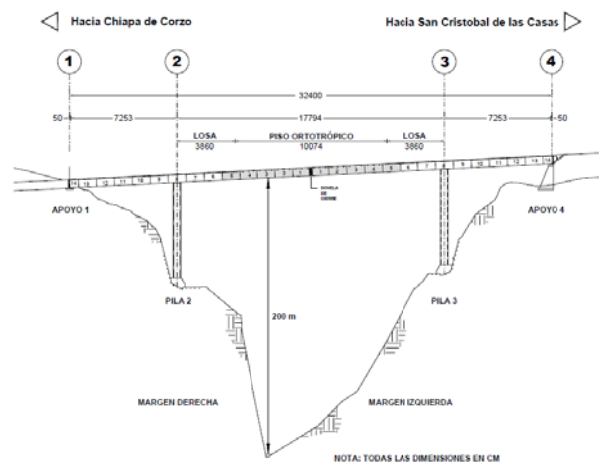


Figura 5.3 Vista en elevación del Puente San Cristóbal

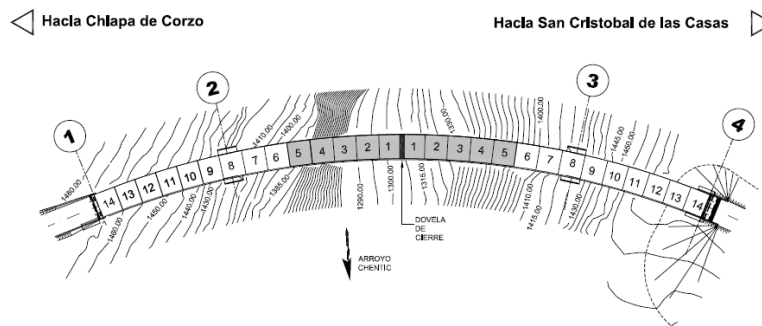


Figura 5.4 Vista en planta del Puente San Cristóbal

La obra se divide en dos conceptos generales: subestructura y superestructura.

La subestructura está compuesta de cuatro apoyos: dos estribos de concreto reforzado convencionales (Estribo 1 y Estribo 4) y dos pilas del mismo material (Pila 2 y Pila 3), de sección transversal rectangular hueca, aproximadamente de 61m de altura cada una. Los apoyos forman 3 claros: 2 en los lados extremos cada uno de 73.0 m, y 1 claro central de 178.0 m. La profundidad máxima, desde la superestructura al fondo de la barranca, es aproximadamente 200 m.

La superestructura consiste en un cajón de acero Grado 50 compuesto de 3 tramos (72.53 m, 177.94 m, 72.53 m), con una longitud total de 324 m. El ancho de la calzada es 13.64 m incorporando dos carriles de tráfico. La superestructura estará formada a base de 28 dovelas de acero estructural A-709 grado 50, de 12.0 metros de longitud cada una y una losa de concreto reforzado de 350 kg/cm^2 de 30 cm de espesor.

La superestructura se construyó mediante el método de empujado desde ambos lados de la barranca, dando lugar a dos vigas en voladizo. Una mitad de la superestructura se lanzó con pendiente hacia arriba y la otra con pendiente hacia abajo. Cada voladizo se forma con 14 segmentos o dovelas: los primeros 5 segmentos de cada voladizo son ortotrópicos (Figura 5.4), mientras que el resto tienen una losa de concreto postensada dando lugar a una sección compuesta acero-concreto (Figura 5.5).

El empujado de los segmentos se realizó con gatos hidráulicos sobre una cremallera de acero y un chasis de empuje. Además, el sistema de empuje empleado garantizó un proceso de lanzamiento estable y controlado, y permitió el alineamiento horizontal y vertical de los segmentos del puente.

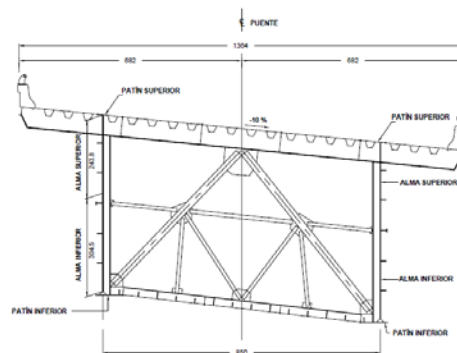


Figura 5.4 Sección transversal con piso ortotrópico

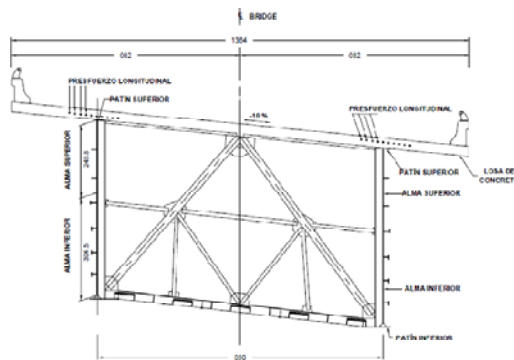


Figura 5.5 Sección transversal con piso de concreto

5.2 Descripción de la instrumentación

A finales de 2002, el Instituto de Ingeniería fue elegido por la SCT como el principal consultor para la construcción del Puente San Cristóbal, para el cual se requirieron de elaborados procesos y sistemas de construcción. Además de estas actividades, el Instituto de Ingeniería participó en la instalación de diferentes tipos de sensores para registro de diversos parámetros, el seguimiento en tiempo real de las condiciones de esfuerzo en diferentes segmentos de los puentes, la construcción de modelos de computadora para calibrar los datos registrados con los diferentes tipos de sensores y la asesoría a ingenieros de la SCT en diversos problemas que se presentan en este tipo de obras (Gómez 2008).

Las actividades del Instituto de Ingeniería en el Puente San Cristóbal se intensificaron a partir del 24 de octubre de 2004, debido a derrumbe de la superestructura correspondiente a la margen derecha (estribo 1 y pila 2), desplomándose siete dovelas y media, con una longitud de 90 m.

Una vez reconstruido el puente y puesto en operación, se implementaron dos tipos de monitoreo: temporal y permanente. El propósito del monitoreo fue hacer un seguimiento de las condiciones de esfuerzo en diferentes dovelas del puente en tiempo real, registrando deformaciones, distribuciones de temperatura y velocidad del viento, entre otros factores que pudieran afectar el comportamiento del puente (Gómez 2009).

La instrumentación implementada en el puente se clasifica en permanente y temporal. La instrumentación permanente se implementó para monitorear los esfuerzos, desplazamientos, deflexiones y temperatura a lo largo del puente; mientras que la instrumentación temporal se usó para pruebas de vibración ambiental.

A fin de identificar la ubicación de la instrumentación, la superestructura se dividió en dos mitades: lado Tuxtla y lado San Cristóbal. Cada mitad compuesta por 14 dovelas, denotadas con la letra D. En la Figura 5.6 se muestra de los segmentos instrumentados.

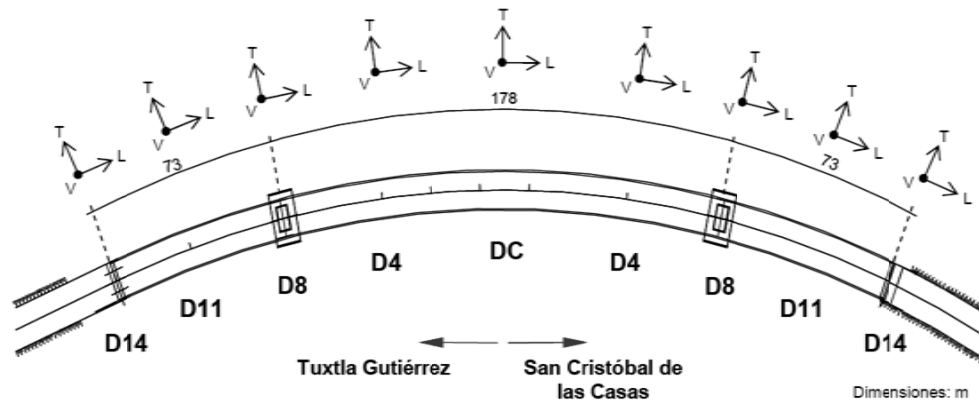


Figura 5.6 Ubicación de las dovelas instrumentadas del puente San Cristóbal (vista en planta)

La instrumentación temporal se compone de (Gómez, 2009):

- 8 servoacelerómetros con una amplitud máxima de medición de 1g, distribuidos en diferentes arreglos sobre la superestructura del puente San Cristóbal y orientados en dirección longitudinal L.
- 4 transductores lineales de desplazamiento.

La distribución de la instrumentación permanente se realizó de la siguiente forma:

- Los acelerómetros fueron colocados en tres segmentos a lo largo del claro central, en las dovelas D4 de ambas márgenes del puente y la dovela central DC. Además se instrumentaron las dovelas D11, D14 y la dovela D8 sobre las pilas 2 y 3.
- Los transductores lineales fueron colocados en los estribos 1 y 4 para medir el desplazamiento en la componente longitudinal L.

La instrumentación permanente se compone de (Gómez 2009):

- 102 cuerdas vibrantes distribuidas a lo largo de la superestructura del puente, orientadas en la dirección longitudinal L (paralelo a la curvatura del puente). Estos sensores registran también temperaturas.
- 6 inclinómetros bidireccionales, distribuidos en el claro central de la superestructura, orientados en dirección L (paralelo a la curvatura del puente) y T (normal a la curvatura del puente).
- 36 deformímetros de resistencia distribuidos en superestructura del puente, orientados en dirección longitudinal L (paralelo a la curvatura del puente).

La distribución de la instrumentación permanente se realizó de la siguiente manera:

- Las cuerdas vibrantes se colocaron en los segmentos D11, D8, D4 de ambos lados de la estructura, así como en el segmento central DC. Además para medir los desplazamientos relativos en la zona de transición (cambio de piso ortotrópico a losa de concreto) se colocaron cuerdas vibrantes en la parte superior de la dovela D6 de ambas márgenes.

- Los inclinómetros bidireccionales se colocaron en las partes superiores de las dovelas D1, D6 y D8 de ambos márgenes del puente, con el objeto de registrar los giros en las direcciones T y L. Con los giros se estimaron las deflexiones del claro central.

La *Figura 5.7* muestra las secciones transversales de las dovelas instrumentadas.

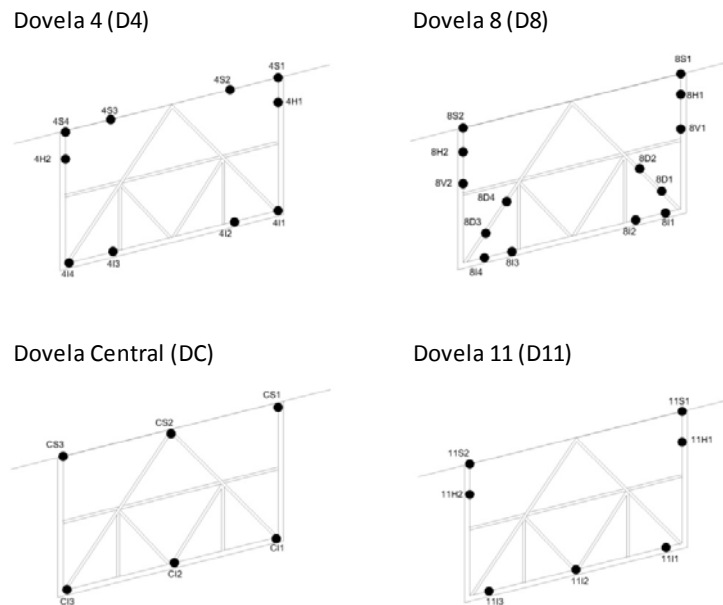


Figura 5.7 Distribución de cuerdas vibrantes en las dovelas más importantes del Puente San Cristóbal

5.3 Elección de la información a utilizar

El monitoreo de los principales parámetros estructurales del puente San Cristóbal mediante la instrumentación temporal y permanente, generó un gran volumen de información. Este trabajo escrito se enfoca a la información producto del monitoreo permanente. Específicamente a la información capturada mediante las cuerdas vibrantes. El parámetro estructural en estudio es el incremento de esfuerzos ($\Delta\sigma \text{ kg/cm}^2$) en diferentes puntos del puente.

Las 102 cuerdas vibrantes generaron registros a partir del 5 de noviembre de 2006 hasta el 2 de enero de 2009. Las primeras mediciones no están separadas entre sí en iguales intervalos de tiempo. No es hasta marzo de 2006 que se registran las mediciones a cada hora.

Con objeto de realizar el análisis de información bajo el enfoque de series de tiempo y en particular bajo la modelación ARIMA, se separaron las mediciones correspondientes a las 8:00 am del 1 de abril del 2007 al 2 de enero del 2009. Esto quiere decir que se conformó una serie de tiempo cuyas observaciones están separadas equidistantemente en el tiempo, cada 24 horas. Hay que recordar que la metodología Box-Jenkins sólo aplica para series de tiempo con intervalos de tiempo discretos.

De forma similar se separaron los datos pertenecientes a las 8:00 pm del 1 de abril del 2007 al 2 de enero del 2009. Las series de tiempo resultantes cuentan con un total de 728 datos.

Posteriormente se graficaron los incrementos de esfuerzos capturados por las 102 cuerdas vibrantes. En la *Figura 5.8* se muestra las series de tiempo de la dovela D4 (parte inferior, lado Tuxtla Gutiérrez) correspondientes a los datos matutinos (8 am). De igual forma se muestran las gráficas (*Figuras 5.9* y *5.10*) de los incrementos de esfuerzos en algunas otras dovelas, las cuales sólo se exponen para fines ilustrativos.

Cabe subrayar que se sometió a corrección algunos registros de esfuerzos, debido a fallas en los instrumentos de captura, los métodos actuales de corrección, quedan fuera del estudio de esta tesis.

Ahora bien, a fin de aplicar la modelación ARIMA y la metodología Box-Jenkins, se seleccionó la información del punto D4-I4 (Dovela 4, punto Inferior 4) (*Figura 5.8*). Este punto se seleccionó debido a que presenta consistencia en las mediciones y requirió mínimas correcciones.

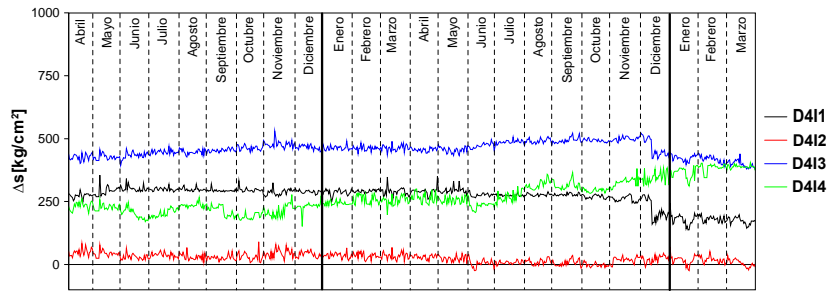


Figura 5.8 Incremento de esfuerzos dovela D4-Inferior lado Tuxtla Gutiérrez (8 am)

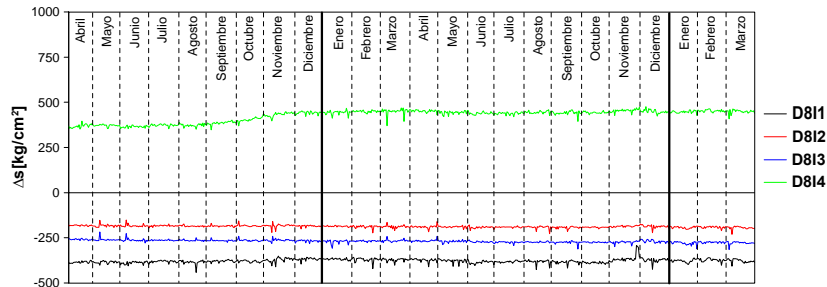


Figura 5.9 Incremento de esfuerzos dovela D8-Inferior lado Tuxtla Gutiérrez (8 am)

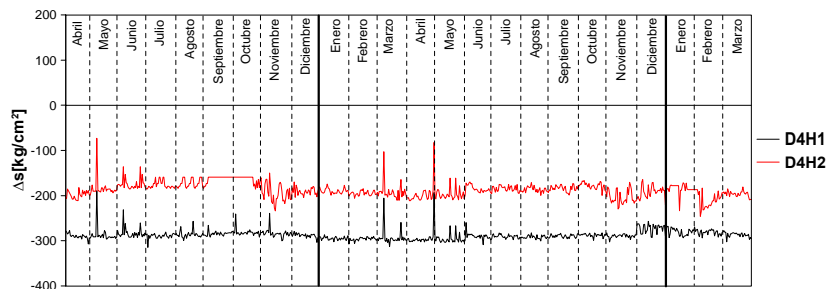


Figura 5.10 Incremento de esfuerzos dovela D4-Horizontal lado San Cristóbal (8 am)

5.4 Ajuste de un modelo ARIMA a un puente instrumentado

En esta sección se ajusta la información perteneciente a la Dovela 4-Inferior 4 (D4-I4) a un modelo ARIMA, mediante la metodología Box-Jenkins. En la *Figura 5.11* se muestra la sección transversal de la Dovela 4 y la distribución de las cuerdas vibrantes.

Ahora bien, en primer lugar se somete la serie de tiempo al análisis clásico con dos propósitos particulares:

1. Identificación de tendencia en la serie
2. Identificación de estacionalidad en la serie y su correspondiente periodo

La principal ventaja de comenzar con un análisis clásico es detectar las características predominantes en la serie. Además, conforme se avance en el análisis clásico, se detectarán las desventajas del método para la modelación y realización de pronósticos satisfactorios.

Antes de someter la información a cualquier análisis, se requiere conocer su gráfico, a fin de realizar una inspección visual. La *Figura 5.12* muestra el gráfico de los incrementos de esfuerzos registrados en la Dovela 4 – Punto Inferior 4, del *1 de abril del 2007* al *28 de marzo de 2009*. Un total de 728 datos registrados. En el **ANEXO 1** se presenta la información completa.

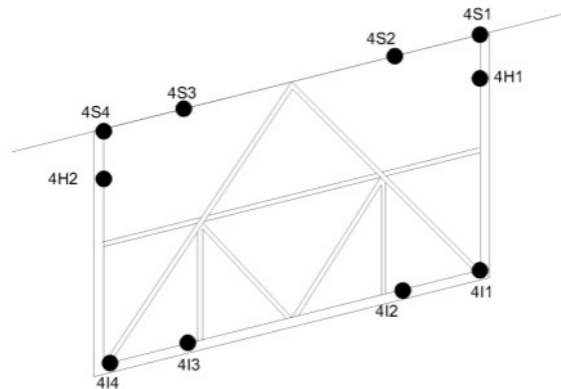


Figura 5.11 Distribución de cuerdas vibrantes en Dovela 4

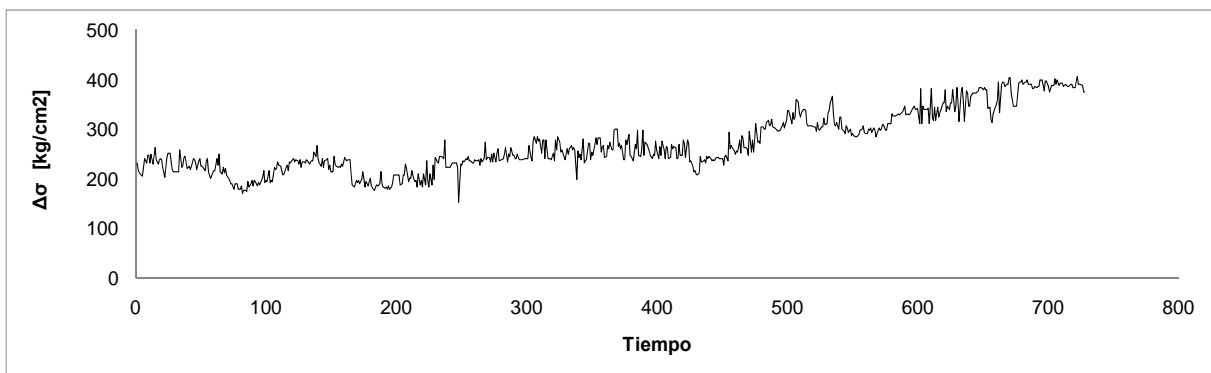


Figura 5.12 Incrementos de Esfuerzos D4-I4

En una primera inspección visual, podemos advertir la presencia de un patrón ascendente en los datos, los valores se incrementan de aproximadamente 200 kg/cm² a casi 400 kg/cm² es decir, existe patrón ascendente (tendencia) en la serie. No así en el caso de estacionalidad, a simple vista no se detectan fluctuaciones periódicas.

El siguiente paso es determinar con exactitud la presencia de tendencia y/o estacionalidad en la serie. Como se vio en el capítulo 3, una forma para determinar con mayor exactitud la presencia de tendencia en la serie, es la suavización por medias móviles. La *Figura 5.13* muestra la aplicación de una media móvil de 20 puntos a la serie.

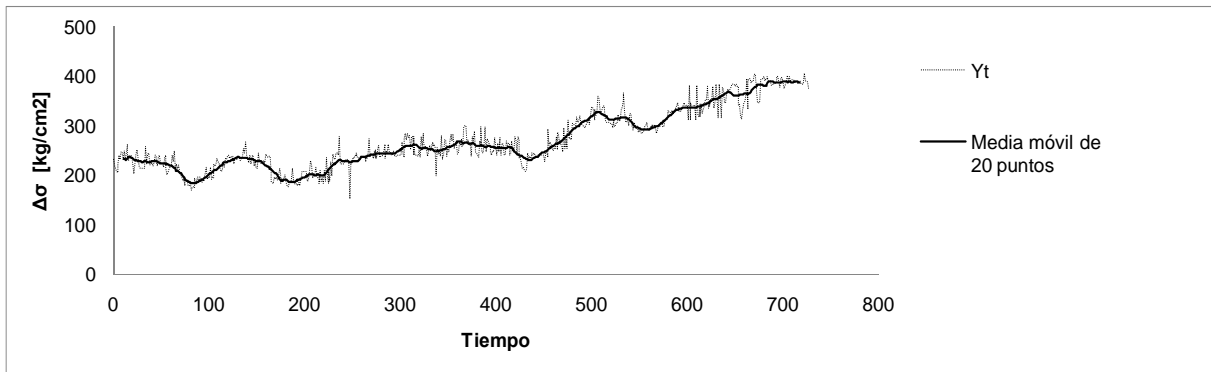


Figura 5.13 Suavización de la serie aplicando una media móvil de 20 puntos

Analizando la figura anterior se comprueba la presencia de tendencia en la serie.

Ahora bien, a fin facilitar la detección de algún patrón estacional en la serie, es recomendable extraer la tendencia de la serie. Supongamos un modelo clásico multiplicativo:

$$Y_t = T_t \times E_t$$

Entonces es necesario ajustar la tendencia T_t a un modelo matemático. Se proponen dos modelos de ajuste de la tendencia: un modelo lineal y uno cuadrático. Se seleccionará el modelo que mejor ajuste según el estadístico R^2 .

El modelo lineal (*Figura 5.14*) se obtiene aplicando las fórmulas vistas en el capítulo 3:

$$b = \frac{\sum(t - \bar{t})(Y_t - \bar{Y}_t)}{\sum(t - \bar{t})^2} = \frac{8084211.09}{32152302} = 0.2514$$

$$a = \bar{Y}_t - b\bar{t} = 269.92 - (0.2514)(365) = 178.272$$

Así, el modelo lineal de la tendencia es:

$$T_t = 178.272 + 0.2514 t$$

$$R^2 = 0.769$$

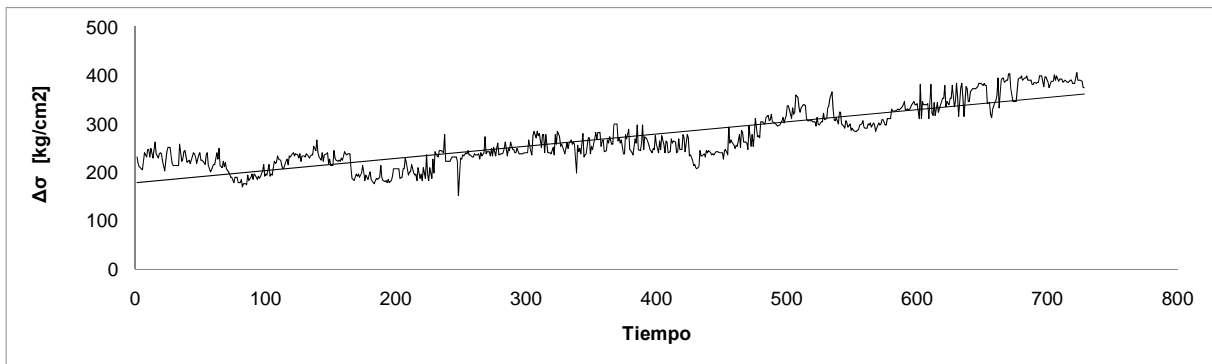


Figura 5.14 Modelo lineal de la tendencia

El modelo cuadrático lo obtenemos mediante las opciones de EXCEL. La Figura 5.15 muestra el ajuste cuadrático a la serie. La ecuación del modelo cuadrático ajustado es:

$$T(t) = 0.0005t^2 - 0.095t + 220.5$$

Su respectivo estadístico R^2 es:

$$R^2 = 0.8665$$

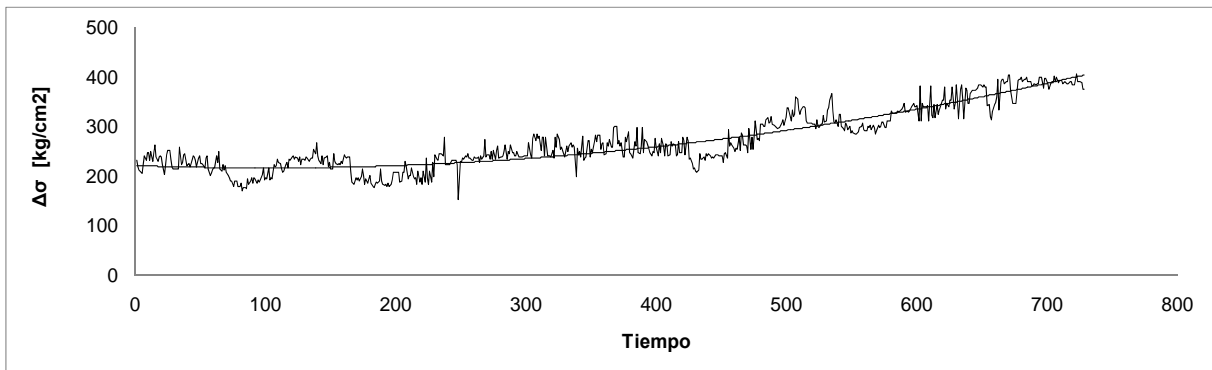


Figura 5.15 Modelo cuadrático de la tendencia

Mediante R^2 determinamos que el modelo cuadrático de la tendencia ajusta mejor a la serie, ya que R^2 es más cercano a 1.

La extracción de la tendencia se realiza por medio de la ecuación:

$$E_t = \frac{Y_t}{T_t}$$

Para ello, se genera una serie de datos mediante:

$$T_t = 0.0005t^2 - 0.095t + 220.5$$

Cada valor de la serie original Y_t se divide entre el valor correspondiente T_t . De esta manera, en teoría, producimos una serie sin tendencia y exclusivamente con patrón estacional y residuos.

La *Figura 5.16* muestra la serie sin tendencia. En una nueva inspección visual no se advierte de la presencia de algún patrón estacional. Se recurre entonces a las dos herramientas presentadas en el capítulo 2: las funciones de autocorrelación y el periodograma.

En primer lugar se examinan las funciones de autocorrelación simple y parcial (FAS y FAP). Estas se calculan para la serie sin tendencia. Las *Figuras 5.17* y *5.18* muestran la FAS y la FAP respectivamente.

La función de autocorrelación parcial muestra un comportamiento de abanico, esto es, los coeficientes de autocorrelación decaen a cero en el retraso 35, y continúan decayendo hasta el retraso 60, posteriormente ascienden hasta valores positivos. Este comportamiento en la FAS podría reflejar la presencia de un patrón estacional, con periodo $T = 60$ días.

Sin embargo, este supuesto patrón estacional, no se ve reflejado en la FAP (*Figura 5.18*). Sólo en los 3 primeros retrasos de tiempo se observan coeficientes significativos (hay que recordar que para considerar un coeficiente como verdaderamente significativo, este debe sobrepasar la línea punteada). Posterior al retraso 3, no se observan espigas que indiquen o comprueben la existencia de un patrón estacional identificable.

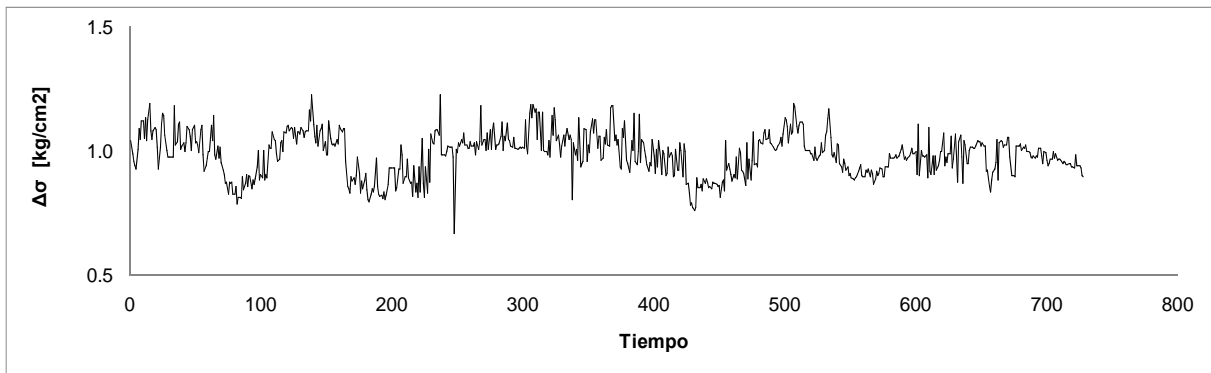


Figura 5.16 Remoción de la tendencia

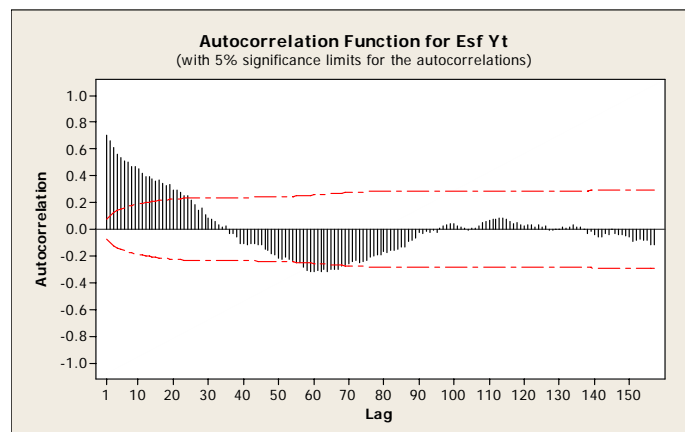


Figura 5.17 Función de autocorrelación simple

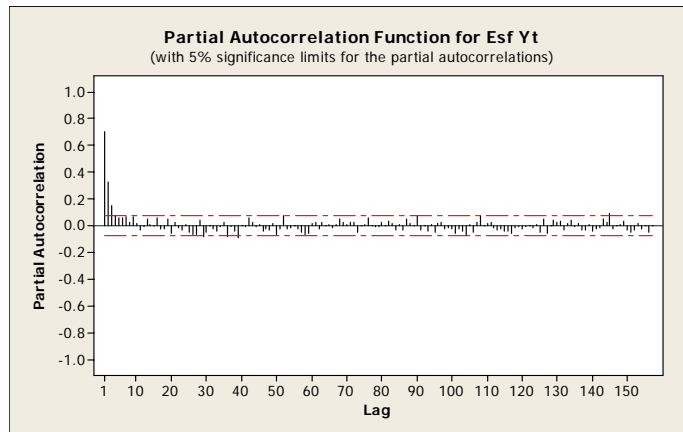


Figura 5.18 Función de autocorrelación parcial

Hasta el momento no se puede realizar una clara identificación de algún patrón estacional en la serie. Se recurre al periodograma de la serie para tratar de esclarecer el problema.

La Figura 5.19 muestra el periodograma correspondiente a los esfuerzos de la Dovela 4-Inferior 4, sin tendencia. Este se obtuvo con el programa *Degtra* (Ordaz y Montoya 2002). Además la Tabla 5.1 muestra las amplitudes destacables del periodograma con su correspondiente frecuencia y periodo asociado.

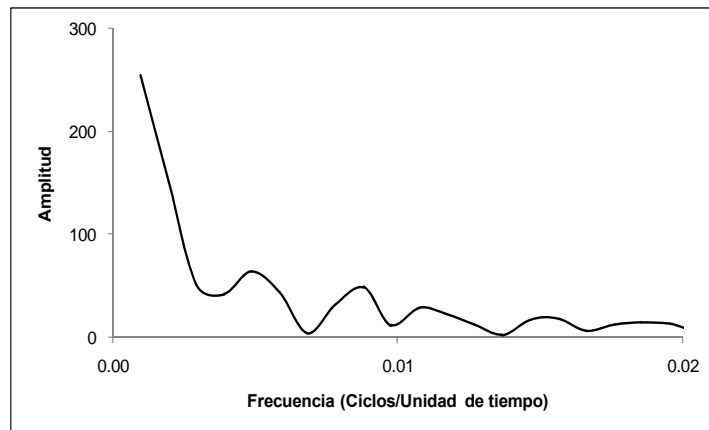


Figura 5.19 Periodograma

Frecuencia	Amplitud	Periodo
0.004883	60.98	204.8
0.007813	33.80	128.0
0.009766	29.16	102.4

Tabla 5.1 Amplitudes destacables - serie incrementos de esfuerzos D4-I4

De acuerdo a la tabla anterior los 3 posibles periodos estacionales de la serie son:

$$T_1 = 204 \text{ días}$$

$$T_2 = 128 \text{ días}$$

$$T_3 = 102 \text{ días}$$

Sin embargo estos periodos no corresponden a algún múltiplo o submúltiplo del periodo antes propuesto ($T = 60 \text{ días}$). Así también, estos periodos no se ven reflejados en la FAS y FAP de la serie. Al no contar con una herramienta extra para la detección de estacionalidad y el periodo correspondiente, se puede suponer que no existe estacionalidad en la serie, o al menos no es identificable por estos medios.

Además, esta conclusión se apoya en el conocimiento de la naturaleza de los datos. Se podría suponer que, una posible causa de un patrón estacional en la serie de esfuerzos, sería debida a la influencia de la variación de temperatura a lo largo del año. Entonces, tanto las funciones de autocorrelación como en el periodograma, deberían de reflejar un patrón estacional anual, con un periodo aproximado $T = 365 \text{ días}$.

Finalmente, después de realizar un análisis clásico a la serie, concluimos que:

- ✓ Existe tendencia en la serie.
- ✓ No existe estacionalidad en la serie, o al menos una estacionalidad claramente detectable.

Antes de pasar a la aplicación de la metodología ARIMA a los datos de la Dovela 4-Inferior 4, cabe resaltar la imposibilidad de realizar pronósticos a la serie tan sólo implementando un análisis clásico para las mediciones de esfuerzos que se intenta estudiar. Esto se debe a que solamente es posible detectar una componente (tendencia) de la serie. Claramente se advierte el gran margen de error que produciría realizar pronósticos utilizando tan sólo un modelo cuadrático de la tendencia. Por esta razón el uso de modelos más avanzados se hace necesario, tal como los ARIMA.

Una vez hecho este paréntesis se inicia con la aplicación de la metodología Box-Jenkins. Se especifica que se trabaja únicamente con los primeros 628 datos. Los 100 datos restantes se utilizarán para probar los pronósticos.

Ahora bien, la primera etapa de la metodología Box-Jenkins es la comprobación de la estacionariedad. Para ello, se concluyó en el análisis previo, la existencia de tendencia en la serie y como se vio en el capítulo 4, la presencia de este patrón en los datos es prueba irrefutable de que la serie es no estacionaria.

Se requiere entonces una transformación, mediante diferencias. Las primeras diferencias de la serie* se presentan en la *Figura 5.20*. A simple vista se puede notar que no existe patrón ascendente o descendente alguno en la serie, todas las observaciones giran alrededor a una media constante de valor cero. Si la afirmación anterior es verdadera, se puede decir que las primeras diferencias transforman la serie en estacionaria, y se puede proseguir con la metodología. Para comprobar esta afirmación, las *Figuras 5.21* y *5.22* muestran las funciones de autocorrelación simple y parcial de la serie de esfuerzos. En la primera figura la FAS se corta después del retraso 1, mientras que la FAP decae gradualmente a cero. Podemos concluir entonces que la serie transformada es estacionaria.

*NOTA: La transformación fue aplicada a la serie original Y_t sin remoción de la tendencia.

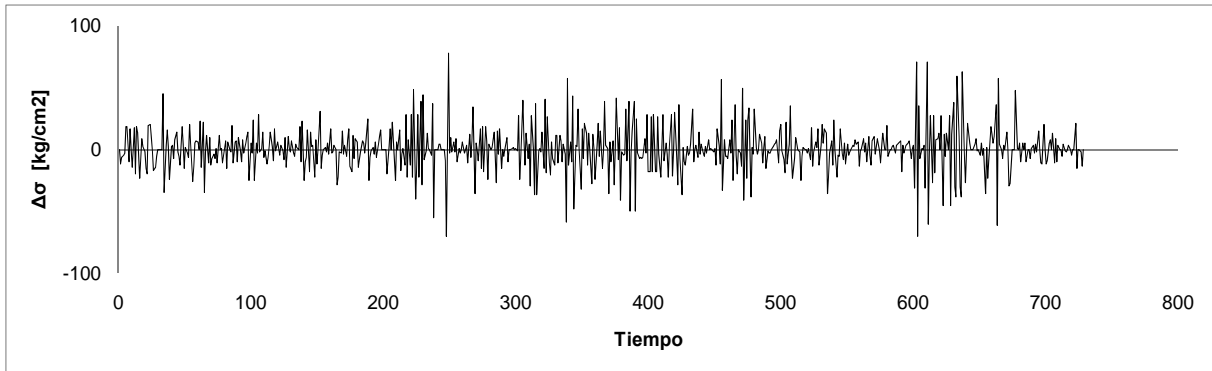


Figura 5.20 Primeras diferencias serie D4-I4

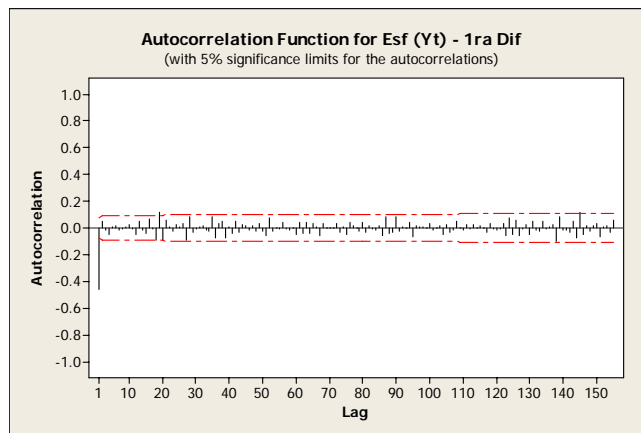


Figura 5.21 Función de autocorrelación simple para D4-I4 (serie diferenciada)

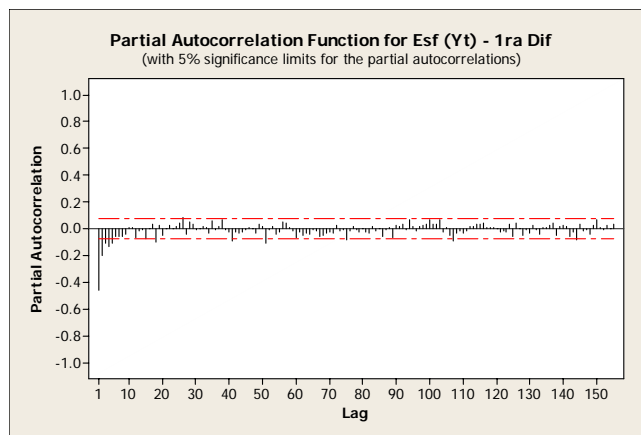


Figura 5.22 Función de autocorrelación parcial para D4-I4 (serie diferenciada)

Una vez inducida la estacionariedad en la serie, se identifica uno o varios modelos ARIMA probables. Se compara las funciones de autocorrelación prácticas obtenidas anteriormente (*Figuras 5.21 y 5.22*) con las funciones de autocorrelación teóricas vistas en la sección 4.3.2.

Al realizar esta comparación, resulta que el patrón observado en las funciones de autocorrelación prácticas calculadas, se asemejan a las funciones teóricas de un modelo MA(1), como se puede apreciar en la *Figura 5.23*.

Entonces, el primer modelo posible sería: ARIMA($p=0,d=1,q=1$), es decir, ARIMA(0,1,1).

Con la ayuda de MINITAB, se estima el correspondiente parámetro θ_1 . No se incluye término constante en el modelo, puesto que la media de los valores de la serie es igual a cero, a casusa de la diferenciación.

La salida de MINITAB para el modelo ARIMA(0,1,1) ajustado a la serie de datos de esfuerzos se muestra en la *Figura 5.24*.

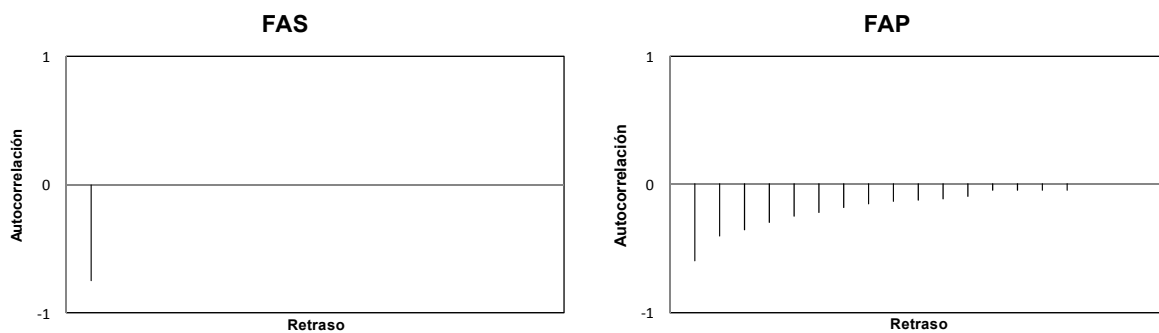


Figura 5.23 Funciones de autocorrelación simple y parcial teóricas para el modelo MA(1)

A partir de analizar la salida de MINITAB, encontramos que $\theta_1 = 0.6286$.

Se analiza si este modelo propuesto es adecuado:

- a) El modelo ARIMA(0,1,1) muestra parsimonia, puesto que sólo consta de un solo coeficiente.
- b) La t asociada al coeficiente tiene un valor igual a 20.23 , por lo tanto $|t| > 2$ y se considera que el coeficiente θ_1 es relevante en el modelo.
- c) De la *Tabla 4.4* el modelo no tiene condiciones de estacionariedad y la condición de invertibilidad es que $|\theta_1| < 1$. Como $\theta_1 = 0.6286$ es menor a la unidad, el modelo cumple con esta condición.

```

ARIMA Model: Esf (Yt)

Estimates at each iteration

Iteration   SSE   Parameters
0   177546   0.100
1   160094   0.250
2   148456   0.400
3   142029   0.550
4   141067   0.613
5   141026   0.625
6   141023   0.628
7   141023   0.628
8   141023   0.629

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type      Coef  SE Coef   T      P
MA 1      0.6286  0.0311  20.23  0.000

Differencing: 1 regular difference
Number of observations: Original series 628, after differencing 627
Residuals:      SS = 140945 (backforecasts excluded)
                MS = 225   DF = 626

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag        12      24      36      48
Chi-Square 12.0   27.3   36.5   52.0
DF         11      23      35      47
P-Value    0.362  0.245  0.397  0.287
    
```

Figura 5.24 Ajuste del modelo ARIMA(0,1,1) con MINITAB

d) El valor de ajuste RMSE es:

$$RMSE = \frac{1}{n - m} \sum \hat{a}_t^2 = \frac{1}{628 - 1} (140945) = \frac{140945}{627} = 225$$

- e) La función de autocorrelación simple de los residuos (Figura 5.25) no muestran espigas significativas, por lo que se concluye que los residuos son completamente aleatorios.
- f) Los valores p asociados al estadístico Ljung-Box son: 0.362, 0.245, 0.397 y 0.287 para los retrasos de tiempo 12, 24, 36 y 48 respectivamente. Todos los valores de p son mayores a 0.05; por lo tanto el modelo es suficiente.

Finalmente se grafica la serie con sus correspondientes valores producidos con el modelo ARIMA (0,1,1). En la Figura 5.26 se observa un gran ajuste con este modelo.

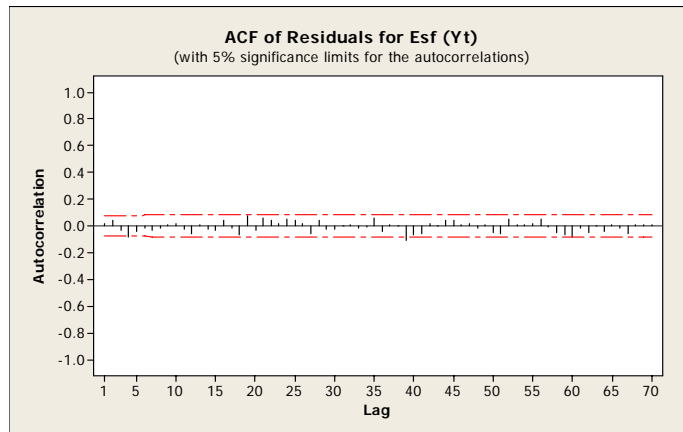


Figura 5.25 Función de autocorrelación simple de los residuos modelo ARIMA(0,1,1)

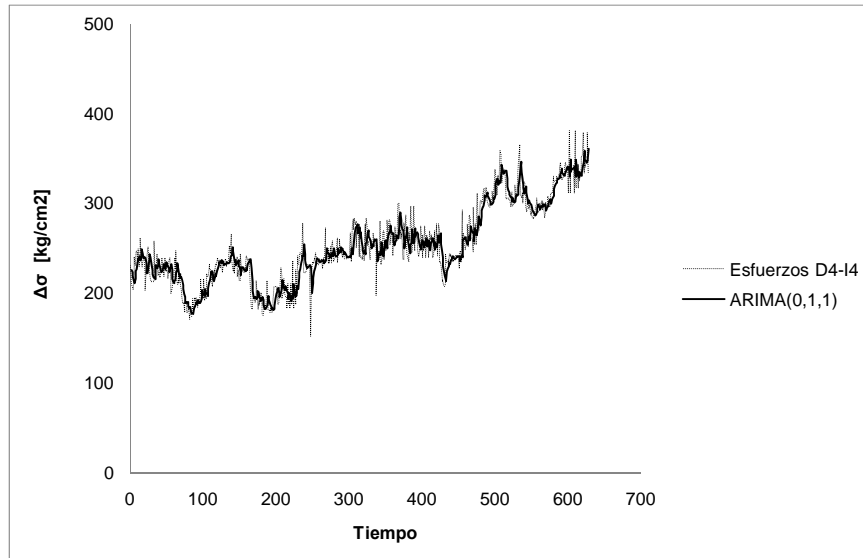


Figura 5.26 Ajuste ARIMA(0,1,1)

En el caso específico de la información de la dovela D4-I4, se encontró un modelo ARIMA adecuado en el primer intento, sin embargo diversos modelos ARIMA podrían ajustarse adecuadamente a la serie. Es conveniente, por tanto, proponer y analizar modelos alternativos a fin de comparar el grado de ajuste que se obtiene, aunque se haya encontrado ya un modelo adecuado.

Para el caso de la información de la Dovela 4-Inferior4, se propone a su vez los modelos: ARIMA(0,1,3), ARIMA(1,1,1), ARIMA(5,1,0) y ARIMA(3,1,4).

Modelo ARIMA(0,1,3): La Figura 5.27 muestra el ajuste realizado por MINITAB. Se tiene que $\theta_1 = 0.6250$, $\theta_2 = -0.0438$ y $\theta_3 = 0.1044$

```

ARIMA Model: Esf (Yt)
Estimates at each iteration

Iteration   SSE      Parameters
0  178832  0.100  0.100  0.100
1  159842  0.250  0.138  0.093
2  147133  0.400  0.126  0.065
3  140747  0.550  0.059  0.048
4  139558  0.616  -0.020  0.085
5  139496  0.622  -0.039  0.101
6  139494  0.625  -0.043  0.104
7  139494  0.625  -0.044  0.104
8  139494  0.625  -0.044  0.104
9  139494  0.625  -0.044  0.104

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type      Coef  SE Coef   T      P
MA 1      0.6250  0.0399  15.66  0.000
MA 2     -0.0438  0.0473  -0.93  0.355
MA 3      0.1044  0.0400   2.61  0.009

Differencing: 1 regular difference
Number of observations: Original series 628, after differencing 627
Residuals:      SS = 139448 (backforecasts excluded)
                MS = 223  DF = 624

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag        12      24      36      48
Chi-Square  5.9    22.4   33.2   46.5
DF          9      21     33     45
P-Value    0.754  0.378  0.459  0.410
    
```

Figura 5.27 Ajuste ARIMA(0,1,3)

Al analizar el ajuste del modelo, de acuerdo a la metodología Box-Jenkins encontramos que:

- El modelo ARIMA(0,1,3) muestra parsimonia.
- La t asociada a los coeficiente tienen un valor igual a $t_{\theta_1} = 15.66$, $t_{\theta_2} = -0.93$ y $t_{\theta_3} = 2.61$. Así $|t_1| > 2$ $|t_2| < 2$ y $|t_3| > 2$, por lo tanto los coeficientes θ_1 y θ_3 son relevantes en el modelo, mientras que θ_2 no lo es.
- El modelo no es adecuado ya que θ_2 no es importante en el modelo. Se rechaza el modelo.

Modelo ARIMA(1,1,1): En este caso $\varphi_1 = 0.0637$ y $\theta_1 = 0.6741$ (Figura 5.28)

De forma análoga:

- El modelo ARIMA(1,1,1) muestra parsimonia.
- La t asociada a los coeficiente tienen un valor igual a $t_{\varphi_1} = 1.02$, $t_{\theta_1} = 14.57$. Así $|t_1| < 2$ y $|t_2| > 2$, por lo tanto el coeficiente θ_1 es relevantes en el modelo, mientras que φ_{21} no lo es.
- El modelo no es adecuado ya que φ_{21} no es importante en el modelo. Se rechaza el modelo.

```

ARIMA Model: Esf (Yt)

Estimates at each iteration

Iteration      SSE      Parameters
0      193249      0.100  0.100
1      155845     -0.050  0.250
2      149087      0.012  0.400
3      142829      0.038  0.550
4      140945      0.034  0.638
5      140876      0.049  0.659
6      140865      0.057  0.668
7      140863      0.061  0.672
8      140862      0.063  0.673
9      140862      0.063  0.674
10     140862      0.064  0.674
11     140862      0.064  0.674

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type      Coef  SE Coef      T      P
AR  1  0.0637  0.0627      1.02  0.310
MA  1  0.6741  0.0463     14.57  0.000

Differencing: 1 regular difference
Number of observations: Original series 628, after differencing 627
Residuals:      SS = 140792 (backforecasts excluded)
                MS = 225  DF = 625

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag      12      24      36      48
Chi-Square  11.0  27.0  36.8  51.6
DF        10      22      34      46
P-Value   0.359  0.212  0.343  0.263
    
```

Figura 5.28 Ajuste ARIMA(1,1,1)

Modelo ARIMA(5,1,0): Mediante MINITAB, los valores de los coeficientes AR son: $\varphi_1 = -0.6024$, $\varphi_2 = -0.3129$, $\varphi_3 = -0.2151$, $\varphi_4 = -0.1958$, $\varphi_5 = -0.1102$ (Figura 5.29)

Este modelo:

- a) El modelo ARIMA(5,1,0) se compone con 5 coeficientes, sigue siendo un modelo con parsimonia.

```

ARIMA Model: Esf (Yt)

Estimates at each iteration

Iteration   SSE                Parameters
0  214235   0.100   0.100   0.100   0.100   0.100
1  186192  -0.050   0.013   0.035   0.039   0.056
2  164810  -0.200  -0.074  -0.032  -0.024   0.011
3  150129  -0.350  -0.163  -0.099  -0.088  -0.034
4  142166  -0.500  -0.251  -0.167  -0.151  -0.080
5  140599  -0.597  -0.309  -0.212  -0.193  -0.109
6  140594  -0.602  -0.313  -0.215  -0.196  -0.110
7  140594  -0.602  -0.313  -0.215  -0.196  -0.110

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type      Coef  SE Coef      T      P
AR 1     -0.6024  0.0400  -15.08  0.000
AR 2     -0.3129  0.0460   -6.80  0.000
AR 3     -0.2151  0.0470   -4.58  0.000
AR 4     -0.1958  0.0461   -4.25  0.000
AR 5     -0.1102  0.0401   -2.75  0.006

Differencing: 1 regular difference
Number of observations: Original series 628, after differencing 627
Residuals:      SS = 140499 (backforecasts excluded)
                MS = 226  DF = 622

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag      12      24      36      48
Chi-Square  9.1    23.9    34.6    49.1
DF         7      19      31      43
P-Value   0.242  0.201  0.301  0.241
    
```

Figura 5.29 Ajuste ARIMA(5,1,0)

b) La t asociada a los coeficientes tienen un valor igual a:

$$t_{\varphi_1} = -15.08$$

$$t_{\varphi_2} = -6.80$$

$$t_{\varphi_3} = -4.58$$

$$t_{\varphi_4} = -4.25$$

$$t_{\varphi_5} = -2.75$$

En este caso:

$$|t_{\varphi_1}| > 2$$

$$|t_{\varphi_2}| > 2$$

$$|t_{\varphi_3}| > 2$$

$$|t_{\varphi_4}| > 2$$

$$|t_{\varphi_5}| > 2$$

Por lo tanto los coeficientes AR son relevantes en el modelo.

- c) Un modelo que utiliza sólo parámetros autorregresivos no tiene condiciones de inversibilidad. La condición de estacionariedad necesaria (pero no suficiente), para un modelo que utiliza sólo parámetros autorregresivos, es que la suma de los valores de los parámetros autorregresivos sea menor que 1.

$$-0.6024 - 0.3129 - 0.2151 - 0.1958 - 0.1102 = -1.436$$

$$-1.436 < 1$$

El modelo cumple con la condición necesaria de estacionariedad.

- d) El valor de ajuste RMSE es:

$$RMSE = \frac{1}{n - m} \sum \hat{a}_t^2 = \frac{1}{628 - 5} (140499) = \frac{140499}{623} = 226$$

La función de autocorrelación simple de los residuos (*Figura 5.30*) no muestra espigas significativas, por lo que se concluye que los residuos son completamente aleatorios.

- e) Los valores p asociados al estadístico Ljung-Box son: 0.242, 0.201, 0.301 y 0.241 para los retrasos de tiempo 12, 24, 36 y 48 respectivamente. Todos los valores de p son mayores a 0.05; por lo tanto el modelo es suficiente.

Gráficamente se observa el ajuste del modelo ARIMA(5,1,0) en la *Figura 5.31*.

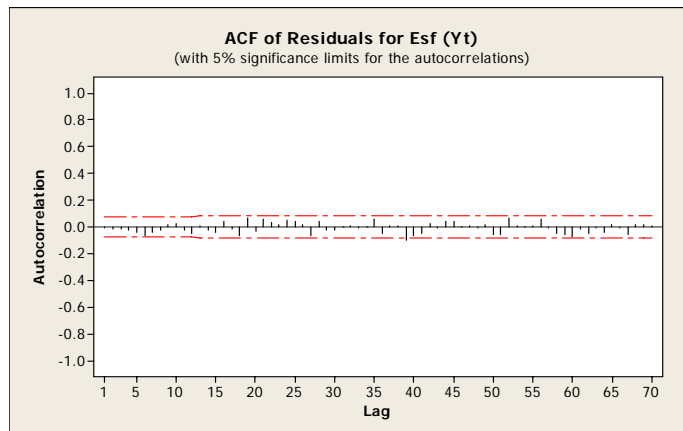


Figura 5.30 Función de autocorrelación simple de los residuos modelo ARIMA(5,1,0)

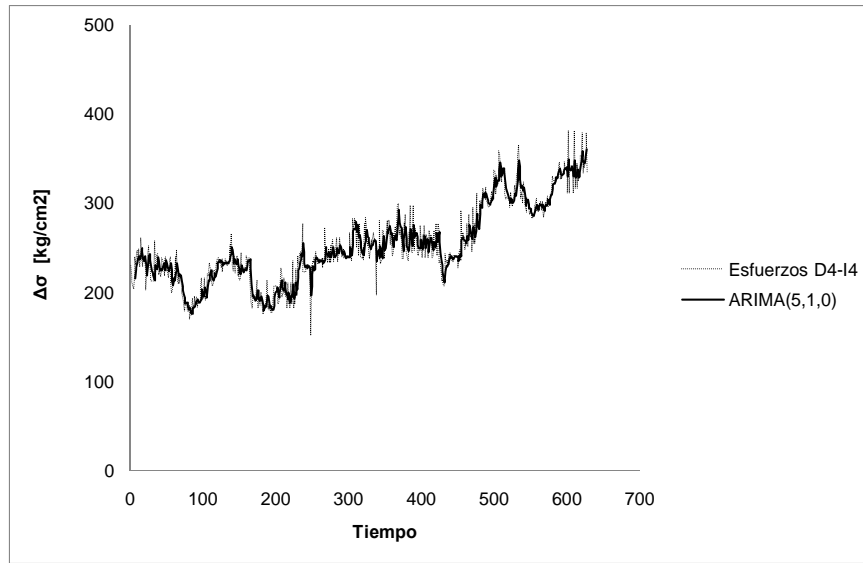


Figura 5.31 Ajuste ARIMA(5,1,0)

Modelo ARIMA(3,1,4): Finalmente en la *Figura 5.32* se puede observar los resultados del ajuste.

Los valores de los coeficientes son:

$$\varphi_1 = 0.4725, \varphi_2 = -0.7169, \varphi_3 = 0.7036, \theta_1 = 1.0827, \theta_2 = -1.0191, \theta_3 = 1.2022, \theta_4 = -0.4178$$

```

Final Estimates of Parameters

Type      Coef  SE Coef    T      P
AR 1      0.4725  0.0365   12.94  0.000
AR 2     -0.7169  0.0566  -12.66  0.000
AR 3      0.7036  0.0414   16.98  0.000
MA 1      1.0827  0.0141   76.60  0.000
MA 2     -1.0191  0.0408  -24.99  0.000
MA 3      1.2022  0.0047  254.20  0.000
MA 4     -0.4178  0.0253  -16.50  0.000

Differencing: 1 regular difference
Number of observations: Original series 628, after differencing 627
Residuals:      SS = 138113 (backforecasts excluded)
                  MS = 223  DF = 620

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag        12      24      36      48
Chi-Square  3.0    17.1   25.7   39.7
DF          5      17     29     41
P-Value     0.695  0.445  0.643  0.528
    
```

Figura 5.32 Ajuste ARIMA(3,1,4)

- a) El modelo ARIMA(3,1,4) se compone con 7 coeficientes, número que puede considerarse un poco alto, sin embargo podrían obtenerse buenos resultados.

b) La t asociada a los coeficientes tienen un valor igual a:

$$t_{\varphi_1} = 12.94$$

$$t_{\varphi_2} = -12.66$$

$$t_{\varphi_3} = 16.98$$

$$t_{\theta_1} = 76.60$$

$$t_{\theta_2} = -24.99$$

$$t_{\theta_3} = 254.20$$

$$t_{\theta_4} = -16.50$$

En este caso:

$$|t_{\varphi_1}| > 2$$

$$|t_{\varphi_2}| > 2$$

$$|t_{\varphi_3}| > 2$$

$$|t_{\theta_1}| > 2$$

$$|t_{\theta_2}| > 2$$

$$|t_{\theta_3}| > 2$$

$$|t_{\theta_4}| > 2$$

Por lo tanto los coeficientes AR y MA son relevantes en el modelo.

c) No se especifica alguna condición de estacionariedad y/o inversibilidad para un modelo ARIMA con parámetros AR y MA . Por consiguiente el modelo hasta este paso, se acepta.

d) El valor de ajuste RMSE es:

$$RMSE = \frac{1}{n - m} \sum \hat{a}_t^2 = \frac{1}{628 - 7} (138113) = \frac{138113}{621} = 222$$

e) La función de autocorrelación simple de los residuos (*Figura 5.33*) no muestran espigas significativas, por lo que se concluye que los residuos son completamente aleatorios.

f) Los valores p asociados al estadístico Ljung-Box son: 0.695, 0.445, 0.643 y 0.528 para los retrasos de tiempo 12, 24, 36 y 48 respectivamente. Todos los valores de p son mayores a 0.05; por lo tanto el modelo es suficiente.

En la gráfica de la *Figura 5.34* se encuentra el respectivo ajuste del modelo ARIMA(3,1,4). Al igual que los modelos ARIMA(0,1,1) y ARIMA(5,1,0) existe un gran ajuste a la serie de datos.

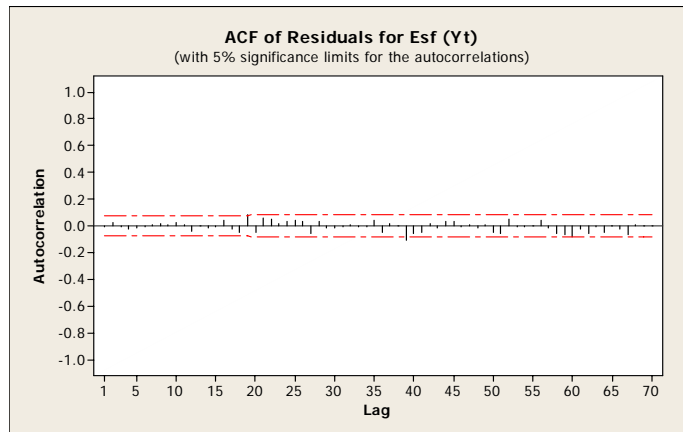


Figura 5.33 Función de autocorrelación simple de los residuos modelo ARIMA(3,1,4)

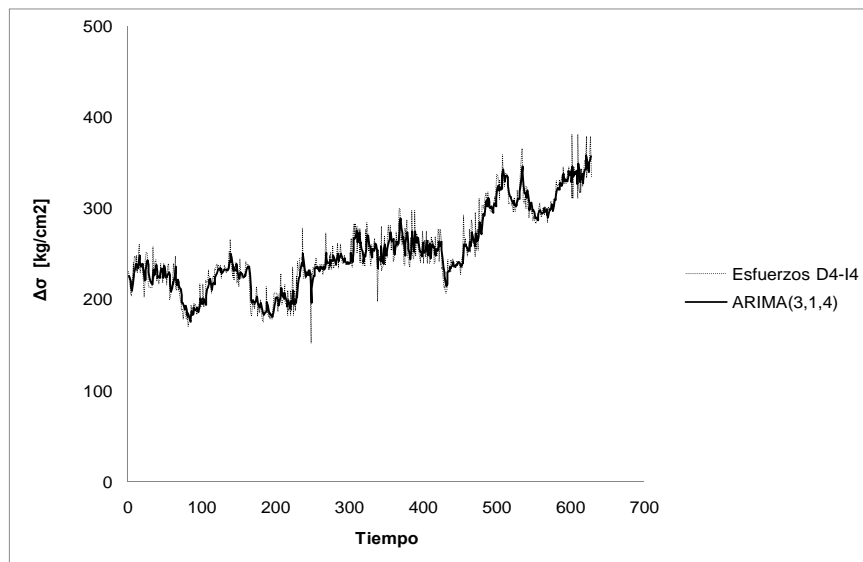


Figura 5.34 Ajuste ARIMA(3,1,4)

Se concluye entonces que los modelos ARIMA(0,1,1), ARIMA(5,1,0) y ARIMA (3,1,4) son adecuados y óptimos para realizar pronósticos.

5.5 Verificación de pronósticos

A partir de los tres modelos ARIMA ajustados previamente, se generan 100 pronósticos puntuales con la ayuda de MINITAB que equivaldría a obtener los incrementos de esfuerzos de 100 días posteriores a Y_{628} . Se comparan los pronósticos generados con los 100 datos restantes reales de los incrementos de esfuerzos.

En las Figuras 5.35, 5.36 y 5.37 se encuentran los pronósticos de \hat{Y}_{629} hasta \hat{Y}_{728} obtenidos con MINITAB de los modelos ARIMA(0,1,1), ARIMA(5,1,0) y ARIMA(3,1,4) respectivamente. Se grafican además los intervalos de confianza del 95%, así como los valores reales medidos.

Ahora bien, al examinar los pronósticos, observamos que existe gran similitud entre los distintos modelos. Los pronósticos convergen a una línea recta, este efecto se debe a la naturaleza de los datos y del modelo ARIMA ajustado. Por ejemplo, en el modelo ARIMA(0,1,1), se obtiene un valor constante desde las primeras predicciones puntuales, mientras que en los dos modelos restantes existen mayores fluctuaciones en un principio, pero al final convergen a un valor constante.

El grado de exactitud de los pronósticos se puede calcular mediante la RMSE:

$$RMSE = \hat{\sigma}_a^2 = \frac{1}{n} \sum \hat{a}_t^2$$

Donde:

\hat{a}_t : residuales

$$\hat{a}_t = Y_t - \hat{Y}_t$$

n : número de observaciones

Para el presente caso $n = 100$. La Tabla 5.2 muestra las RMSE calculada para cada modelo. Se muestra que la mejor aproximación la da el modelo ARIMA(0,1,1). Sin embargo, este modelo en particular proyecta un valor constante para todas las predicciones, lo que podría resultar no muy eficiente o atractivo, para ciertos propósitos.

Modelo	$\sum \hat{a}_t^2$	RMSE
(0,1,1)	114397	1144
(5,1,0)	117235	1172
(3,1,4)	125618	1256

Tabla 5.2 Grado de exactitud de los pronósticos

Más allá de observar tan sólo las predicciones puntuales \hat{Y} , conviene fijar la atención al intervalo de confianza de las predicciones: en los 3 modelos ajustados la combinación intervalo de predicción-valor futuro da resultados satisfactorios. Es decir, los valores reales se encuentran dentro de los intervalos de confianza.

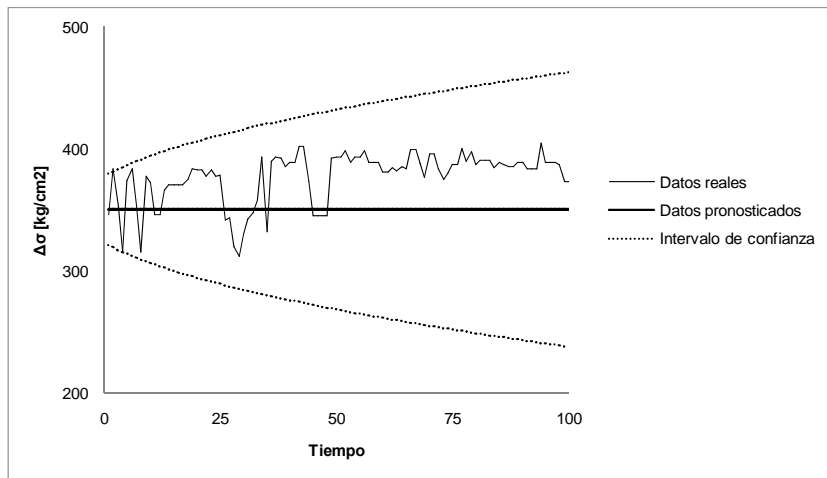


Figura 5.35 Pronósticos ARIMA(0,1,1)

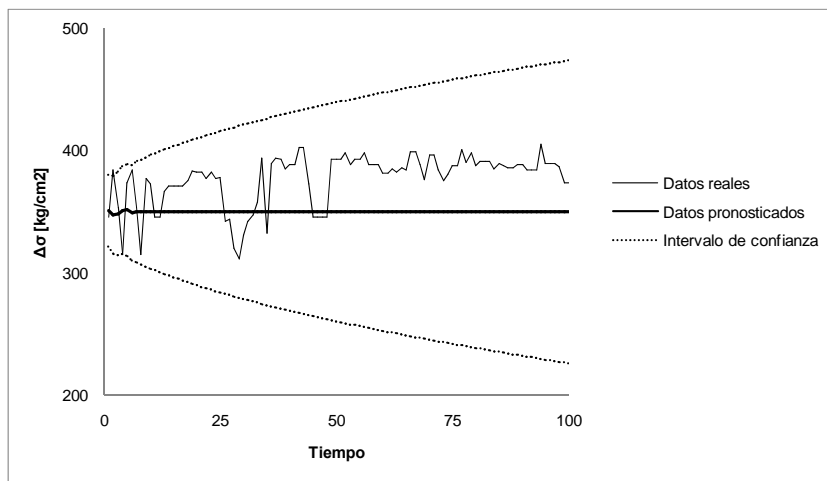


Figura 5.36 Pronósticos ARIMA(5,1,0)

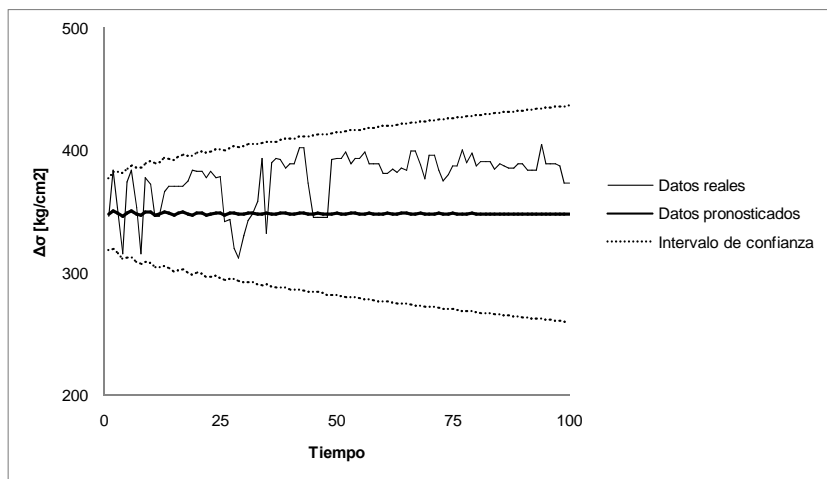


Figura 5.37 Pronósticos ARIMA(3,1,4)