



UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO



FACULTAD DE INGENIERÍA
INSTITUTO DE INGENIERÍA

“APLICACIÓN DE SERIES DE TIEMPO
EN EL MONITOREO ESTRUCTURAL”

T E S I S

PARA OBTENER EL TÍTULO DE:
INGENIERO CIVIL

P R E S E N T A:
JOEL MAURICIO CORTÉS PATIÑO

DIRECTOR DE TESIS:
M.I. GERARDO RODRÍGUEZ GUTIÉRREZ

MÉXICO D.F. 2011



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dedicatorias

A mis padres y hermano

A mis amigos de toda la vida

Agradecimientos

*En primer lugar agradezco a mi **mamá**, a mi **papá** y a mi **hermano** por siempre estar al tanto de mí. Gracias por brindarme todo su amor, confianza e incondicional apoyo que siempre he recibido y que me alienta a esforzarme día a día. Sus valiosas enseñanzas las llevaré siempre conmigo.*

Agradezco el apoyo y sus buenos deseos a mis amigos de secundaria, prepa y universidad, con quien he recorrido este camino en diferentes etapas de mi vida y con quienes he pasado gratos momentos.

No puede dejar de agradecer al M. I. Gerardo Rodríguez Gutiérrez por su apoyo en la realización del servicio social y en este trabajo de tesis. También por toda la confianza brindada a lo largo de año y medio y permitirme colaborar de manera activa en proyectos del instituto.

A la Universidad Nacional Autónoma de México por brindarme la oportunidad de entrar en sus aulas.

A la Facultad de Ingeniería por la formación profesional recibida a lo largo de todo este tiempo.

Al Instituto de Ingeniería de la UNAM por abrirme sus puertas y proporcionarme los medios para realizar mi trabajo de tesis y permitirme participar en proyectos.

A todos ellos, Gracias!

Joel Mauricio Cortés Patiño

CONTENIDO

INTRODUCCIÓN	1
1 SALUD ESTRUCTURAL	3
1.1 ANTECEDENTES	3
1.2 EL MONITOREO DE LA SALUD ESTRUCTURAL	4
1.3 ETAPAS DEL MONITOREO DE LA SALUD ESTRUCTURAL	5
1.4 MANEJO DE LA INFORMACIÓN	6
2 CONCEPTOS FUNDAMENTALES DE SERIES DE TIEMPO	8
2.1 DEFINICIÓN DE UNA SERIE DE TIEMPO	8
2.2 TRATAMIENTO DE SERIES DE TIEMPO	9
2.3 HERRAMIENTAS BÁSICAS PARA EL ANÁLISIS DE SERIES DE TIEMPO	10
2.3.1 <i>Gráfico de la serie</i>	11
2.3.2 <i>Funciones de autocorrelación</i>	11
2.3.3 <i>Periodograma</i>	23
3 MODELACIÓN CLÁSICA	25
3.1 ENFOQUE DETERMINÍSTICO	25
3.2 ESQUEMAS DE INTEGRACIÓN	27
3.3 COMPONENTES DE UNA SERIE DE TIEMPO	28
3.3.1 <i>Tendencia</i>	28
3.3.2 <i>Estacionalidad</i>	34
3.3.3 <i>Ciclo</i>	38
3.3.4 <i>Aleatoriedad</i>	39
3.4 VENTAJAS Y DESVENTAJAS MODELACIÓN CLÁSICA	39
4 MODELACIÓN ARIMA	41
4.1 ENFOQUE Y PROCESOS ESTOCÁSTICOS	42
4.2 MODELOS AR, MA Y MIXTOS	42
4.3 METODOLOGÍA BOX-JENKINS NO ESTACIONAL	45
4.3.1 <i>Estacionariedad</i>	47
4.3.2 <i>Identificación</i>	55
4.3.3 <i>Estimación</i>	62
4.3.4 <i>Evaluación del modelo</i>	66
4.3.5 <i>Pronóstico</i>	69
4.4 METODOLOGÍA BOX-JENKINS ESTACIONAL	74
4.5 VENTAJAS Y DESVENTAJAS DE LA MODELACIÓN ARIMA	77
5 APLICACIÓN DE LA MODELACIÓN ARIMA A UN PUENTE INSTRUMENTADO	79
5.1 EL CASO DE UN PUENTE EN CURVA	79
5.2 DESCRIPCIÓN DE LA INSTRUMENTACIÓN	82
5.3 ELECCIÓN DE LA INFORMACIÓN A UTILIZAR	84

5.4 AJUSTE DE UN MODELO ARIMA A UN PUENTE INSTRUMENTADO	86
5.5 VERIFICACIÓN DE PRONÓSTICOS.....	102
6 HERRAMIENTAS PARA IDENTIFICAR CAMBIOS SIGNIFICATIVOS EN EL COMPORTAMIENTO ESTRUCTURAL DE UN PUENTE INSTRUMENTADO.	105
6.1 PROCESOS DE CONTROL ESTADÍSTICO	106
6.2 ALGORITMO CLASIFICADOR	112
7 CONCLUSIONES	116
ANEXO 1 - INCREMENTOS DE ESFUERZOS D4-I4	119
ANEXO 2 - AJUSTE ARIMA(5,1,0) 628 DATOS D4-I4.....	123
ANEXO 3 - AJUSTE ARIMA(5,1,0) 100 DATOS D4-I4	128
ANEXO 4 - AJUSTE ARIMA(5,1,0) 100 DATOS D4-I4 / ALGORITMO CLASIFICADOR	129
REFERENCIAS	130

Introducción

El constante incremento de los niveles de seguridad requeridos en las distintas obras civiles y las implicaciones financieras que conlleva, ha creado la necesidad de observar el desempeño de las estructuras frente a las acciones o cargas cotidianas y frente a eventos extraordinarios tales como sismos, huracanes, inundaciones, etc. La implementación de sistemas de monitoreo se ha convertido en una alternativa práctica en la evaluación continua del rendimiento de las obras de infraestructura, además de complementar eficientemente la inspección visual de las mismas.

En los últimos años la aplicación de estos sistemas ha despertado gran interés entre académicos e investigadores, debido al gran potencial que muestra en la temprana detección de daños y en los beneficios que se obtienen; esencialmente en materia de seguridad y economía.

No obstante, convertir grandes volúmenes de información resultado del monitoreo en información útil, representa un gran reto que necesita la aplicación de técnicas especiales de procesamiento de datos. En este campo se ha concentrado las investigaciones en diversas partes del mundo en las últimas décadas. La mayoría de las técnicas y métodos de análisis se han desarrollado bajo el marco de *series de tiempo* y con un enfoque *estadístico* de procesamiento de datos.

En esta tesis se presenta el uso de modelos autorregresivos y de media móvil en el monitoreo estructural. A partir de estos modelos se desarrollan dos herramientas de identificación de cambios en el comportamiento estructural. Los conceptos expuestos se aplican al caso particular de un puente en curva: el Puente San Cristóbal.

De esta manera, el objetivo del presente trabajo es proponer una técnica y metodología de análisis de la información (basado en el uso de modelos matemáticos). Con ello se pretende en primera instancia, describir el estado actual de una estructura y así establecer un punto inicial de comparación con futuros estados de la misma. En segunda instancia, se pretende establecer las bases para pronosticar posibles comportamientos futuros para poder implementar un adecuado plan de acciones preventivas y/o correctivas.

A través de siete capítulos se intenta explicar de forma clara y sencilla los conceptos necesarios para el análisis de la evolución del estado de las obras de infraestructura. Se busca que en el transcurso de la

lectura se adquieran las habilidades necesarias para adentrarse en los entresijos de las series de tiempo y formar una visión clara en su aplicación en el monitoreo estructural.

Para ello, en el primer capítulo se exponen las actividades referentes al monitoreo estructural, sin profundizar demasiado en ellas, destacando el manejo de la información que da pie al desarrollo de la parte fundamental de la tesis.

El segundo capítulo se enfoca al tratamiento de la información producto del monitoreo estructural bajo el concepto de series de tiempo. Se presentan las nociones básicas que el analista debe tener presentes para un correcto análisis de la información, así como las herramientas fundamentales para lograr un entendimiento adecuado de la evolución

En el capítulo tres se presentan las componentes de una serie de tiempo vistas bajo la modelación clásica. En este capítulo se busca entender el desarrollo de la serie a través del tiempo y detectar patrones repetitivos presentes en ella, que son de vital importancia para la aplicación de los modelos autorregresivos y de media móvil, además de ayudar al analista en la correcta interpretación de resultados.

El capítulo cuatro por su parte entra de lleno a los modelos ARIMA y a la aplicación de la metodología Box-Jenkins para ajustar un modelo matemático a una serie de tiempo. Se exponen de forma práctica los pasos a seguir en el modelado de series de datos con las consideraciones más importantes que se deben tomar en cuenta.

En el capítulo cinco se ajustan tres modelos ARIMA que muestran suficiencia a los *incrementos de esfuerzos* provenientes del monitoreo permanente implementado en el Puente San Cristóbal del 1 de abril del 2007 al 18 de diciembre de 2008. Al final del capítulo se pone a prueba la capacidad del modelo matemático ajustado con los datos capturados del 19 de abril del 2007 al 28 de marzo de 2009.

En el capítulo seis se utiliza uno de los tres modelos ajustados en el capítulo cinco para desarrollar dos herramientas que sirven para el control y detección de cambios significativos en el comportamiento estructural: el cuadro de control y el algoritmo clasificador.

Finalmente en el capítulo siete se presentan las conclusiones y recomendaciones derivadas del desarrollo de este trabajo escrito referentes a la aplicación de series de tiempo en el monitoreo estructural.

1 Salud estructural

1.1 Antecedentes

Las obras de infraestructura de gran magnitud tales como puentes carreteros, túneles, presas y líneas de transporte colectivo, cumplen un papel estratégico en el desarrollo de muchas actividades que se llevan a cabo dentro de un entorno determinado. Estas estructuras benefician al ser humano en diversas formas. Sin embargo, cuando una obra de este tipo envejece o se ve sometida a acciones extraordinarias como sismos, estas pueden presentar fallas o un inadecuado funcionamiento, afectando de forma significativa las actividades culturales, sociales y económicas del ser humano.

La falta de acciones preventivas ante el deterioro y daño estructural en las obras civiles, trae diversas consecuencias negativas. La más importante es sin duda donde se involucra la integridad física de los miembros de una población. En general, aun cuando no se presenten pérdidas humanas, la población se ve afectada de gran manera ante una estructura parcial o totalmente fuera de servicio. El daño económico que se presenta también es de consideración. Esta puede ser directa e indirecta. En forma directa debido a los altos costos de reparación y de manera indirecta en las actividades económicas humanas que pudieran realizarse dentro de la estructura o cercana a ella.

Es por esto que las principales metas de la ingeniería estructural son lograr un buen diseño, calidad de construcción, durabilidad y seguridad de las estructuras. En la actualidad es reconocido el hecho que las construcciones más seguras y durables son aquellas que son vigiladas y controladas, es en este sentido que el monitoreo de la salud estructural suele jugar un papel importante para cumplir este propósito.

El estudio de la salud estructural en la ingeniería civil ha tomado gran importancia en los últimos años, llamando la atención del mundo académico y de investigación. El gran desarrollo y avance que ha experimentado en las últimas décadas se debe en parte a dos motivos principales. El primero es la gran aceptación en el sector de la construcción por los beneficios potenciales que puede llegar a reportar y el

segundo motivo es el surgimiento de nueva tecnología que hace posible el monitoreo de forma eficiente.

Los beneficios que representa el uso de sistemas de monitoreo se basan principalmente en la optimización de la operación, el mantenimiento y reparación de la infraestructura, lo cual ayuda a minimizar el riesgo de posibles colapsos repentinos que pudieran poner en riesgo vidas humanas. Una temprana detección de fallas en la estructura, permite que los costos de mantenimiento y reparación sean menores que si se detectan en una etapa futura cuando el problema haya agudizado.

1.2 El monitoreo de la salud estructural

En el comportamiento humano, el sistema nervioso detecta anomalías en el funcionamiento del cuerpo y manda dicha información al cerebro. El paciente acude al médico para ser examinado, obtener un diagnóstico y finalmente una cura. De manera similar, lo mismo ocurre en una estructura monitoreada, ya que el objetivo del monitoreo es detectar comportamientos inusuales que indiquen el inadecuado funcionamiento de una estructura. Esta detección de comportamientos inusuales requerirá una inspección, un diagnóstico y finalmente reparación (Glisic e Inaudi 2007).

El monitoreo de la salud estructural se define como un proceso cuyo objetivo es recabar información acerca de la condición y comportamiento de una estructura a través del tiempo. Este proceso consiste en la continua recopilación de los parámetros más representativos que indican el estado de una estructura.

La selección de los parámetros a recopilar dependen de varios factores, como el tipo de estructura, el propósito de la misma, los materiales de construcción, ciertas condiciones ambientales, etc. En términos generales estos factores pueden ser mecánicos, físicos o químicos. Algunos de estos parámetros se presentan en la *Tabla 1.1*.

La observación y recopilación de parámetros puede realizarse tanto de forma local como de forma global. Se refiere un monitoreo en forma local, a la observación del comportamiento de un material en específico, y un monitoreo en forma global a la observación de la estructura en conjunto. El uso del monitoreo en forma local o global dependerá del tipo de estructura. Normalmente estructuras innovadoras que implementen nuevos materiales de construcción requerirán un monitoreo local, mientras que estructuras convencionales así como construcciones antiguas, requieran un monitoreo global.

El tiempo de duración del monitoreo también se selecciona de acuerdo al tipo de estructura, así, este puede ser de corto, mediano o largo plazo. Al monitoreo de corto y largo plazo se les suele denominar también temporal y permanente, respectivamente.

El monitoreo permanente se utiliza cada vez con más frecuencia inclusive implementándose desde la etapa constructiva debido a que ayuda a los ingenieros a conocer y entender el comportamiento real de la estructura. En ciertos casos ayuda también a detectar errores constructivos que pueden inferir en el

funcionamiento futuro de la estructura, permitiendo aplicar las acciones correspondientes para corregir el problema.

Tipo	Parámetro
Mecánicas	Esfuerzo
	Defromación
	Desplazamiento
	Cargas actuantes
Físicas	Temperatura
	Humedad
	Presión de poro
Químicas	Penetración de cloro, sulfatos
	pH
	Oxidación del acero

Tabla 1.1 Parámetros más representativos de una estructura

Recientes tecnologías surgidas de las industrias aeronáutica, espacial y automotriz principalmente, aplicadas al monitoreo ha permitido un gran avance avances en este campo. No obstante, sólo un limitado número de aplicaciones son llevadas a la práctica. Esta situación se debe principalmente a tres problemas. El primero de ellos es que las técnicas de monitoreo estructural no son aún integradas en los programas de estudio de la ingeniería civil, por lo que existe cierto desconocimiento del tema. El segundo problema es que la eficiencia de los sistemas de monitoreo estructural sólo se puede apreciar a mediano plazo. El tercer problema es el manejo e interpretación de los grandes volúmenes de información que suelen obtenerse al aplicar un monitoreo a una estructura.

1.3 Etapas del monitoreo de la salud estructural

Las principales actividades que conforman el monitoreo estructural son: selección de la estrategia del monitoreo, instalación y mantenimiento del sistema seleccionado, procesamiento de la información y en caso de que sea interrumpido el monitoreo se procederá a desinstalar el equipo y a su almacenamiento. Cada una de estas etapas dividiéndose a su vez en sub-actividades que complementan el monitoreo de la salud estructural.

La *Figura 1.1* muestra un resumen de las actividades que conforman el monitoreo estructural y las sub-actividades que conlleva (Glisic e Inaudi 2007).

Como se puede apreciar en dicha figura, el desarrollo que presenta el monitoreo de estructuras lleva consigo la participación de un número importante de disciplinas, es decir, la intervención de distintos especialistas de los diversos campos de la ingeniería, así como de otras áreas. Resultan de ello, una gran cantidad de aspectos de interés a ser estudiados.

Sin dejar de lado los problemas de carácter operativo y de integración, que están relacionados con la selección los sistemas de adquisición, almacenamiento y transmisión de datos, así como las formas de comunicación entre los distintos componentes del sistema, es en el manejo y selección de información

y en el desarrollo de algoritmos para la evaluación del estado que guarda una estructura, en los cuales se pretende profundizar en esta propuesta de tesis.

En este sentido y de manera particular se busca indagar en la aplicación de algoritmos, que partiendo del hecho de que la información que se obtiene de manera experimental puede ser tratada bajo el campo de las series de tiempo, desembocan en la elaboración de modelos estadísticos que permitan tanto describir, como llegar a predecir, el comportamiento de una obra civil determinada, a través de los parámetros más representativos que lo permitan.

Estrategia de monitoreo	Instalación del sistema de monitoreo	Mantenimiento del sistema de monitoreo	Manejo de la información	Actividades finales
<ul style="list-style-type: none"> Objetivos del monitoreo Selección de los parámetros a monitorear Selección del sistema de monitoreo Ubicación de sensores Calendario de actividades Costos 	<ul style="list-style-type: none"> Instalación de sensores Instalación de accesorios Instalación de programas Instalación de interfaces 	<ul style="list-style-type: none"> Abastacimient de energía eléctrica Mantenimiento de líneas de comunicación Planes de mantenimiento Reparaciones 	<ul style="list-style-type: none"> Ejecución de mediciones Almacenamiento de la información Acceso a la información Visualización Exportación de información Interpretación Análisis Uso de la información 	<ul style="list-style-type: none"> Finalización del monitoreo Retiro del sistema de monitoreo Almacenamiento del sistema de monitoreo

Figura 1.1 Etapas del monitoreo estructural

1.4 Manejo de la información

El manejo de la información producto del monitoreo estructural puede ser básico o avanzado. El manejo básico de la información consiste en la captura y almacenamiento de la información. La captura de datos puede ser manual, semi-automática o automática, en sitio o remotamente, periódica o continuamente, estáticamente o dinámicamente. Estas opciones pueden ser combinadas de diferentes maneras según sea necesario. Por ejemplo, en el caso del monitoreo permanente es preferible una captura automática y continua, cuya información se pueda acceder de forma remota sin necesidad de presencia humana en el sitio. Algunos posibles métodos de captura se presentan en la *Figura 1.2*.

El almacenamiento de la información puede ser en forma de tablas, diagramas, etc., guardados en distintos medios electrónicos (CD, DVD, etc.) o impresos en papel. El medio de almacenamiento debe ser seguro y contar con copias de seguridad a fin de garantizar que la información nunca se pierda, además de ser de fácil acceso. La *Figura 1.3* muestra los dos tipos de almacenamiento que se puede tener.

El manejo de la información producto de la captura requiere el uso de programas y sistemas de cómputo ya que sin ellos sería difícil el manejo de grandes volúmenes de información.

Los programas utilizados para el manejo de la información son herramientas muy completas que ayudan a la interpretación, visualización, exportación análisis y uso de la información. El uso de estas herramientas ayudan a transformar la datos obtenidos en información útil que darán cuenta del estado de la estructura; sin éstas, las capturas sólo representarían una lista interminable de números. Igualmente que en la captura de datos, la transformación de la información puede ser manual, semi-automática o automática.

El tratamiento de la información de forma manual consiste en la interpretación, visualización, exportación y análisis. Es práctico cuando la cantidad de información es limitada. El tratamiento de forma semi-automática consiste normalmente en la exportación de datos en forma manual y un análisis automático de la información. El tratamiento de forma automática es la más conveniente debido a que puede ser realizada de forma rápida y no depende de la cantidad de información recopilada.

Finalmente el tipo de manejo de la información a utilizar deberá planearse en conjunto con la selección de la estrategia de monitoreo. Los procedimientos y herramientas a utilizar deberán ser compatibles con el sistema de monitoreo seleccionado.

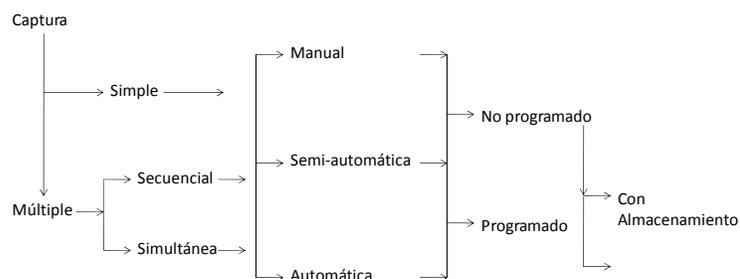


Figura 1.2 Métodos de captura

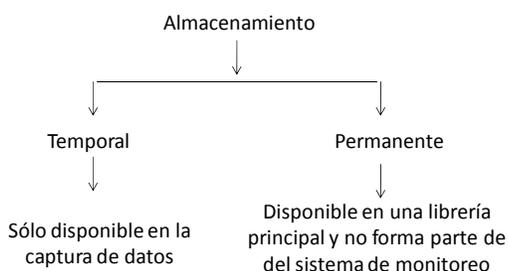


Figura 1.3 Tipos de almacenamiento

2 Conceptos fundamentales de series de tiempo

2.1 Definición de una serie de tiempo

Una serie de tiempo es el conjunto de mediciones que describen la evolución de un fenómeno o variable a lo largo del tiempo (Pepió 2001).

De manera formal, es una la secuencia cronológicamente ordenada de valores de medición sobre el estado de una variable cuantitativa de un fenómeno o proceso. Dichas mediciones están ordenadas respecto al tiempo y son generalmente dependientes entre sí. Esta dependencia entre las observaciones jugará un papel importante en el análisis de la serie.

Así entonces, una serie de tiempo puede representar desde los precios de un artículo, las tasas de desempleo, la temperatura máxima diaria, la velocidad del viento, hasta los esfuerzos y temperaturas en diversos puntos de una obra civil instrumentada.

Las observaciones de una serie de tiempo pueden ser denotadas como:

$$Y_1, Y_2, Y_3, Y_4, \dots, Y_{t-1}, Y_t$$

Donde:

Y_t : es el valor tomado por el proceso en el instante t

$t = 1, 2, 3, \dots, etc.$

Al referirnos a las series tiempo cabe distinguir entre aquellas cuyas observaciones son tomadas en cada punto de tiempo o series simultáneas y las que no tienen un valor instantáneo en cada punto, dado que son acumulación de valores, como las series temporales acumuladas. Esta distinción no tiene efectos sobre el análisis, puesto el método y la teoría de las series de tiempo se aplican por igual, a ambas modalidades (Arнау 2001).

Los valores de una serie de tiempo se clasifican en continuos o discretos. Los valores continuos son aquellos que se miden con cifras decimales. Por ejemplo: $Y_1 = 12.35$, $Y_2 = 45.67$, etc. Valores discretos son aquellos que sólo se miden en enteros. Por ejemplo: $Y_1 = 24$, $Y_2 = 56$, etc. (Pankratz 1983).

Por otro lado los valores de una serie de tiempo se pueden medir en periodos de tiempo continuos o discretos. Una serie medida en periodos de tiempo continuos se presenta cuando los valores de estado de la variable se obtienen de forma permanente o fluida en el tiempo. Ofrece la medida instantánea del estado de la variable observada, teóricamente para momentos de tiempo tan pequeños como se quiera. Por ejemplo: $Y_{1.1}$, $Y_{1.2}$, $Y_{1.3}$, etc.

En una serie medida en periodos de tiempo discretos, los valores se obtienen para intervalos de tiempo usualmente homogéneos y representa una magnitud acumulada del estado de la variable durante cada uno de esos intervalos. Por ejemplo: Y_1 , Y_2 , Y_3 , etc.

Conviene resaltar que en la mayoría de la bibliografía disponible se utiliza el término *serie discreta* para cuando los valores se miden en periodos de tiempo discretos y serie continua cuando los valores se miden en periodos de tiempo continuos.

2.2 Tratamiento de series de tiempo

Como se mencionó en el capítulo anterior, el manejo de la información producto de un monitoreo es de suma importancia para poder conocer el estado que guarda una estructura. Si se aprovecha al máximo esta información, se obtendrán resultados útiles y se podrán alcanzar los objetivos planteados al implementar un sistema de monitoreo estructural.

La manera más común de tratar la información es en forma directa. Esto significa que se analiza la información de manera visual, gráfica y con la obtención de los parámetros estadísticos más representativos tales como la media, varianza, desviación estándar, etc. De esta manera se obtienen estimaciones confiables acerca del estado actual de una estructura. No obstante, la interpretación final depende de la experiencia adquirida del analista a través del tiempo. De esta manera, aunado al gran volumen de datos recopilados, en algunas ocasiones puede ser difícil el análisis e interpretación y se puede inducir con ello a decisiones equivocadas.

Por estas razones se hace necesario y se propone algún tipo de tratamiento que ayude a extraer resultados y conclusiones útiles que complementen el análisis actual de la información, especialmente cuando se tiene gran cantidad de datos.

Las características de la información capturada en un monitoreo estructural permite que su análisis sea bajo el marco de *series de tiempo*, ya que los parámetros recopilados están asociados a un momento del tiempo determinado, es decir, se conoce el valor de la variable en momento de tiempo específico. Asimismo, estas capturas de información normalmente se hacen separadas equidistantemente en el tiempo, cada hora, cada día, cada semana, etc., según se requiera. Característica que como se verá en próximos capítulos, será importante para el análisis de la serie.

Dentro del marco de series de tiempo, resalta el disponer de información histórica para poder intentar predecir, aunque sea de forma aproximada, un comportamiento futuro que facilite la toma de decisiones cuando pretendemos modificar alguna determinada realidad, sobre todo con el fin de planificar, prever o prevenir.

Es claro que ante una toma de decisión existen diversas alternativas posibles. Sin embargo estas se encuentran en un ambiente de fluctuaciones de forma permanente, se hace presente, por tanto, una gran incertidumbre ante el futuro.

Para reducir precisamente este grado de incertidumbre se suele recurrir a la elaboración de previsiones que tratan de anticipar la evolución de algún fenómeno. Justamente bajo el marco de series temporales, el objetivo principal al analizar las series de tiempo es reducir la incertidumbre sobre el futuro e inclusive en el mismo presente del fenómeno que se desee estudiar. Es decir, el análisis de series de tiempo se puede utilizar también con el único propósito de describir la evolución de cierto fenómeno o conocer el estado actual de cierta variable, que a simple vista resultaría difícil de detectar.

La estadística ha desarrollado teoría y métodos que apuntan a resolver el problema de predicción mediante el uso de series de tiempo. Sin embargo, este no puede ser resuelto por argumentos puramente matemáticos, debe ser el resultado de una combinación matemático-especialista.

Existen distintos tipos de análisis que han surgido a lo largo del tiempo. Dos de ellos resaltan por su eficacia y su amplia divulgación: *La modelación clásica* y *La modelación ARIMA*. Cada uno de ellos trabajan bajo enfoques distintos: el enfoque determinista y el enfoque estocástico, respectivamente. Ambos se detallarán en los capítulos 3 y 4 respectivamente.

2.3 Herramientas básicas para el análisis de series de tiempo

Para entrar a detalle al estudio de series de tiempo es necesario aprender el uso de algunas herramientas que proporcionan información importante acerca de ciertas particularidades presentes en series de datos. Podemos mencionar, entre algunas otras particularidades, la detección patrones ascendentes o descendentes, oscilatorios, periódicos, etc.

Las herramientas que se presentan a continuación sirven tanto para la modelación clásica como para la modelación ARIMA.

2.3.1 Gráfico de la serie

Por muy simple y obvio que pudiera parecer, la exploración gráfica es (o debe ser) el punto de partida para cualquier análisis que se pretenda realizar a una serie de datos. Mediante el gráfico de la serie se puede detectar si los valores siguen alguna regularidad o algún movimiento armónico, también cambios sistemáticos en la media e inclusive cambios en la varianza.

Aunque la exploración del gráfico de una serie no es prueba definitiva para comprobar la existencia de algún tipo de patrón (tendencia, estacionalidad, ciclo, etc.), sí puede proporcionar una idea primaria sobre alguna de estas características.

2.3.2 Funciones de autocorrelación

Una parte fundamental en el estudio de series de tiempo, bajo la modelación clásica y especialmente en la modelación ARIMA (que se estudiará posteriormente) gira en torno a la relación que pueda existir entre los valores de una serie de tiempo. A esta relación o correlación entre los valores de una misma variable se le llama autocorrelación.

La autocorrelación se define como la relación mutua existente entre valores de una serie de tiempo en diferentes periodos y describe lo que tiende a sucederle a un valor si se da un cambio en el otro

Se distinguen dos tipos de autocorrelación que sirven para conocer tanto la existencia de correlación entre los datos de la serie como el grado o intensidad y el tipo de la misma (Bowerman, O'Connell, Koehler 2007).

Autocorrelación simple: Mide la relación lineal entre las observaciones de una serie de datos Y_t , distanciados en un lapso de tiempo k . A este lapso de tiempo k se le conoce como retardo o retraso. Este retardo denota el periodo de tiempo entre los valores de la serie, para el cual se mide el tipo y grado de correlación de la variable considerada.

Cabe aclarar que la autocorrelación simple ofrece una medida de la relación para todos los valores de la serie respecto a los valores de la serie observados k unidades de tiempo antes, comparando la variación conjunta de todos los valores de Y_t en intervalos de k unidades de tiempo, con la variación de la secuencia respecto a esos valores.

Autocorrelación parcial: Es una medida asociada a la autocorrelación simple. De forma sencilla se puede describir la autocorrelación parcial como la estimación de la autocorrelación simple, para el mismo retardo k , con la eliminación del efecto producido por las autocorrelaciones para retardos menores a k , las cuales están presentes en la estimación de la autocorrelación simple. Dicho de otra forma, la autocorrelación parcial no considera las autocorrelaciones acumuladas para el retardo k para el que se estima (Aguirre 1994).

La diferencia específica entre estos dos tipos de autocorrelación se halla en que la autocorrelación simple brinda para un retardo k tanto la relación entre las observaciones con una diferencia de k retardos de tiempo, como la relación para retardos menores, mientras que la autocorrelación parcial brinda sólo la relación para la diferencia estricta en k retrasos de tiempo (Aguirre 1994).

El grado de relación, tanto para la autocorrelación simple y parcial, se mide mediante el coeficiente de autocorrelación, que también se distingue entre simple y parcial. Ambos coeficientes de autocorrelación varían entre +1 y -1.

Estos coeficientes además de detectar la existencia de relación ó correlación, miden el tipo y grado de esta. Por ejemplo, un valor aproximado de +1 implica una fuerte relación entre dos observaciones separadas k unidades de tiempo. El signo positivo (+), denota una relación directa entre las dos observaciones, es decir, que cuando el valor de alguna de ellas aumenta, el otro también tenderá a aumentar; o viceversa, si disminuye un valor, el otro tenderá a disminuir también. Un coeficiente de autocorrelación cercano a -1 expresa que la relación entre las observaciones es inversa, el aumento del valor de la variable se asocia a la disminución de la otra. Un coeficiente de 0 indica que los dos valores de una misma serie de tiempo no están relacionados (Makridakis y Wheelwright 2000).

Los patrones de datos, que incluyen componentes como la tendencia y estacionalidad, se pueden estudiar si se usa esta herramienta. Las autocorrelaciones proporcionan información importante acerca de la estructura de un conjunto de datos y de sus patrones. Por ejemplo, en un conjunto de datos completamente aleatorios la autocorrelación entre valores sucesivos estará cercana a 0, o será igual a 0, pero los valores de datos de fuerte naturaleza estacional o cíclica estarán sumamente autocorrelacionados (Hanke 2006).

El cálculo del coeficiente de autocorrelación simple (r_k) entre observaciones que se encuentran a k periodos de distancia se realiza por medio de la siguiente expresión (Bowerman, O'Connell, Koehler 2007):

$$r_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

Donde:

r_k : *coeficiente de autocorrelación simple para un retraso de k periodos*

\bar{Y} : *media de los valores de la serie*

Y_t : *observación en el periodo t*

Y_{t+k} : *observación en el periodo con k retrasos*

n : *número total de periodos*

Una vez conocida la expresión para el cálculo de este coeficiente, es necesario aclarar que en la práctica cuando se tienen datos sin correlación, r_k puede ser diferente de cero, con lo que surge el problema de establecer un intervalo alrededor de $r_k = 0$ que ayude a juzgar si la diferencia a cero del coeficiente r_k es significativa o no.

El intervalo de significación para r_k se construye usualmente con el error típico:

$$S_{r_k} = \begin{cases} \frac{1}{\sqrt{n}} & \text{si } k = 1 \\ \sqrt{\frac{1 + 2 \sum_{i=1}^{k-1} r_i^2}{n}} & \text{si } k = 2, 3 \dots \end{cases}$$

En conjunto con este valor, se puede estimar la estadística t_{r_k} del error típico que esta asociada a la estimación del coeficiente de autocorrelación simple, para juzgar sobre la significación del mismo:

$$t_{r_k} = \frac{r_k}{S_{r_k}}$$

Por el momento no se profundizará acerca del uso de S_{r_k} y t_{r_k} , tan sólo se presentan para referencia futura.

De manera similar, el cálculo del coeficiente de autocorrelación parcial (r_{kk}) entre observaciones que se encuentran a k periodos de distancia se realiza por medio de la siguiente expresión:

$$r_{kk} = \begin{cases} r_1 & \text{si } k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1j} r_j} & \text{si } k = 2, 3 \dots \end{cases}$$

Donde:

$$r_{kj} = r_{k-1j} - r_{kk} r_{k-1k-j} \quad \text{para } j = 1, 2, \dots, k-1$$

El error estándar de r_{kk} es:

$$S_{r_{kk}} = \frac{1}{\sqrt{n}}$$

La estadística $t_{r_{kk}}$ es:

$$t_{r_{kk}} = \frac{r_{kk}}{S_{r_{kk}}}$$

Los coeficientes de autocorrelación, ya sea simple o parcial, que se obtienen para los diferentes periodos $k = 1,2,3 \dots etc$ se grafican de la manera en que se muestra en la *Figura 2.1*. A esta gráfica se le conoce como correlograma o función de autocorrelación.

El correlograma o función de autocorrelación es una gráfica de las autocorrelaciones para varios retrasos de tiempo de una serie de tiempo.

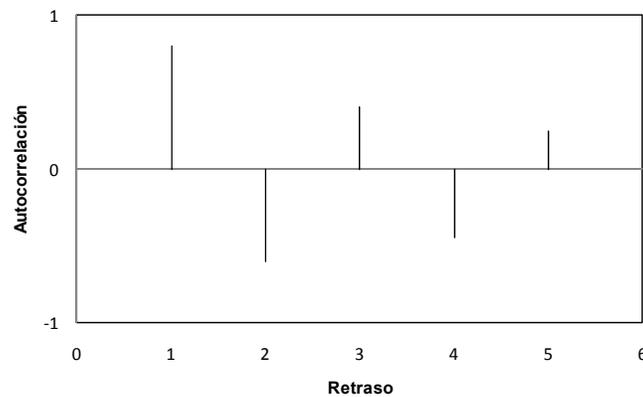


Figura 2.1 Correlograma o función de autocorrelación

En el correlograma de la *Figura 2.1*, la escala horizontal en la parte de debajo de la gráfica muestra cada uno de los retrasos de tiempo que son de interés ($k = 1,2,3 \dots et$), mientras que la escala vertical del lado izquierdo muestra el posible rango de un coeficiente de autocorrelación de -1 a +1. La línea horizontal a la mitad de la gráfica representa autocorrelaciones de cero.

Para simplificar el uso de estos términos, la *función de autocorrelación simple* se abreviará con las siglas **FAS** y a la *función de autocorrelación parcial* como **FAP**.

Algunos autores como Box y Jenkins sugieren que el número máximo de autocorrelaciones calculadas (tanto simple y parcial) sea de $n/4$, donde n es el número total de observaciones.

Para ejemplificar el cálculo de los coeficientes de autocorrelación simple y parcial, la obtención de los errores estándar, la estadística t , así como la obtención de la función de autocorrelación (correlograma) se hace referencia a los datos de la *Tabla 2.1* y su correspondiente gráfico que se muestra en la *Figura 2.2*.

Tiempo t	Y_t
1	123
2	130
3	125
4	138
5	145
6	142
7	141
8	146
9	147
10	157
11	150
12	160

Tabla 2.1 Datos serie de tiempo

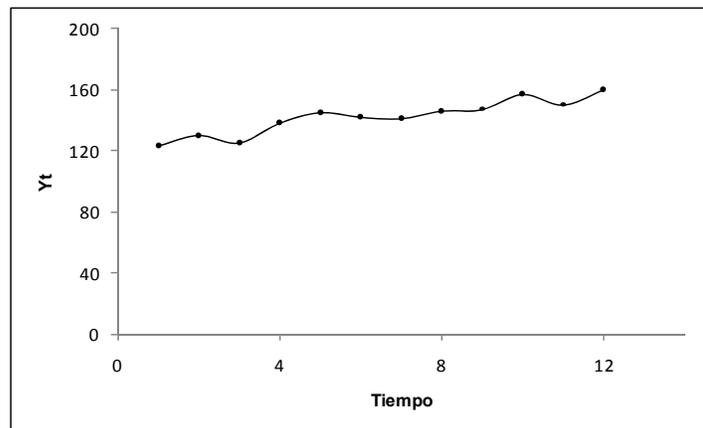


Figura 2.2 Gráfico de la serie de tiempo

Se comenzará por el cálculo de los coeficientes de autocorrelación simple, recordando el número máximo de coeficientes $n/4$. Para la serie de este ejemplo $n = 12$ datos, así se calcularán tan sólo 3 coeficientes de autocorrelación (r_1, r_2, r_3).

Tomando en cuenta que:

$$r_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

Se tiene entonces que para el coeficiente r_1 , $n = 12$ y $k = 1$ y a partir de los datos de la *Tabla 2.1*:

$$r_1 = \frac{\sum_{t=1}^{11} (Y_t - \bar{Y})(Y_{t+1} - \bar{Y})}{\sum_{t=1}^{12} (Y_t - \bar{Y})^2}$$

También:

$$\bar{Y} = \left(\frac{1}{n}\right) \sum_{t=1}^n Y_t = \left(\frac{1}{12}\right) 1740 = 142$$

Desarrollando tenemos que:

$$r_1 = \frac{(Y_1 - \bar{Y})(Y_2 - \bar{Y}) + (Y_2 - \bar{Y})(Y_3 - \bar{Y}) + (Y_3 - \bar{Y})(Y_4 - \bar{Y}) + \dots + (Y_{11} - \bar{Y})(Y_{12} - \bar{Y})}{(Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2 + (Y_3 - \bar{Y})^2 + \dots + (Y_{12} - \bar{Y})^2}$$

Donde:

$$Y_1 = 123$$

$$Y_2 = 130$$

$$Y_3 = 125$$



$$Y_{12} = 160$$

Sustituyendo:

$$r_1 = \frac{(123 - 142)(130 - 142) + (130 - 142)(125 - 142) + \dots + (150 - 142)(160 - 142)}{(123 - 142)^2 + (130 - 142)^2 + (125 - 142)^2 + \dots + (160 - 142)^2}$$

$$r_1 = \frac{843}{1474}$$

$$r_1 = 0.5719$$

El coeficiente de autocorrelación r_2 para el segundo retardo de tiempo se calcula entonces con $n = 12$ y $k = 2$:

$$r_2 = \frac{\sum_{t=1}^{10} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{12} (Y_t - \bar{Y})^2}$$

Así:

$$r_2 = \frac{(Y_1 - \bar{Y})(Y_3 - \bar{Y}) + (Y_2 - \bar{Y})(Y_4 - \bar{Y}) + (Y_3 - \bar{Y})(Y_5 - \bar{Y}) + \dots + (Y_{10} - \bar{Y})(Y_{12} - \bar{Y})}{(Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2 + (Y_3 - \bar{Y})^2 + \dots + (Y_{12} - \bar{Y})^2}$$

$$r_2 = \frac{(123 - 142)(125 - 142) + (130 - 142)(138 - 142) + \dots + (157 - 142)(160 - 142)}{(123 - 142)^2 + (130 - 142)^2 + (125 - 142)^2 + \dots + (160 - 142)^2}$$

$$r_2 = \frac{682}{1474}$$

$$r_2 = 0.4627$$

De forma similar se procede para el cálculo de r_3 , así entonces:

$$r_3 = \frac{(Y_1 - \bar{Y})(Y_4 - \bar{Y}) + (Y_2 - \bar{Y})(Y_5 - \bar{Y}) + (Y_3 - \bar{Y})(Y_6 - \bar{Y}) + \dots + (Y_9 - \bar{Y})(Y_{12} - \bar{Y})}{(Y_1 - \bar{Y})^2 + (Y_2 - \bar{Y})^2 + (Y_3 - \bar{Y})^2 + \dots + (Y_{12} - \bar{Y})^2}$$

$$r_3 = \frac{(123 - 142)(138 - 142) + (130 - 142)(145 - 142) + \dots + (147 - 142)(160 - 142)}{(123 - 142)^2 + (130 - 142)^2 + (125 - 142)^2 + \dots + (160 - 142)^2}$$

$$r_3 = \frac{163}{1474}$$

$$r_3 = 0.1106$$

Una vez calculados los coeficientes, creamos la función de autocorrelación simple (FAS) o correlograma, como se muestra en la *Figura 2.3*.

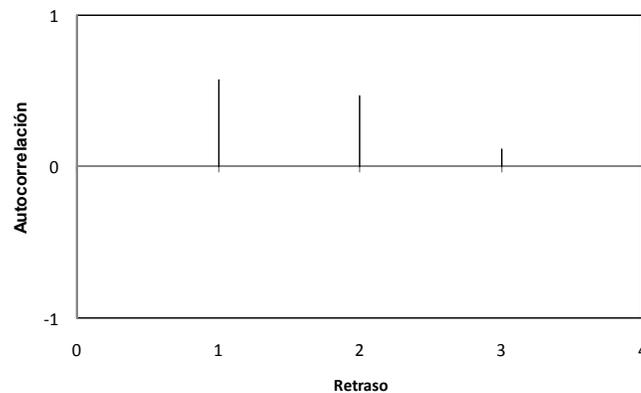


Figura 2.3 Función de autocorrelación simple (FAS)

Ahora bien, a partir de r_1, r_2, r_3 , obtenemos sus correspondientes errores estándares y su estadístico t asociado.

Recordando las fórmulas:

$$S_{r_k} = \begin{cases} \frac{1}{\sqrt{n}} & \text{si } k = 1 \\ \sqrt{\frac{1 + 2 \sum_{i=1}^{k-1} r_i^2}{n}} & \text{si } k = 2, 3 \dots \end{cases}$$

$$t_{r_k} = \frac{r_k}{S_{r_k}}$$

Para $r_1, k = 1$ y $n = 12$; entonces:

$$S_{r_1} = \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{12}}$$

$$S_{r_1} = 0.2887$$

$$t_{r_1} = \frac{r_1}{S_{r_1}} = \frac{0.5719}{0.2887}$$

$$t_{r_1} = 1.98$$

Para $r_2, k = 2$ y $n = 12$; entonces:

$$S_{r_2} = \sqrt{\frac{1 + 2 \sum_{i=1}^{2-1} r_i^2}{12}} = \sqrt{\frac{1 + 2 r_1^2}{12}} = \sqrt{\frac{1 + 2 (0.5719^2)}{12}}$$

$$S_{r_2} = 0.3713$$

$$t_{r_2} = \frac{r_2}{S_{r_2}} = \frac{0.4627}{0.3713}$$

$$t_{r_2} = 1.25$$

Para r_3 , $k = 3$ y $n = 12$; entonces:

$$S_{r_2} = \sqrt{\frac{1 + 2 \sum_{i=1}^{3-1} r_i^2}{12}} = \sqrt{\frac{1 + 2 (r_1^2 + r_2^2)}{12}} = \sqrt{\frac{1 + 2 (0.5719^2 + 0.4627^2)}{12}}$$

$$S_{r_2} = 0.4166$$

$$t_{r_3} = \frac{r_3}{S_{r_3}} = \frac{0.1106}{0.4166}$$

$$t_{r_1} = 0.27$$

Siguiendo en el ejemplo, se calcularán ahora los coeficientes de autocorrelación parcial r_{11} , r_{22} , r_{33} , sus respectivos errores y estadísticos t . Se procederá de forma similar a la mostrada en el cálculo de autocorrelaciones simples.

Teniendo en cuenta que:

$$r_{kk} = \begin{cases} r_1 & \text{si } k = 1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1-j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1-j} r_j} & \text{si } k = 2, 3 \dots \end{cases}$$

Entonces para $k = 1$:

$$r_{11} = r_1$$

$$r_{11} = 0.5719$$

Para $k = 2$:

$$r_{22} = \frac{r_2 - \sum_{j=1}^{2-1} r_{2-1-j} r_{2-j}}{1 - \sum_{j=1}^{2-1} r_{2-1-j} r_j}$$

$$r_{22} = \frac{r_2 - r_{11} r_1}{1 - r_{11} r_1} = \frac{0.4627 - (0.5719)(0.5719)}{1 - (0.5719)(0.5719)}$$

$$r_{22} = 0.2015$$

Para $k = 3$:

$$r_{33} = \frac{r_3 - \sum_{j=1}^{3-1} r_{3-1j} r_{3-j}}{1 - \sum_{j=1}^{3-1} r_{3-1j} r_j}$$

$$r_{33} = \frac{r_3 - (r_{21} r_2 + r_{22} r_1)}{1 - (r_{21} r_1 + r_{22} r_2)}$$

Donde:

$$r_{kj} = r_{k-1j} - r_{kk} r_{k-1k-j}$$

$$r_{21} = r_{11} - r_{22} r_{11}$$

$$r_{21} = 0.5719 - (0.2015)(0.5719)$$

$$r_{21} = 0.4567$$

Entonces:

$$r_{33} = \frac{0.1106 - [(0.4567)(0.4627) + (0.2015)(0.5719)]}{1 - [(0.4567)(0.5719) + (0.2015)(0.4627)]}$$

$$r_{33} = -0.3345$$

La respectiva función de autocorrelación parcial (FAP) se observa en la *Figura 2.4*.

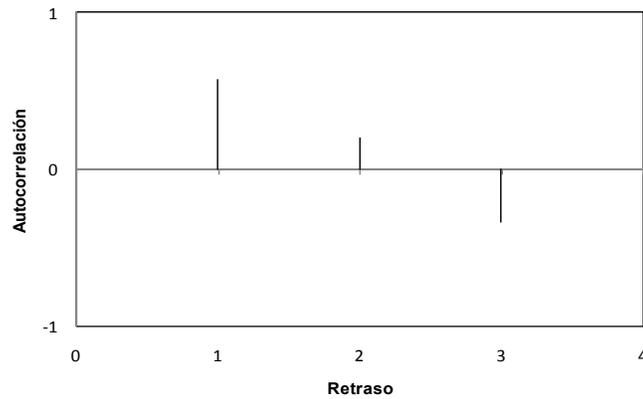


Figura 2.4 Función de autocorrelación parcial (FAP)

Los errores estándar para cada uno de los coeficientes calculados son:

$$S_{r_{kk}} = \frac{1}{\sqrt{n}}$$

$$S_{r_{11}} = S_{r_{22}} = S_{r_{33}} = \frac{1}{\sqrt{12}} = 0.2887$$

La estadística $t_{r_{kk}}$ para cada uno de los coeficientes es:

$$t_{r_{kk}} = \frac{r_{kk}}{S_{r_{kk}}}$$

$$t_{r_{11}} = \frac{r_{11}}{S_{r_{11}}} = \frac{0.5719}{0.2887}$$

$$t_{r_{11}} = 1.98$$

$$t_{r_{22}} = \frac{r_{22}}{S_{r_{22}}} = \frac{0.2015}{0.2887}$$

$$t_{r_{22}} = 0.70$$

$$t_{r_{33}} = \frac{r_{33}}{S_{r_{33}}} = \frac{-0.3345}{0.2887}$$

$$t_{r_{33}} = -1.16$$

Para finalizar este ejemplo de aplicación, en la *Figura 2.5* se resumen los cálculos realizados. De acuerdo con los coeficientes de autocorrelación simple se puede inferir (para la serie de ejemplo) que existe una mediana relación de los valores respecto a aquellos distanciados un retardo de tiempo y la relación disminuye para los retardos 2 y 3. Por otra parte los coeficientes de autocorrelación parcial indican una mediana relación directa de los valores con aquellos separados 1 y 2 retrasos de tiempo y una relación inversa con el valor separado 3 retrasos de tiempo.

Coefficiente	Valor	Error estándar	Valor	Estadístico t	Valor
r_1	0.572	S_{r_1}	0.289	t_{r_1}	1.981
r_2	0.463	S_{r_2}	0.371	t_{r_2}	1.246
r_3	0.111	S_{r_3}	0.417	t_{r_3}	0.265
r_{11}	0.572	$S_{r_{11}}$	0.289	$t_{r_{11}}$	1.981
r_{22}	0.202	$S_{r_{22}}$	0.289	$t_{r_{22}}$	0.698
r_{33}	-0.335	$S_{r_{33}}$	0.289	$t_{r_{33}}$	-1.159

Figura 2.5 Cuadro resumen

Un comentario que resulta interesante e importante destacar, producto del ejemplo anterior, es que para una serie con $n = 12$ datos, la obtención de los coeficientes de autocorrelación simples y parciales puede resultar un trabajo laborioso. Es intuitivo que para una serie de tiempo con más observaciones, la dificultad de esta labor se complique aún más. El problema que ello conduce es que la tarea de analizar una serie de tiempo se vea mermada y se retrase de manera significativa. No es práctico detenerse demasiado tiempo en la obtención de las funciones de autocorrelación simple y parcial (FAS y FAP), debido al hecho de que aunque son una herramienta fundamental para el análisis de series de tiempo, de ella no se obtienen directamente los resultados que se buscan al estudiar una serie de datos. Por ese motivo es recomendable y más conveniente usar programas de cómputo estadístico que ayude en la obtención de dichos coeficientes. Algunos paquetes estadísticos pueden ser MINITAB, SAS, algún complemento para MATLAB o EXCEL, entre otros. Estos paquetes además ofrecen diversas opciones para el análisis y tratamiento de series de tiempo (algunas de estas opciones se mencionarán en otros capítulos del presente texto).

Para introducir el uso de los programas de cómputo al tratamiento de las series de tiempo, se compara la salida del programa MINITAB con los resultados del ejemplo expuesto anteriormente. La salida del programa se muestra en la *Figura 2.6* donde se observa el valor de los coeficientes de autocorrelación simple. En la *Figura 2.7* se muestran los coeficientes de autocorrelación parcial.

Autocorrelation Function: Datos Serie

Lag	ACF	T	LBQ
1	0,571913	1,98	5,00
2	0,462687	1,25	8,59
3	0,110583	0,27	8,82

Figura 2.6 Coeficientes de autocorrelación simple

Partial Autocorrelation Function: Datos Serie

Lag	PACF	T
1	0,571913	1,98
2	0,201514	0,70
3	-0,334512	-1,16

Figura 2.7 Coeficientes de autocorrelación parcial

Podemos observar que la salida de MINITAB contiene, además del cálculo de los coeficientes, el estadístico t , mismo que se obtuvo en el ejemplo de los datos de la *Tabla 2.1*. Se obtienen los mismos resultados que en el ejemplo.

MINITAB también ofrece la posibilidad de obtener las gráficas de la FAS y la FAP, tal como se muestra en las *Figuras 2.8* y *2.9* respectivamente. Las dos bandas de puntos que se observan en la gráfica representan el valor de dos desviaciones estándar para cada retraso. Es decir $2(S_{r_k})$ por arriba y $2(S_{r_k})$ por debajo del eje central. Más adelante analizaremos la ayuda que ofrecen estas bandas.

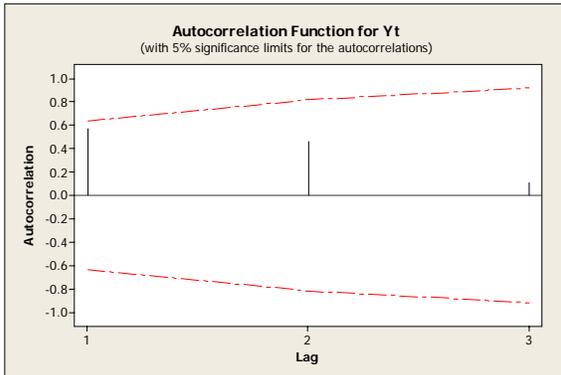


Figura 2.8 Función de autocorrelación simple

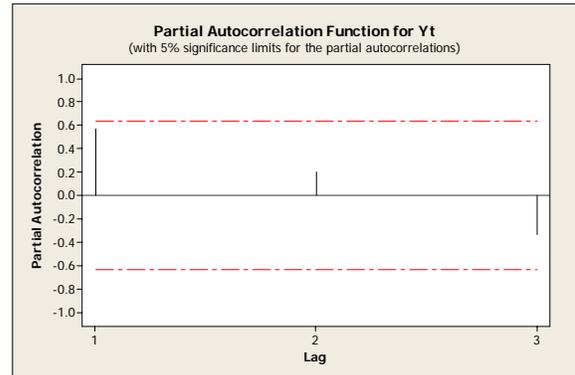


Figura 2.9 Función de autocorrelación parcial

2.3.3 Periodograma

El periodograma es un instrumento que sirve para detectar estacionalidad en una serie y determinar su periodo (Aguirre 1994). Estas cuestiones se estudiarán en los próximos capítulos con más de detalle, por el momento sólo se mencionarán las características básicas de este instrumento.

Para comprender el funcionamiento de esta herramienta y su utilidad, supongamos que una señal en el tiempo presenta movimientos oscilatorios. Así, si establecemos una unidad arbitraria de tiempo, se puede medir la cantidad de ondulaciones o impulsos producidos por la señal durante esa unidad de tiempo. A esta cantidad de ondulaciones o fluctuaciones en la unidad de tiempo se le denomina frecuencia de la señal. Esta frecuencia se mide en ciclos en una unidad de tiempo. Ahora bien, si en vez de medir la cantidad de ciclos o impulsos producidos por la señal en una unidad de tiempo, medimos el tiempo invertido por esa señal en completar un ciclo, estamos obteniendo el periodo de la señal. Como se verá un poco más adelante, es ahí donde hallamos la utilidad de esta herramienta en una serie de tiempo, donde se requiere conocer el periodo estacional; es decir, se desea conocer el tiempo invertido por la variable en realizar un ciclo regular o repetitivo de manera completa. El periodo se mide en unidades de tiempo.

De la definición de frecuencia y periodo se deducen las siguientes relaciones:

$$Periodo = \frac{1}{Frecuencia}$$

y

$$Frecuencia = \frac{1}{Periodo}$$

Asociado a estos términos se incluye otro que resulta de interés: la amplitud. La amplitud en una señal es la distancia desde la base hasta la cima o cresta de un ciclo de esa señal.

El periodograma se construye a partir de transformar la serie de tiempo, llevando la variable del dominio del tiempo (su forma natural) al dominio de frecuencias. Esto se realiza aplicando una transformada a los valores de la serie. Existen diferentes transformadas para la construcción de un periodograma, la más usual y aplicada es la transformada rápida de Fourier. En la *Figura 2.10* se muestra el periodograma para una serie de tiempo. En el presente texto no se detalla el desarrollo de la transformada rápida de Fourier, sólo se mencionan los programas de cómputo *Degtra* (Ordaz y Montoya 2002) y *Scope* que proveen esta herramienta de transformación.

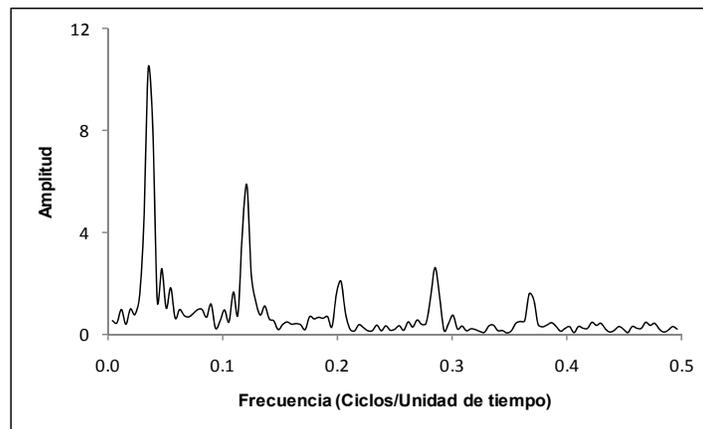


Figura 2.10 Periodograma

Ahora bien, la interpretación de un periodograma se realiza mediante la búsqueda de amplitudes destacables en frecuencias bajas, es decir, picos al inicio del gráfico. Una vez detectadas dichas amplitudes se busca la frecuencia correspondiente y se calcula el periodo correspondiente mediante:

$$p = \frac{1}{f}$$

De esta manera se obtiene un *periodo estacional posible* de la serie de tiempo. Decir que es un periodo *posible* hace referencia al hecho que el periodograma también puede mostrar amplitudes destacables no sólo en las frecuencias correspondientes a los verdaderos periodos estacionales, sino también en los múltiplos y submúltiplos de dichas frecuencias. Para realizar la discriminación de los verdaderos periodos se debe comparar los resultados con los obtenidos de las funciones de autocorrelación simple y parcial.

Otra consideración es que para una serie extensa, un periodograma puede mostrar varios patrones estacionales y cíclicos. Este último se presenta para frecuencias mucho más bajas que para un patrón estacional (Aguirre 1994).

3 Modelación clásica

Es a finales del siglo XVII que se tienen los primeros indicios de estudios a conductas sociales, movimientos demográficos, enfermedades, así como tasas de nacimientos a partir de registros históricos. Se sentaron algunas de las bases de la metodología que está vigente en la actualidad. No es hasta el siglo XIX cuando se realizan los primeros pronósticos para enfermedades ajustando curvas a series de tiempo y extrapolando valores. Ya en el siglo XX se proponen las ideas de la modelación clásica en el tratamiento de series de tiempo. En un principio se desarrollo este método para el análisis de series económicas y al ir perfeccionándose, el método se empezó a aplicar en diversos ámbitos como en la salud, transportes, contaminación, etc., pudiéndose aplicar a la mayoría de series de tiempo de diversos campos de conocimiento (Aguirre 1994).

3.1 Enfoque determinístico

La principal característica de la modelación clásica es el enfoque determinístico con que se estudia y analiza una serie de tiempo.

El enfoque determinístico se refiere a que cualquier acontecimiento, responde a una causa, y así, una vez dada la causa, el acontecimiento ha de seguir sin posible variación y con ello supone la existencia de leyes físicas estrictas que rigen el comportamiento de las cosas (Rodríguez 2000).

Por medio de dichas leyes es posible predecir la totalidad de los acontecimientos futuros, esto sin aceptar cambios o variaciones, ya que no se acepta la existencia de hechos aleatorios o caóticos en la realidad. Estos últimos se consideran como errores (por fallas humanas y/o imperfecciones de los instrumentos de medición), o simplemente como una parte de la serie imposible de estudiar y modelar.

Específicamente en una serie de tiempo, el enfoque determinístico dicta que la variable observada presenta un patrón de comportamiento fijo, con lo cual las irregularidades de la serie se contemplan como una desviación respecto a una pauta de comportamiento sistemática (Rodríguez 2000).

Es importante conocer el enfoque determinístico para comprender cabalmente la metodología expuesta en las siguientes páginas y así identificar fácilmente las ventajas y desventajas que se derivan de ella.

La modelación clásica bajo el enfoque determinista, parte de considerar a la serie como un proceso dependiente del tiempo el cual rige el comportamiento de la variable en la serie. En otros términos, el modelo buscado para representar al proceso que gobierna el comportamiento de la serie es una función del tiempo (Aguirre 1994).

Una función del tiempo es una relación en la cual entra una unidad de tiempo y de la cual sale el valor correspondiente de la serie, con un cierto margen de error.

Esquemáticamente se puede representar la idea de función del tiempo de la siguiente manera:



Donde Y_t representa a los valores observados de la variable Y en el momento t y ϵ_t representa, según este enfoque determinista, un margen de error producido por la existencia en la serie de una perturbación aleatoria imposible de modelar.

Como se ha mencionado previamente, esta perturbación se atribuye a imperfecciones en la observación, recopilación y transmisión de la información obtenida, todas ellas presentes en cualquier sistema de medición. Como se observa en el esquema, la variable t es la variable independiente y la variable resultante Y_t es a su vez la variable dependiente. De forma abreviada se puede representar el esquema con la notación:

$$Y_t = f(t)$$

Con este concepto en mente, podemos decir que modelación clásica se centra en la búsqueda de la expresión analítica del proceso que gobierna el comportamiento observado en la serie.

Dicha función $Y_t = f(t)$ puede construirse por la composición de cuatro elementos que pueden aparecer como patrones básicos en una serie. A estos patrones o regularidades también se les suele denominar componentes, descriptores o determinantes. El procedimiento de análisis se suele llamar *descomposición* (Rodríguez 2000).

Las cuatro componentes de una serie son (Uriel 1995):

Tendencia (T_t). Movimiento ó dirección general de la variable en periodos prolongados de tiempo. *Figura 3.1*

Estacionalidad (E_t). Fluctuaciones periódicas de la variable, más o menos regulares, en periodos relativamente cortos de tiempo, con una oscilación repetitiva para lapsos de tiempo contiguos. *Figura 3.2*

Ciclo (C_t). Movimientos de la variable similares a la estacionalidad, pero relativos a periodos de tiempo mucho más prolongados. Sólo se detecta en series suficientemente largas. *Figura 3.3*

Aleatoriedad (a_t). Es el movimiento irregular de la variable, determinado por el azar, impredecible de forma determinística. A esta componente se le suele llamar residuo aleatorio o ruido. *Figura 3.4*

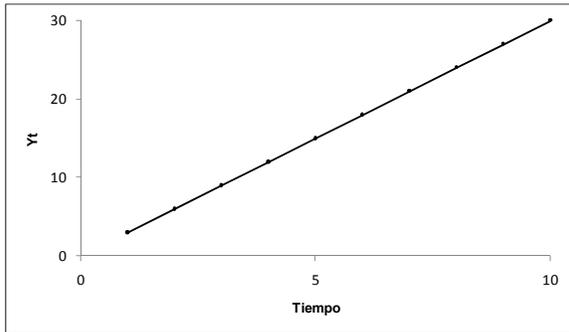


Figura 3.1 Tendencia

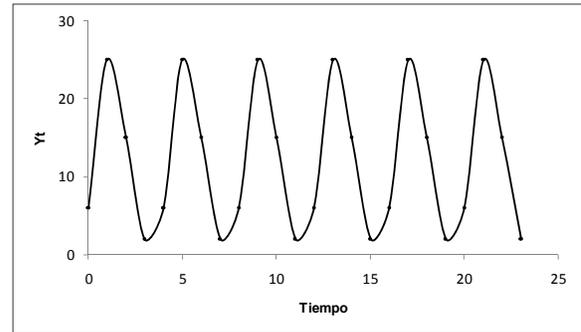


Figura 3.2 Estacionalidad

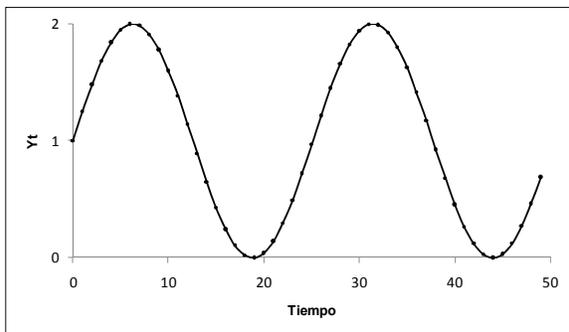


Figura 3.3 Ciclo

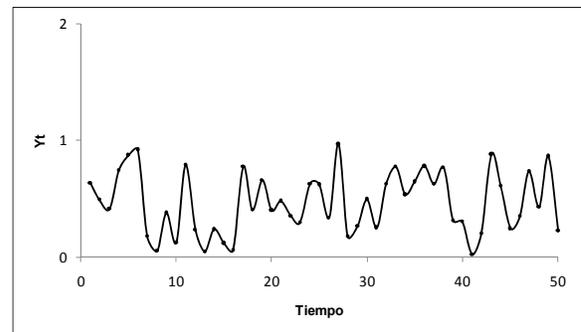


Figura 3.4 Aleatoriedad

3.2 Esquemas de integración

Cada componente detectada en la serie es modelada por separado y una vez estimadas las expresiones analíticas más apropiadas para cada una, el modelo que engloba a cada componente se construye como una integración o composición de los modelos individuales.

Los esquemas de integración tienen dos objetivos principales:

- Ayudar a comprender el comportamiento del fenómeno mostrando de manera simple, es decir, el crecimiento o declinación de la serie
- Producir pronósticos de corto y largo plazo de valores futuros de una serie de tiempo

La integración de las componentes de la serie puede ser mediante (Rodríguez 2000):

1. Esquema aditivo. Los modelos de cada componente se suman entre sí, para formar la función f como $T_t + E_t + C_t + a_t$

2. Esquema multiplicativo. Los modelos de cada componente se multiplican entre sí, para formar la función f como $T_t \times E_t \times C_t \times a_t$

3. Esquema mixto. Surge a partir de la combinación de las dos anteriores. Se forma la función f como $(T_t \times E_t \times C_t) + a_t$

El esquema multiplicativo es el más utilizado en la práctica y significa que el valor de la serie en cualquier momento de tiempo, viene dado por el producto de los valores correspondientes a las cuatro componentes. Este esquema suele aplicarse cuando los valores de la serie muestra oscilaciones cada vez mayores o menores, es decir, la variabilidad de la serie no se mantiene constante.

Por el contrario, cuando la variabilidad de la serie permanece constante, un esquema aditivo da mejores resultados.

El caso mixto se utiliza cuando las oscilaciones estacionales tienden a crecer con el tiempo, y la variabilidad de estas oscilaciones se espera sean constantes.

En la mayoría de las series de tiempo, es difícil manejar los componentes cíclicos de las series de tiempo, por eso dentro de la modelación clásica se considera de poca relevancia esta componente. Además no se cuenta con las herramientas necesarias para la modelación de la componente aleatoria (Aguirre 1994). Por estas razones se estudian con mayor detalle la tendencia y estacionalidad.

3.3 Componentes de una serie de tiempo

3.3.1 Tendencia

La tendencia es la componente de la serie que describe la dirección general hacia la cual se extreman sus valores, como movimiento de dirección global (Aguirre 1994).

En una serie de tiempo no siempre se encuentran las cuatro componentes presente, por esta razón el primer paso para el tratamiento de la tendencia es identificar su presencia. Una vez detectada, el segundo paso en su tratamiento es la modelación mediante el ajuste de una ecuación que relacione los valores de la serie y al tiempo.

La identificación de la tendencia principalmente se hace realizando una inspección visual al gráfico de la serie con el fin de detectar inclinaciones o elevaciones de los valores de la serie que revelarán la presencia de esta componente en la serie. En la mayoría de los casos, la simple inspección visual confirma la existencia de tendencia en la serie.

Si en una serie la identificación no fuera posible mediante la exploración del gráfico o simplemente exista duda de la presencia de tendencia, existen métodos de suavizamiento que ayudan a ver de forma clara la existencia de algún patrón de tendencia en los datos. Dichos métodos son llamados: *suavizamiento por media móvil*, *exponencial de Brown*, *exponencial de Holt* y *estacional de Winter*. Los últimos tres diseñados con fines de pronóstico, pero pueden utilizarse para detectar tendencia (Aguirre 2004). Sólo se estudia el suavizamiento por medias móviles, por ser el de más difusión y cuyos resultados despejan la duda de la presencia de la tendencia en una serie de tiempo.

Suavizamiento por medias móviles: La media móvil se aplica a una serie marcadamente irregular para detectar su tendencia, determinando la forma de la tendencia. La media móvil es un promedio de los valores de la serie de datos desplazándose de un extremo a otro de la misma y sustituyendo esos valores por el resultado promedio (Hanke y Wichern 2006).

La cantidad de valores de la variable a considerar en la media móvil y los coeficientes utilizados para ponderar esos valores, definen la clase de media móvil empleada para el suavizamiento de la serie.

Una media móvil simple de tres puntos se obtiene mediante:

$$Y_t = \frac{Y_{t-1} + Y_t + Y_{t+1}}{3}$$

Cuando la cantidad de puntos en la media móvil es par, se toma la mitad de los valores extremos. Esto se hace para centrar el valor Y_t que será sustituido por la media móvil en cada momento. Así, la media móvil de cuatro puntos es:

$$Y_t = \frac{1/2(Y_{t-2}) + Y_{t-1} + Y_t + Y_{t+1} + 1/2(Y_{t+2})}{4}$$

Como ejemplo de la aplicación de este método, se hace referencia a la serie de datos que se muestra en la *Tabla 3.1* de la cual se obtendrá la media móvil de 4 puntos. Su gráfico correspondiente se muestra en la *Figura 3.5*

t	Y_t	t	Y_t	t	Y_t	t	Y_t
1	500	8	300	15	250	22	500
2	350	9	350	16	550	23	400
3	250	10	200	17	550	24	650
4	400	11	150	18	400	25	850
5	450	12	400	19	350	26	600
6	350	13	550	20	600	27	450
7	200	14	350	21	750	28	700

Tabla 3.1 Serie de tiempo

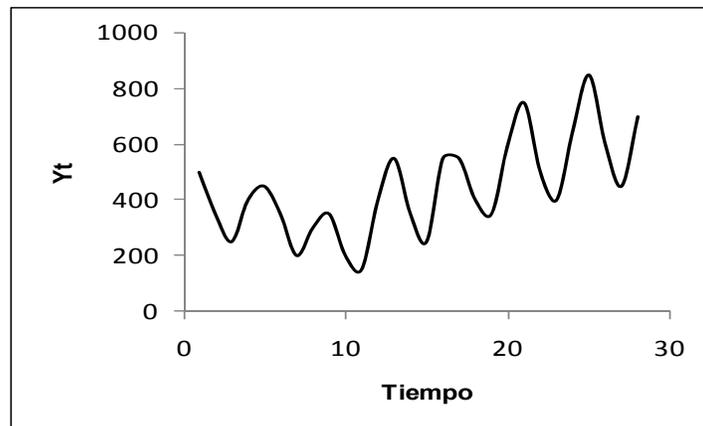


Figura 3.5 Gráfico de la serie

La media móvil de 4 puntos se puede calcular a partir de Y_3 :

$$Y_3 = \frac{1/2(Y_{3-2}) + Y_{3-1} + Y_3 + Y_{3+1} + 1/2(Y_{3+2})}{4}$$

$$Y_3 = \frac{1/2(Y_1) + Y_2 + Y_3 + Y_4 + 1/2(Y_5)}{4}$$

$$Y_3 = \frac{1/2(500) + 350 + 250 + 400 + 1/2(450)}{4}$$

$$Y_3 = 368.75$$

Aplicando el mismo procedimiento, se obtienen las medias móviles restantes. Los resultados se muestran en la *Tabla 3.2* Gráficamente el resultado se presenta en la *Figura 3.6*.

t	Y _t	Y _t	t	Y _t	Y _t	t	Y _t	Y _t	t	Y _t	Y _t
1	500		8	300	281.3	15	250	425.0	22	500	568.8
2	350		9	350	256.3	16	550	431.3	23	400	587.5
3	250	368.8	10	200	262.5	17	550	450.0	24	650	612.5
4	400	362.5	11	150	300.0	18	400	468.8	25	850	631.3
5	450	356.3	12	400	343.8	19	350	500.0	26	600	643.8
6	350	337.5	13	550	375.0	20	600	537.5	27	450	
7	200	312.5	14	350	406.3	21	750	556.3	28	700	

Tabla 3.2 Obtención media móvil de 4 puntos

Como se puede observar en la gráfica, la media o promedio móvil de 4 puntos muestra un incremento gradual en los valores de la serie. Mediante este procedimiento comprobamos la presencia de la componente tendencial en la serie.

Una vez que se ha detectado la existencia de tendencia en la serie y nos hemos hecho una idea de su forma, es necesario obtener un modelo matemático para representar este componente. El método más usual para la modelación de la tendencia consiste en ajustar a la serie a una determinada curva geométrica cuya expresión algebraica sea conocida. Los modelos algebraicos más conocidos y utilizados para modelar la tendencia son los modelos lineales, cuadráticos, exponenciales, potenciales y logarítmicos.

Se ahondará sólo en los modelos lineales y cuadráticos por ser los más sencillos de obtener y que con ellos se obtienen buenos resultados.

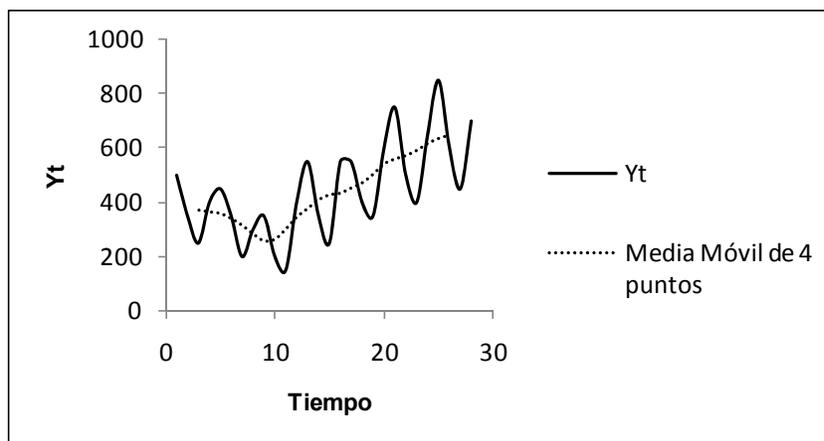


Figura 3.6 Suavizado con media móvil de 4 puntos

Modelo lineal: El modelo lineal de la tendencia representa a una línea recta como se muestra en la *Figura 3.7*. Esta línea puede ser ascendente o descendente. Es muy buena aproximación cuando la serie muestra una curvatura pequeña.

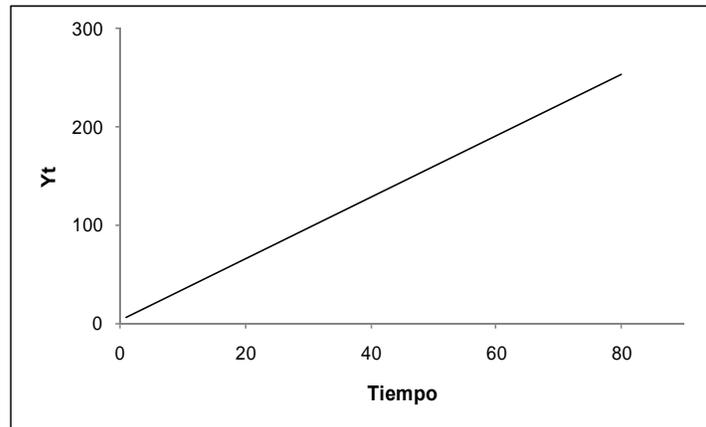


Figura 3.7 Modelo lineal

La forma general de este modelo es:

$$T_t = a + bt$$

Donde

T_t : es el valor predicho para la tendencia en el tiempo t

t : es la variable independiente que representa el tiempo y supone valores enteros (1, 2, 3, etc.)

a : representa la intersección con el eje T_t

b : representa el incremento promedio o decremento en T_t para cada incremento de tiempo.

Por medio del método de mínimos cuadrados puede deducirse la ecuación del modelo lineal que represente a la tendencia. Este método selecciona los valores de los coeficientes en la ecuación de la tendencia (a y b), de forma que los valores de la tendencia estimada T_t son los más cercanos a los valores reales de Y_t .

La proximidad se mide mediante el criterio de la suma del cuadrado de los errores residuales SSE (Bowerman, O'Connell, Koehler 2007):

$$SSE = \sum (Y_t - T_t)^2$$

La línea que mejor se ajusta a un conjunto de datos T_t , t es aquella que minimiza la suma del cuadrado de las distancias desde los puntos hasta la línea, midiéndola en dirección vertical ó T_t . Se conoce como recta de mínimos cuadrados o recta de regresión ajustada.

Así, el método de mínimos cuadrados elige los valores para a y b para reducir el error de la suma de los cuadrados:

$$SSE = \sum (Y_t - T_t)^2 = \sum (Y_t - a - bt)^2$$

Usando un poco de cálculo diferencial, se pueden hacer derivaciones sobre expresiones algebraicas específicas para obtener los valores de mínimos cuadrados.

Específicamente:

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

$$a = \bar{Y}_t - b\bar{t}$$

El ajuste lineal de la tendencia es el método más usado debido a lo práctico que resulta su manejo.

No obstante, existen modelos no lineales para el ajuste de la tendencia que permiten un mejor ajuste, pero que a cambio su uso resulta en ocasiones no muy práctico o difícil de realizar. Para este tipo de modelos es aconsejable el uso de programas de cómputo que realicen el ajuste correspondiente.

Modelo cuadrático de la tendencia: El modelo cuadrático se utiliza cuando la tendencia se asemeja a un movimiento curvilíneo. Este modelo se representa en la *Figura 3.8* El modelo cuadrático general es:

$$T_t = a + bt + ct^2$$

Modelo exponencial de la tendencia: El modelo exponencial caracteriza a la tendencia con curvas similares a la cuadrática, pero de aumento o disminución más fuerte (*Figura 3.9*). La expresión general es:

$$T_t = ab^t$$

Modelo logarítmico de la tendencia: Este modelo se caracteriza por iniciar con curvatura más pronunciada que la exponencial, pero más suave en su crecimiento posterior (*Figura 3.10*). Su forma general es:

$$T_t = a + b \log t$$

Para los modelos no lineales, al igual que para el modelo lineal, la manera de obtener una expresión matemática es mediante el método de mínimos cuadrados.

El juicio y la experiencia podrían desempeñar una función significativa en la selección y uso de una curva de tendencia. Para utilizar una curva de tendencia para pronosticar, se debe tener la capacidad de argumentar que se ha seleccionado la tendencia adecuada y que, con toda probabilidad, el futuro será como el pasado.

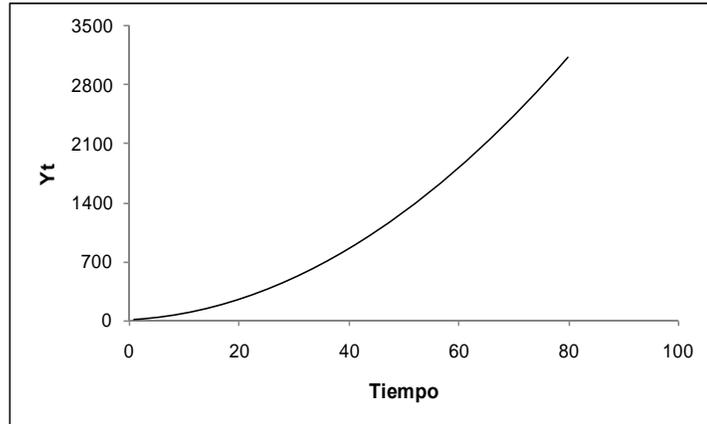


Figura 3.8 Modelo cuadrático

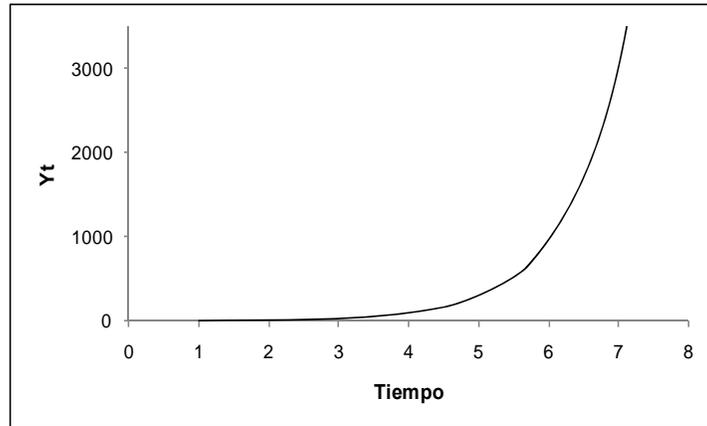


Figura 3.9 Modelo exponencial

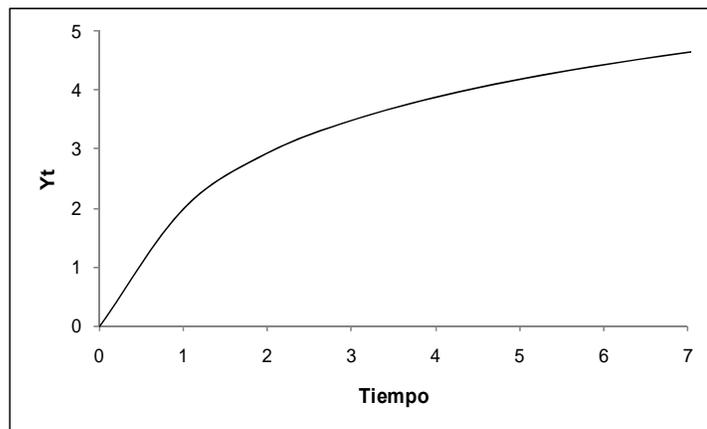


Figura 3.10 Modelo logarítmico

3.3.2 Estacionalidad

Es aquella componente que se caracteriza por una oscilación o fluctuación de tipo local en las observaciones de la serie de tiempo en periodos relativamente cortos de tiempo, manifestada con cierta regularidad. Estos efectos en una serie de tiempo se deben a factores recurrentes cada determinado tiempo (Aguirre 1994).

Al igual que en el tratamiento de la tendencia, el primer paso a seguir es identificar la presencia de esta componente, para luego encontrar un modelo matemático que la represente adecuadamente.

La determinación de existencia de estacionalidad en una serie va ligada a definir el lapso de tiempo en el cual se produce este movimiento regular y reincidente de la variable. A este lapso de tiempo se le conoce con el nombre de *periodo estacional*.

Al igual que en la detección de la tendencia, identificar la estacionalidad en una serie comienza con la inspección visual de su gráfico. Este primer reconocimiento consiste en identificar los grandes movimientos de tipo más general y repetitivos con cierta regularidad.

Por ejemplo, en la *Figura 3.11* se muestran los datos de una serie con marcada estacionalidad. Claramente se aprecian fluctuaciones en los valores de la serie que se repiten constantemente e inclusive aumentan gradualmente al avanzar en el tiempo. Este patrón observado indica la presencia de estacionalidad. Cabe destacar que junto al reconocimiento visual, es importante conocer la naturaleza del proceso estudiado y los factores que de alguna manera pudieran afectar a la variable de manera periódica, como pudiera ser las estaciones del año.

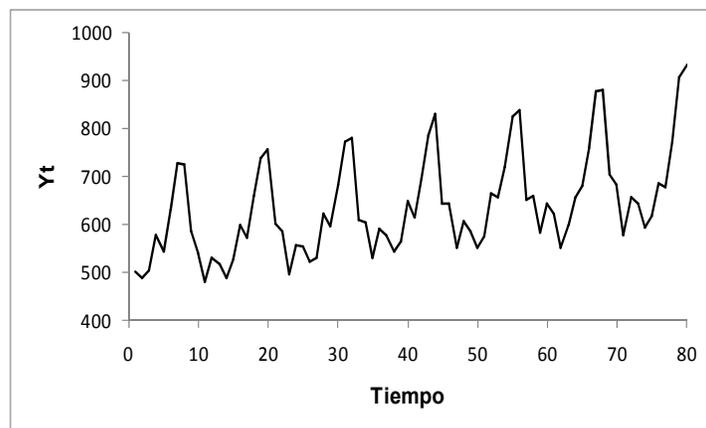


Figura 3.11 Estacionalidad

Ahora bien, aunque en diversas ocasiones es posible distinguir la presencia de estacionalidad en una serie e inclusive determinar su periodo de ocurrencia simplemente observando el gráfico de la serie, en diversas series de tiempo la inspección visual no es prueba definitiva para comprobar la existencia de estacionalidad.

Por ese motivo usualmente se recurre a algunos instrumentos que sirven para comprobar de manera más exacta y confiable la existencia de este patrón. Las más comunes son las funciones de autocorrelación y el periodograma. Las funciones de autocorrelación miden para cada lapso de tiempo, el tipo y grado de correlación existente entre los valores de la serie separados entre sí por un periodo de

tiempo igual a este retardo. De acuerdo con esto, si existe estacionalidad en una serie, los valores separados entre sí por lapsos de tiempo iguales al periodo estacional, estos deben estar correlacionados de alguna forma, pues tienden a repetir un similar movimiento en instantes iguales correspondientes a cada periodo estacional (Hanke y Wichern 2006).

Para mostrar cómo la función de autocorrelación simple y la función de autocorrelación parcial sirven como instrumento en la detección confiable de la estacionalidad y determinación de su periodo, supóngase la serie presentada anteriormente en el gráfico de la *Figura 3.11*.

En primera instancia, una rápida inspección visual advirtió la presencia de estacionalidad. Se observan fluctuaciones que se repiten continuamente en el tiempo en un periodo aproximado de cada 12 o 13 datos.

Ahora bien, se obtienen las funciones de autocorrelación simple y parcial. Al prestar especial atención a la *Figura 3.12* que muestra la FAS de la serie, se observa un valor inicial de 0.68 para el primer retardo y posteriormente decaen los valores subsecuentes a negativo hasta el sexto retardo donde vuelven a aumentar hasta el retardo 12 y así sucesivamente. En este análisis se busca detectar una fuerte autocorrelación en el retardo (Lag) 12, lo cual indicaría que los valores separados por 12 retardos están fuertemente correlacionados. Se distingue el coeficiente del retardo 12 respecto a otros debido a su magnitud (0.78) y a que rebasa la línea punteada. Esta espiga es una fuerte evidencia de la existencia de un patrón que se repite constantemente en el tiempo con periodo igual a 12.

Esta conclusión se puede confirmar al analizar la FAP (*Figura 3.13*) donde se aprecia que en el retardo 12, se presenta un valor alto positivo (también llamada espiga). Lo cual confirma una alta relación entre los valores separados por 12 unidades de tiempo. En esa misma figura en el retardo 13 existe un valor importante que resalta de los demás, pero que a diferencia del valor del retardo 12, el retardo 13 es un valor negativo y este no figura de manera significativa en la FAS.

A modo de resumen, se puede concluir que la detección de la estacionalidad en una serie mediante la FAS y la FAP se realiza observando los valores significativamente altos (espigas) en comparación con otros en la misma gráfica (que en la mayoría de las veces resaltan a primera vista) y que se presentan en ambas funciones de autocorrelación simple y parcial. En un amplio número de series de tiempo la estacionalidad puede ser reconocida con gran precisión utilizando esta herramienta y con frecuencia suelen presentar periodos trimestrales, semestrales y anuales.

El uso de la FAS y la FAP en el reconocimiento de la estacionalidad y la determinación de su periodo para el analista puede presentar un poco de dificultad en un principio, pero al transcurso del tiempo, se desarrolla la habilidad necesaria para detectar patrones de este tipo.

Sin embargo, aún cuando se adquiriera cierta habilidad para detectar patrones estacionales, existen series en las cuales aún con las funciones de autocorrelación se hace difícil determinar con precisión la existencia de esta componente y determinar el periodo con que se repiten. A veces también sólo existe incertidumbre acerca de la presencia de estacionalidad y se requiere una herramienta que confirme los resultados que arrojan la FAS y FAP. Frecuentemente se hace uso del periodograma para realizar dicha función. El periodograma es una herramienta que sirve para detectar la presencia de estacionalidad y determinar su periodo con mayor precisión.

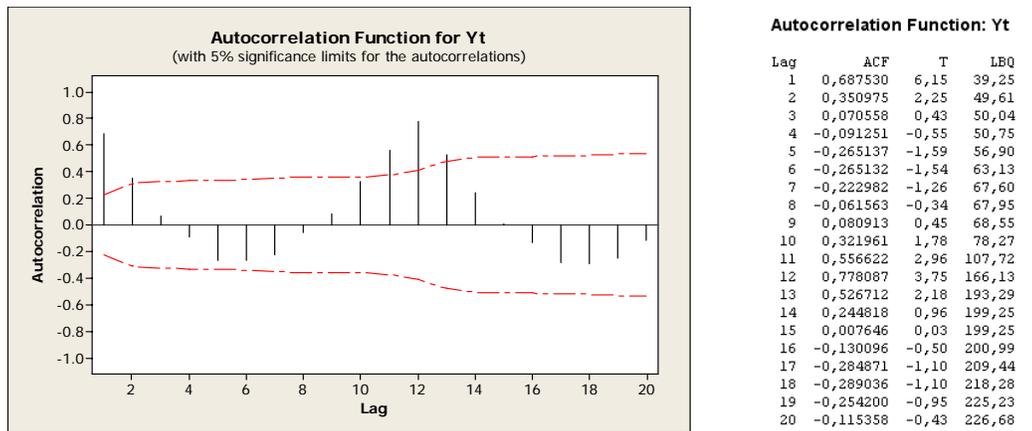


Figura 3.12 Función de autocorrelación simple

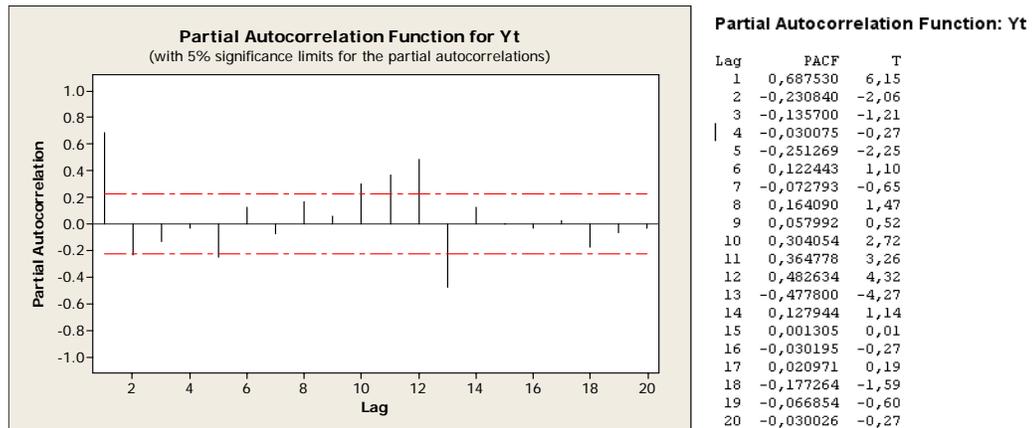


Figura 3.13 Función de autocorrelación parcial

Por medio del programa *Degtra* se obtiene el periodograma perteneciente a la serie de ejemplo (Figura 3.14).

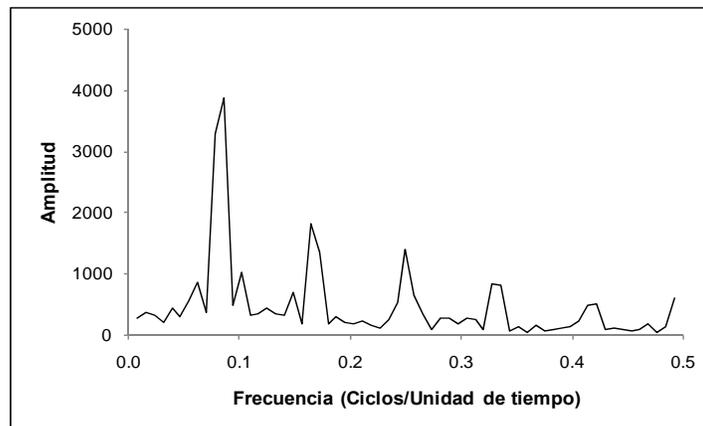


Figura 3.14 Periodograma

De acuerdo al capítulo anterior, la interpretación del periodograma consiste en la localización de amplitudes destacables preferentemente en las primeras frecuencias del gráfico. En este caso se presentan 5 amplitudes que resaltan de las demás. Son de interés las primeras tres para determinar la periodicidad de la serie. Para esta labor normalmente se utiliza la tabla adjunta de datos de salida del programa que elabora el periodograma. Para nuestro periodograma de ejemplo, la *Figura 3.15* muestra la lista de las amplitudes destacables con su respectiva frecuencia.

Frecuencia	Amplitud
0.0833	3864
0.1641	1811
0.2500	1401

Figura 3.15 Amplitudes destacables

Ahora bien, con la relación:

$$p = \frac{1}{f}$$

Se obtienen los respectivos periodos:

$$p_1 = \frac{1}{0.0833} = 12$$

$$p_2 = \frac{1}{0.1641} = 6$$

$$p_3 = \frac{1}{0.2500} = 4$$

Del análisis del periodograma finalmente resalta el periodo igual a 12, el cual confirma que este es el correspondiente al patrón estacional de la serie, tal y como se concluyó en la inspección de la serie y con las funciones de autocorrelación. Los periodos 6 y 4 corresponden a submúltiplos del periodo estacional.

Una vez detectada la presencia de estacionalidad en una serie y definido su periodo, el siguiente paso es la obtención del modelo E_t para esta componente. Se han desarrollado diversos métodos para modelar la variación estacional. La idea básica de todos ellos es en primer lugar estimar y eliminar la tendencia de la serie original, dejando únicamente las componentes estacionales y aleatorias. Después, los valores estacionales se recopilan y se resumen para producir un número para cada intervalo observado del año.

La eliminación de la tendencia en una serie se realiza por medio de los esquemas de integración. Si el esquema elegido es aditivo, entonces la tendencia se remueve mediante la ecuación:

$$Y_t - T_t = E_t + a_t$$

Si el esquema elegido es multiplicativo, la remoción de la tendencia es:

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

En ambos casos la tendencia se elimina para dejar exclusivamente la componente estacional (si es que la hay) y la componente aleatoria. Así se facilita la detección y modelado de la estacionalidad.

A diferencia de la tendencia que se representa por una curva mejor ajustada, la estacionalidad tiene que ser calculada mediante un conjunto de valores para cada uno de los intervalos observado del año (semana, mes o trimestre) y con frecuencia se presenta en la forma de un número índice. Estos son porcentajes que muestran los cambios a través del tiempo, es decir, muestran cuales periodos del año son relativamente bajos y cuáles son relativamente altos. Los índices estacionales en conjunto describen el patrón estacional.

La variación estacional es expresado mediante *variables ficticias* (Bowerman, O'Connell, Koehler 2007):

$$E_t = \beta_{s1} X_{s1,t} + \beta_{s2} X_{s2,t} + \dots + \beta_{s(L-1)} X_{s(L-1),t}$$

Donde:

L: es el periodo estacional

$\beta_{s1}, \beta_{s2}, \beta_{s(L-1)}$: son los números representativos de cada uno de los intervalos observados

$$X_{s1,t} = \begin{cases} 1 & \text{si el periodo t es la estación 1} \\ 0 & \text{si no es así} \end{cases}$$

$$X_{s2,t} = \begin{cases} 1 & \text{si el periodo t es la estación 2} \\ 0 & \text{si no es así} \end{cases}$$

$$X_{s(L-1),t} = \begin{cases} 1 & \text{si el periodo t es la estación L - 1} \\ 0 & \text{si no es así} \end{cases}$$

El objetivo de la variable ficticia es asegurar que se incluye en el modelo, un parámetro estacional apropiado para cada periodo. Usualmente los parámetros estacionales o números índice correspondientes a cada intervalo observado se obtienen por medio de programas de computadora tales como SAS y MINITAB. Este trabajo de tesis se limita a la identificación de la estacionalidad y obtención del periodo estacional, elementos importantes en la modelación ARIMA que se estudiará en el siguiente capítulo. Si se desea conocer más acerca de la modelación estacional, puede recurrirse al libro de Bowerman, O'Connell y Koehler (2007) y al de Uriel Jiménez Ezequiel (1995).

3.3.3 Ciclo

El ciclo es una componente similar a la estacionalidad, pero manifestada en periodos considerablemente más extensos de tiempo. El grado en que pueden determinarse los ciclos a partir de datos históricos, tanto sus longitudes, como sus magnitudes (diferencias entre altos y bajos), están lejos de ser constantes. Esta carencia de un patrón de oleaje constante, hace difícil que los ciclos se distingan de las tendencias que evolucionan con suavidad. En general los ciclos se presentan en periodos que van desde año y medio hasta los diez años.

Igualmente que en la estacionalidad, el ciclo se detecta por medio de las funciones de autocorrelación y el periodograma. Por una parte se busca que en retardos altos existan coeficientes de correlación significativos (espigas) tanto en la FAS como en la FAP. Por otro lado en el periodograma se analiza si existen amplitudes destacables en frecuencias bajas.

Sin embargo, cuando existe ciclo en una serie de tiempo pero no contiene una cantidad suficiente de valores observados, su detección puede llegar a confundirse con la tendencia y ser modelada junto a ella. Además, la modelación se torna también difícil o imposible. En consecuencia, para mantener las cosas relativamente simples, se supone que cualquier ciclo en los datos, es parte de la tendencia.

En series suficientemente largas donde se ha detectado el patrón cíclico, la modelación de esta componente se realiza comúnmente con la serie de Fourier.

3.3.4 Aleatoriedad

Son oscilaciones no sistemáticas que en general afectan a la serie en el momento en que ocurren y normalmente tienen una estructura puramente aleatoria. En la modelación clásica esta componente se considera de poca relevancia ya que la considera como un mero factor errático, de igual importancia a los errores de observación, recopilación y transmisión de la información obtenida, presentes en cualquier sistema de medición. Es decir, considera la aleatoriedad como hechos aislados que no repercutirán en el comportamiento futuro de una variable.

No es posible dar un tratamiento a esta componente debido a que no se dispone de las herramientas necesarias para su modelación, además de que la aleatoriedad no contiene información de cara a la predicción. Es por ello que en la modelación clásica una serie de tiempo se concibe en términos globales como la integración de una parte sistémica que pretendemos modelar y otra aleatoria.

3.4 Ventajas y desventajas modelación clásica

A pesar de que el enfoque clásico puede resolver satisfactoriamente un grupo de problemas bastante amplio relativo al tratamiento de series de tiempo, tanto para finalidades descriptivas como de pronóstico, ésta ha caído en desuso, debido al menor grado de desarrollo de las técnicas de modelación en comparación con otros enfoques disponibles en la actualidad.

La modelación clásica se ve seriamente limitada en series cuyas componentes son difíciles de detectar y por tanto modelar adecuadamente con las herramientas desarrolladas para tales propósitos. Específicamente, en ocasiones se torna complicado detectar el patrón estacional, principalmente cuando su periodo es irregular e inconstante. Caso similar sucede con la componente cíclica, donde en primera lugar se depende de la suficiencia de datos contenida en la serie, seguido de la común irregularidad de su periodicidad. Finalmente, la modelación clásica se ve sumamente limitada en series donde existe una fuerte cantidad de ruido (aleatoriedad).

Otra desventaja es la forma en que se ajustan los datos. En la modelación clásica se ajusta la serie a una forma predeterminada y a un modelo matemático preestablecido (modelo lineal, cuadrático, logarítmico, etc.). Contrario a ello, métodos actuales ajustan un modelo, proveniente de una clase general, a la serie de tiempo, es decir, se crea un modelo específico para la serie de tiempo.

Es por esta razón que las nuevas formas de tratamiento de información (analizada bajo el concepto de series de tiempo), proponen modelos más flexibles y con mejor ajuste a las series de tiempo.

No obstante, en la modelación clásica se establecen las herramientas necesarias para comprender las nuevas metodologías para el manejo de información, que se difunden día a día con mayor auge. Desde la inspección visual, la detección de la tendencia, estacionalidad, ciclo, etc., hasta la determinación de los periodos respectivos, facilitan de gran manera la aplicación de otros enfoques a las series de tiempo.

Particularmente la *modelación ARIMA* que se verá en el capítulo siguiente, se vale de conceptos pertenecientes al enfoque clásico que son fundamentales para el desarrollo de la misma. La experiencia en el reconocimiento de componentes presentes en la serie de tiempo (especialmente en lo que se refiere a la tendencia y la estacionalidad) juega un papel importante en el correcto ajuste y modelación de la serie.

4 Modelación ARIMA

A finales del siglo XX diversos investigadores comenzaron a desarrollar y aplicar nuevas propuestas para la modelación de series de tiempo, las cuales lograron probar su eficacia, ya que son capaces de tratar con cualquier patrón de datos y producir pronósticos precisos en base a la descripción de patrones históricos en los datos (Aguirre 1994). Las nuevas propuestas de análisis de series de tiempo se empezaron a implementar en un principio a problemas de contaminación, en la economía, a enfermedades epidemiológicas y en la actualidad a fenómenos físicos e inclusive fenómenos sociales.

Específicamente los modelos desarrollados en las últimas décadas son los llamados autorregresivos (AR), de medias móviles (MA), integrados (I), así como las posibles combinaciones entre estos (ARMA y ARIMA).

La principal diferencia entre estos modelos y los clásicos es el enfoque estocástico que se le da a las series de tiempo, en vez de tratarlas de forma determinística. Bajo este enfoque se concibe a la serie de tiempo como un conjunto de valores de tipo aleatorio, generados a partir de un proceso totalmente desconocido, es decir, se concibe a la serie como un proceso estocástico.

Así mismo, derivado del desconocimiento del proceso generador de los datos, el objetivo de este enfoque es tratar de identificar el modelo probabilístico que represente las características principales del comportamiento de la serie (Aguirre 1994).

El desarrollo y aplicación de este tipo de modelos de series de tiempo en un principio estuvieron gravemente limitados, fundamentalmente debido a razones de cómputo. Actualmente con la amplia disponibilidad de las computadoras, el uso de los modelos ARIMA se ha tornado posible. En el presente texto se hace imprescindible el uso de paquetes informáticos especializados en la modelación ARIMA. En especial se hacen necesarios debido a la frecuencia con que se manejan grandes volúmenes de información. De otra forma, utilizar métodos tradicionales haría que la tarea de modelar bajo este enfoque fuera casi imposible de realizar.

4.1 Enfoque y procesos estocásticos

En general las series de interés llevan asociados fenómenos aleatorios. Por este motivo el estudio de su comportamiento pasado sólo permite acercarse a la estructura o modelo probabilístico para la predicción del futuro. La teoría de los procesos estocásticos se centra en el estudio y modelización de sistemas que evolucionan a lo largo del tiempo, o del espacio, de acuerdo a unas leyes no determinísticas, esto es, de carácter aleatorio y probabilístico.

Para aclarar esta idea definimos un proceso estocástico como una familia de variables aleatorias asociadas a un conjunto índice de números reales, de tal forma que a cada elemento del conjunto le corresponda una y sólo una variable aleatoria que evoluciona con el tiempo:

$$\text{Variables aleatorias } \{Y_t\}$$

$$\text{Con } t = 1, 2, 3, \dots, n$$

Bajo este enfoque, una serie de tiempo se define como una observación particular sobre el estado dinámico de una variable de un proceso de naturaleza aleatoria. Simultáneamente se presupone que los procesos de series de tiempo siguen un comportamiento aleatorio y debido a la propia dinámica de dicho proceso, tales impulsos pueden generar un conjunto de datos serialmente dependientes.

Así, mediante la identificación de la estructura de dependencia latente existe entre las observaciones, se puede modelar la serie.

Una de las principales ventajas de este enfoque aplicado a las series de tiempo, es la gran flexibilidad que se logra para representar un buen número fenómenos reales mediante una sola clase general de modelos. Otra ventaja es la facilidad y precisión para realizar pronósticos.

4.2 Modelos AR, MA y mixtos

La estructura existente de dependencia entre los datos en un proceso estocástico, puede ajustarse a seis modelos que pueden describir una gran variedad de fenómenos reales. La principal característica de dichos modelos es que no involucran a las variables independientes en su construcción, es decir, a la variable del tiempo t , como se hace en el análisis clásico. En cambio, emplean la información que se encuentra en la serie misma para generar los pronósticos (Makridakis y Wheelwright 2000).

Los modelos son los que se describen a continuación (Arnau 2001):

Proceso de ruido blanco – (a): Es un proceso formado por una secuencia de variables aleatorias mutuamente independientes e idénticamente distribuidas.

Modelo no estacionario de corrido aleatorio - (I): Correspondiente a aquellas situaciones en las que los impulsos aleatorios tienden a sumarse o a integrarse en el tiempo. La integración refleja la presencia de un componente de tendencia.

Modelo autorregresivo de orden p - AR(p): Se trata de un modelo en el que una determinada observación es predecible a partir de la observación inmediatamente anterior (modelo autorregresivo de primer orden) o a partir de las dos observaciones que le preceden (modelo autorregresivo de segundo orden). En este caso, la observación actual se define como la suma ponderada de una cantidad finita p de observaciones precedentes más un impulso aleatorio independiente.

Matemáticamente, un modelo autorregresivo tiene la forma siguiente:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + \dots + \varphi_p Y_{t-p} + a_t$$

En donde Y_t es la variable dependiente y $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ son las variables independientes.

En este caso, las variables independientes son valores de la misma variable (de aquí el nombre de auto), pero de periodos anteriores ($t - 1, t - 2, t - 3, \dots, t - p$). Por último, a_t es el error, o término residual, que representa perturbaciones aleatorias que no pueden ser explicadas por el modelo.

El modelo se llama autorregresivo porque se asemeja a la ecuación de regresión:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_p X_p + e$$

Las variables independientes son simplemente valores rezagados de la variable dependiente con rezagos de tiempo 1, 2, 3, ..., p periodos.

Modelo de Medias Móviles de orden q - MA(q): En este modelo, una determinada observación está condicionada por los “impulsos aleatorios” de las observaciones anteriores. De esta forma la observación actual se define como la suma del impulso actual y de los impulsos aleatorios anteriores con un determinado peso.

El modelo de medias móviles tiene la siguiente forma:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Donde a_t es el residuo o error en el periodo t y $a_{t-1}, a_{t-2}, a_{t-3}, \dots, a_{t-q}$ son los valores anteriores del error. La ecuación es semejante a la anterior, con la excepción de que implica que la variable dependiente Y_t depende de valores previos del término residual más que de la variable misma.

Modelo autorregresivo de medias móviles de orden p,q - ARMA(p,q): Este modelo es la combinación de las estructuras anteriores: modelo autorregresivo y modelo de medias móviles. Así, una observación está determinada tanto por observaciones anteriores así como por “impulsos aleatorios” o también llamados “errores” de observaciones pasadas.

La forma general de un modelo autorregresivo de medias móviles es:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + \dots + \varphi_p Y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Se observa que la ecuación es sencillamente una combinación de las ecuaciones del modelo AR y MA.

En la ecuación anterior p simboliza los valores tomados por la variable durante los p periodos anteriores ($t - 1, t - 2, t - 3, \dots, t - p$). El grado de influencia de cada valor anterior sobre el valor considerado de la variable viene dado por el parámetro correspondiente ($\varphi_1, \varphi_2, \dots, \varphi_p$). Por su parte q representa los errores o residuos de la variable durante los q momentos anteriores ($t - 1, t - 2, t - 3, \dots, t - q$). El grado de influencia de cada uno de ellos viene ponderado por el parámetro correspondiente ($\theta_1, \theta_2, \dots, \theta_q$).

Un modelo ARMA puede ajustarse a cualquier patrón de datos. Sin embargo, los valores de p y q se deben de especificar.

Por ejemplo si $p = 1$ y $q = 0$, el modelo ARMA es:

$$Y_t = \varphi_1 Y_{t-1} + a_t$$

El modelo derivado de esta ecuación es llamado modelo AR de primer orden y se escribe como AR(1) o ARMA(1,0).

Si $p = 2$ y $q = 0$, el modelo correspondiente es un AR(2) o un modelo ARMA(2,0) de la forma:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + a_t$$

La ecuación es un proceso AR de segundo orden.

Cuando $p = 0$ y $q = 1$, se trata de un modelo MA de primer orden y se escribe como MA(1) o ARMA(0,1):

$$Y_t = a_t - \theta_1 a_{t-1}$$

Cuando $p = 0$ y $q = 2$, el modelo es un MA(2) o ARMA(0,2) de la forma:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

Por último p y q pueden cada una ser diferentes de 0, en cuyo caso existe un modelo ARMA mixto. Por ejemplo, cuando $p = 1$ y $q = 1$, el modelo es un modelo ARMA(1,1) y es de la forma:

$$Y_t = \varphi_1 Y_{t-1} + a_t - \theta_1 a_{t-1}$$

Es interesante darse cuenta que esta ecuación es no lineal y por lo tanto suficientemente general para describir una amplia variedad de patrones de datos.

Obviamente p y q pueden adoptar muchos otros valores, aun cuando raramente son mayores que 2.

Modelo autorregresivo integrado de medias móviles de orden p,d,q - ARIMA(p,d,q): Al igual que un modelo ARMA, es la combinación de los modelos autorregresivo y el de medias móviles, con la particularidad de incluir un proceso de restablecimiento (el cual se denomina integración) de

inestabilidad original presente en una serie de tiempo. La idea de inestabilidad es una característica que se presenta más adelante en la sección 4.3.1

La forma general de un modelo ARIMA es semejante al de un modelo ARMA:

$$Y'_t = \varphi_1 Y'_{t-1} + \varphi_2 Y'_{t-2} + \varphi_3 Y'_{t-3} + \dots + \varphi_p Y'_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Donde:

Y'_t : Es la serie inducida a la estabilidad

En ocasiones es apropiado incluir un término constante δ a los modelos AR, MA y ARIMA. En un modelo que contiene sólo parámetros de media móvil, el valor de la constante es la media de los valores de la serie ($\delta = \mu$). Mientras tanto en un modelo autorregresivo $\delta = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$ (Arнау 2001). La inclusión del término depende de la serie en estudio. Concretamente, depende del valor de la media. Si la media de todas las observaciones es cero o muy cercana, no se incluye término constante. Si la media es significativamente distinta de cero, se incluye el término.

Un rápido indicador, es conocer si la serie ha sido diferenciada (este concepto se estudia posteriormente). Si la serie ha sido diferenciada, la constante es eliminada, se puede probar que en una serie diferenciada $\mu = 0$. Por el otro lado si la serie no es diferenciada, se debe verificar el valor de μ (Arнау 2001).

4.3 Metodología Box-Jenkins No estacional

Existen distintas metodologías que ayudan a elegir el modelo que mejor se represente a la serie de tiempo de entre los diversos modelos existentes. La metodología más usada y difundida es la que propusieron los profesores G.E.P Box y G.M Jenkins en la década de los años 1970, en la cual lograron grandes avances en identificar, ajustar y verificar los modelos ARIMA adecuados (Hanke y Wichern 2006). Comúnmente se le conoce como *Metodología Box-Jenkins*. Se especifica en esta metodología únicamente el uso de los modelos ARIMA ya que los modelos autorregresivos (AR) y los de medias móviles (MA) se dicen que son las formas particulares de la clase general de modelos ARMA y este a su vez del modelo ARIMA.

Esta metodología se basa en tratar de determinar cuál es el modelo probabilístico que rige el comportamiento del proceso a lo largo del tiempo. Cabe apuntar en este momento la diferencia entre proceso y modelo. Se puede decir que un *proceso* es lo real, es decir, el fenómeno en sí, del cual se desconoce su mecanismo generador. Por el otro lado un *modelo* es solo la imitación o representación del proceso (Pankratz 1983).

En esta sección se estudiará el caso particular de la aplicación de modelos ARIMA para datos no estacionales, es decir, datos que no presentan fluctuaciones periódicas a lo largo del tiempo. Además en cuanto al tipo de datos que son aptos para someterse al análisis mediante la modelación ARIMA, se puede decir debido a la experiencia, que la mayoría de las series de tiempo se pueden modelar mediante estos métodos. Por ese motivo esta metodología es ampliamente utilizada.

Sólo se destacan en seguida algunas consideraciones importantes acerca de la naturaleza de los datos en la modelación ARIMA (Pankratz 1983):

- Los modelos ARIMA aplican tanto para datos discretos o continuos. Los datos discretos son aquellos que son medidos solamente en números enteros, nunca con cifras decimales. Mientras tanto los datos continuos son medidos en intervalos fraccionarios, es decir, con cifras decimales.
- Aunque la metodología ARIMA trata tanto con datos discretos y continuos, sólo se puede aplicar a datos espaciados equidistantemente en el tiempo, en intervalos discretos de tiempo. Los datos medidos en intervalos discretos de tiempo pueden clasificarse en dos tipos: 1) datos que son producto de la acumulación durante un periodo de tiempo, por ejemplo los ahorros de una persona en un mes. 2) datos que son producto de la medición instantánea periódicamente, por ejemplo la medición de la presión en una tubería en intervalos de una hora.
- Para elaborar un modelo ARIMA se requiere una cierta cantidad de datos mínimo. Los profesores Box y Jenkins sugieren un mínimo de 50 observaciones. Un modelo ARIMA se puede aplicar a una serie con menor tamaño, realizando con especial cautela su interpretación. Para series con patrones estacionales se aconseja una serie con gran número de muestras observadas.
- Los modelos ARIMA son especialmente útiles en el tratamiento de series que presentan patrones estacionales.
- Los métodos Box-Jenkins aplican a series estacionarias y no estacionarias. Una serie estacionaria es aquella cuya media, varianza y función de autocorrelación permanecen constantes en el tiempo.
- Se asume que las perturbaciones aleatorias (a_t) presentes en la serie, son independientes entre sí. No existe correlación entre ellas, por lo tanto ningún patrón modelable.

Una vez hechas estas anotaciones, se puede entrar de lleno a la metodología. La *Figura 4.1* muestra las etapas esenciales en la aplicación de esta metodología (Aguirre 1994). De la figura se destacan en especial 5 etapas principales:

1. Estacionariedad
2. Identificación
3. Estimación
4. Evaluación
5. Pronóstico

Cada una de estas etapas se desarrolla en las siguientes secciones.

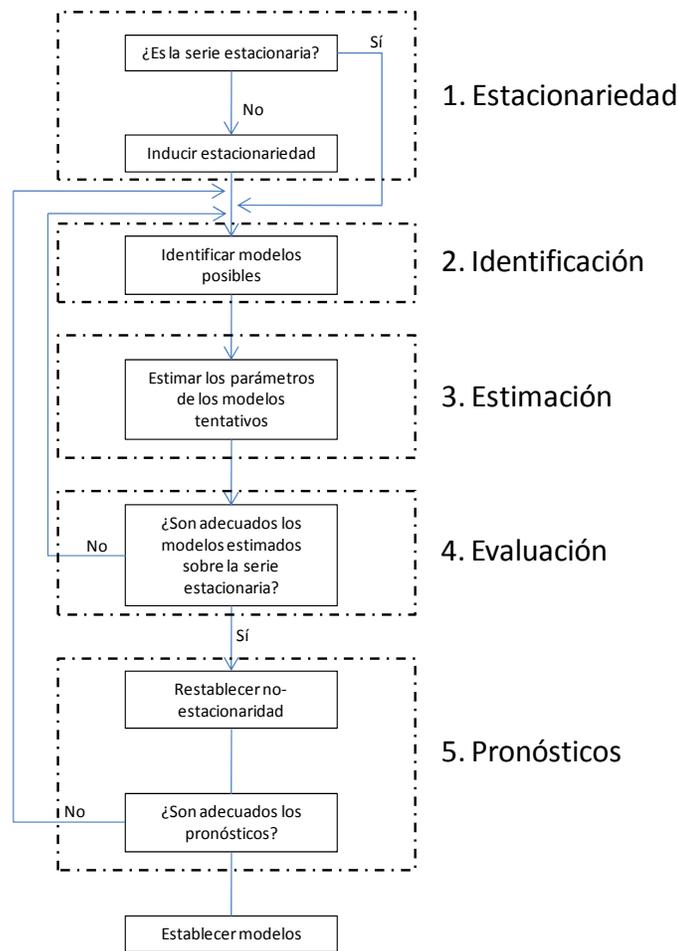


Figura 4.1 Metodología Box – Jenkins

4.3.1 Estacionariedad

Los procesos estocásticos se clasifican en estacionarios y no estacionarios. La idea de estacionariedad está relacionada con la estabilidad de la serie. Un proceso estacionario se describe como una secuencia de datos o valores que no presentan cambio sistemático en la media, ni cambio sistemático alguno en la varianza, se dice entonces que la serie es estable. De forma parecida, un proceso no estacionario es aquel que presenta cambios sistemáticos en la media y/o varianza; la serie de datos es inestable en el tiempo.

Aplicado dicho concepto a una secuencia cronológica y explicado en forma intuitiva, una serie de tiempo es estacionaria si sus propiedades estadísticas (media, varianza) son esencialmente constantes a través del tiempo. Una serie cuya media y/o varianza cambian a través del tiempo es una serie no estacionaria (Bowerman, O’Connell, Koehler 2007).

A partir de la definición previa, se puede advertir que una serie cuyos valores fluctúan respecto de una media constante, es decir, sin tendencia, es una serie estacionaria.

La *Figura 4.2* muestra la gráfica de una serie de tiempo, en ella se observa que los valores de una serie de tiempo fluctúan respecto de una media constante, entonces es razonable pensar que la serie temporal es estacionaria. Si los valores no fluctúan respecto de una media constante o fluctúan con variación constante, entonces es razonable pensar que la serie de tiempo es no estacionaria, como se observa en la *Figura 4.3*.

Este concepto es importante ya que el primer paso en la metodología Box-Jenkins es determinar si la serie de tiempo es estacionaria. En caso de que la serie de datos no sea estable a lo largo del tiempo, es necesario aplicar una transformación a la serie para inducirlo a ello. Este paso es indispensable para poder aplicar las demás etapas de esta metodología.

Existen dos formas de conocer si una serie es estacionaria: por medio del gráfico de la serie y mediante la exploración de función de autocorrelación simple.

1. Por medio de inspección gráfica

Visualmente se puede advertir si la serie es estacionaria si se detectan elevaciones o inclinaciones en los datos. Cualquier patrón de este tipo expresa que la serie es inestable. Por ejemplo la *Figura 4.2* previamente analizada, muestra una serie de datos estable, se observa que los valores giran alrededor de una media constante, sin incrementos o decrementos. Por otro lado la *Figura 4.3* muestra una serie con una fuerte tendencia ascendente, indicador inmediato de inestabilidad.

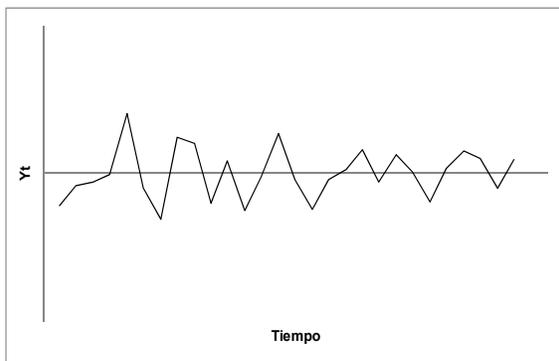


Figura 4.2 Serie de tiempo estacionaria

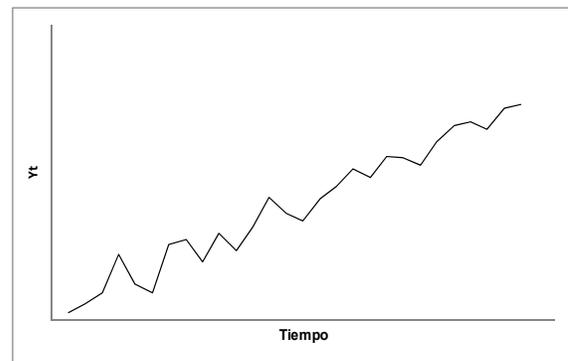


Figura 4.3 Serie de tiempo no estacionaria

2. Por medio de las funciones de autocorrelación

Cuando visualmente no es factible determinar si la serie es estacionaria, usualmente se recurre a la función de autocorrelación simple (FAS), en especial en aquellas series con tendencia poco remarcada. El reconocimiento de la estabilidad se logra a partir de la variedad de comportamientos que esta función que puede mostrar (Bowerman, O'Connell, Koehler 2007).

En primer lugar, la función de autocorrelación simple se puede cortar ó truncar. Para explicar este término, se dice que existe una espiga en el desfaseamiento k en la FAS si r_k es estadísticamente grande.

Se considera que existe una espiga en la función si el valor absoluto de:

$$t_{r_k} = \frac{r_k}{s_{r_k}}$$

es mayor que 2. La función de autocorrelación se trunca después del desfaseamiento k si no hay espigas en los desfaseamientos mayores que k en la función. Por ejemplo, si

$$|t_{r_1}| = 3.671$$

$$|t_{r_2}| = 2.873$$

$$|t_{r_3}| = 0.517$$

Entonces la función de autocorrelación simple se trunca o se corta después del desfaseamiento 2. Esto se ilustra en la siguiente figura (Figura 4.4).

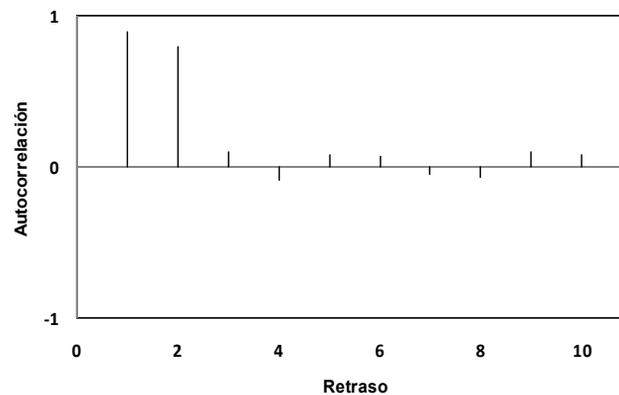


Figura 4.4 La función de autocorrelación se trunca

La salida de MINITAB muestra automáticamente el valor de t asociado a cada coeficiente y además muestra en la gráfica (Figura 4.5) un par de bandas que representan el valor de dos desviaciones estándar $2(s_{r_k})$. Si un coeficiente sobrepasa estas bandas ya sea en la parte de arriba como en la de abajo, significa que el valor de $|t_{r_k}|$ es mayor que 2.

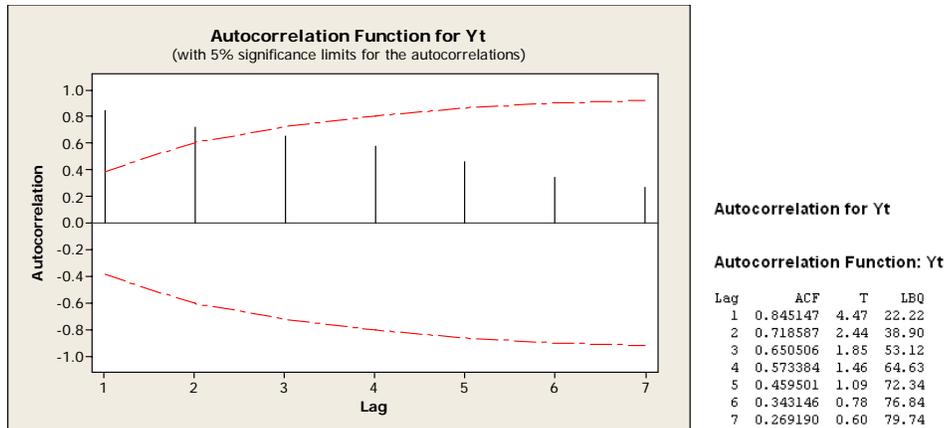


Figura 4.5 Salida MINITAB función de autocorrelación simple

Segundo, la función de autocorrelación simple decae ó se extingue si esta no se corta o muere sino que decrece en una forma permanente. Puede extinguirse de tres maneras:

- En forma exponencial (Figura 4.6)
- En seno amortiguada (Figura 4.7)
- En combinación de 1 y 2 (Figura 4.8)

Además puede cortarse con gran rapidez como se muestra en la Figura 4.9 o puede extinguirse con lentitud extrema como se muestra en la Figura 4.10.

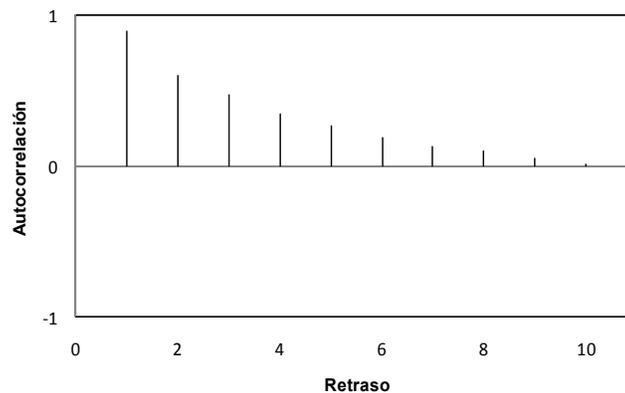


Figura 4.6 Se extingue en forma exponencial

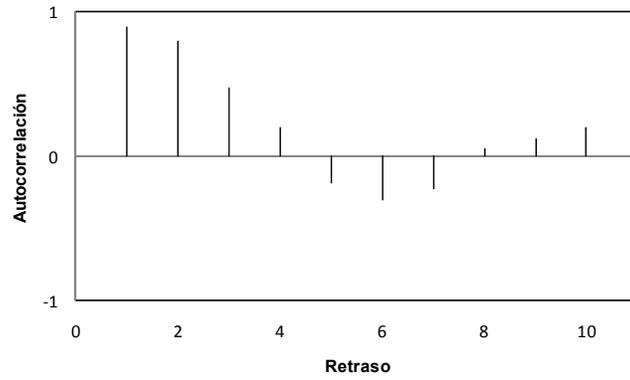


Figura 4.7 Se extingue en forma de seno amortiguada

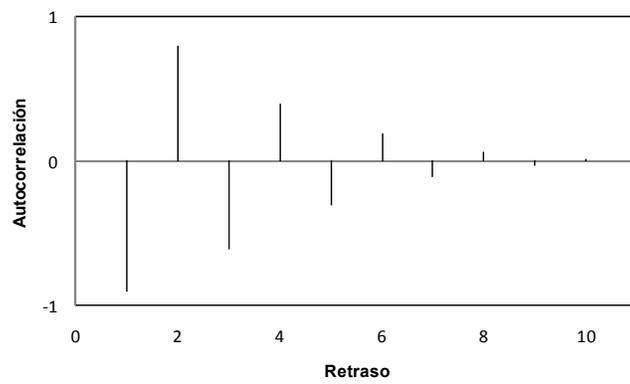


Figura 4.8 Se extingue en forma en forma combinada

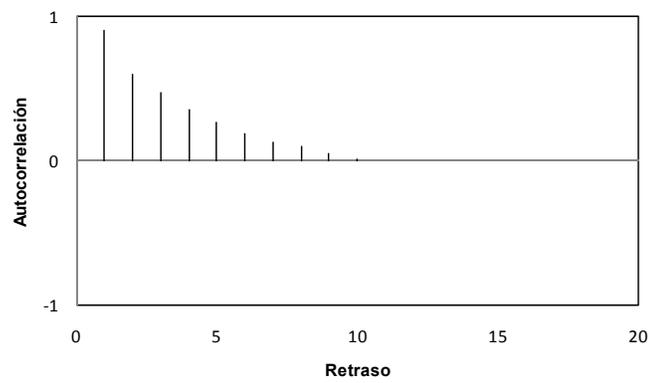


Figura 4.9 Se extingue con gran rapidez

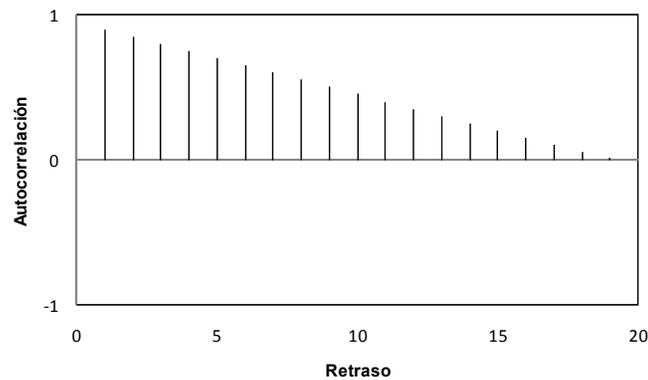


Figura 4.10 Se extingue con lentitud extrema

En general, se puede demostrar que:

Si la función de autocorrelación simple de los valores Y_1, Y_2, etc de la serie de tiempo se corta claramente con rapidez o si se corta rápidamente, entonces se deben considerar que los valores de la serie temporal son estacionarios.

Si la función de autocorrelación simple de los valores Y_1, Y_2, etc de la serie de tiempo se corta con lentitud extrema, entonces se deben considerar que los valores de la serie de tiempo son no estacionarios

Cabe aclarar que los términos “claramente con rapidez” y con “lentitud extrema” es algo arbitrario, y se puede determinar mejor a través de la experiencia. Además, la experiencia indica que en lo que se refiere a datos no estacionales, si la función de autocorrelación simple se trunca claramente con rapidez, casi siempre sucede así, después de un desfase k que es menor o igual a 2 (Bowerman, O’Connell, Koehler 2007).

De esta simple forma se determina por medio de la FAS si la serie es estacionaria.

Ahora bien, la realidad es que las series de tiempo normalmente no se comportan establemente, es decir, no son procesos estocásticos estacionarios. Cuando se presenta esa situación entonces es preciso aplicar una transformación a los valores de la serie para tener una serie temporal estacionaria.

Existen distintas maneras de inducir la estacionariedad en una serie de tiempo, la más usada es el método llamado de construcción de diferencias. Este método consiste, como su nombre lo dice, en obtener diferencias entre los mismos valores de la serie, para remover cualquier patrón de tendencia.

Las primeras diferencias de los valores de la serie de tiempo Y_1, Y_2, \dots, Y_n son:

$$Y'_t = Y_t - Y_{t-1}$$

$$t = 2, 3, \dots, n$$

La *Tabla 4.1* muestra cómo se hace la diferenciación y cómo los datos con una tendencia lineal se convierten en estacionarios (horizontales) después de diferenciar. Gráficamente se puede observar esta transformación en la *Figura 4.11*

Serie de datos	Primeras diferencias	Nueva serie
2	4-2 = 2	2
4	6-4 = 2	2
6	8-6 = 2	2
8	10-8 = 2	2
10	12-10 = 2	2
12	14-12 = 2	2
14	-	-

Tabla 4.1 Obtención de las primeras diferencias en una serie

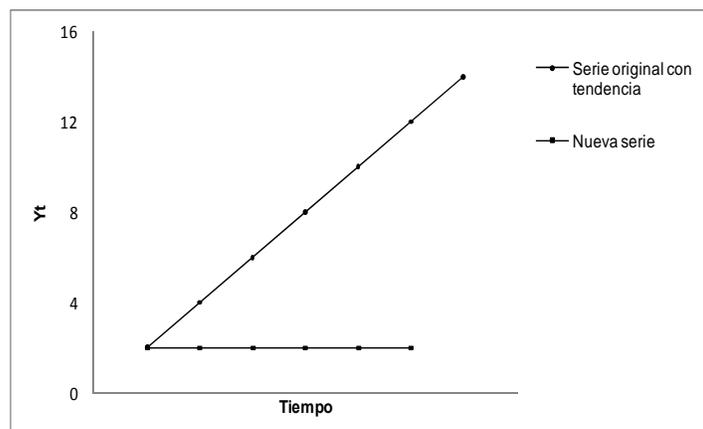


Figura 4.11 Transformación de la serie de tiempo mediante las primeras diferencias

En ocasiones es necesario realizar más de una diferenciación para generar valores de series de tiempo estacionarias. Si los valores originales no son estacionarios y las primeras diferencias de los valores originales de la serie de tiempo tampoco son estacionarios, entonces podemos obtenerlas segundas diferencias de los valores originales de la serie de tiempo.

Las segundas diferencias de los valores de la serie de tiempo son:

$$Y'_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$$

$$Y'_t = Y_t - 2 Y_{t-1} + Y_{t-2}$$

$$t = 3, 4, \dots, n$$

La *Tabla 4.2* muestra cómo se obtienen las segundas diferencias a una serie de tiempo, se puede apreciar que las primeras diferencias siguen siendo no estacionarias, al aplicar una nueva diferenciación se obtiene una serie estacionaria. Cabe aclarar que la notación Y'_t tan sólo se usa para indicar que la serie se transformó, ya sea mediante las primeras o las segundas diferencias.

En la gráfica de la *Figura 4.12* se observa ambas transformaciones a la serie original.

Serie de datos	Primeras diferencias	Nueva serie	Segundas diferencias	Nueva serie
5.6	$15.5 - 5.6 = 9.9$	9.9	$13.5 - 9.9 = 3.6$	3.6
15.5	$29 - 15.5 = 13.5$	13.5	$17.1 - 13.5 = 3.6$	3.6
29.0	$46 - 29 = 17$	17.1	$20.7 - 17.1 = 3.6$	3.6
46.0	$66.7 - 46 = 20.7$	20.7	$24.3 - 20.7 = 3.6$	3.6
66.7	$91 - 66.7 = 24.3$	24.3	$27.9 - 24.3 = 3.6$	3.6
91.0	$118.9 - 91 = 27.9$	27.9	-	-
118.9	-	-	-	-

Tabla 4.2 Obtención de las segundas diferencias de una serie

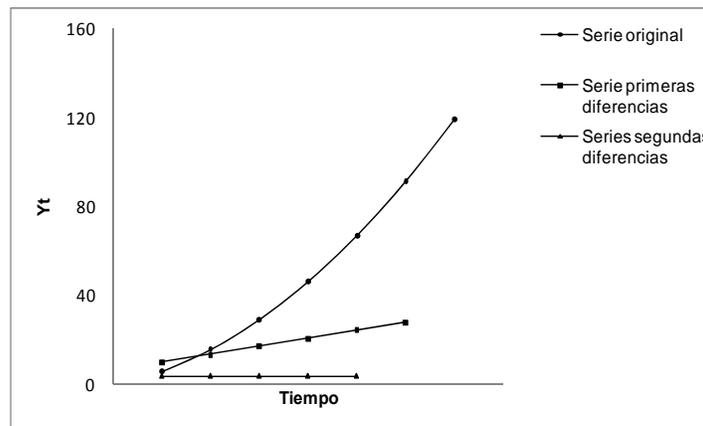


Figura 4.12 Transformación de la serie mediante primeras y segundas diferencias

La experiencia indica que si los valores de la serie de tiempo original son no estacionarios y no estacionales, entonces si se aplica la primera transformación por diferencias o la segunda transformación por diferencias por lo general se producen valores estacionarios de la serie de tiempo.

Una vez introducido el tema de estabilidad en una serie de tiempo, es turno de hacer un pequeño paréntesis en el curso de la metodología Box-Jenkins, a fin esclarecer la diferencia entre un modelo ARMA y un modelo ARIMA. La diferencia se halla en que un modelo ARMA(p,q) es capaz de operar únicamente sobre series estacionarias, mientras que un modelo ARIMA(p,d,q) es capaz de operar tanto sobre series de tiempo no estacionarias, como en series estacionarias.

El modelo ARIMA(p,d,q) es capaz de operar sobre cualquier ya que se incluye el proceso de estabilización de la serie, por medio del parámetro d , que simboliza el grado de diferenciación aplicado a la serie de tiempo. Por ejemplo si $d = 1$, significa que se esta trabajando con las primeras diferencias de la serie; si $d = 2$, se trabaja con las segundas diferencias y así sucesivamente.

4.3.2 Identificación

Una vez que se asegura que la serie es estable en el tiempo. El siguiente paso en la metodología Box-Jenkins es la identificación del modelo probable que rige el proceso de la serie de tiempo.

Las ideas básicas de esta fase son las siguientes:

- La serie de tiempo que se encuentre en proceso de estudio cuenta con sus respectivas funciones de autocorrelación simple y parcial (FAS y FAP), que se denominan prácticas o calculadas.
- Por otra parte, cada una de las distintas configuraciones ARMA posee su FAS y FAP teóricas asociadas al modelo.
- Si la FAS y FAP calculadas de la serie a la que deseamos ajustar un modelo se asemeja a alguna o varias FAS y FAP teóricas, entonces podemos decir que el modelo ARMA teórico es un modelo tentativo para la serie.

Queda claro entonces que la identificación del modelo probable se realiza por medio de la comparación de las funciones de autocorrelación calculadas contra las teóricas (tanto simples como parciales).

Las funciones de autocorrelación calculadas sólo se usan como guías para escoger uno o varios modelos ARMA que parezcan apropiados, ya que tan sólo brindan una aproximación a la estructura más adecuada que debe considerarse.

Las funciones de autocorrelación teóricas por su parte se derivan de una familia de modelos ARMA propuestos por Box y Jenkins. Cabe aclarar que aunque existen una infinidad de posibles procesos dentro de estos modelos ARMA, la mayoría de los procesos que ocurren normalmente en la práctica caben dentro de un número reducido de modelos. Es por esto que los modelos más comunes en la práctica, suelen ser precisamente los propuestos y desarrollados por estos profesores (Pankratz 1983)

De esta manera entonces los profesores Box y Jenkins sugirieron una completa familia de modelos teóricos de los cuales se puede escoger: AR(1), AR(2), MA(1), MA(2), ARMA(1,1).

Las principales características que se observan de las funciones de autocorrelación teóricas son:

Modelo	FAS	FAP
AR	Decae a cero	Se trunca o se corta (después del retraso p)
MA	Se trunca o se corta (después del retraso q)	Decae a cero
ARMA	Decae a cero	Decae a cero

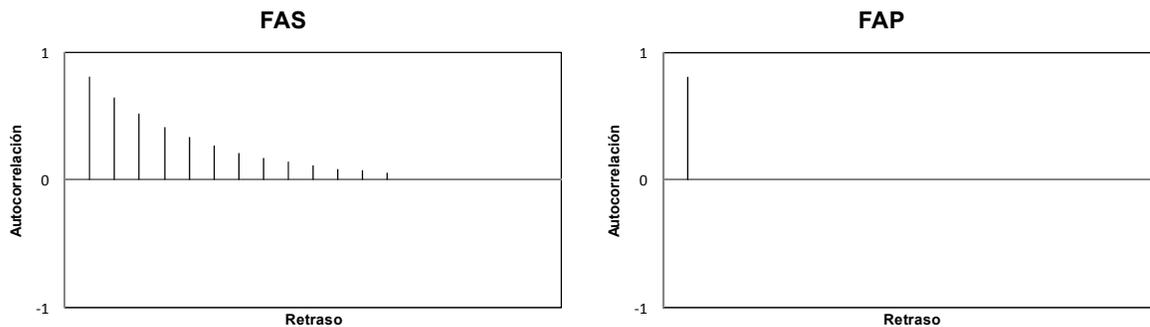
Tabla 4.3 Características FAS y FAP Teóricas (Pankratz)

A partir de la tabla anterior se describe en seguida a detalle las características de cada uno de los posibles modelos ARMA (Pankratz 1983).

Modelo AR: Las funciones de autocorrelación simples de los procesos autorregresivos decaen a cero. Es común que se usen también las expresiones: “se extingue” o “se muere”. La forma en que se extingue la FAS puede ser de forma exponencial, en forma de seno amortiguado, etc. Ahora bien, las funciones de autocorrelación parcial contienen espigas significativas hasta el retraso p , después del cual la función “se corta”. El número de espigas indica por tanto, el orden del proceso autorregresivo. Usualmente p no sobrepasa el valor de 2 o 3.

La *Figura 4.13* (Aguirre 1994) muestra dos configuraciones posibles para un modelo AR(1) o bien ARMA(1,0). De esta figura se puede destacar que la FAS de un proceso autorregresivo de primer orden decae exponencialmente hacia cero, mientras que existe una espiga significativa en la FAP en el primer retraso. Además si $\varphi_1 > 0$, La FAS decae positivamente hacia cero y la espiga en la FAP es positiva también. Si $\varphi_1 < 0$, la FAS decae hacia cero alternando signos, mientras que la espiga significativa en la FAP es negativa.

a) $\varphi_1 > 0$



b) $\varphi_1 < 0$

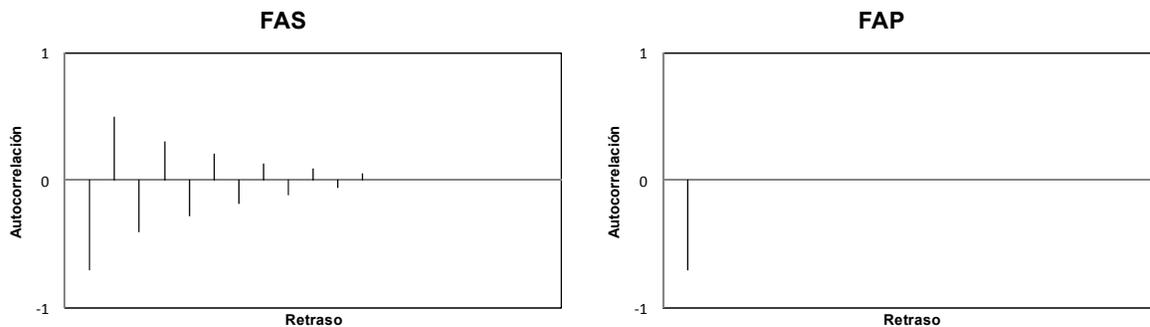
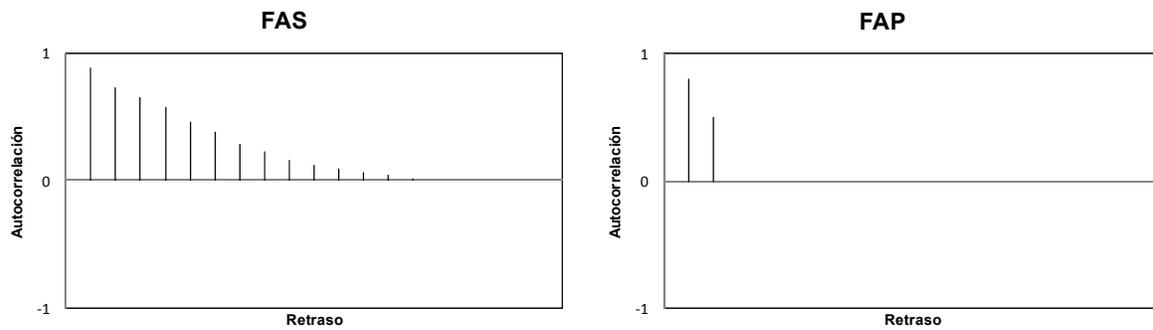


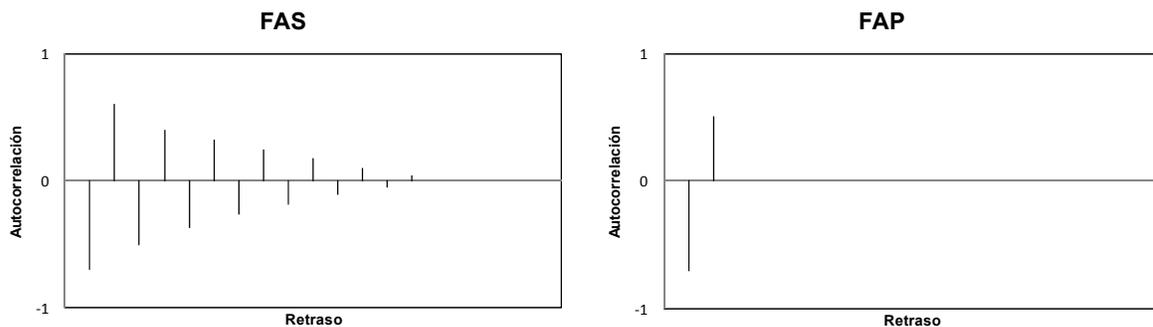
Figura 4.13 Funciones de autocorrelación simple y parcial teóricas para el modelo AR(1)

Para un modelo AR(2) o ARMA (2,0) existen una mayor variedad de patrones posibles que para un modelo AR(1). En general un proceso autorregresivo de segundo orden cuenta con una FAS que decae a cero tanto de forma exponencial como de seno amortiguado ó bien una mezcla de ambos. La FAP se caracteriza por tener espigas en los retrasos 1 y 2. El patrón exacto depende en las posibles combinación de signos para ϕ_1 y ϕ_2 . En la *Figura 4.14a* y *4.14b* se muestran cuatro posibles combinaciones.

a)



b)



c)

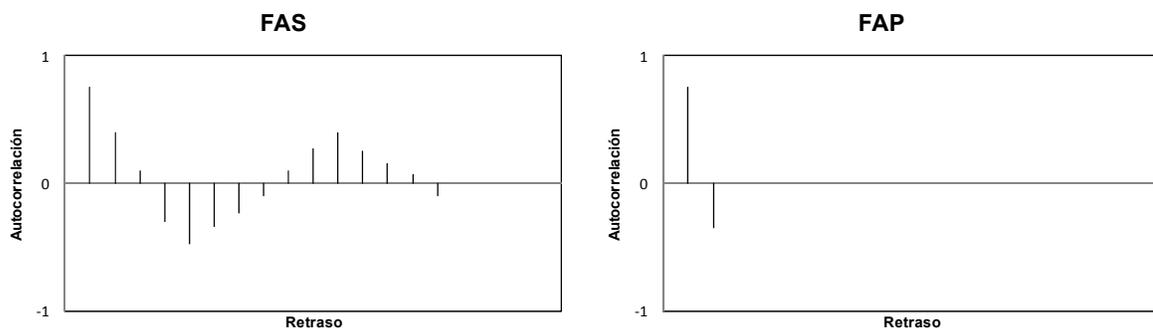


Figura 4.14a Funciones de autocorrelación simple y parcial teóricas para el modelo AR(2)

d)

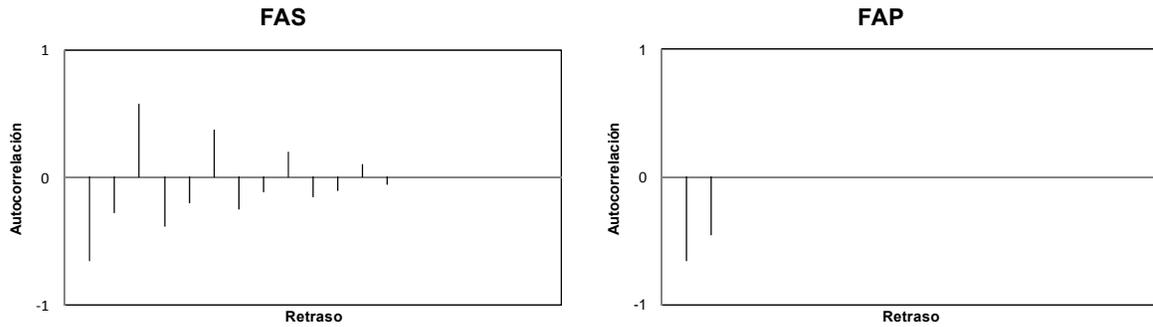
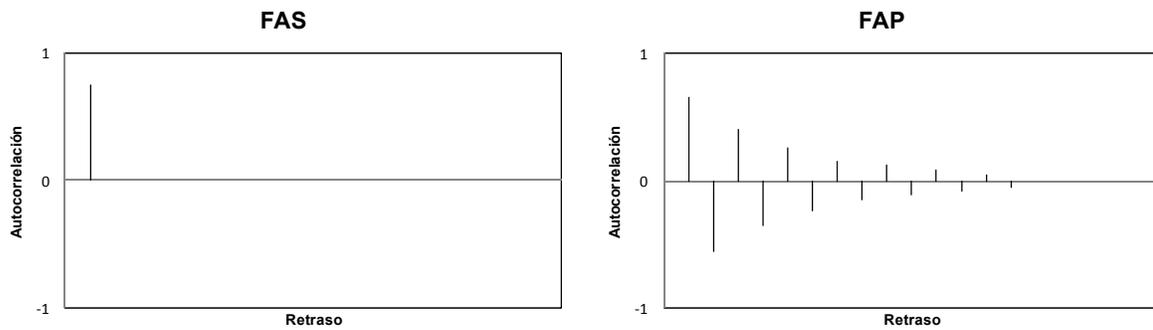


Figura 4.14b Funciones de autocorrelación simple y parcial teóricas para el modelo AR(2)

Modelo MA: En general la FAS de un proceso de medias móviles cuenta con espigas significativas hasta el retraso q , después del cual se corta la función. La FAP por su lado decae de forma exponencial o en forma de seno amortiguado. Normalmente el valor de q no es mayor que 2 o 3.

La Figura 4.15 muestra las configuraciones de un modelo MA(1). Se ilustra como si $\theta < 0$ en la FAS se presenta una espiga positiva mientras que la FAP decae exponencialmente a cero alternando signos. Si $\theta > 0$ la espiga en la FAS es negativa, la FAP decae exponencialmente a cero con signo negativo.

a) $\theta < 0$



b) $\theta > 0$

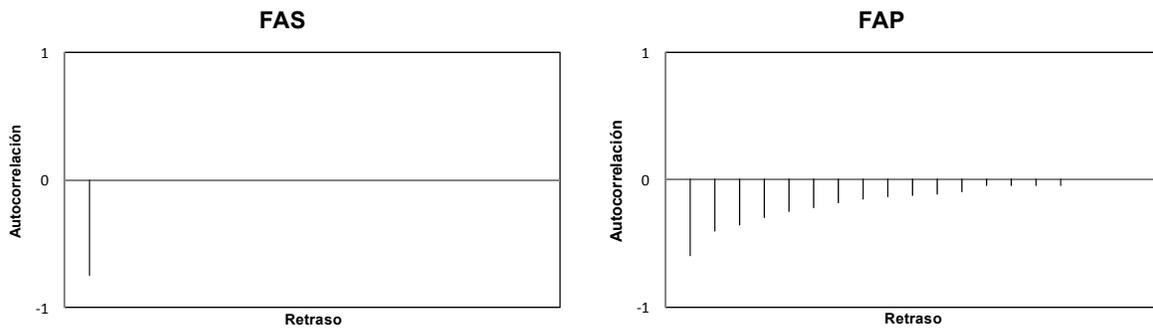
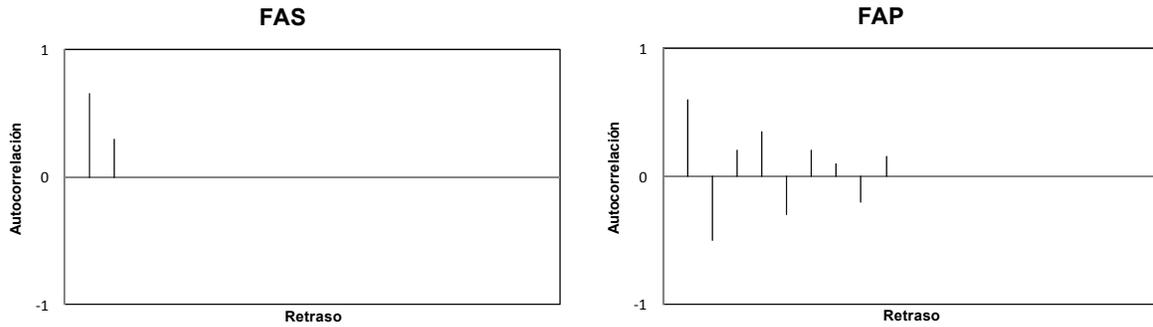


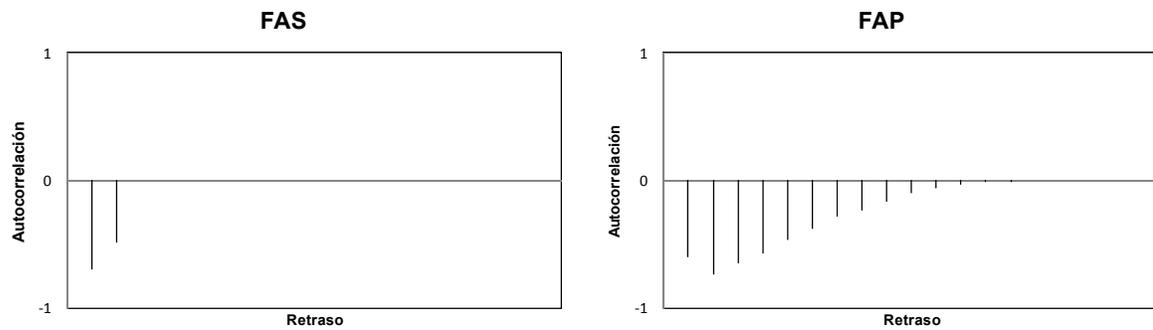
Figura 4.15 Funciones de autocorrelación simples y parcial teóricas para el modelo MA(1)

Para un modelo MA(2) la *Figura 4.16a* y *4.16b* muestra algunas formas que pueden tomar las funciones de autocorrelación. Se muestra que existen espigas en la FAS hasta el retraso q , en este caso hasta $q = 2$ y después de este retraso la función de corta o se extingue. La FAP decae a cero con distintas configuraciones.

a)



b)



c)

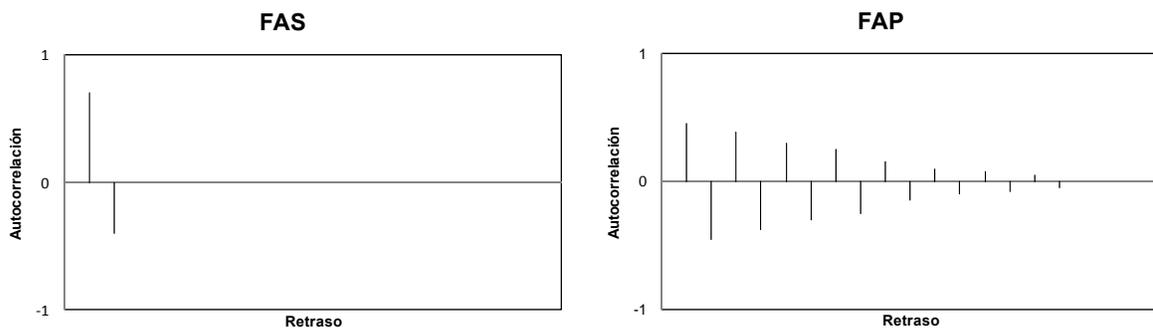


Figura 4.16a Funciones de autocorrelación simples y parcial teóricas para el modelo MA(2)

d)

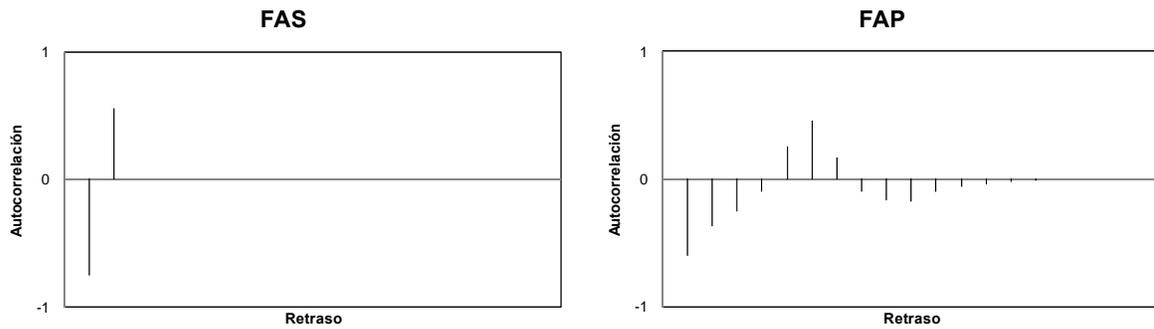
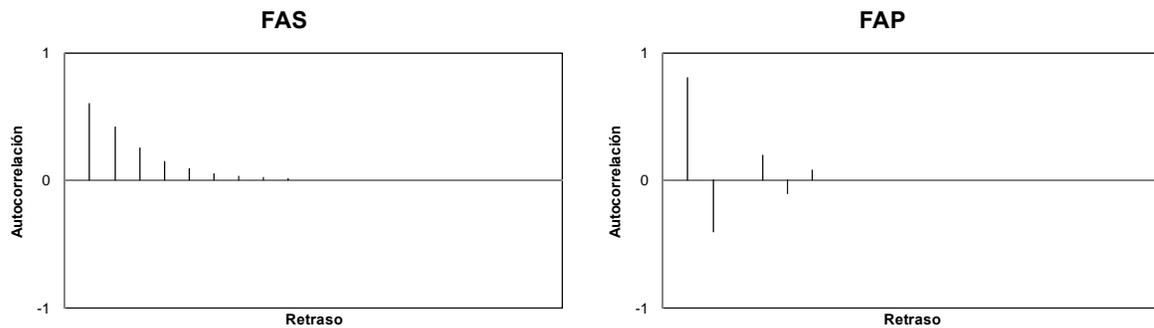


Figura 4.16b Funciones de autocorrelación simple y parcial teóricas para el modelo MA(2)

Modelos ARMA: Las funciones de autocorrelación combinan las características que se presentan en los modelos AR y los MA. En general la FAS decae a cero tanto exponencialmente como en forma de seno amortiguado después del retraso de tiempo $p - q$. La FAP teórica también decae a cero después del retraso $p - q$. Se muestran en la Figura 4.17a y 4.17b seis posibles combinaciones para el modelo teórico ARMA(1,1)

a)



b)

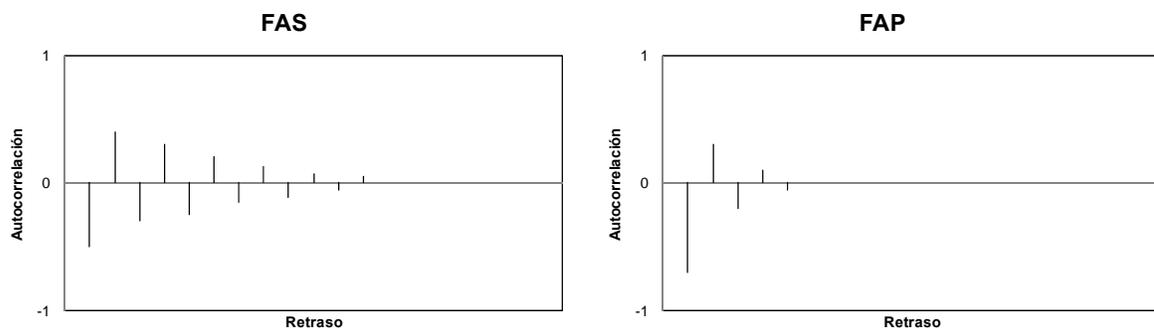


Figura 4.17a Funciones de autocorrelación simple y parcial teóricas para el modelo ARMA(1,1)

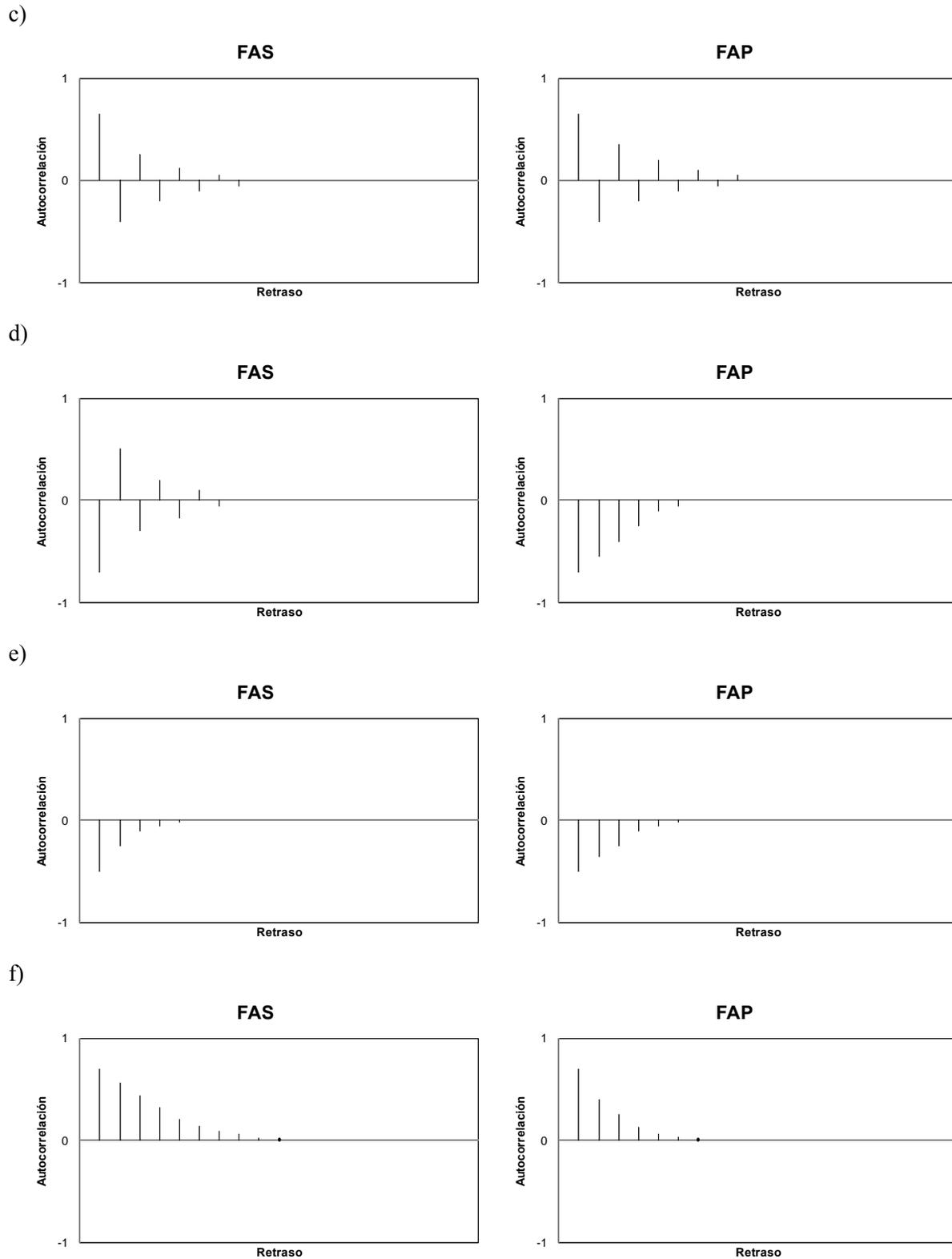


Figura 4.17b Funciones de autocorrelación simple y parcial teóricas para el modelo ARMA(1,1)

Es probable que exista cierta ambigüedad al determinar un modelo ARIMA apropiado a partir de los patrones que provienen de las FAS y FAP parciales de la muestra. Cualquier modelo que se escoja, se considera simplemente tentativo: sólo es un candidato para ser considerado como el modelo ideal. Si no fuera el caso, se deberá intentar con un modelo alterno.

4.3.3 Estimación

En esta etapa se estiman los coeficientes del modelo escogido tentativamente en el paso anterior, existen distintos criterios para dicho propósito. Usualmente se realiza por medio de paquetes informáticos. Además en esta etapa se puede pre-visualizar la adecuación del modelo a la serie de tiempo. Particularmente si los coeficientes no satisfacen ciertas condiciones matemáticas, el modelo es rechazado.

El principal criterio para el cálculo de los parámetros es el llamado *Estimación de máxima verosimilitud* ó *Maximum Likelihood* (ML por sus siglas en ingles). Se ha demostrado que los parámetros estimados a través de este criterio reflejan con gran exactitud las características presentes en los datos de la serie de tiempo. Sin embargo éste método se ve limitado debido al gran tiempo que conlleva realizar la estimación, aún con el uso de la computadora (Pankratz 1983).

Por esta razón la mayoría de la literatura existente sugiere el uso del criterio de *mínimos cuadrados* o *least squares* (LS por sus siglas en ingles). Se ha demostrado que las estimaciones hechas mediante este criterio son muy cercanas a las del criterio de máxima verosimilitud. La estimación de los parámetros del modelo ARMA seleccionado, se realiza por medio de minimizar la suma de los cuadrados de los residuales SSR (sum of squared residuals). En general, estos estimados de los mínimos cuadrados deben obtenerse mediante un procedimiento no lineal de mínimos cuadrados. Un procedimiento no lineal de mínimos cuadrados es, sencillamente, un algoritmo que encuentra el mínimo de la suma de la función de errores cuadrados, mediante un proceso iterativo.

Los residuales de cualquier modelo ARMA se definen como:

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

Donde:

\hat{a}_t : *residual de la serie*

Y_t : *serie original*

\hat{Y}_t : *serie calculada con los parámetros estimados*

Entonces, la suma de los cuadrados de los residuales se define como:

$$SSR = \sum \hat{a}_t^2$$

Esta última función es la que se debe minimizar, a través de la búsqueda de los parámetros θ y φ del modelo ARMA que se desea ajustar.

Finalmente obtenidos los parámetros θ y φ del modelo ARMA propuesto, deben de cubrir en conjunto los siguientes requerimientos (Pankratz 1983):

1. El modelo debe tener parsimonia: Box y Jenkins enfatizaron la importancia del principio de parsimonia. Un modelo con parsimonia ajusta la información disponible sin usar coeficientes innecesarios. Es decir, se opta siempre por el modelo con el menor número de coeficientes. Este principio es importante ya que en la práctica un modelo con esta característica produce mejores predicciones. El objetivo en la modelación ARIMA no es encontrar el modelo exacto que represente al proceso generador de las observaciones, es más bien encontrar el modelo que se aproxime al verdadero proceso y que explique el comportamiento de la variable en estudio de forma adecuada y práctica.
2. El modelo debe ser estacionario e inversible: La metodología Box-Jenkins requiere que el modelo que se utiliza para la descripción y pronóstico de una serie de tiempo sea tanto estacionario como inversible.

La inversibilidad se refiere a que cualquier modelo ARIMA puede expresar a la serie Y_t en función de las Y observaciones pasadas, es decir $Y_{t-1}, Y_{t-2}, Y_{t-3}, etc.$ Esto es obvio para un modelo autorregresivo de orden p :

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + \dots + \varphi_p Y_{t-p} + a_t$$

No es obvio para un modelo de medias móviles de orden q :

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

No obstante, se puede demostrar que un modelo de medias móviles puede expresar a la serie Y_t como una serie *infinita* de Y observaciones pasadas. Es decir, un modelo de medias móviles se puede transformar o invertir en un modelo autorregresivo de orden infinito.

La *Tabla 4.4* (Bowerman, O'Connell, Koehler 2007) muestra las condiciones de estacionariedad e inversibilidad para los modelos ARMA más comunes. Las condiciones para modelos generales son complicadas por lo que no se estudian en este texto. Sin embargo, se puede decir que si un modelo utiliza parámetros autorregresivos, una condición de estacionariedad necesaria (pero no suficiente) es que la suma de los valores de los parámetros autorregresivos es menor que 1. Si un modelo utiliza parámetros de media móvil, una condición de inversibilidad necesaria (pero no suficiente) es que la suma de los valores de los parámetros de la media móvil es menor que 1. Además, un modelo que utiliza sólo parámetros autorregresivos no tiene condiciones de inversibilidad y un modelo que utiliza sólo parámetros de media móvil no tiene condiciones de estacionariedad.

Modelo	Condiciones de estacionariedad	Condiciones de inversibilidad
AR(1)	$ \varphi_1 < 1$	Ninguna
AR(2)	$\varphi_1 + \varphi_2 < 1$ $\varphi_2 - \varphi_1 < 1$ $ \varphi_2 < 1$	Ninguna
MA(1)	Ninguna	$ \theta_1 < 1$
MA(2)	Ninguna	$\theta_1 + \theta_2 < 1$ $\theta_2 - \theta_1 < 1$ $ \theta_2 < 1$
ARMA(1,1)	$ \varphi_1 < 1$	$ \theta_1 < 1$

Tabla 4.4 Condiciones de estacionariedad e invertibilidad

3. Los coeficientes del modelo deben ser estadísticamente significativos: Se debe determinar si los coeficientes son importantes en el modelo o se deben excluir del mismo. Se hace uso de los valores del error estándar y el valor t asociado a cada uno de los coeficientes del modelo.

El valor de t se define como:

$$t = \frac{\text{coeficiente calculado}}{\text{error estándar asociado al coeficiente calculado}}$$

Ambos valores normalmente son calculados automáticamente por los programas estadísticos tales como MINITAB o SAS.

La exclusión de aquellos coeficientes que no son importantes en el modelo se hace mediante la hipótesis de que el coeficiente no es significativo, por tanto si se rechaza esta hipótesis, el coeficiente es significativo y debe permanecer en el modelo. Como regla práctica se dice que es razonable incluir en el modelo un parámetro cuya estadística t absoluta es mayor que 2. Si el valor de t absoluto de un coeficiente es menor a 2, entonces debemos considerar seriamente, excluir el parámetro del modelo.

La Figuras 4.18 y 4.19 muestra la salida de dos ajustes realizados por MINITAB. En la primer figura se observa que los valores t asociados a los coeficientes calculados son menores a dos, por lo que supone que el modelo propuesto no es apropiado y debe considerarse excluir algún parámetro del modelo. En la segunda figura se ha excluido un parámetro y el valor t en valor absoluto es mayor a dos, para el coeficiente MA. En este caso el parámetro es importante en el modelo y no debe excluirse.

```

ARIMA Model: Yt
Estimates at each iteration
Iteration      SSE      Parameters
0 169.473      0.100  0.100
1 161.440      0.195  0.005
2 160.266      0.057  -0.145
3 159.645      -0.075  -0.295
4 159.497      -0.037  -0.286
5 159.497      -0.048  -0.298
6 159.496      -0.036  -0.286
7 159.496      -0.045  -0.295
8 159.496      -0.038  -0.288
9 159.496      -0.043  -0.294
10 159.496     -0.039  -0.290
11 159.495     -0.042  -0.293
12 159.495     -0.040  -0.290
13 159.495     -0.042  -0.292
14 159.495     -0.040  -0.291
15 159.495     -0.041  -0.292
16 159.495     -0.041  -0.291
17 159.495     -0.041  -0.292
18 159.495     -0.041  -0.292
19 159.495     -0.041  -0.291
20 159.495     -0.041  -0.291

Relative change in each estimate less than 0.0010

Final Estimates of Parameters
Type      Coef SE Coef      T      P
AR 1      -0.0409  0.3656  -0.11  0.911
MA 1      -0.2914  0.3504  -0.83  0.407

Differencing: 1 regular difference
Number of observations: Original series 120, after differencing 119
Residuals:      SS = 159.468 (backforecasts excluded)
                MS = 1.363  DF = 117

Modified Box-Pierce (Ljung-Box) Chi-Square statistic
Lag      12      24      36      48
Chi-Square 14.0  27.6  43.6  57.4
DF        10      22      34      46
P-Value   0.172  0.188  0.125  0.121
    
```

Figura 4.18 Ajuste MINITAB – modelo inadecuado

```

ARIMA Model: Yt
Estimates at each iteration
Iteration      SSE      Parameters
0 178.398      0.100
1 166.087      -0.050
2 160.064      -0.200
3 159.534      -0.254
4 159.533      -0.256
5 159.533      -0.256

Relative change in each estimate less than 0.0010

Final Estimates of Parameters
Type      Coef SE Coef      T      P
MA 1      -0.2561  0.0892  -2.87  0.005

Differencing: 1 regular difference
Number of observations: Original series 120, after differencing 119
Residuals:      SS = 159.505 (backforecasts excluded)
                MS = 1.352  DF = 118

Modified Box-Pierce (Ljung-Box) Chi-Square statistic
Lag      12      24      36      48
Chi-Square 13.8  27.4  43.2  56.8
DF        11      23      35      47
P-Value   0.243  0.241  0.160  0.154
    
```

Figura 4.19 Ajuste MINITAB – modelo adecuado

4. El modelo debe proporcionar un adecuado ajuste: Aplicar la modelación ARIMA a una serie de tiempo no es garantía de que se obtendrá un modelo que ajuste a la perfección, en especial en aquellas series donde existe gran cantidad de perturbaciones aleatorias que no puedan ser ajustadas al modelo. Suele suceder que para una misma serie de tiempo, existen varios modelos que cumplen con las

características de un modelo adecuado y proporcionan resultados similares, sin embargo debe escogerse aquel que se ajuste mejor a la serie.

Una medida útil que ayuda a conocer el grado de ajuste del modelo a la serie es la **RMSE** (root-mean-squared error). Esta medida da a conocer la desviación estándar de los residuales \hat{a}_t , y se calcula como:

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

Donde:

\hat{a}_t : residuales del modelo

n : número de residuales

m : número de coeficientes del modelo

Generalmente se utiliza el valor de RMSE para compararlo con el de otros modelos estimados para la misma serie. Se preferirá aquel modelo cuyo RMSE tienda a tener menor valor.

4.3.4 Evaluación del modelo

Una vez estimados los coeficientes del modelo propuesto, la siguiente etapa en la metodología Box-Jenkins es la evaluación ó verificación del mismo. En este paso se comprueba la eficiencia del modelo y se decide si es estadísticamente adecuado.

Un modelo estadísticamente adecuado es aquel cuyos residuales son independientes entre sí. Es decir, si los residuales son completamente aleatorios.

Para comprender mejor el objetivo de esta etapa, es conveniente recordar algunas ideas importantes:

1. La idea básica de la metodología es modelar aquellos datos que estén correlacionados entre sí, mediante la combinación de términos AR y MA.
2. Los fenómenos reales siempre presentan **perturbaciones aleatorias** (a_t), también llamados **choques aleatorios** ó **proceso de ruido blanco**.
3. La metodología Box-Jenkins asume que las perturbaciones aleatorias (a_t) pertenecientes a un proceso Y_t son independientes entre sí. No existe correlación alguna entre ellas. Por tanto estas perturbaciones no se modelan mediante términos AR o MA.

En la práctica las **perturbaciones aleatorias** no se pueden identificar directamente en una serie de tiempo. Sin embargo, los residuales \hat{a}_t del modelo ARIMA, nos proporcionan un cálculo aproximado de las perturbaciones reales a_t (Pankratz 1983).

Recordemos que los residuales de cualquier modelo ARIMA se definen como:

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

Donde:

\hat{a}_t : residual en el tiempo t

Y_t : serie original

\hat{Y}_t : serie calculada con los parámetros estimados

Entonces, si los residuales muestran estar correlacionados entre sí, significa que existe un patrón que aun no ha sido tomado en cuenta por los términos autorregresivos y/o de medias móviles del modelo propuesto.

Por lo tanto, se concluye que si los residuales están correlacionados de alguna manera, estos no son ruido blanco y se debe buscar otro modelo cuyos residuales sean completamente aleatorios.

La **función de autocorrelación simple de los residuales** es el instrumento que se utiliza para determinar si el modelo es estadísticamente adecuado. El cálculo de los coeficientes de autocorrelación se realiza por medio de la misma fórmula vista en el capítulo 2 pero aplicada a los residuales (\hat{a}_t) del modelo ajustado (Pankratz 1983):

$$r_k(\hat{a}) = \frac{\sum_{t=1}^{n-k} (\hat{a}_t - \bar{a})(\hat{a}_{t+k} - \bar{a})}{\sum_{t=1}^n (\hat{a}_t - \bar{a})^2}$$

Donde:

$r_k(\hat{a})$: coeficiente de autocorrelación simple residual para un retraso de k periodos

\bar{a} : media de los residuales

\hat{a}_t : residual en el periodo t

\hat{a}_{t+k} : residual en el periodo con k retrasos

n : número total de residuales

Mediante el análisis de esta función se busca que los coeficientes de autocorrelación sean cero o muy cercanos a él. Específicamente se pretende que los coeficientes no sean significativos. Esto se determina a través del estadístico t asociado a la función de autocorrelación simple de los residuales:

$$t = \frac{r_k(\hat{a})}{s[r_k(\hat{a})]}$$

$$s[r_k(\hat{a})] = \sqrt{\frac{1 + 2 \sum_{j=1}^{k-1} r_j(\hat{a})^2}{n}}$$

La mayoría de los programas de computadora realiza automáticamente el cálculo del estadístico t . MINITAB por su parte despliega el gráfico de la función de autocorrelación de los residuales, tal como se observa en la *Figura 4.20*.

Por medio del gráfico y de las bandas se puede establecer si el valor de un coeficiente es significativo. Si uno o varios valores de autocorrelación sobrepasan las bandas de dos desviaciones estándar, significa que el ó los coeficientes son significativamente distintos de cero y por lo tanto, los residuales no son independientes entre sí. El modelo no es estadísticamente adecuado. De modo contrario, si los valor de los coeficientes no sobrepasan la banda de puntos, indica que dichos valores no son significativo y por tanto se asume que los residuales son independientes entre si y se acepta el modelo.

También una regla común es que para los tres primeros rezagos de tiempo, si el valor del estadístico t calculado es menor que 1.25 y menor que 1.6 para rezagos posteriores, se concluya que las perturbaciones aleatorias son independientes entre sí (Pankratz 1983).

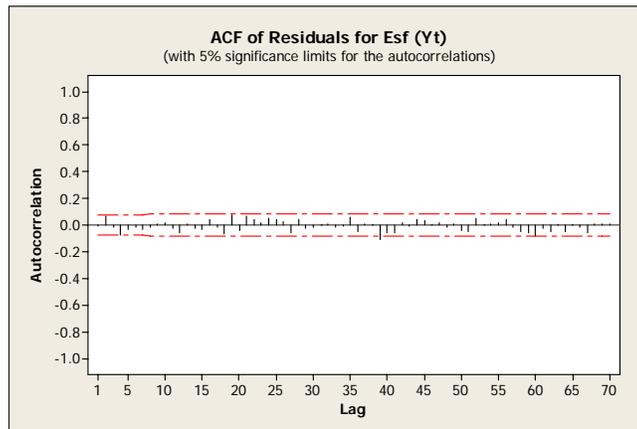


Figura 4.20 Función de autocorrelación simple de los residuales mediante MINITAB

Es necesario precisar que el cálculo del estadístico t tan sólo es una aproximación del valor real, lo que podría provocar que se consideren los residuos como aleatorios, cuando en realidad no lo son, o viceversa. Por esta razón se recomienda el uso del estadístico **Ljung-Box**, el cual también se enfoca al análisis de los coeficientes de autocorrelación de los residuales, pero lo hace de manera grupal. Es una prueba que considera un conjunto de K coeficientes de autocorrelación.

El estadístico **Ljung-Box** se define como:

$$Q^* = n'(n' + 2) \sum_{k=1}^K (n' - k)^{-1} r_k (\hat{a}_t)^2$$

Donde:

$$n' = n - d$$

n : es el número de observaciones de la serie

d : es el grado de diferenciación

$r_k (\hat{a}_t)^2$: es el cuadrado de $r_k (\hat{a}_t)$

K : número de autocorrelaciones

El valor de K se fija arbitrariamente, la mayoría de los programas de computadora establecen $K = 6, 12, 18$ y 24 .

La forma en que se prueba la eficiencia de un modelo a través de este estadístico sigue el mismo razonamiento, es decir, se trata de comprobar que los residuos no están correlacionados y, por tanto, las autocorrelaciones de los residuos deben ser pequeñas. Por consiguiente, Q^* debe ser pequeña. Ahora bien, juzgar que el valor de Q^* es un valor pequeño, es sumamente subjetivo. Es por eso que se utilizan ciertas condiciones para establecer un punto de rechazo de la suficiencia del modelo. Se utiliza una escala de distribución *chi-cuadrado*, a la cual se asocia un valor p que representa la probabilidad de juzgar de manera correcta el valor de Q^* . Se puede demostrar que si el valor de p es pequeño, el valor de Q^* es grande y el modelo no es adecuado y viceversa (Bowerman, O’Connell, Koehler 2007).

De forma práctica, si el valor de p es mayor que 0.01 pero menor que 0.05, se considera que el modelo es inaceptable sin lugar a dudas. Si el valor de p es mayor que 0.05 entonces se dice que es razonable concluir que el modelo es suficiente. No es del alcance de esta tesis demostrar los valores anteriores, pero se consideran aceptables y válidos debido a la amplia difusión que tienen en la literatura existente.

También cabe destacar que la razón de mencionar el valor de p para analizar la suficiencia de un modelo es que en los paquetes informáticos se incluye con frecuencia este valor. Además es una manera práctica y rápida de evaluar la suficiencia del modelo. Una manera de comprobar que el modelo es estadísticamente adecuado.

Para ejemplificar lo enunciado anteriormente, consideremos la salida de MINITAB para dos modelos propuestos (Figuras 4.21 y 4.22). En primera figura, la tabla se muestra los valores del estadístico *Ljung-Box* y el valor p . Los valores p son menores que 0.05, en este caso, se rechaza la suficiencia del modelo y debe proponerse un nuevo modelo. En la segunda, los valores de p son mayores a 0.05 por lo que se acepta la suficiencia del modelo y puede procederse a la realización de pronósticos.

Modified Box-Pierce (Ljung-Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	20.4	40.5	49.5	73.4
DF	8	20	32	44
P-Value	0.009	0.004	0.025	0.004

Figura 4.21 Salida MINITAB 1

Modified Box-Pierce (Ljung-Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	13.8	27.4	43.2	56.8
DF	11	23	35	47
P-Value	0.243	0.241	0.160	0.154

Figura 4.22 Salida MINITAB 2

4.3.5 Pronóstico

La fase final de la metodología Box-Jenkins es pronosticar valores futuros de la serie de tiempo. En primer lugar se debe comprender cómo realizar *predicciones puntuales* a partir de un modelo ARIMA ya estimado. En segundo lugar se debe estudiar cómo estimar límites de probabilidad alrededor de una estimación puntual, conocidos particularmente como *intervalos de confianza*. Y por último se debe conocer y probar la *suficiencia de los pronósticos*.

A partir de este punto y con el fin de explicar la generación de pronósticos, se asumirá que se conoce al modelo ARIMA en su totalidad, es decir, se conoce sus coeficientes y se ha probado que el modelo es adecuado. Se hace notar que en la práctica, estas consideraciones son verdaderas si se ha identificado y estimado los parámetros del modelo conforme a la aplicación de las etapas anteriores.

El pronóstico de una serie de tiempo se hace por medio de predicciones ó estimaciones puntuales. Así, una **predicción puntual** (\hat{Y}_t) se define como el valor de la variable Y en el tiempo t , calculado por medio del modelo ARIMA ajustado a la serie.

La forma más conveniente para empezar a realizar predicciones puntuales es escribir el modelo incluyendo el restablecimiento de la no estacionariedad. Hay que recordar que en el primer paso de la metodología, se induce la estacionariedad en la serie de tiempo (sólo en caso que se requiera), por medio de diferencias. Para restablecer las características originales de la serie, se realiza la operación inversa, es decir, se *integra* a la serie. A este proceso inverso se le conoce como **integración**. Se advierte la importancia del restablecimiento de condiciones iniciales, ya que es lógico que se requieran para obtener pronósticos adecuados y apegados a la realidad.

Para ejemplificar cómo se escribe un modelo considerando la integración de la serie, tomemos un modelo ARIMA(1,1,0) cuya ecuación es:

$$Y'_t = \varphi_1 Y_{t-1} + a_t$$

Donde:

$$Y'_t = Y_t - Y_{t-1}$$

Y'_t : primeras diferencias de la serie

Simplemente se sustituye Y'_t en la ecuación del modelo:

$$Y_t - Y_{t-1} = \varphi_1 Y_{t-1} + a_t$$

Despejando Y_t :

$$Y_t = Y_{t-1} + \varphi_1 Y_{t-1} + a_t$$

Ahora bien, para explicar la forma en que se realizan las predicciones puntuales tomemos como ejemplo el modelo ARIMA(0,1,1) ajustado a la serie que se muestra en la *Tabla 4.5*. (Se explicará de esta forma debido a que es práctico y fácilmente entendible).

t	Y_t	t	Y_t	t	Y_t	t	Y_t
1	15.0000	31	10.7752	61	-1.3173	91	10.5502
2	14.4064	32	10.1129	62	-0.6021	92	11.4741
3	14.9383	33	9.9330	63	0.1400	93	11.5568
4	16.0374	34	11.7435	64	1.4030	94	11.7986
5	15.6320	35	12.2590	65	1.9280	95	11.8867
6	14.3975	36	12.5009	66	3.5626	96	11.2951
7	13.8959	37	11.5378	67	1.9615	97	12.7847
8	14.0765	38	9.6649	68	4.8463	98	13.9435
9	16.3750	39	10.1043	69	6.5454	99	13.6859
10	16.5342	40	10.3452	70	8.0141	100	14.1136
11	16.3839	41	9.2835	71	7.9746	101	13.8949
12	14.1006	42	7.7219	72	8.4959	102	14.2853
13	17.7876	43	6.8300	73	8.4539	103	16.3867
14	17.7354	44	8.2046	74	8.7114	104	17.0884
15	17.0010	45	8.5289	75	7.3780	105	15.8861
16	17.7485	46	8.8733	76	8.1905	106	14.8227
17	18.1888	47	8.7948	77	9.9720	107	15.9479
18	18.5997	48	8.1577	78	9.6930	108	15.0982
19	17.5859	49	7.9128	79	9.4506	109	13.8770
20	15.7389	50	8.7978	80	11.2088	110	14.2746
21	13.6971	51	9.0775	81	11.4986	111	15.1682
22	15.0059	52	9.3234	82	13.2778	112	15.3818
23	16.2574	53	10.4739	83	13.5910	113	14.1863
24	14.3506	54	10.6943	84	13.4297	114	13.9996
25	11.9515	55	9.8367	85	13.3125	115	15.2463
26	12.0328	56	8.1803	86	12.7445	116	17.0179
27	11.2142	57	7.2509	87	11.7979	117	17.2929
28	11.7023	58	5.0814	88	11.7319	118	16.6366
29	12.5905	59	1.8313	89	11.6523	119	15.3410
30	12.1991	60	-0.9127	90	11.3718	120	15.6453

Tabla 4.5 Serie de tiempo ejemplo

La ecuación del modelo, considerando la integración de la serie es:

$$Y_t = Y_{t-1} + a_t - \theta_1 a_{t-1}$$

Una predicción puntual para este modelo ARIMA(0,1,1) particular se calcularía a partir de:

$$\hat{Y}_t = Y_{t-1} + \hat{a}_t - \hat{\theta}_1 \hat{a}_{t-1}$$

Donde:

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

$$\hat{a}_{t-1} = Y_{t-1} - \hat{Y}_{t-1}$$

Al realizar predicciones puntuales (específicamente para este ejemplo) se considera lo siguiente:

- La predicción puntual de \hat{a}_t de la perturbación futura es cero.
- La predicción puntual de \hat{a}_{t-1} de la perturbación aleatoria pasada es el residuo $(t - 1)$ -ésimo si podemos calcular \hat{Y}_{t-1} y es cero si no la podemos determinar.

El coeficiente del modelo calculado mediante MINITAB es:

$$\hat{\theta}_1 = -0.3534$$

A partir del valor del coeficiente se trata de predecir el valor de la variable en el tiempo $t = 1$, la ecuación queda:

$$\hat{Y}_1 = Y_0 + \hat{a}_1 - (-0.2561)\hat{a}_0$$

Puesto que no hay valor para Y_0 , no se puede calcular el valor de \hat{Y}_1 , ni el primer residuo ($\hat{a}_1 = Y_1 - \hat{Y}_1$) y este se considera 0.

Luego para intentar predecir \hat{Y}_2 a partir del tiempo de origen 1 se tiene:

$$\hat{Y}_2 = Y_1 + \hat{a}_2 - (-0.3534)\hat{a}_1$$

$$\text{Como } \hat{a}_1 = \hat{a}_2 = 0$$

$$\hat{Y}_2 = 15 + 0 + 0.3534(0) = 15$$

En este caso $Y_1 = 15$ de la *Tabla 4.5*

Al usar el valor $Y_2 = 14.4064$ se puede calcular:

$$\hat{a}_2 = Y_2 - \hat{Y}_2 = 14.4064 - 15 = -0.5936$$

Con este residuo se puede calcular Y_3 a partir del origen 2 como:

$$\hat{Y}_3 = Y_2 + \hat{a}_3 - (-0.3534)\hat{a}_2$$

$$\hat{Y}_3 = 14.4064 + 0 - (-0.3534)(-0.5936) = 14.1966$$

Con este valor se calcula a su vez:

$$\hat{a}_3 = Y_3 - \hat{Y}_3 = 14.9383 - 14.1966 = 0.7417$$

Al continuar con este proceso con cada uno de los siguientes periodos, se obtiene que la predicción puntual $\hat{Y}_{120} = 14.956$. El residuo para el periodo 120 es:

$$\hat{a}_{120} = Y_{120} - \hat{Y}_{120} = 15.6443 - 14.956 = 0.6893$$

Ahora bien, hasta ahora se han predicho valores hasta el periodo 120, a partir del origen inmediato anterior, es decir, se calculó el valor de \hat{Y}_{120} a partir del origen 119, el valor \hat{Y}_{119} a partir del origen 118, y así sucesivamente. Sin embargo como no se cuenta con las observaciones para los periodos subsecuentes 121, 122, 123, etc., las predicciones puntuales $\hat{Y}_{121}, \hat{Y}_{122}, \hat{Y}_{123}, etc.$ se obtienen a partir del origen 120. Entonces el pronóstico puntual para Y_{121} es:

$$\hat{Y}_{121} = Y_{120} + \hat{a}_{121} - (-0.3534)\hat{a}_{120}$$

$$\hat{Y}_{121} = 15.6453 + 0.3534(0.6893)$$

$$\hat{Y}_{121} = 15.8889$$

Puesto que no se cuenta con la observación Y_{121} , no se puede obtener el valor \hat{a}_{121} por lo que se supone un valor de 0. Entonces la predicción puntual para Y_{122} a partir del origen 120 es:

$$\hat{Y}_{122} = Y_{121} + \hat{a}_{122} - (-0.3534)\hat{a}_{121}$$

Se asume que el valor Y_{121} es igual al valor de \hat{Y}_{121} puesto que no existe valor observado en ese tiempo.

$$\hat{Y}_{122} = 15.8889 + 0 - (-0.3534)(0)$$

$$\hat{Y}_{122} = 15.8889$$

Esta operación se realiza sucesivamente hasta el periodo de tiempo que se desee pronosticar.

Por otra parte cada pronóstico posee un intervalo de confianza asociado a la probabilidad de que el valor futuro observado se encuentre dentro de dicho intervalo. Normalmente se establece que la probabilidad sea del 95%. A esta combinación probabilidad-intervalo normalmente se le conoce como *intervalo de confianza del 95%*.

Por medio de esta manera se puede conocer si los pronósticos son satisfactorios o no. La generación de pronósticos es satisfactoria si los valores futuros se encuentran dentro del intervalo de confianza. De manera inversa serán insatisfactorios si estos se hallan fuera del intervalo.

El cálculo de un intervalo de confianza del 95% para un pronóstico, se realiza por medio de la ecuación:

$$[\hat{Y}_{n+t} \pm 1.96 SE_{n+t}]$$

Donde:

\hat{Y}_{n+t} : pronóstico en el periodo $(n + t)$ a partir del origen n

SE_{n+t} : error estandar del error del pronóstico del periodo $(n + t)$

El cálculo de SE_{n+t} quedará fuera del alcance de este texto. No obstante, las salidas de los programas estadísticos contienen los valor SE_{n+t} correspondientes a cada predicción puntual.

Para el ejemplo anterior donde $\hat{Y}_{121} = 15.8889$ y $\hat{Y}_{122} = 15.8889$ se tiene que

$$SE_{120+1} = 1.0394$$

$$SE_{120+2} = 1.7491$$

El intervalo de confianza del 95% para Y_{121} es:

$$[15.8889 \pm 1.96 (1.0394)] = [13.8517, 17.9261]$$

Para Y_{122} es:

$$[15.8889 \pm 1.96 (1.7491)] = [12.4608, 19.3170]$$

Después de que se ha encontrado un modelo adecuado, se pueden llevar a cabo los pronósticos para un periodo, o varios, en el futuro.

Si el patrón de la serie parece cambiar con el tiempo, los nuevos datos podrían usarse para volver a estimar los parámetros del modelo o, de ser necesario, desarrollar un modelo completamente nuevo.

Vigilar los errores de pronóstico es una buena idea. Si las magnitudes de los errores más recientes tienden a ser considerablemente mayores que los anteriores, quizá sea hora de evaluar nuevamente el modelo.

4.4 Metodología Box-Jenkins Estacional

La metodología vista hasta ahora, está enfocada a series de tiempo no estacionales. Sin embargo, en la práctica se presentan frecuentemente series con estacionalidad. Este hecho no se había remarcado anteriormente debido a que las etapas de la metodología para series estacionales no varían en su concepto y siguen las mismas pautas.

Las variaciones que deben tenerse en cuenta al aplicar modelos ARIMA a series estacionales consisten principalmente en tres aspectos: el reconocimiento del periodo estacional, la estabilización de la serie y la adición de términos autorregresivos y/o de medias móviles referidos a L periodos de tiempo anteriores donde $L = \text{periodo estacional}$. A continuación se presentan aquellas consideraciones extras que deben hacerse en cada una de las etapas.

Estacionariedad: Uno de los requisitos para poder iniciar la identificación y estimación del modelo es que la serie debe ser estacionaria. Recordemos que si la serie presenta algún tipo de patrón ascendente o descendente, la serie es inestable o no estacionaria, por lo que la serie debe ser transformada para lograr su estabilización. La *Figura 4.23* muestra una serie estacional donde se muestra un aumento en los valores de la serie.

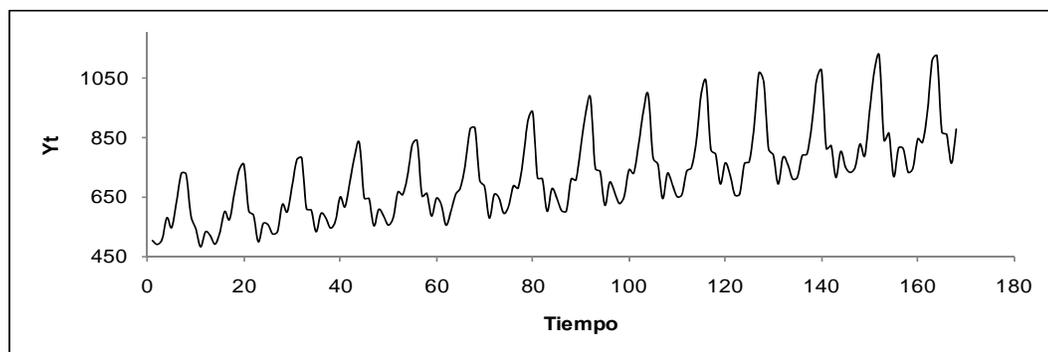


Figura 4.23 Serie de tiempo no estacionaria

Para series estacionales, la estabilidad se logra por medio de las transformaciones siguientes:

- Diferenciación regular: Es la el mismo tipo de diferenciación presentado en la sección. Al igual que en series no estacionales, se puede requerir las primeras o las segundas diferencias de la serie.

$$Y'_t = Y_t - Y_{t-1} \quad \text{Primeras diferencias}$$

- Diferencias estacionales: Se obtienen a partir de restar el valor de la variable L periodos anteriores.

$$Y^*_t = Y_t - Y_{t-L}$$

- Combinación de ambas:

$$Y^*_t = Y_t - Y_{t-1} - Y_{t-L} - Y_{t-L-1}$$

El periodo estacional L se obtiene de acuerdo a lo visto en la modelación clásica, con la ayuda de las funciones de autocorrelación o en su caso con el periodograma.

Para conocer si una serie estacional es estacionaria desde un principio o bien una vez aplicada alguna de las transformaciones anteriores, se sigue el mismo razonar que el visto en la *sección 4.3.1*.

Si la función de autocorrelación simple de los valores $Y_1, Y_2, etc.$ de la serie de tiempo se corta claramente con rapidez o si se corta rápidamente para los primeros retrasos (1,2,3) y para los retrasos de L periodos de tiempo ($L, 2L, 3L, etc.$), entonces se deben considerar que los valores de la serie estacional son estacionarios. (Bowerman, O'Connell, Koehler 2007)

Si la función de autocorrelación simple de los valores Y_1, Y_2, etc para los primeros retrasos (1,2,3) y para los retrasos de L periodos de tiempo ($L, 2L, 3L, etc.$) de la serie de tiempo, se cortan con lentitud extrema, entonces se deben considerar que los valores de la serie de tiempo son no estacionarios

Además de estas transformaciones, normalmente en series estacionales, se suelen aplicar transformaciones del tipo:

$$Y^*_t = Ln Y_t$$

$$Y^*_t = Y_t^{0.25}$$

ó

$$Y^*_t = Y_t^{0.5}$$

Estas transformaciones se deben a la variabilidad que puede existir entre distintos periodos de tiempo. La gráfica de la figura anterior (*Figura 4.23*) muestra el ejemplo de una serie estacional con

variabilidad estacional, es decir, se tiene un incremento en el rango de valores en cada periodo de tiempo.

Ahora bien, las Figuras 4.24, 4.25 y 4.26 muestran la transformación de la misma serie, utilizando $Y^*_t = Y_t^{0.5}$, $Y^*_t = Y_t^{0.25}$ y $Y^*_t = \text{Ln } Y_t$ respectivamente. Se puede observar que la última transformación quita totalmente la variación estacional presente en la serie, mientras que las otras dos tan sólo la atenúan ligeramente.

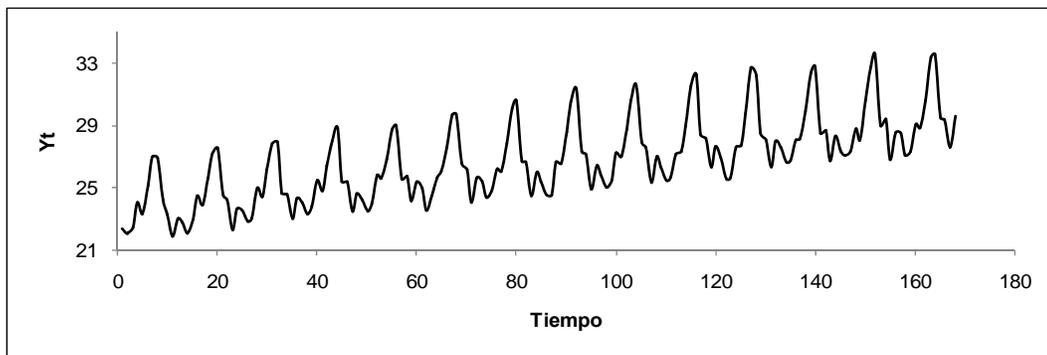


Figura 4.24 Raíz cuadrada de Y_t

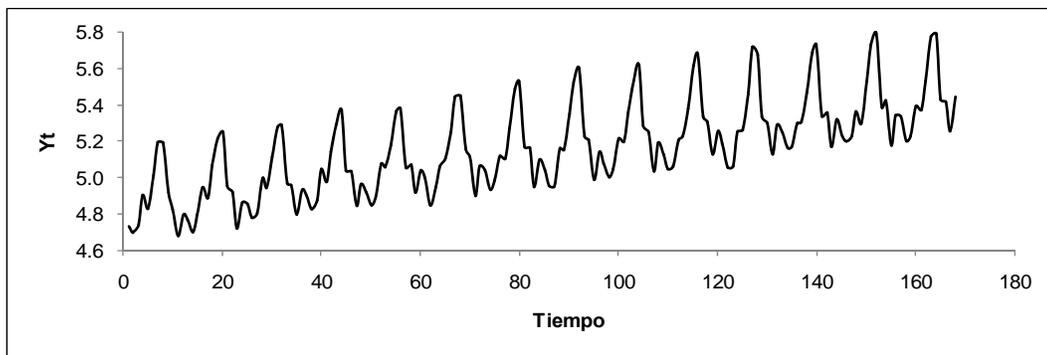


Figura 4.25 Raíz cuarta de Y_t

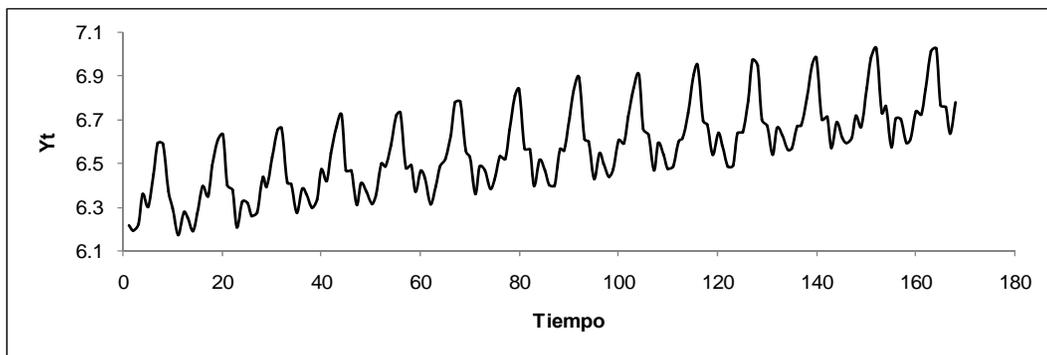


Figura 4.26 Logaritmo natural de Y_t

Identificación: Para series estacionales, se realiza de la misma forma que para una serie no estacional, es decir, observando el comportamiento de la FAS y la FAP, tanto para los primeros periodos de tiempo ($t = 1, 2, 3, 4, 5$) como para aquellos alejados L periodos de tiempo ($t = L, 2L, 3L, 4L$).

El modelo propuesto para series estacionales estará conformado tanto por coeficientes autorregresivos y de media móvil ordinarios, como de coeficientes autorregresivos y medias móviles estacionales de orden P y Q respectivamente.

Un modelo autorregresivo estacional de orden P sería:

$$Y_t = \varphi_L Y_{t-L} + \varphi_{2L} Y_{t-2L} + \dots + \varphi_{PL} Y_{t-PL} + a_t$$

Se puede demostrar que la FAS se extingue en los retrasos $L, 2L, 3L, \dots, etc.$, mientras que la FAP tiene autocorrelaciones significativas (espigas) en los retrasos $L, 2L, \dots, PL$.

De forma similar, un modelo de media móvil estacional de orden Q sería:

$$Y_t = a_t - \theta_L a_{t-L} + \theta_{2L} a_{t-2L} + \dots + \theta_{PL} a_{t-PL}$$

En este modelo la FAS tiene autocorrelaciones significativas (espigas) en los retrasos $L, 2L, \dots, QL$ y autocorrelaciones no significativas en los demás. La FAP se extingue en los retrasos de tiempo $L, 2L, 3L, \dots, etc$ (Bowerman, O'Connell, Koehler 2007).

Estimación, Evaluación y Pronóstico: Las tres etapas finales se efectúan de la misma manera que el visto para series no estacionales. Únicamente en el estadístico Ljung-Box, se calcula n' mediante la fórmula:

$$n' = n - (d + LD)$$

Donde:

d: grado de diferenciación regular

D: grado de diferenciación estacional

L: periodo estacional

4.5 Ventajas y desventajas de la modelación ARIMA

La modelación ARIMA presenta importantes ventajas frente a la modelación clásica. La principal de ellas es el gran grado de ajuste que proporciona a la mayoría de las series de tiempo. A diferencia de la modelación clásica en donde se ajusta una serie a un modelo matemático ya establecido, los modelos ARIMA se ajustan a una serie en particular.

Utilizando la metáfora, como lenguaje, la modelación ARIMA viene a ser como la confección de un traje que debe ser ajustado a una serie de tiempo particular. De acuerdo a esta analogía, el analista de series de tiempo utiliza instrumentos (los modelos estadísticos), materiales (los datos) y planes (estrategias de construcción del modelo) para la definitiva especificación del modelo.

Otras ventajas de los modelos ARIMA sobre los métodos clásicos de tratamiento de series de tiempo son:

- Los conceptos que se utilizan para la modelación ARIMA, se derivan de sólidas teorías de la probabilidad clásica y de la estadística matemática
- Los modelos ARIMA son una familia de modelos, no simplemente un único modelo
- Box y Jenkins desarrollaron una completa estrategia que sirve de guía para escoger un apropiado modelo dentro de esta gama de modelos
- Un apropiado modelo ARIMA produce óptimas predicciones

Por otro lado, dentro de las dificultades que se pueden presentar en la modelación ARIMA, se encuentra el problema que supone para el analista, recién iniciado en la teoría y la ejecución de la metodología Box-Jenkins, la correcta elección de un modelo adecuado en los primeros intentos. Es por ello que la modelación de un proceso ARIMA ha sido referida por algunos autores como un arte. Quizá sería preferible utilizar el término habilidad, debido a que las técnicas básicas en la modelación de los procesos ARIMA son fácilmente accesibles y ajustables tras un aprendizaje relativamente mediano.

Una desventaja más es que, aunque se utilizan programas de cómputo en el cálculo de los coeficientes de un modelo ARIMA probable, la aplicación de la metodología Box-Jenkins sigue siendo una labor manual, es decir, requiere ser ejecutada por el analista. Hay que recordar que no cualquier modelo propuesto proporciona necesariamente el ajuste adecuado.

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

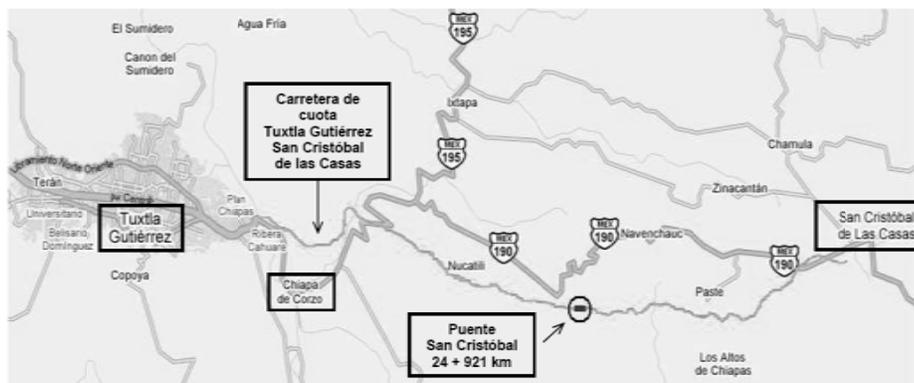


Figura 5.1 Ubicación Puente San Cristóbal

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).

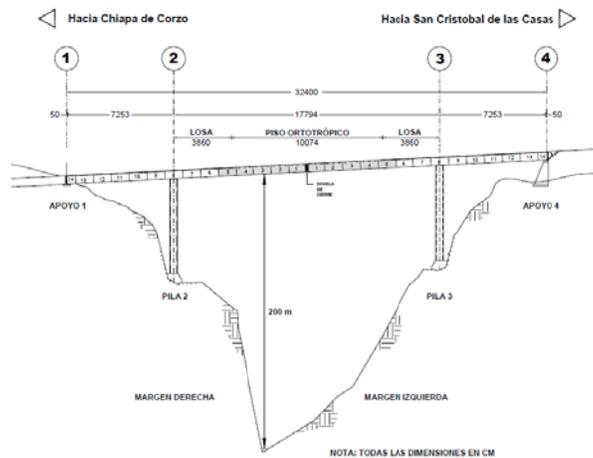


Figura 5.3 Vista en elevación del Puente San Cristóbal

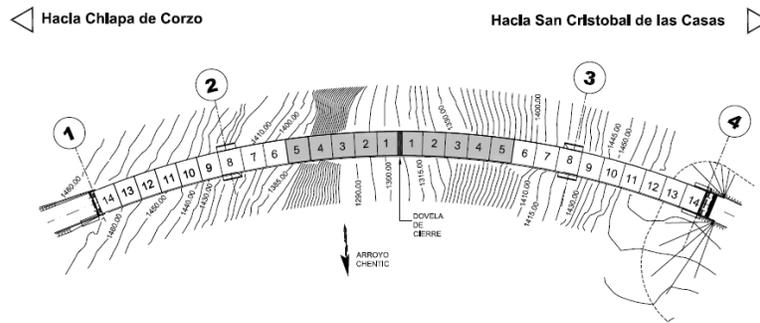


Figura 5.4 Vista en planta del Puente San Cristóbal

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.

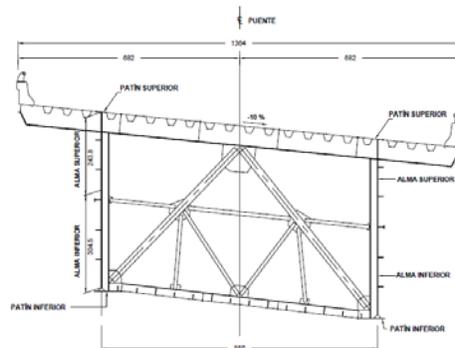


Figura 5.4 Sección transversal con piso ortotrópico

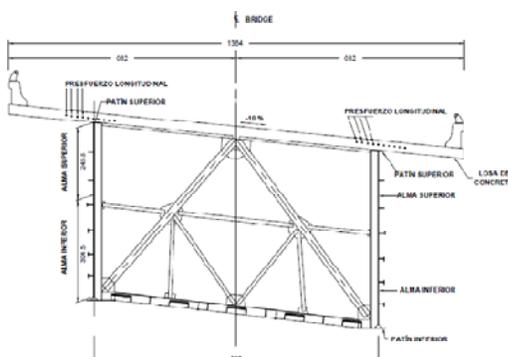


Figura 5.5 Sección transversal con piso de concreto

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.

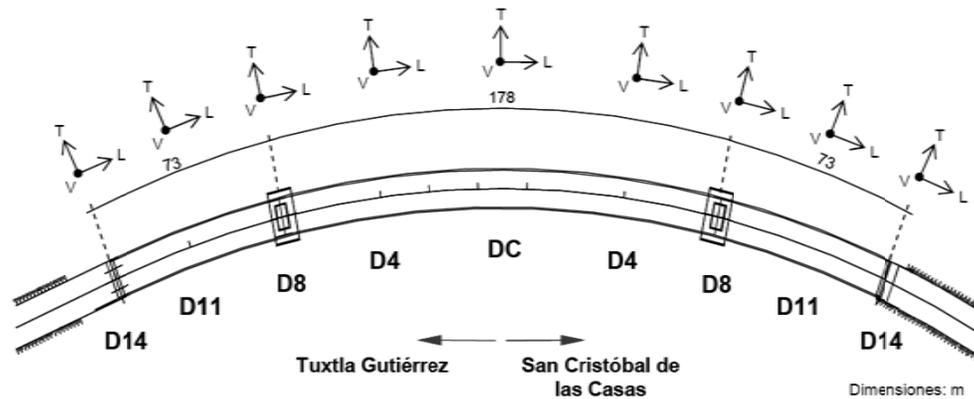


Figura 5.6 Ubicación de las dovelas instrumentadas del puente San Cristóbal (vista en planta)

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.

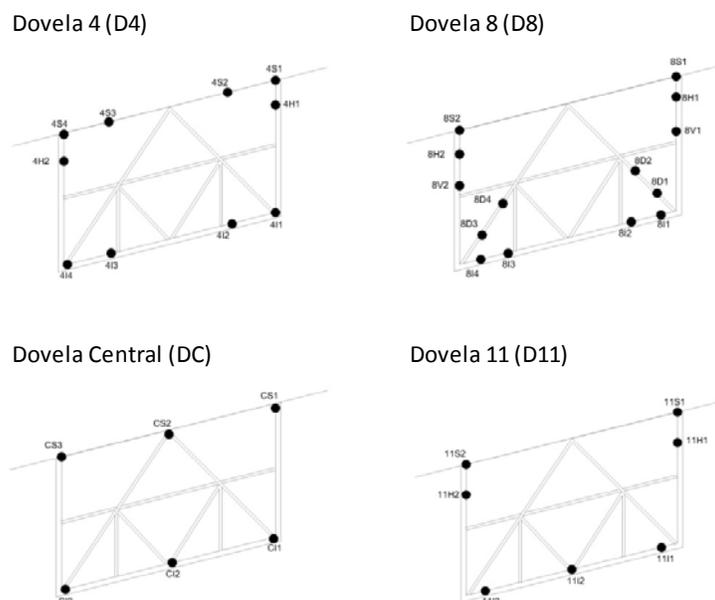


Figura 5.7 Distribución de cuerdas vibrantes en las dovelas más importantes del Puente San Cristóbal

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.

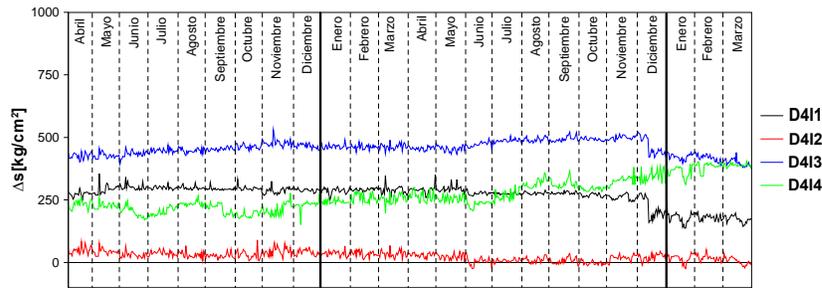


Figura 5.8 Incremento de esfuerzos dovela D4-Inferior lado Tuxtla Gutiérrez (8 am)

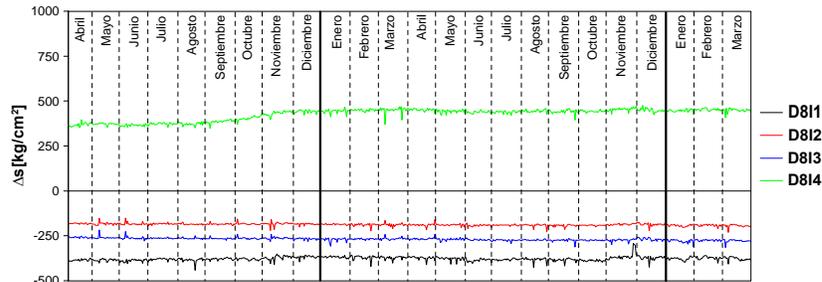


Figura 5.9 Incremento de esfuerzos dovela D8-Inferior lado Tuxtla Gutiérrez (8 am)

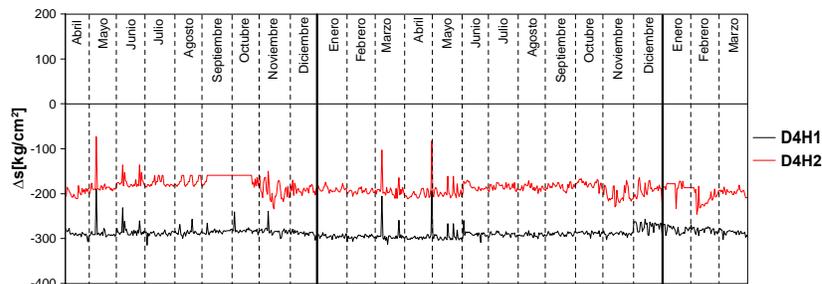


Figura 5.10 Incremento de esfuerzos dovela D4-Horizontal lado San Cristóbal (8 am)

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.

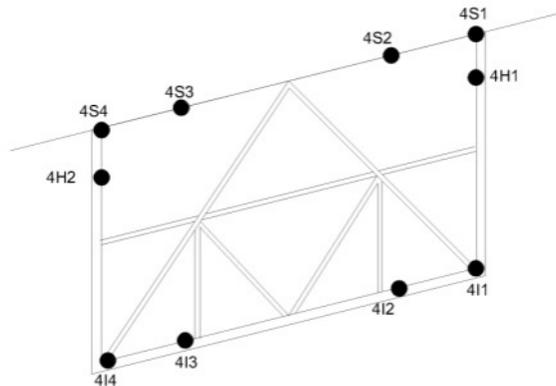


Figura 5.11 Distribución de cuerdas vibrantes en Dovela 4

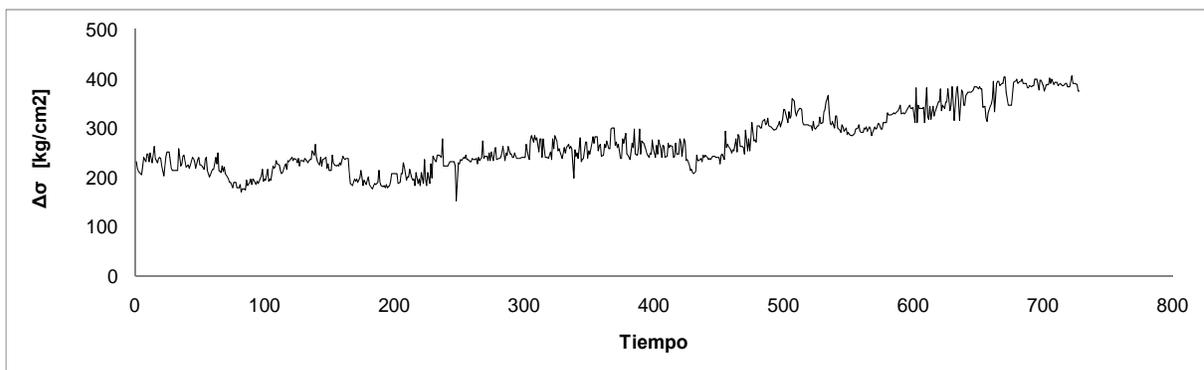


Figura 5.12 Incrementos de Esfuerzos D4-I4

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.

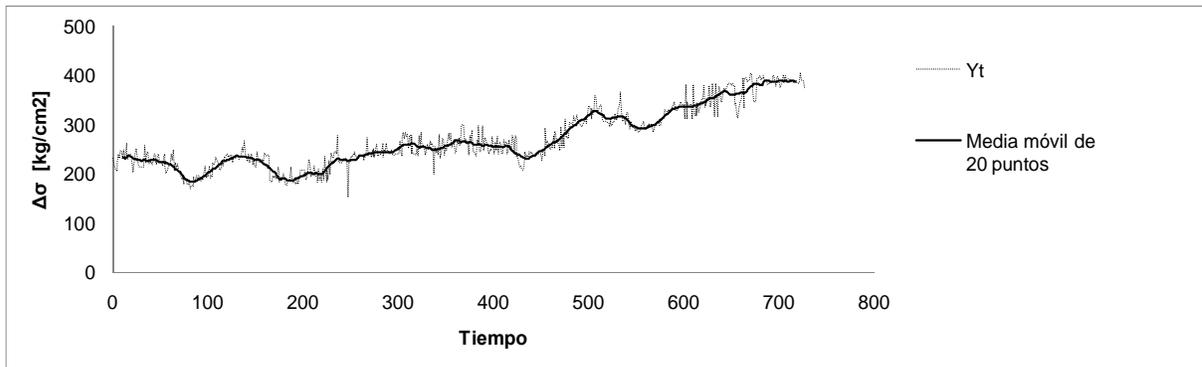


Figura 5.13 Suavización de la serie aplicando una media móvil de 20 puntos

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:

$$T_t = 178.272 + 0.2514 t$$

$$R^2 = 0.769$$

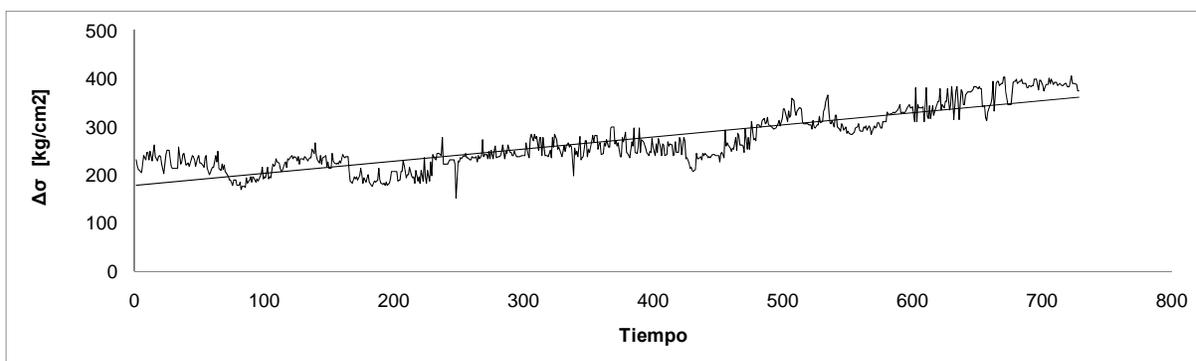


Figura 5.14 Modelo lineal de la tendencia

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:

$$R^2 = 0.8665$$

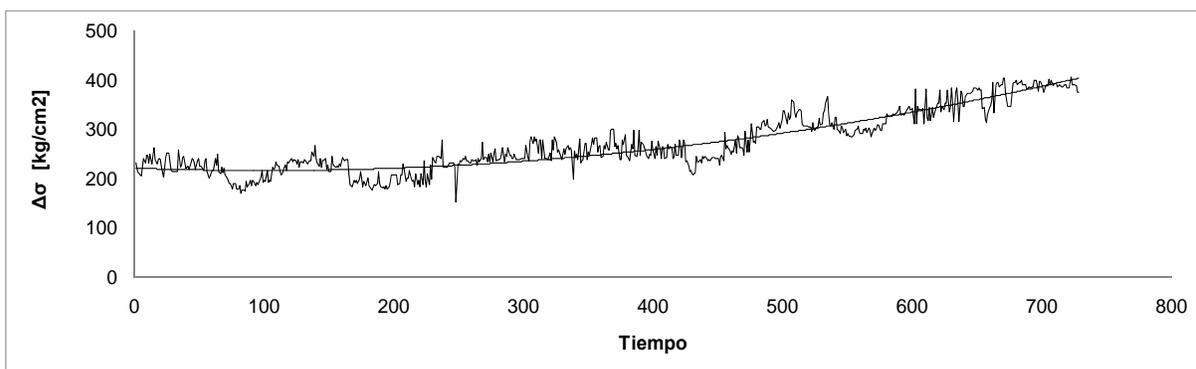


Figura 5.15 Modelo cuadrático de la tendencia

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.

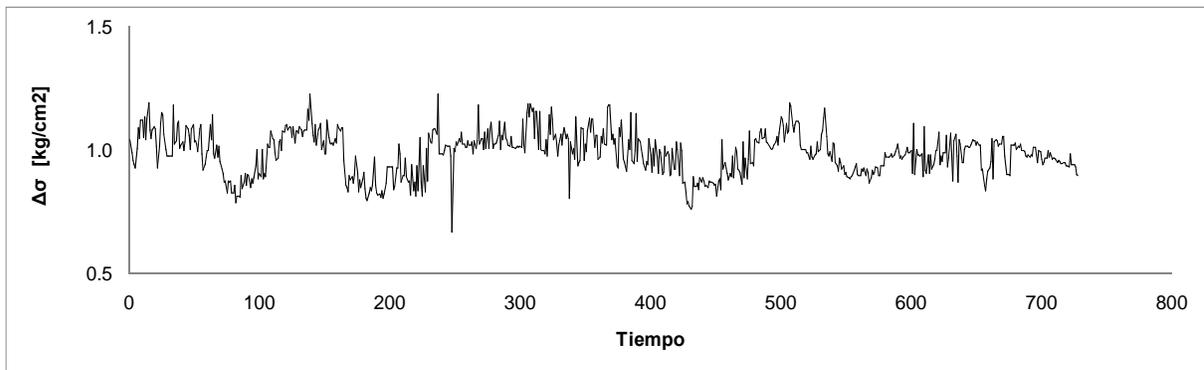


Figura 5.16 Remoción de la tendencia

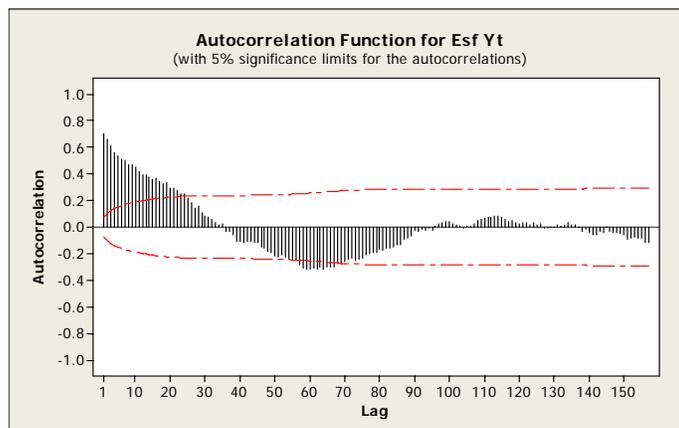


Figura 5.17 Función de autocorrelación simple

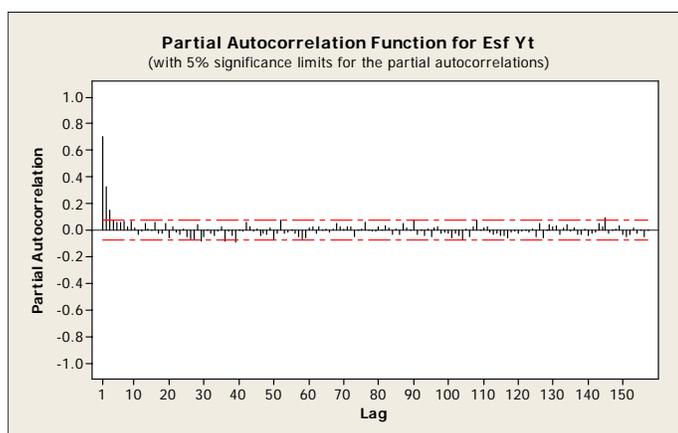


Figura 5.18 Función de autocorrelación parcial

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.

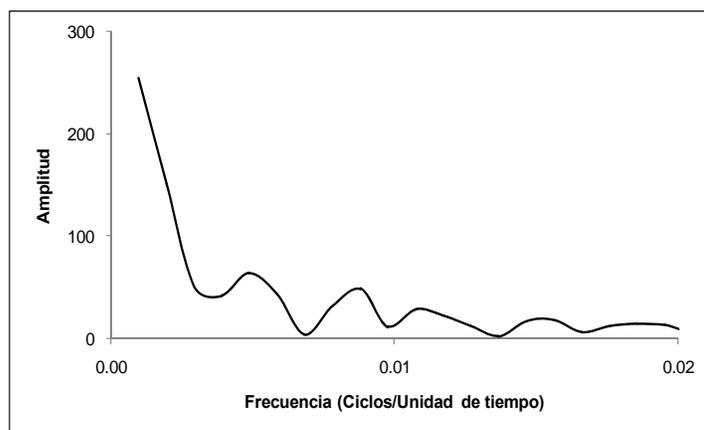


Figura 5.19 Periodograma

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.

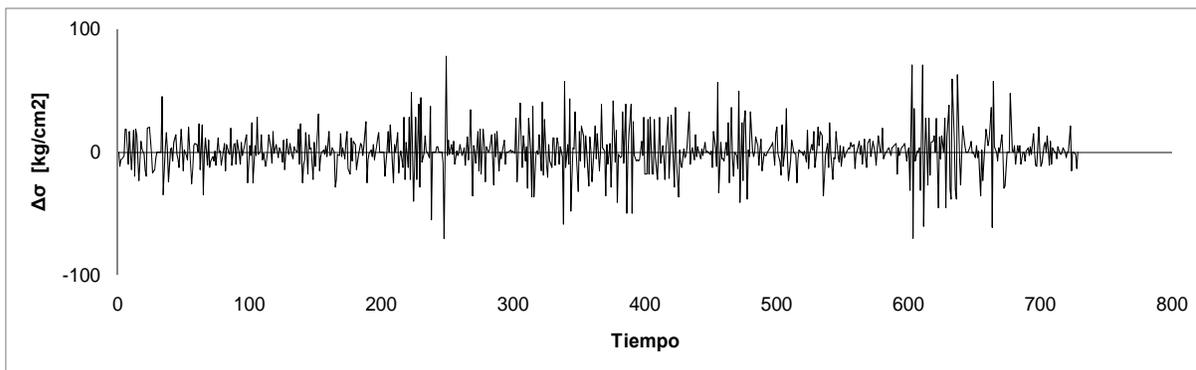


Figura 5.20 Primeras diferencias serie D4-I4

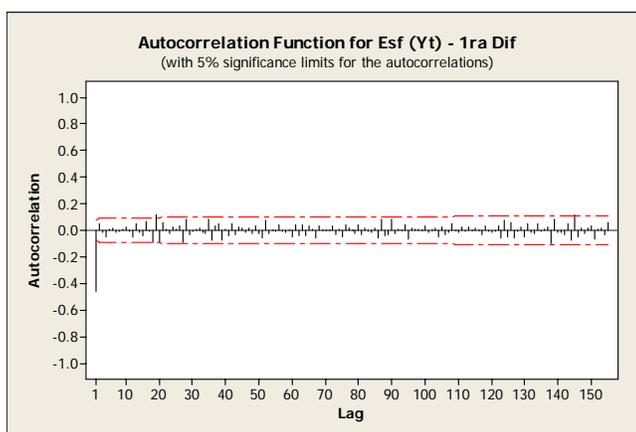


Figura 5.21 Función de autocorrelación simple para D4-I4 (serie diferenciada)

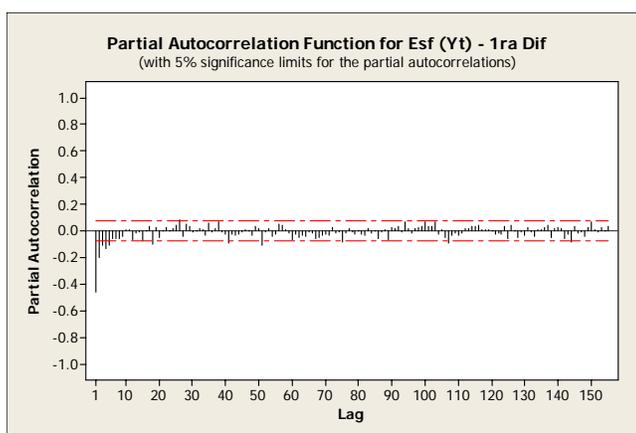


Figura 5.22 Función de autocorrelación parcial para D4-I4 (serie diferenciada)

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*.

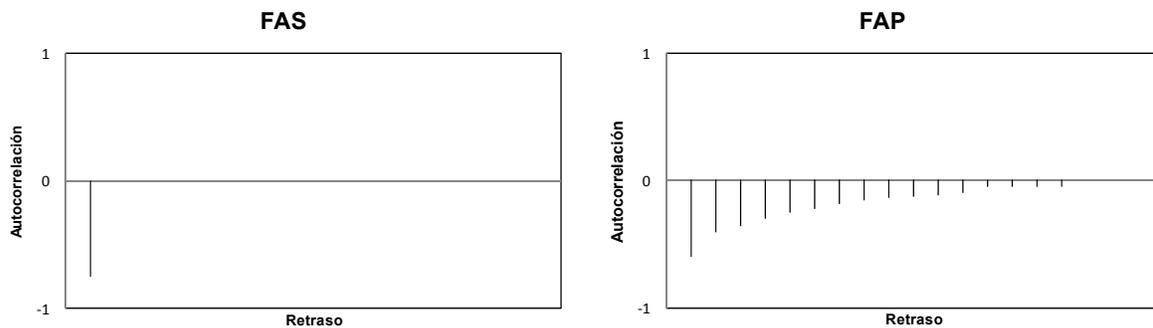


Figura 5.23 Funciones de autocorrelación simple y parcial teóricas para el modelo MA(1)

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.

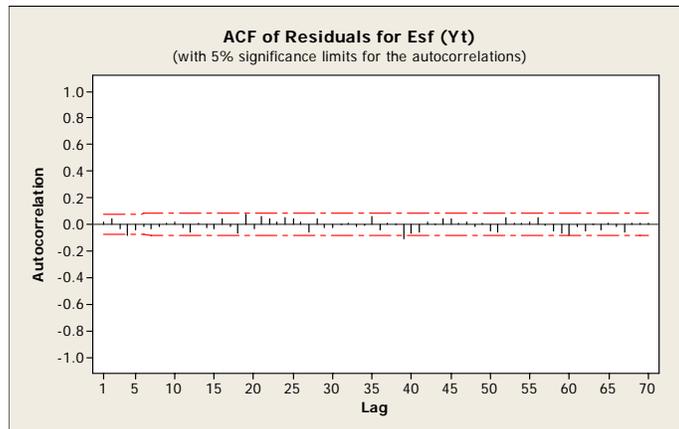


Figura 5.25 Función de autocorrelación simple de los residuos modelo ARIMA(0,1,1)

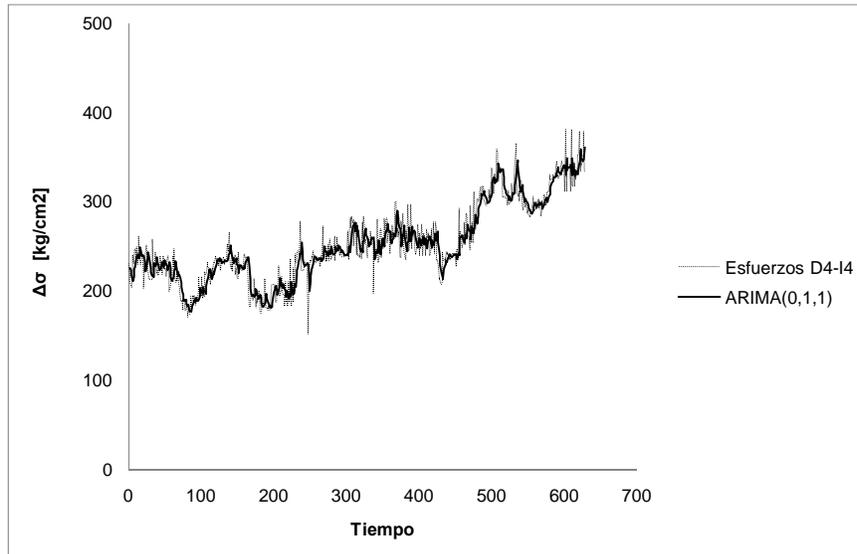


Figura 5.26 Ajuste ARIMA(0,1,1)

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*.

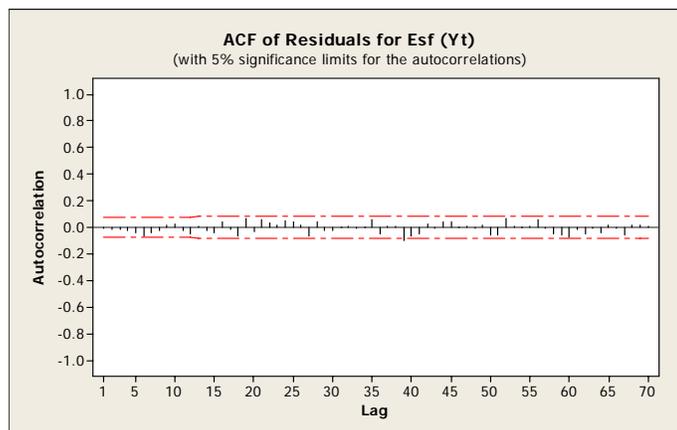


Figura 5.30 Función de autocorrelación simple de los residuos modelo ARIMA(5,1,0)

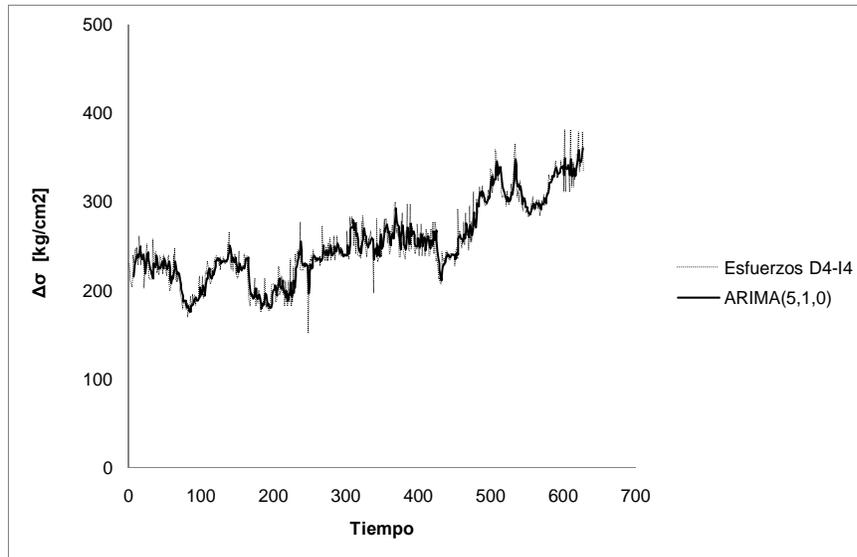


Figura 5.31 Ajuste ARIMA(5,1,0)

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.

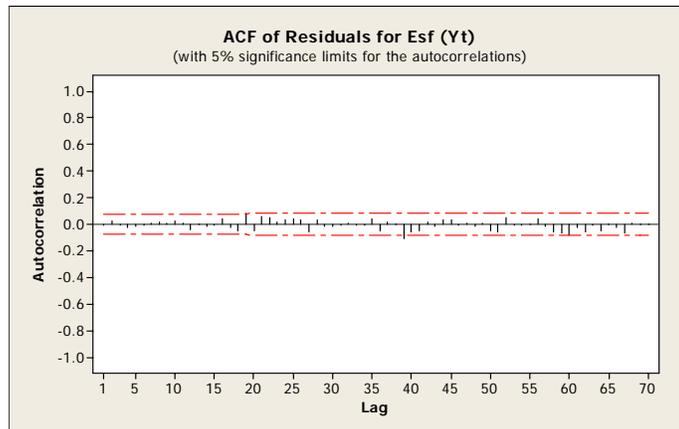


Figura 5.33 Función de autocorrelación simple de los residuos modelo ARIMA(3,1,4)

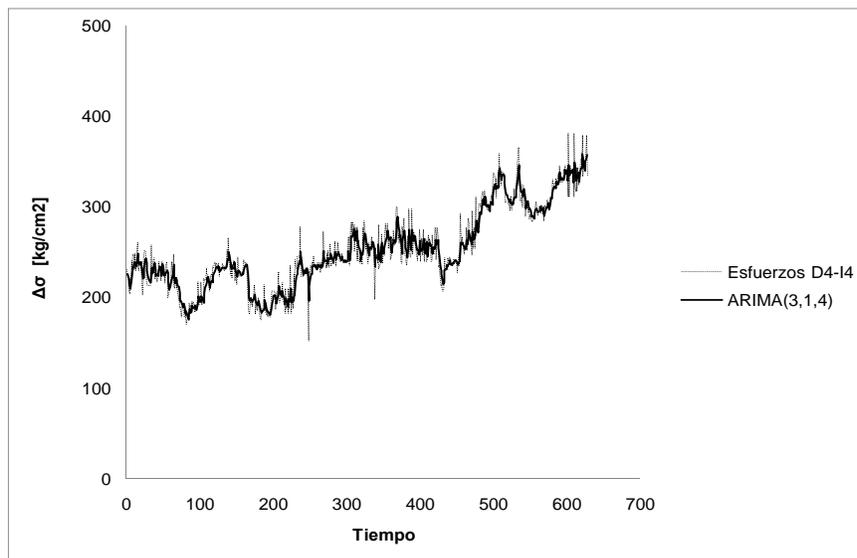


Figura 5.34 Ajuste ARIMA(3,1,4)

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.

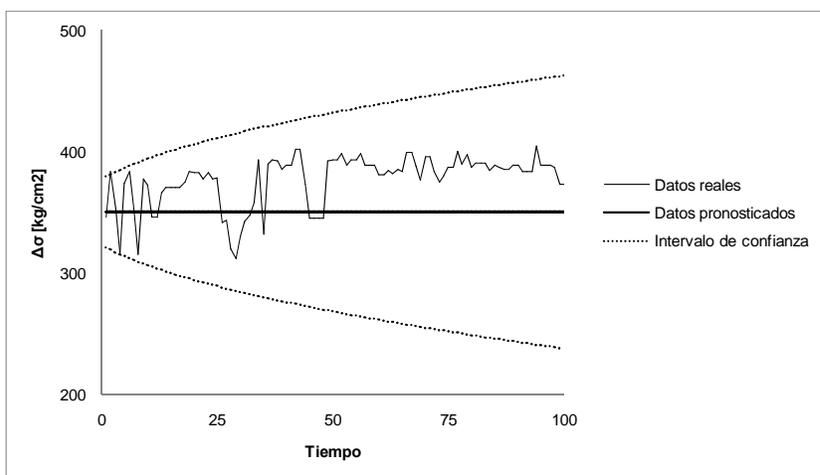


Figura 5.35 Pronósticos ARIMA(0,1,1)

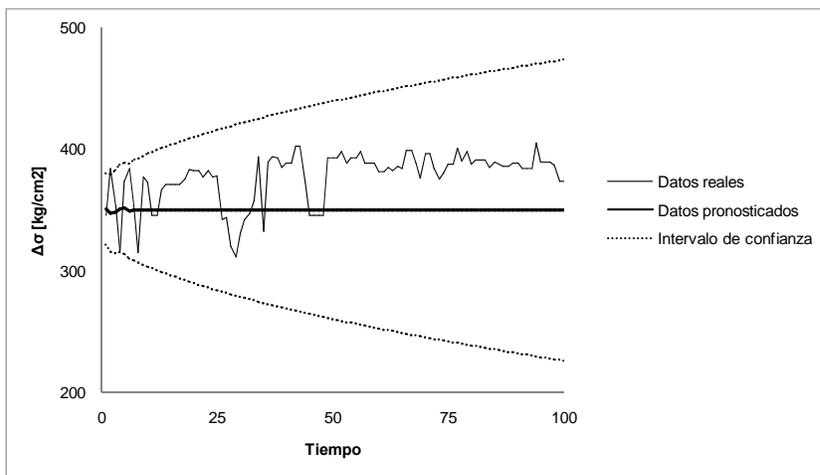


Figura 5.36 Pronósticos ARIMA(5,1,0)

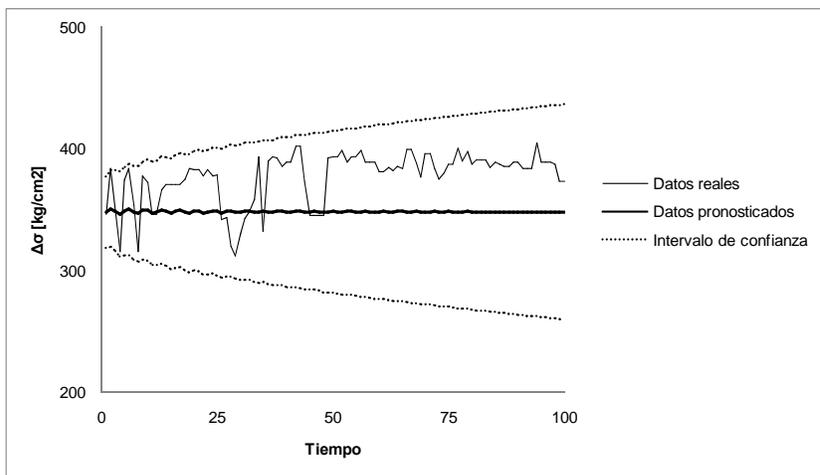


Figura 5.37 Pronósticos ARIMA(3,1,4)

6 Herramientas para identificar cambios significativos en el comportamiento estructural de un puente instrumentado

Como se ha visto hasta ahora, los modelos ARIMA son una herramienta eficaz para la modelación y pronóstico de un gran número de fenómenos físicos. Su aplicación en años recientes en el campo de la ingeniería civil (particularmente en el monitoreo de la salud estructural), ha llevado consigo numerosos avances y logros referentes al modelado de fenómenos así como pronóstico de situaciones futuras. No obstante, en la actualidad aún existe el reto de determinar anticipadamente la presencia de daño, su magnitud y la ubicación del mismo, por medio del análisis de parámetros estructurales.

La teoría vista hasta este punto, tan sólo es capaz de modelar y proporcionar pronósticos. Sin embargo, no se expone algún punto de partida que determine si cambios en las mediciones son verdaderamente importantes e indiquen deterioro o daño en el sistema estructural. Se hace indispensable desarrollar herramientas que clasifiquen de algún modo, diferentes estados de la estructura, es decir, que identifique cambios significativos en los parámetros estructurales.

En este capítulo se presentan dos herramientas que ayudan a monitorear el estado de una estructura por medio de modelos ARIMA ajustados a registros provenientes de acelerómetros, deformímetros, cuerdas vibrantes, etc. Particularmente se trabaja con el modelo **ARIMA(5,1,0)** de las **628 observaciones de incrementos de esfuerzos** de la dovela **D4-I4** del Puente San Cristóbal.

Antes de continuar, cabe mencionar (a fin de obtener una idea general) las premisas en que se basan ambas herramientas:

- Los registros actuales (aceleraciones, esfuerzos, deformaciones, etc.) representan el estado saludable de la estructura.
- Los registros son generados por un proceso desconocido representado por un modelo ARIMA determinado.
- Un cambio en las propiedades estructurales del sistema, se reflejará indudablemente en los registros, esto significa que el proceso se deteriora, consecuentemente se verá un cambio en los coeficientes del modelo.
- No cualquier cambio es significativo o relevante. Los coeficientes de un modelo tienen cierta variabilidad, pero siguen representando el mismo proceso.
- Si nuevos registros presentan cambios significativos en los coeficientes del modelo ARIMA, esto puede significar deterioro y/o daño en el sistema estructural.

Se aclara que el uso de estas herramientas en el presente trabajo se limita a establecer las condiciones iniciales consideradas como el estado saludable de la estructura (puente San Cristóbal).

Así mismo, no es del alcance de este texto analizar la magnitud del posible daño y la identificación del mismo, debido a las siguientes razones:

- La estructura en cuestión (Puente San Cristóbal) es una obra recientemente construida y por lo tanto saludable
- Consecuentemente no se cuenta con registros de algún posible estado deteriorado de la estructura
- No se cuentan con registros posteriores a marzo del 2009

6.1 Procesos de control estadístico

El proceso de control estadístico (SPC por sus siglas en inglés) ó también llamado control estadístico de procesos, es una herramienta que trata de determinar y eliminar variaciones en un proceso mientras este se lleva a cabo (Everett y Ronald 1992). Es decir, busca que el proceso se encuentre bajo ciertos límites de referencia establecidos como óptimos o permisibles.

En el campo de la ingeniería civil, se propone que el proceso de control estadístico sea utilizado para proveer una herramienta en el monitoreo de la salud estructural. Específicamente se expone el cuadro de control (o gráfico de control) de media muestral \bar{X} para monitorear la media de los principales parámetros estructurales. El objetivo principal de un cuadro de control es comparar registros nuevos con anteriores e identificar inconsistencias entre ellos.

El gráfico de control \bar{X} es construido a partir de la información generada en el inicio de la vida útil de la estructura, lo que representa el estado cero ó inicial de la estructura. Comúnmente se le denomina como el estado saludable de la estructura.

En una idea general, la construcción de un cuadro de control se efectúa dibujando una línea central (Center Line CL) y dos líneas horizontales extras que representan los límites superior e inferior (Upper Control Limits UCL y Lower Control Limits LCL). En la *Figura 6.1* se muestra gráficamente los elementos de un cuadro de control.

Los límites se calculan mediante:

$$CL = \mu$$

$$UCL = \mu + Z_{\alpha/2} \sigma$$

$$LCL = \mu - Z_{\alpha/2} \sigma$$

Donde:

μ : media de la serie

σ : desviación estándar de la serie

$1 - \alpha$: probabilidad de decir que el proceso cambió cuando en realidad no lo ha hecho



Figura 6.1 Gráfico de control \bar{X}

La teoría del proceso de control estadístico, específicamente en la construcción de gráficos de control, requiere de datos no correlacionados. Es decir, datos sin tendencia. No obstante, la naturaleza de la mayoría de los procesos físicos en la ingeniería civil, tales como los incrementos de esfuerzos, presentan correlación. Por este motivo, se propone el uso de los residuales del modelo ARIMA ajustado a las observaciones. Como se explica en el capítulo 4, los residuales del modelo ARIMA ajustado son independientes entre sí, no existe correlación entre ellos. Se sugiere además el ajuste de un modelo puramente autorregresivo, debido a la facilidad para manejar este tipo de modelos.

Para aplicar los modelos AR y el proceso de control estadístico al monitoreo estructural, se procede de la siguiente manera (Fugate, Sohn y Farrar 2001):

- Se ajusta un modelo AR de orden p a la información del estado inicial de la estructura (estado saludable)
- Se obtienen los residuales mediante $\hat{a}_t = Y_t - \hat{Y}_t$
- Se normalizan los residuales extrayendo a cada valor la media μ y dividiendo entre la desviación estándar σ
- Se construye el cuadro de control \bar{X} con los residuales normalizados calculados anteriormente
- Nuevos registros se ajustan al mismo modelo AR de orden p calculado para el estado inicial y se calculan los respectivos residuos
- Los nuevos residuales calculados se normalizan con la media μ y desviación estándar σ pertenecientes a la primera muestra de datos
- Los nuevos datos normalizados se introducen en el gráfico de control construido previamente
- Si la condición de la estructura no ha cambiado, entonces se debe esperar que los nuevos residuales en su mayoría se encuentren dentro de los límites superior e inferior
- Si ha ocurrido un cambio significativo en la estructura, se espera que un alto porcentaje de residuos se encuentren fuera de los límites de control

Cabe mencionar en este último caso, si se presenta una gran cantidad de mediciones fuera de los límites de control, no necesariamente representa que la estructura se encuentre dañada, más bien significa que factores operativos y/o climatológicos han ocasionado que la distribución de los datos cambie. Sin embargo, si las mediciones fuera de los límites de control no son debidas a estos factores, la estructura debe ser inspeccionada por posible daño.

La validez del uso de modelos ARIMA en la construcción de los cuadros de control, se basa en la idea que si existe algún daño o cambio en la estructura, el modelo AR ajustado a la serie inicial, no debe ajustarse con los nuevos registros y existirá entonces un porcentaje alto de mediciones (residuales) fuera de los límites de control.

En cuanto a la normalización, se puede decir que simplifica el uso de la distribución normal en el cálculo de los límites de control.

Construcción del gráfico de media muestral \bar{X} : A continuación se detalla la construcción del gráfico de control \bar{X} con los datos pertenecientes a los incrementos de esfuerzos del Puente San Cristóbal en el punto D4-I4, del 1 de abril del 2007 al 18 de diciembre de 2008. Además se utiliza el modelo ARIMA(5,1,0) ajustado a la serie en el capítulo 4, cuyos coeficientes son:

$$\varphi_1 = -0.6024$$

$$\varphi_2 = -0.3129$$

$$\varphi_3 = -0.2151$$

$$\varphi_4 = -0.1958$$

$$\varphi_5 = -0.1102$$

Para comenzar la construcción del gráfico, se recurre a la tabla presentada en el ANEXO 2, donde se muestra el ajuste del modelo ARIMA a las primeras 628 observaciones de la serie.

Producto de la diferenciación de la serie y del ajuste se obtiene un total de 622 residuales. El cálculo de la media y la desviación estándar resulta entonces:

$$\mu = 0.53$$

$$\sigma = 15.007$$

Con estos valores los residuales se normalizan mediante la ecuación:

$$\hat{a}_{tN} = \frac{\hat{a}_t - \mu}{\sigma}$$

Donde:

\hat{a}_{tN} : residual normalizado

A partir de la normalización de los residuales, la nueva media μ y la desviación estándar σ tienen el valor de 0 y 1, respectivamente, por tanto, los límites de control quedan:

$$CL = 0$$

$$UCL = 0 + Z_{\alpha/2} \sigma$$

$$LCL = 0 - Z_{\alpha/2} \sigma$$

En lo que se refiere al valor de $Z_{\alpha/2}$, se probaron distintos valores de α y se observó que un valor de 0.027 es conveniente a fin de probar la eficacia del cuadro de control, que se verá un poco más adelante.

Para un valor $\alpha = 0.027$ y a partir de la distribución normal estándar, se obtiene que el valor de $Z_{\alpha/2}$ es igual a 3.

Una vez establecido este valor, obtenemos:

$$\mu = 0$$

$$\sigma = 1$$

$$Z_{\alpha/2} = 3$$

Por lo tanto:

$$CL = 0$$

$$UCL = 0 + 3(1) = 3$$

$$LCL = 0 - 3(1) = -3$$

El cuadro de control se construye, como se aprecia en la figura siguiente (Figura 6.2).

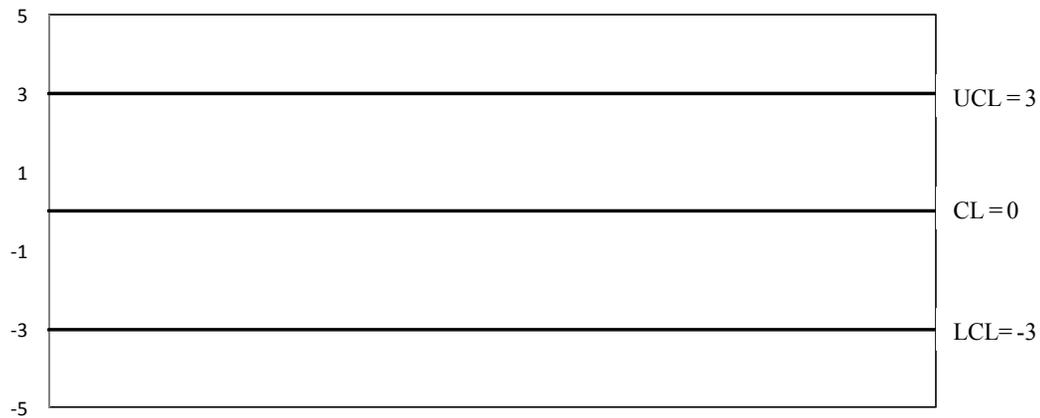


Figura 6.2 Gráfico de control \bar{X} para la Dovela D4-I4 del Puente San Cristóbal

Finalmente se grafican los residuales normalizados en el cuadro de control:

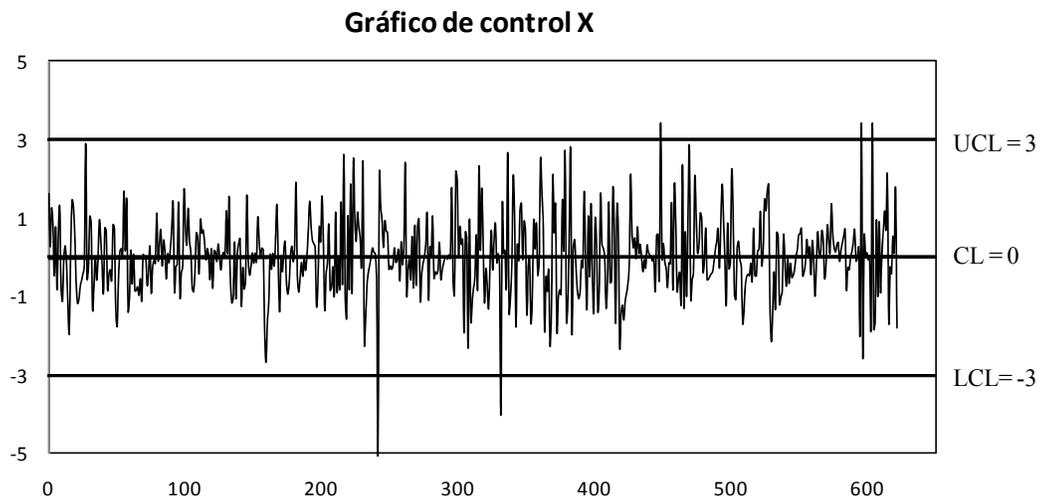


Figura 6.3 Residuales de la Dovela D4-I4 en el Gráfico de control \bar{X}

Se observa que los residuales normalizados del modelo ARIMA, en su mayoría se encuentran dentro de los límites de control. Tan sólo 5 mediciones de un total de 628 quedan fuera de estos límites, lo cual no es indicio de que exista un cambio significativo en el parámetro en estudio (incremento de esfuerzo).

A fin de evitar falsas alarmas (indicios de cambios significativos en el parámetro estructural cuando no en realidad no se han presentado), se comprueba la validez del cuadro por medio de los datos pertenecientes del 19 de diciembre de 2008 al 28 de marzo del 2009. Es decir, los últimos 100 datos capturados. Se utilizan estos datos debido a que con certeza se puede decir que representan el estado saludable de la estructura, al igual que los datos utilizados para la construcción del cuadro de control. Por tanto, al ingresar esta nueva información al gráfico de control se espera que no exista un número

elevado de datos fuera de los límites de control. Por medio de esta prueba se examina la eficacia de los cuadros de control en el monitoreo estructural.

En primer lugar se ajusta la nueva información al mismo modelo ARIMA, establecido en un principio: ARIMA(5,1,0) con los mismo coeficientes calculados:

$$\varphi_1 = -0.6024$$

$$\varphi_2 = -0.3129$$

$$\varphi_3 = -0.2151$$

$$\varphi_4 = -0.1958$$

$$\varphi_5 = -0.1102$$

El ajuste de dicha información se presenta en el **ANEXO 3**. Producto de la diferenciación de esta serie, y del ajuste se obtiene un total de 94 residuales. Estos nuevos residuales se normalizan mediante la media y desviación estándar calculadas para la primera serie:

$$\mu = 0.53$$

$$\sigma = 15.007$$

Se grafican en el cuadro de control construido anteriormente:

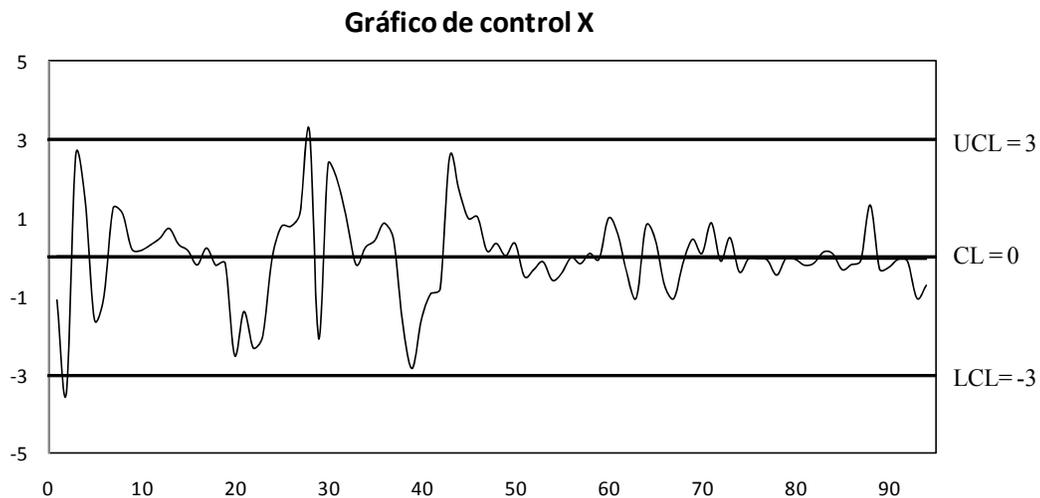


Figura 6.4 Residuales del ajuste del modelo ARIMA(5,1,0) a 100 datos en el Gráfico de control \bar{X}

Se observa en la figura anterior que los nuevos residuales se encuentran dentro de los límites de control establecidos, tan sólo 2 datos de un total de 94 se encuentran fuera de estos límites, lo cual tampoco es indicio que haya ocurrido un cambio significativo en el parámetro estructural. Se comprueba entonces que el cuadro de control no emite falsas alarmas, ya que ambas series pertenecientes al estado saludable de la estructura, se encuentran dentro de los límites permisibles. Se infiere que el cuadro es confiable y puede ser utilizado para el monitoreo estructural.

Finalmente vale la pena destacar un par de comentarios respecto al proceso de control estadístico aplicado al monitoreo estructural. El primero gira entorno al valor de $Z_{\alpha/2}$, el cual en resumidas cuentas representa la rigurosidad del control que se requiera tener sobre algún parámetro estructural en la estructura. Para probar la eficacia de los cuadros de control en el monitoreo estructural, en primera instancia, se establece el valor de $\alpha = 0.027$, es decir, se espera que tan sólo el 2.7% de los residuales se encuentren fuera de los límites de control y que esta situación no representa un cambio significativo en la estructura (Fugate, Sohn y Farrar 2001). Sin embargo este límite podría ser un tanto riguroso considerando que actualmente los esfuerzos actuantes en el punto D4-I4 del puente, son aproximadamente el 30% de los esfuerzos permisibles. El esfuerzo permisible de diseño para el puente San Cristóbal es de 1936 kg/cm². Por esta razón podría ser conveniente establecer límites de control menos conservadores que contemplen el inevitable incremento de esfuerzos actuantes sobre la estructura debido a la fatiga de los materiales y al desgaste natural de la estructura y que por supuesto estos que no rebasen los esfuerzos permisibles.

6.2 Algoritmo clasificador

Un algoritmo es una secuencia de actividades o pasos que conducen a la realización de una tarea o acción determinada. El clasificador que se presenta a continuación es capaz de identificar cambios significativos en algún parámetro estructural.

Esta herramienta puede complementar al proceso de control estadístico en el monitoreo de la salud estructural basándose en la idea que si existe algún cambio en el sistema estructural, entonces también existirá un cambio en los coeficientes del modelo ARIMA. Así, es lógico pensar que una manera para conocer el estado de una estructura es mediante la clasificación de los coeficientes ARIMA.

Sin embargo no cualquier cambio en el valor de los coeficientes del modelo representará un cambio en el sistema estructural. Esto se debe a que los coeficientes de un modelo tienen cierta variabilidad natural, debida principalmente a la longitud de la serie. La variabilidad de los coeficientes es menor para series extensas y mientras la serie tienda a infinito, el valor de los coeficientes calculados convergerá en los valores reales del proceso.

Es difícil deducir expresiones exactas para determinar la variación natural de los coeficientes ARIMA. No obstante, Box y Jenkins demostraron que el intervalo de confianza de estos coeficientes está dado por la suma de los cuadrados de los residuales:

$$S(\beta) < S(\hat{\beta}) \left[1 + \frac{\chi_{\alpha}^2(k)}{n} \right]$$

Donde:

$S(\cdot)$ es la suma de los cuadrados de los residuales

β son el conjunto de coeficientes calculados a partir de una serie de longitud n

$\hat{\beta}$ son los coeficientes iniciales conocidos previamente (punto de comparación)

$\chi_{\alpha}^2(k)$ es el valor de la variable chi – cuadrado con k grados de libertad que deja a su derecha un área (probabilidad) de valor α

La desigualdad anterior permite la comparación en conjunto de todos los coeficientes ARIMA simultáneamente. Si se considera que la serie de tiempo se genera a partir de un proceso con los verdaderos coeficientes o parámetros ARIMA $\hat{\beta}$, en un estado inicial donde la estructura se encuentra en óptimas condiciones, entonces la ecuación anterior prueba si los parámetros calculados β son estadísticamente similares o diferentes a $\hat{\beta}$, en otras palabras, saludable o no saludable.

Es relevante mencionar que la desigualdad descrita funciona únicamente para series de igual número de datos n . No obstante, es práctico obtener una expresión que además sirva para la comparación entre series de diferente longitud, por ello se propone la siguiente expresión:

$$\frac{S(\beta)}{n_2} < \frac{S(\hat{\beta})}{n_1} \left[1 + \frac{\chi_{\alpha}^2(k)}{n} \right]$$

Donde:

n_1 es el número de residuales de la serie inicial

n_2 es el número de residuales de la serie a comparar

El valor de alfa se establece como $\alpha = 0.01$ y k es igual al orden del modelo ARIMA

A partir de las desigualdades presentadas, se describe a continuación el algoritmo clasificador para el monitoreo estructural (Brownjohn y Carden 2007):

Paso 1. Estimar un modelo ARIMA de orden (p,d,q) y sus correspondientes coeficientes β para la primer serie de tiempo

Paso 2. Establecer la clase 1 (inicial) con los coeficientes β , así: $\beta = \hat{\beta}_1$

Paso 3. Calcular $S(\hat{\beta}_c)$, $c = 1, 2, \dots, m$ en donde m es el número de clases formadas previamente por el algoritmo

Paso 4. Ajustar una nueva serie de tiempo al modelo ARIMA del mismo orden (p,d,q) ajustado en el paso 1, se obtiene así β

Paso 5. Calcular $S(\beta)$

Paso 6. Comparar $S(\beta)$ y $S(\hat{\beta}_c)$ mediante alguna de las desigualdades y determinar si esta permanece

Paso 7. Si la desigualdad permanece, clasificar la serie de tiempo en clase c

Paso 8. Si la desigualdad no permanece formar una nueva clase $c = m + 1$ y β calculado en el paso 4 pasa a ser $\hat{\beta}_{m+1}$, es decir: $\beta = \hat{\beta}_{m+1}$

Paso 9. Repetir pasos 3 al 8 para nuevas series de tiempo

Bajo este algoritmo, si la respuesta estructural presenta algún cambio importante debido a cargas extraordinarias o inclusive debido a cambios climatológicos, el algoritmo debiera generar diversas

clases. Para fines del monitoreo estructural es conveniente separar los efectos ambientales tales como la temperatura. Para este objetivo, se recomienda la creación de una base de datos inicial en la cual se incluyan diversas condiciones ambientales actuando sobre la estructura. A partir de esta base de datos se compara las nuevas series de tiempo capturadas.

En el uso del algoritmo se advierte que en un principio se pueden crear diversas clases aun cuando la información pertenezca a un mismo estado saludable de la estructura. Esto se debe a la longitud de la serie de tiempo inicial, hay que recordar que en series cortas, existe una mayor variabilidad de los coeficientes del modelo ARIMA.

Con motivo de probar la efectividad del algoritmo presentado, se utiliza nuevamente la información del monitoreo del Puente San Cristóbal: los incrementos de esfuerzos de la dovela D4 punto Inferior 4 (D4-I4). Los 728 datos pertenecientes del 1 de abril del 2007 al 28 de marzo de 2009 se dividen en dos partes: una primera serie de tiempo con 628 datos y una segunda serie con los 100 datos restantes.

Al aplicar el algoritmo a ambas series de tiempo pertenecientes al mismo estado saludable de la estructura, se espera que no se genere una clase nueva, es decir, que ambas series sean clase 1.

Paso 1. Para comenzar, se utiliza el modelo ARIMA(5,1,0) ajustado en el capítulo 4 a la primera serie de tiempo (628 datos), se utiliza este modelo debido a la facilidad que implica el manejo de coeficientes autorregresivos.

El conjunto de coeficientes β son:

$$\varphi_1 = -0.6024$$

$$\varphi_2 = -0.3129$$

$$\varphi_3 = -0.2151$$

$$\varphi_4 = -0.1958$$

$$\varphi_5 = -0.1102$$

Paso 2. Se establece $\beta = \hat{\beta}_1$, es decir, la clase 1.

Paso 3. Se calcula $S(\hat{\beta}_c)$ donde $c = 1$, ya que se ha formado hasta el momento una sola clase. De acuerdo a la información del **ANEXO 2**, la suma de los cuadrados de los residuos es:

$$S(\hat{\beta}_c) = S(\hat{\beta}_1) = 140027.26$$

Paso 4. Se ajusta la segunda serie de tiempo (100 datos) a un modelo del mismo orden, es decir, ARIMA(5,1,0). El conjunto de coeficientes β resultan:

$$\varphi_1 = -0.3201$$

$$\varphi_2 = -0.3425$$

$$\varphi_3 = -0.0891$$

$$\varphi_4 = 0.1842$$

$$\varphi_5 = -0.1349$$

Paso 5. Se calcula $S(\beta)$. En el ANEXO 4 se presenta los residuos del modelo ajustado.

$$S(\beta) = 21624.48$$

Paso 6. Se comparan $S(\beta)$ y $S(\hat{\beta}_c)$ mediante la desigualdad:

$$\frac{S(\beta)}{n_2} < \frac{S(\hat{\beta}_c)}{n_1} \left[1 + \frac{\chi_{\alpha}^2(k)}{n} \right]$$

En este caso $n = 628$ datos iniciales; $n_1 = 622$ y $n_2 = 94$ ya que es el número de residuos resultantes de la diferenciación a la serie y el ajuste del modelo ARIMA(5,1,0).

$$\frac{21624.48}{94} < \frac{140027.26}{622} \left[1 + \frac{15.0863}{628} \right]$$

$$230.048 < 230.532$$

Paso 7. La desigualdad permanece, por lo tanto se clasifica la serie de tiempo 2 en clase 1

Finalmente la *Tabla 6.1* muestra un resumen del algoritmo desarrollado.

		Coeficientes ARIMA Clase		
Serie 1	Incremento de esfuerzos (1 Abr 2007-18 Dic 2008)	β_1	$\varphi_1 = -0.6024$	1
			$\varphi_2 = -0.3129$	
			$\varphi_3 = -0.2151$	
			$\varphi_4 = -0.1958$	
			$\varphi_5 = -0.1102$	
Serie 2	Incremento de esfuerzos (19 Dic 2008-28 Mar 2009)	β	$\varphi_1 = -0.3201$	1
			$\varphi_2 = -0.3425$	
			$\varphi_3 = -0.0891$	
			$\varphi_4 = 0.1842$	
			$\varphi_5 = -0.1349$	

Tabla 6.2 Resumen algoritmo clasificador aplicado a dos series de tiempo

Las dos series de tiempo pertenecientes al mismo estado saludable de la estructura, se catalogaron en la misma clase 1, como se esperaba que sucediera. Así, el algoritmo muestra consistencia y representa una opción viable en el monitoreo de la salud estructural.

7 Conclusiones

El desarrollo de este trabajo escrito y la experiencia ganada a través del proceso de búsqueda de información, aplicación de metodologías e interpretación de resultados trajo consigo el surgimiento de muchas reflexiones y recomendaciones acerca de la implementación de sistemas de monitoreo estructural en la actualidad, específicamente en nuestro país. En primera instancia, se debe impulsar la aplicación de sistemas de monitoreo estructural tanto en el sector público como en el sector privado, ya que frecuentemente se presta poca atención a observar la evolución de las obras civiles en su vida útil. La mayor parte de los recursos humanos y económicos se centran en la parte joven de un proyecto (planeación y construcción), olvidándose de la vida de servicio, hasta que la fatiga de los materiales y envejecimiento paulatino suponen un peligro inminente para la población así como para las actividades económicas y sociales que dependen de esa infraestructura.

A pesar de que en un principio puede ser poco atractivo la implementación de sistemas de monitoreo, es un hecho que en fechas recientes el monitoreo de la salud estructural se abre paso con gran auge debido a los beneficios que reporta. El principal sin duda, son los niveles de seguridad. En segundo lugar el beneficio económico que se obtiene. Una anticipada estrategia y acciones correctivas ante cambios estructurales detectados en temprana edad reducen de manera importante los costos de reparación.

Ahora bien, en cuanto al manejo de la información, se debe continuar en la búsqueda y perfeccionamiento de metodologías que maximicen los beneficios del monitoreo estructural. Esta búsqueda debe estar enfocada al manejo fluido y eficaz de grandes volúmenes de información, así como a la automatización de las actividades que conlleva. Aprovechar las tecnologías presentes, especialmente el potencial de sistemas de cómputo, puede representar un ahorro significativo de esfuerzos y tiempo al analista. Además de ayudar a la comparación e intercambio de información de distintos proyectos de forma rápida y sencilla, inclusive a distancias lejanas.

Profundizar en conocimientos relacionados con sistemas estructurales, los procesos físicos subyacentes a los cambios que se presentan en el sistema y las relaciones existentes entre ellos, juega un papel importante en el desarrollo o perfeccionamiento de los métodos de análisis. También es de gran utilidad la investigación profunda en este campo de la ingeniería civil, para adoptar medidas estratégicas de monitoreo hechas a la medida de las necesidades propias de cada obra y organismo ejecutor (ya sea público o privado). Estas últimas elaboradas a partir de las posibilidades económicas y los estándares de seguridad que deban cumplirse en cada obra.

La formación de técnicas de análisis innovadoras y revolucionarias debe ser un revulsivo en la aplicación de sistemas de monitoreo. La base de los beneficios futuros se encuentra precisamente en la optimización de los procesos de análisis y en la capacidad de observar a través de los datos, patrones que indiquen cambios significativos. Sería ideal no solo aplicar los avances probados en países donde se ha logrado una mayor penetración en la materia, si no también generar propuestas propias derivadas de la gran experiencia de los analistas adquirida en el paso del tiempo.

No obstante, se advierte no dejar de lado los métodos actuales de análisis y mucho menos la inspección visual en la obra civil. Estas labores no deben por ningún motivo sustituirse.

En cuanto a la aplicación de la metodología clásica y ARIMA a información del Puente San Cristóbal, debo mencionar la razón por la cual en este trabajo de tesis se estudió 1 solo punto de los 102 totales medidos con las cuerdas vibrantes. Aunque podría parecer poco para un trabajo de final de carrera, se debían establecer las bases y consideraciones de ambas metodologías (clásico y ARIMA), las cuales no existían previamente, particularmente en su aplicación en el monitoreo estructural. Por tanto, la dificultad en la delimitación de los alcances del tema supuso un reto importante, ya que se buscó plasmar de forma clara y entendible la teoría elemental de las metodologías expuestas. Además la forma de operar sobre los demás puntos es el mismo que para el punto D4-I4 desarrollado en capítulos 5 y 6. Analizar más puntos de la estructura sin concluir en una o varias aplicaciones que reflejen un análisis global de la estructura (en este caso el puente San Cristóbal), resultaría redundante en el desarrollo de la tesis.

También es importante mencionar que la propuesta presentada, sienta apenas las primeras bases del amplio potencial del enfoque de series de tiempo aplicado en el monitoreo estructural.

Acercas de los resultados obtenidos, se concluye que los modelos ARIMA proveen un excepcional ajuste a cualquier serie de tiempo. Puntualmente los 3 ajustes a la serie en estudio (incremento de esfuerzos en la Dovela 4 punto Inferior 4) son muy buenos. El grado de ajuste de los modelos no se hubiera alcanzado con un análisis clásico, en el cual la única componente detectable y modelada fue la tendencia.

Pese a ello, aun existen mejoras que pueden hacerse en la metodología, especialmente en lo que refiere a la identificación de la estacionalidad en la serie. En la aplicación de la metodología Box-Jenkins, se tuvo la dificultad de definir un periodo estacional, al final inexistente, pero que generó incertidumbre en el correcto análisis de la información. Es necesario hallar una herramienta más eficaz para la adecuada identificación de esta componente y la obtención del periodo correspondiente. Se debe tener

precaución en la identificación de esta componente, a fin de evitar crear una estacionalidad artificial que resulta por demás inapropiada.

En lo que se refiere a los pronósticos obtenidos, las estimaciones de valores futuros parecen no ser tan atractivos debido a que, a largo plazo, convergen a un valor constante. Un modelo puramente de media móvil (MA) desde un principio produce valores constantes. Un modelo puramente autorregresivo (AR) y un modelo combinado (ARMA y/o ARIMA), producen pronósticos no constantes en un principio. No obstante, los valores pronosticados por ambos modelos convergen en una línea recta a largo plazo. Con ello se confirma que los modelos ARIMA son adecuados para pronósticos a corto plazo. Los modelos ARIMA muestran efectividad en 20 o 30 periodos de tiempo posteriores a la última observación. Este periodo es bueno si se consideran datos mensuales, puesto que se tendría de entre 20 o 30 meses de buenas estimaciones futuras. No sucede de la misma forma si se usa información diaria, en donde se tendría 20 o 30 días de valores predichos. No obstante si se desea trabajar con datos mensuales se necesita una extensa base de datos para describir adecuadamente el proceso. Por su cuenta, los intervalos de confianza muestran proveer un rango confiable de posibles valores futuros. Se puede tener la certeza que los valores futuros se encontrarán en este rango.

Finalmente las herramientas para la detección de cambios significativos en los parámetros estructurales, presentadas en el capítulo 6, mostraron efectividad y consistencia en los resultados arrojados. El control estadístico de procesos a través del cuadro de control es una opción viable y sencilla para el monitoreo rápido de las condiciones de una estructura. Una vez construido el cuadro, nueva información puede ser analizada en un corto periodo de tiempo. Por su parte el algoritmo clasificador complementa al proceso de control estadístico, al separar por clases las diferentes condiciones de una estructura.

Las energías futuras deben enfocarse en la automatización de ambas herramientas y en la definición de límites razonables en los cambios que indudablemente ocurrirán en el Puente San Cristóbal. A partir de los límites se podrán establecer medidas preventivas que asegurarán las condiciones de seguridad necesarias en este tipo de infraestructura.

Esta última tarea puede representar un reto importante ya que el monitoreo permanente instalado en el Puente San Cristóbal fue parcialmente robado y por esta razón no existen registros posteriores a marzo de 2009. Este último, sin duda, es un aspecto que también debe ser tomado en cuenta en la instalación de futuros sistemas de monitoreo en otras obras de infraestructura o el Puente San Cristóbal.

Anexo 1

Incrementos de esfuerzos Dovela 4 – Inferior 4 (D4-I4)

Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)
01/04/2007	1	230.034	17/05/2007	47	217.917	02/07/2007	93	183.540	17/08/2007	139	266.112
02/04/2007	2	218.253	18/05/2007	48	235.914	03/07/2007	94	190.785	18/08/2007	140	240.765
03/04/2007	3	211.596	19/05/2007	49	239.673	04/07/2007	95	187.383	19/08/2007	141	230.202
04/04/2007	4	206.745	20/05/2007	50	224.049	05/07/2007	96	192.150	20/08/2007	142	224.049
05/04/2007	5	203.847	21/05/2007	51	225.330	06/07/2007	97	202.566	21/08/2007	143	239.316
06/04/2007	6	221.844	22/05/2007	52	222.096	07/07/2007	98	216.384	22/08/2007	144	221.235
07/04/2007	7	240.345	23/05/2007	53	215.943	08/07/2007	99	191.205	23/08/2007	145	234.885
08/04/2007	8	230.370	24/05/2007	54	235.914	09/07/2007	100	196.077	24/08/2007	146	237.027
09/04/2007	9	247.002	25/05/2007	55	239.589	10/07/2007	101	193.179	25/08/2007	147	240.513
10/04/2007	10	246.141	26/05/2007	56	213.129	11/07/2007	102	216.384	26/08/2007	148	218.337
11/04/2007	11	231.315	27/05/2007	57	199.059	12/07/2007	103	191.205	27/08/2007	149	225.855
12/04/2007	12	248.871	28/05/2007	58	203.763	13/07/2007	104	196.077	28/08/2007	150	213.822
13/04/2007	13	228.921	29/05/2007	59	210.756	14/07/2007	105	193.179	29/08/2007	151	213.822
14/04/2007	14	247.086	30/05/2007	60	216.552	15/07/2007	106	220.983	30/08/2007	152	244.104
15/04/2007	15	261.156	31/05/2007	61	216.552	16/07/2007	107	218.400	31/08/2007	153	228.921
16/04/2007	16	237.615	01/06/2007	62	239.589	17/07/2007	108	218.862	01/09/2007	154	223.713
17/04/2007	17	229.005	02/06/2007	63	225.414	18/07/2007	109	232.575	02/09/2007	155	223.125
18/04/2007	18	237.447	03/06/2007	64	247.506	19/07/2007	110	226.191	03/09/2007	156	224.490
19/04/2007	19	240.093	04/06/2007	65	212.877	20/07/2007	111	225.330	04/09/2007	157	222.768
20/04/2007	20	238.140	05/06/2007	66	209.643	21/07/2007	112	213.570	05/09/2007	158	228.144
21/04/2007	21	222.516	06/06/2007	67	220.983	22/07/2007	113	206.745	06/09/2007	159	224.574
22/04/2007	22	202.902	07/06/2007	68	209.979	23/07/2007	114	209.475	07/09/2007	160	241.038
23/04/2007	23	222.348	08/06/2007	69	219.702	24/07/2007	115	223.041	08/09/2007	161	237.951
24/04/2007	24	242.403	09/06/2007	70	206.829	25/07/2007	116	225.582	09/09/2007	162	235.410
25/04/2007	25	251.685	10/06/2007	71	199.752	26/07/2007	117	215.943	10/09/2007	163	238.392
26/04/2007	26	250.152	11/06/2007	72	195.909	27/07/2007	118	232.848	11/09/2007	164	238.392
27/04/2007	27	232.764	12/06/2007	73	188.391	28/07/2007	119	231.903	12/09/2007	165	209.559
28/04/2007	28	218.505	13/06/2007	74	189.252	29/07/2007	120	238.035	13/09/2007	166	187.299
29/04/2007	29	213.045	14/06/2007	75	178.332	30/07/2007	121	238.728	14/09/2007	167	184.905
30/04/2007	30	213.045	15/06/2007	76	189.168	31/07/2007	122	233.772	15/09/2007	168	181.650
01/05/2007	31	213.045	16/06/2007	77	188.391	01/08/2007	123	237.279	16/09/2007	169	195.993
02/05/2007	32	213.045	17/06/2007	78	189.252	02/08/2007	124	236.334	17/09/2007	170	192.234
03/05/2007	33	213.045	18/06/2007	79	178.332	03/08/2007	125	227.892	18/09/2007	171	195.489
04/05/2007	34	257.586	19/06/2007	80	178.500	04/08/2007	126	237.111	19/09/2007	172	189.588
05/05/2007	35	223.041	20/06/2007	81	185.241	05/08/2007	127	222.768	20/09/2007	173	197.442
06/05/2007	36	226.443	21/06/2007	82	170.058	06/08/2007	128	233.100	21/09/2007	174	213.570
07/05/2007	37	241.710	22/06/2007	83	176.106	07/08/2007	129	233.688	22/09/2007	175	199.752
08/05/2007	38	243.579	23/06/2007	84	176.106	08/08/2007	130	231.126	23/09/2007	176	181.650
09/05/2007	39	219.618	24/06/2007	85	174.237	09/08/2007	131	237.531	24/09/2007	177	193.011
10/05/2007	40	222.264	25/06/2007	86	193.599	10/08/2007	132	234.465	25/09/2007	178	185.661
11/05/2007	41	225.162	26/06/2007	87	182.259	11/08/2007	133	229.173	26/09/2007	179	193.767
12/05/2007	42	217.308	27/06/2007	88	188.475	12/08/2007	134	233.352	27/09/2007	180	199.920
13/05/2007	43	224.826	28/06/2007	89	195.300	13/08/2007	135	234.465	28/09/2007	181	185.493
14/05/2007	44	238.896	29/06/2007	90	185.157	14/08/2007	136	234.465	29/09/2007	182	176.967
15/05/2007	45	236.418	30/06/2007	91	194.796	15/08/2007	137	252.798	30/09/2007	183	175.266
16/05/2007	46	230.622	01/07/2007	92	193.515	16/08/2007	138	243.075	01/10/2007	184	181.566

Incrementos de esfuerzos Dovela 4-Inferior 4 (D4-I4) [Anexo 1]

Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)
02/10/2007	185	186.858	17/11/2007	231	232.407	02/01/2008	277	251.685	17/02/2008	323	257.586
03/10/2007	186	184.044	18/11/2007	232	231.399	03/01/2008	278	258.510	18/02/2008	324	284.109
04/10/2007	187	188.664	19/11/2007	233	244.188	04/01/2008	279	234.549	19/02/2008	325	273.189
05/10/2007	188	213.570	20/11/2007	234	245.133	05/01/2008	280	238.476	20/02/2008	326	252.294
06/10/2007	189	188.475	21/11/2007	235	245.133	06/01/2008	281	238.476	21/02/2008	327	257.838
07/10/2007	190	181.839	22/11/2007	236	240.513	07/01/2008	282	239.841	22/02/2008	328	249.984
08/10/2007	191	180.285	23/11/2007	237	277.620	08/01/2008	283	247.674	23/02/2008	329	237.363
09/10/2007	192	181.755	24/11/2007	238	222.516	09/01/2008	284	261.408	24/02/2008	330	248.619
10/10/2007	193	178.080	25/11/2007	239	222.516	10/01/2008	285	234.045	25/02/2008	331	259.455
11/10/2007	194	183.960	26/11/2007	240	222.516	11/01/2008	286	248.451	26/02/2008	332	248.619
12/10/2007	195	177.660	27/11/2007	241	222.516	12/01/2008	287	244.188	27/02/2008	333	259.455
13/10/2007	196	181.398	28/11/2007	242	226.611	13/01/2008	288	260.736	28/02/2008	334	266.532
14/10/2007	197	191.394	29/11/2007	243	231.042	14/01/2008	289	245.301	29/02/2008	335	256.389
15/10/2007	198	206.661	30/11/2007	244	231.042	15/01/2008	290	245.301	01/03/2008	336	260.652
16/10/2007	199	206.661	01/12/2007	245	231.042	16/01/2008	291	239.757	02/03/2008	337	256.389
17/10/2007	200	206.661	02/12/2007	246	231.042	17/01/2008	292	239.232	03/03/2008	338	197.274
18/10/2007	201	206.661	03/12/2007	247	222.012	18/01/2008	293	248.619	04/03/2008	339	254.331
19/10/2007	202	206.661	04/12/2007	248	151.557	19/01/2008	294	238.140	05/03/2008	340	241.626
20/10/2007	203	186.606	05/12/2007	249	229.257	20/01/2008	295	238.140	06/03/2008	341	247.506
21/10/2007	204	188.916	06/12/2007	250	226.275	21/01/2008	296	238.140	07/03/2008	342	236.943
22/10/2007	205	205.632	07/12/2007	251	235.221	22/01/2008	297	238.560	08/03/2008	343	280.014
23/10/2007	206	205.632	08/12/2007	252	233.016	23/01/2008	298	239.757	09/03/2008	344	231.399
24/10/2007	207	227.724	09/12/2007	253	239.064	24/01/2008	299	239.757	10/03/2008	345	234.885
25/10/2007	208	219.282	10/12/2007	254	236.670	25/01/2008	300	240.597	11/03/2008	346	237.195
26/10/2007	209	194.019	11/12/2007	255	245.133	26/01/2008	301	239.757	12/03/2008	347	269.514
27/10/2007	210	199.563	12/12/2007	256	234.801	27/01/2008	302	266.784	13/03/2008	348	269.514
28/10/2007	211	199.563	13/12/2007	257	235.053	28/01/2008	303	242.403	14/03/2008	349	237.363
29/10/2007	212	215.523	14/12/2007	258	231.903	29/01/2008	304	234.969	15/03/2008	350	253.911
30/10/2007	213	198.387	15/12/2007	259	231.903	30/01/2008	305	274.386	16/03/2008	351	246.582
31/10/2007	214	194.460	16/12/2007	260	237.951	31/01/2008	306	282.912	17/03/2008	352	267.897
01/11/2007	215	194.880	17/12/2007	261	234.633	01/02/2008	307	270.207	18/03/2008	353	281.043
02/11/2007	216	182.175	18/12/2007	262	234.633	02/02/2008	308	283.080	19/03/2008	354	267.897
03/11/2007	217	209.811	19/12/2007	263	237.951	03/02/2008	309	275.919	20/03/2008	355	281.043
04/11/2007	218	186.942	20/12/2007	264	226.863	04/02/2008	310	280.098	21/03/2008	356	281.043
05/11/2007	219	194.880	21/12/2007	265	238.728	05/02/2008	311	250.068	22/03/2008	357	253.218
06/11/2007	220	182.175	22/12/2007	266	232.239	06/02/2008	312	277.284	23/03/2008	358	264.999
07/11/2007	221	209.811	23/12/2007	267	238.896	07/02/2008	313	277.284	24/03/2008	359	265.860
08/11/2007	222	186.942	24/12/2007	268	272.685	08/02/2008	314	240.765	25/03/2008	360	241.374
09/11/2007	223	235.578	25/12/2007	269	236.775	09/02/2008	315	277.284	26/03/2008	361	244.188
10/11/2007	224	194.880	26/12/2007	270	241.878	10/02/2008	316	240.765	27/03/2008	362	265.251
11/11/2007	225	182.175	27/12/2007	271	241.710	11/02/2008	317	239.589	28/03/2008	363	259.119
12/11/2007	226	209.811	28/12/2007	272	232.491	12/02/2008	318	239.841	29/03/2008	364	273.777
13/11/2007	227	186.942	29/12/2007	273	248.703	13/02/2008	319	237.447	30/03/2008	365	258.006
14/11/2007	228	225.855	30/12/2007	274	233.436	14/02/2008	320	251.181	31/03/2008	366	258.006
15/11/2007	229	197.106	31/12/2007	275	251.685	15/02/2008	321	235.998	01/04/2008	367	296.478
16/11/2007	230	240.870	01/01/2008	276	233.436	16/02/2008	322	276.780	02/04/2008	368	299.376

Incrementos de esfuerzos Dovela 4-Inferior 4 (D4-I4) [Anexo 1]

Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)
03/04/2008	369	299.964	19/05/2008	415	239.148	04/07/2008	461	255.780	19/08/2008	507	358.575
04/04/2008	370	264.222	20/05/2008	416	241.290	05/07/2008	462	252.714	20/08/2008	508	354.564
05/04/2008	371	270.123	21/05/2008	417	269.346	06/07/2008	463	276.696	21/08/2008	509	331.443
06/04/2008	372	260.127	22/05/2008	418	247.338	07/07/2008	464	251.433	22/08/2008	510	323.862
07/04/2008	373	266.784	23/05/2008	419	247.338	08/07/2008	465	287.007	23/08/2008	511	333.249
08/04/2008	374	237.951	24/05/2008	420	277.200	09/07/2008	466	282.492	24/08/2008	512	338.436
09/04/2008	375	237.027	25/05/2008	421	269.955	10/07/2008	467	262.689	25/08/2008	513	338.436
10/04/2008	376	277.956	26/05/2008	422	241.290	11/07/2008	468	262.689	26/08/2008	514	335.895
11/04/2008	377	269.430	27/05/2008	423	277.200	12/07/2008	469	260.043	27/08/2008	515	310.632
12/04/2008	378	287.091	28/05/2008	424	269.955	13/07/2008	470	245.973	28/08/2008	516	304.752
13/04/2008	379	246.141	29/05/2008	425	233.520	14/07/2008	471	295.533	29/08/2008	517	306.117
14/04/2008	380	243.936	30/05/2008	426	235.305	15/07/2008	472	253.995	30/08/2008	518	306.117
15/04/2008	381	239.316	31/05/2008	427	225.330	16/07/2008	473	277.200	31/08/2008	519	306.117
16/04/2008	382	235.305	01/06/2008	428	212.457	17/07/2008	474	253.995	01/09/2008	520	303.135
17/04/2008	383	268.065	02/06/2008	429	214.410	18/07/2008	475	277.200	02/09/2008	521	303.387
18/04/2008	384	258.510	03/06/2008	430	209.643	19/07/2008	476	310.968	03/09/2008	522	295.281
19/04/2008	385	297.066	04/06/2008	431	206.577	20/07/2008	477	272.853	04/09/2008	523	312.249
20/04/2008	386	247.002	05/06/2008	432	210.672	21/07/2008	478	274.050	05/09/2008	524	298.515
21/04/2008	387	244.188	06/06/2008	433	243.327	22/07/2008	479	270.459	06/09/2008	525	296.730
22/04/2008	388	258.510	07/06/2008	434	233.436	23/07/2008	480	302.526	07/09/2008	526	302.694
23/04/2008	389	297.066	08/06/2008	435	233.436	24/07/2008	481	303.555	08/09/2008	527	303.807
24/04/2008	390	247.002	09/06/2008	436	236.775	25/07/2008	482	299.712	09/09/2008	528	320.439
25/04/2008	391	271.908	10/06/2008	437	229.950	26/07/2008	483	299.712	10/09/2008	529	307.986
26/04/2008	392	269.010	11/06/2008	438	243.411	27/07/2008	484	312.165	11/09/2008	530	311.325
27/04/2008	393	261.156	12/06/2008	439	239.757	28/07/2008	485	316.869	12/09/2008	531	331.191
28/04/2008	394	254.688	13/06/2008	440	243.936	29/07/2008	486	305.697	13/09/2008	532	335.706
29/04/2008	395	246.834	14/06/2008	441	240.681	30/07/2008	487	307.482	14/09/2008	533	352.359
30/04/2008	396	240.429	15/06/2008	442	235.137	31/07/2008	488	317.982	15/09/2008	534	365.652
01/05/2008	397	249.228	16/06/2008	443	237.363	01/08/2008	489	302.967	16/09/2008	535	330.078
02/05/2008	398	246.414	17/06/2008	444	234.297	02/08/2008	490	302.022	17/09/2008	536	308.763
03/05/2008	399	274.638	18/06/2008	445	241.794	03/08/2008	491	298.956	18/09/2008	537	305.340
04/05/2008	400	256.725	19/06/2008	446	241.374	04/08/2008	492	296.142	19/09/2008	538	312.081
05/05/2008	401	238.980	20/06/2008	447	241.458	05/08/2008	493	295.617	20/09/2008	539	299.712
06/05/2008	402	264.915	21/06/2008	448	240.765	06/08/2008	494	297.675	21/09/2008	540	323.694
07/05/2008	403	246.414	22/06/2008	449	239.064	07/08/2008	495	301.329	22/09/2008	541	323.001
08/05/2008	404	274.638	23/06/2008	450	239.757	08/08/2008	496	305.865	23/09/2008	542	300.069
09/05/2008	405	256.725	24/06/2008	451	227.304	09/08/2008	497	313.110	24/09/2008	543	295.701
10/05/2008	406	238.980	25/06/2008	452	243.852	10/08/2008	498	302.442	25/09/2008	544	289.737
11/05/2008	407	264.915	26/06/2008	453	247.254	11/08/2008	499	317.793	26/09/2008	545	306.201
12/05/2008	408	261.408	27/06/2008	454	235.305	12/08/2008	500	337.428	27/09/2008	546	301.938
13/05/2008	409	239.148	28/06/2008	455	292.131	13/08/2008	501	336.315	28/09/2008	547	293.328
14/05/2008	410	241.290	29/06/2008	456	258.951	14/08/2008	502	328.041	29/09/2008	548	298.347
15/05/2008	411	269.346	30/06/2008	457	266.868	15/08/2008	503	309.351	30/09/2008	549	286.755
16/05/2008	412	259.707	01/07/2008	458	260.736	16/08/2008	504	331.191	01/10/2008	550	291.102
17/05/2008	413	264.915	02/07/2008	459	255.360	17/08/2008	505	319.683	02/10/2008	551	285.894
18/05/2008	414	261.408	03/07/2008	460	247.863	18/08/2008	506	323.862	03/10/2008	552	283.605

Incrementos de esfuerzos Dovela 4-Inferior 4 (D4-I4) [Anexo 1]

Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)	Fecha	No.	Δ Esf (kg/cm ²)
04/10/2008	553	282.912	19/11/2008	599	338.793	04/01/2009	645	371.028	19/02/2009	691	382.200
05/10/2008	554	286.146	20/11/2008	600	342.111	05/01/2009	646	375.375	20/02/2009	692	385.959
06/10/2008	555	288.624	21/11/2008	601	310.884	06/01/2009	647	383.565	21/02/2009	693	384.174
07/10/2008	556	296.142	22/11/2008	602	381.171	07/01/2009	648	382.536	22/02/2009	694	399.189
08/10/2008	557	299.964	23/11/2008	603	310.884	08/01/2009	649	382.536	23/02/2009	695	399.189
09/10/2008	558	305.613	24/11/2008	604	345.786	09/01/2009	650	377.601	24/02/2009	696	388.185
10/10/2008	559	291.879	25/11/2008	605	338.268	10/01/2009	651	382.536	25/02/2009	697	376.488
11/10/2008	560	290.598	26/11/2008	606	338.793	11/01/2009	652	377.601	26/02/2009	698	396.459
12/10/2008	561	291.186	27/11/2008	607	338.793	12/01/2009	653	378.189	27/02/2009	699	396.018
13/10/2008	562	299.880	28/11/2008	608	342.111	13/01/2009	654	342.027	28/02/2009	700	384.258
14/10/2008	563	297.339	29/11/2008	609	310.884	14/01/2009	655	343.812	01/03/2009	701	375.207
15/10/2008	564	302.022	30/11/2008	610	381.171	15/01/2009	656	319.935	02/03/2009	702	380.247
16/10/2008	565	291.879	01/12/2008	611	320.712	16/01/2009	657	311.829	03/03/2009	703	387.660
17/10/2008	566	302.022	02/12/2008	612	316.512	17/01/2009	658	330.435	04/03/2009	704	387.660
18/10/2008	567	296.982	03/12/2008	613	343.308	18/01/2009	659	342.531	05/03/2009	705	400.806
19/10/2008	568	283.857	04/12/2008	614	316.512	19/01/2009	660	347.739	06/03/2009	706	390.138
20/10/2008	569	290.850	05/12/2008	615	343.308	20/01/2009	661	358.071	07/03/2009	707	398.244
21/10/2008	570	301.161	06/12/2008	616	324.198	21/01/2009	662	393.729	08/03/2009	708	387.744
22/10/2008	571	297.066	07/12/2008	617	331.527	22/01/2009	663	332.556	09/03/2009	709	391.083
23/10/2008	572	299.460	08/12/2008	618	339.717	23/01/2009	664	389.550	10/03/2009	710	391.083
24/10/2008	573	307.818	09/12/2008	619	352.443	24/01/2009	665	393.729	11/03/2009	711	391.083
25/10/2008	574	306.978	10/12/2008	620	352.443	25/01/2009	666	392.952	12/03/2009	712	385.014
26/10/2008	575	296.310	11/12/2008	621	379.134	26/01/2009	667	385.371	13/03/2009	713	389.109
27/10/2008	576	297.759	12/12/2008	622	333.501	27/01/2009	668	388.941	14/03/2009	714	387.660
28/10/2008	577	311.073	13/12/2008	623	345.786	28/01/2009	669	388.941	15/03/2009	715	385.791
29/10/2008	578	311.073	14/12/2008	624	339.717	29/01/2009	670	402.507	16/03/2009	716	385.791
30/10/2008	579	311.073	15/12/2008	625	352.443	30/01/2009	671	402.507	17/03/2009	717	388.857
31/10/2008	580	329.994	16/12/2008	626	352.443	31/01/2009	672	373.002	18/03/2009	718	388.857
01/11/2008	581	324.702	17/12/2008	627	379.134	01/02/2009	673	345.261	19/03/2009	719	384.006
02/11/2008	582	324.702	18/12/2008	628	333.501	02/02/2009	674	345.261	20/03/2009	720	384.006
03/11/2008	583	324.702	19/12/2008	629	345.786	03/02/2009	675	345.261	21/03/2009	721	384.006
04/11/2008	584	327.264	20/12/2008	630	383.901	04/02/2009	676	345.261	22/03/2009	722	404.985
05/11/2008	585	330.687	21/12/2008	631	353.451	05/02/2009	677	392.532	23/03/2009	723	389.109
06/11/2008	586	327.264	22/12/2008	632	315.252	06/02/2009	678	393.204	24/03/2009	724	389.109
07/11/2008	587	330.687	23/12/2008	633	374.010	07/02/2009	679	393.204	25/03/2009	725	389.109
08/11/2008	588	334.446	24/12/2008	634	383.901	08/02/2009	680	398.496	26/03/2009	726	387.156
09/11/2008	589	339.297	25/12/2008	635	353.451	09/02/2009	681	388.689	27/03/2009	727	373.758
10/11/2008	590	345.786	26/12/2008	636	315.252	10/02/2009	682	393.204	28/03/2009	728	373.758
11/11/2008	591	327.264	27/12/2008	637	377.433	11/02/2009	683	393.204			
12/11/2008	592	330.687	28/12/2008	638	372.729	12/02/2009	684	398.496			
13/11/2008	593	327.264	29/12/2008	639	345.786	13/02/2009	685	388.689			
14/11/2008	594	330.687	30/12/2008	640	345.786	14/02/2009	686	388.689			
15/11/2008	595	334.446	31/12/2008	641	366.345	15/02/2009	687	388.689			
16/11/2008	596	339.297	01/01/2009	642	371.028	16/02/2009	688	381.276			
17/11/2008	597	345.786	02/01/2009	643	371.028	17/02/2009	689	381.360			
18/11/2008	598	338.268	03/01/2009	644	371.028	18/02/2009	690	384.930			

Anexo 2

Ajuste ARIMA (5,1,0) a 628 datos de los incrementos de esfuerzos D4-I4

No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t
1	230.03	-	-	47	217.92	231.26	-13.34	93	183.54	191.43	-7.89
2	218.25	-	-	48	235.91	224.33	11.58	94	190.79	189.11	1.67
3	211.60	-	-	49	239.67	229.23	10.44	95	187.38	189.05	-1.67
4	206.75	-	-	50	224.05	235.92	-11.87	96	192.15	188.50	3.65
5	203.85	-	-	51	225.33	231.54	-6.21	97	202.57	190.88	11.69
6	221.84	-	-	52	222.10	226.52	-4.42	98	216.38	195.21	21.17
7	240.35	215.55	24.79	53	215.94	224.28	-8.34	99	191.21	203.64	-12.44
8	230.37	225.87	4.50	54	235.91	223.03	12.88	100	196.08	199.25	-3.17
9	247.00	227.82	19.18	55	239.59	227.98	11.61	101	193.18	195.49	-2.31
10	246.14	232.92	13.22	56	213.13	232.94	-19.81	102	216.38	194.96	21.42
11	231.32	237.99	-6.68	57	199.06	225.19	-26.13	103	191.21	205.67	-14.47
12	248.87	236.85	12.02	58	203.76	211.79	-8.03	104	196.08	201.56	-5.48
13	228.92	240.96	-12.04	59	210.76	208.10	2.65	105	193.18	196.06	-2.88
14	247.09	236.97	10.12	60	216.55	212.87	3.68	106	220.98	194.59	26.39
15	261.16	241.61	19.55	61	216.55	215.53	1.02	107	218.40	206.46	11.94
16	237.62	249.48	-11.87	62	239.59	213.86	25.72	108	218.86	213.70	5.16
17	229.01	245.46	-16.45	63	225.41	222.58	2.84	109	232.58	213.44	19.13
18	237.45	237.18	0.27	64	247.51	224.84	22.67	110	226.19	219.60	6.59
19	240.09	235.36	4.73	65	212.88	233.04	-20.16	111	225.33	223.09	2.24
20	238.14	240.77	-2.63	66	209.64	225.36	-15.72	112	213.57	225.09	-11.52
21	222.52	240.95	-18.44	67	220.98	217.91	3.07	113	206.75	219.56	-12.82
22	202.90	231.27	-28.36	68	209.98	219.85	-9.87	114	209.48	214.46	-4.99
23	222.35	218.58	3.77	69	219.70	218.10	1.60	115	223.04	213.37	9.67
24	242.40	220.22	22.18	70	206.83	219.30	-12.47	116	225.58	217.88	7.70
25	251.69	231.73	19.96	71	199.75	212.04	-12.29	117	215.94	221.85	-5.91
26	250.15	241.20	8.95	72	195.91	206.86	-10.95	118	232.85	218.26	14.59
27	232.76	242.21	-9.45	73	188.39	202.52	-14.13	119	231.90	222.18	9.73
28	218.51	235.65	-17.15	74	189.25	197.09	-7.84	120	238.04	227.26	10.77
29	213.05	228.84	-15.79	75	178.33	194.72	-16.39	121	238.73	232.61	6.12
30	213.05	223.81	-10.77	76	189.17	187.79	1.38	122	233.77	234.35	-0.58
31	213.05	221.39	-8.35	77	188.39	187.77	0.62	123	237.28	233.54	3.74
32	213.05	218.93	-5.88	78	189.25	188.48	0.78	124	236.33	235.47	0.86
33	213.05	215.69	-2.64	79	178.33	188.69	-10.36	125	227.89	236.06	-8.17
34	257.59	213.65	43.94	80	178.50	183.89	-5.39	126	237.11	233.41	3.70
35	223.04	230.75	-7.71	81	185.24	180.59	4.65	127	222.77	234.26	-11.49
36	226.44	229.91	-3.47	82	170.06	183.39	-13.33	128	233.10	230.14	2.96
37	241.71	225.63	16.08	83	176.11	179.10	-3.00	129	233.69	231.14	2.55
38	243.58	230.16	13.42	84	176.11	176.94	-0.83	130	231.13	232.31	-1.18
39	219.62	238.80	-19.18	85	174.24	176.14	-1.90	131	237.53	232.06	5.48
40	222.26	233.33	-11.06	86	193.60	176.29	17.31	132	234.47	233.91	0.56
41	225.16	224.40	0.76	87	182.26	183.01	-0.75	133	229.17	233.60	-4.43
42	217.31	225.69	-8.38	88	188.48	182.77	5.71	134	233.35	232.38	0.97
43	224.83	225.05	-0.22	89	195.30	184.48	10.82	135	234.47	232.18	2.29
44	238.90	224.25	14.64	90	185.16	188.10	-2.94	136	234.47	233.52	0.95
45	236.42	228.90	7.52	91	194.80	187.88	6.91	137	252.80	234.59	18.21
46	230.62	233.11	-2.49	92	193.52	190.73	2.79	138	243.08	241.28	1.80

No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t
139	266.11	242.52	23.60	185	186.86	182.28	4.57	231	232.41	216.56	15.84
140	240.77	251.21	-10.45	186	184.04	185.32	-1.28	232	231.40	224.89	6.51
141	230.20	247.33	-17.12	187	188.66	184.00	4.66	233	244.19	226.58	17.61
142	224.05	239.43	-15.38	188	213.57	184.58	28.99	234	245.13	233.22	11.91
143	239.32	233.07	6.24	189	188.48	195.99	-7.52	235	245.13	237.61	7.52
144	221.24	236.74	-15.50	190	181.84	194.77	-12.93	236	240.51	243.22	-2.70
145	234.89	233.54	1.35	191	180.29	187.74	-7.45	237	277.62	240.70	36.92
146	237.03	231.41	5.62	192	181.76	183.31	-1.55	238	222.52	255.12	-32.60
147	240.51	233.04	7.47	193	178.08	184.95	-6.87	239	222.52	244.99	-22.47
148	218.34	236.66	-18.33	194	183.96	184.23	-0.27	240	222.52	232.69	-10.17
149	225.86	229.47	-3.61	195	177.66	182.29	-4.63	241	222.52	227.61	-5.09
150	213.82	225.59	-11.77	196	181.40	180.29	1.11	242	226.61	229.22	-2.60
151	213.82	222.57	-8.75	197	191.39	180.41	10.98	243	231.04	230.22	0.82
152	244.10	219.93	24.18	198	206.66	184.81	21.85	244	231.04	227.09	3.95
153	228.92	229.42	-0.50	199	206.66	194.12	12.54	245	231.04	228.77	2.27
154	223.71	230.12	-6.41	200	206.66	199.70	6.96	246	231.04	229.29	1.75
155	223.13	226.42	-3.29	201	206.66	201.01	5.65	247	222.01	229.72	-7.71
156	224.49	222.44	2.05	202	206.66	202.57	4.09	248	151.56	226.96	-75.41
157	222.77	224.61	-1.84	203	186.61	204.98	-18.37	249	229.26	196.83	32.43
158	228.14	226.20	1.95	204	188.92	198.69	-9.77	250	226.28	206.44	19.84
159	224.57	225.84	-1.27	205	205.63	193.80	11.83	251	235.22	220.68	14.55
160	241.04	225.21	15.83	206	205.63	199.15	6.48	252	233.02	228.85	4.17
161	237.95	231.27	6.68	207	227.72	203.83	23.89	253	239.06	224.74	14.33
162	235.41	234.56	0.85	208	219.28	212.58	6.70	254	236.67	226.21	10.46
163	238.39	234.47	3.92	209	194.02	213.93	-19.91	255	245.13	235.27	9.86
164	238.39	235.22	3.17	210	199.56	205.29	-5.72	256	234.80	238.93	-4.13
165	209.56	236.79	-27.24	211	199.56	201.62	-2.06	257	235.05	237.95	-2.90
166	187.30	227.13	-39.83	212	215.52	202.48	13.04	258	231.90	236.12	-4.21
167	184.91	209.43	-24.52	213	198.39	210.59	-12.21	259	231.90	234.55	-2.65
168	181.65	199.19	-17.54	214	194.46	205.41	-10.95	260	237.95	233.92	4.03
169	195.99	194.79	1.20	215	194.88	198.15	-3.27	261	234.63	236.07	-1.44
170	192.23	196.42	-4.19	216	182.18	196.42	-14.24	262	234.63	235.33	-0.70
171	195.49	193.63	1.86	217	209.81	192.14	17.67	263	237.95	234.72	3.23
172	189.59	192.52	-2.93	218	186.94	199.71	-12.76	264	226.86	235.48	-8.62
173	197.44	190.48	6.96	219	194.88	195.15	-0.27	265	238.73	232.49	6.24
174	213.57	193.01	20.56	220	182.18	193.75	-11.58	266	232.24	234.70	-2.46
175	199.75	202.44	-2.69	221	209.81	188.25	21.56	267	238.90	234.17	4.73
176	181.65	202.14	-20.49	222	186.94	196.86	-9.92	268	272.69	236.17	36.52
177	193.01	192.52	0.49	223	235.58	195.77	39.81	269	236.78	250.54	-13.77
178	185.66	190.78	-5.12	224	194.88	209.10	-14.22	270	241.88	246.37	-4.49
179	193.77	191.35	2.41	225	182.18	205.08	-22.91	271	241.71	242.19	-0.48
180	199.92	193.81	6.11	226	209.81	193.54	16.27	272	232.49	240.59	-8.10
181	185.49	195.03	-9.53	227	186.94	198.89	-11.95	273	248.70	240.31	8.40
182	176.97	190.70	-13.74	228	225.86	197.41	28.44	274	233.44	244.82	-11.38
183	175.27	184.52	-9.25	229	197.11	210.60	-13.49	275	251.69	239.01	12.67
184	181.57	179.96	1.60	230	240.87	203.16	37.71	276	233.44	243.81	-10.37

No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t
277	251.69	239.84	11.84	323	257.59	254.45	3.14	369	299.96	287.06	12.90
278	258.51	243.68	14.83	324	284.11	257.23	26.88	370	264.22	292.17	-27.95
279	234.55	250.72	-16.17	325	273.19	266.83	6.36	371	270.12	277.41	-7.29
280	238.48	244.49	-6.01	326	252.29	269.28	-16.99	372	260.13	272.82	-12.69
281	238.48	240.58	-2.10	327	257.84	261.86	-4.02	373	266.78	271.55	-4.77
282	239.84	239.05	0.79	328	249.98	260.31	-10.32	374	237.95	271.57	-33.62
283	247.67	242.11	5.56	329	237.36	256.69	-19.33	375	237.03	258.17	-21.14
284	261.41	244.40	17.01	330	248.62	251.53	-2.91	376	277.96	246.48	31.47
285	234.05	249.96	-15.91	331	259.46	248.69	10.76	377	269.43	259.59	9.84
286	248.45	244.28	4.17	332	248.62	253.05	-4.43	378	287.09	266.87	20.22
287	244.19	243.70	0.49	333	259.46	252.67	6.78	379	246.14	273.68	-27.54
288	260.74	244.58	16.16	334	266.53	253.17	13.36	380	243.94	259.21	-15.27
289	245.30	252.85	-7.55	335	256.39	257.85	-1.46	381	239.32	251.44	-12.12
290	245.30	250.53	-5.23	336	260.65	258.88	1.77	382	235.31	249.08	-13.77
291	239.76	245.82	-6.06	337	256.39	258.81	-2.42	383	268.07	245.71	22.35
292	239.23	243.65	-4.41	338	197.27	257.22	-59.95	384	258.51	255.52	2.99
293	248.62	242.48	6.14	339	254.33	234.51	19.82	385	297.07	256.02	41.04
294	238.14	246.02	-7.88	340	241.63	239.66	1.97	386	247.00	271.08	-24.08
295	238.14	242.71	-4.57	341	247.51	244.50	3.00	387	244.19	261.18	-16.99
296	238.14	240.11	-1.97	342	236.94	247.71	-10.77	388	258.51	251.52	6.99
297	238.56	238.61	-0.05	343	280.01	239.54	40.47	389	297.07	255.03	42.03
298	239.76	239.32	0.43	344	231.40	252.31	-20.91	390	247.00	275.51	-28.51
299	239.76	240.06	-0.30	345	234.89	249.73	-14.84	391	271.91	268.09	3.82
300	240.60	239.29	1.30	346	237.20	240.16	-2.96	392	269.01	261.79	7.22
301	239.76	239.75	0.01	347	269.51	237.90	31.62	393	261.16	264.60	-3.44
302	266.78	239.72	27.06	348	269.51	253.34	16.17	394	254.69	266.99	-12.30
303	242.40	250.45	-8.05	349	237.36	263.58	-26.22	395	246.83	262.31	-15.47
304	234.97	248.65	-13.68	350	253.91	248.95	4.97	396	240.43	253.10	-12.67
305	274.39	241.34	33.05	351	246.58	247.42	-0.84	397	249.23	249.99	-0.77
306	282.91	253.01	29.90	352	267.90	249.17	18.73	398	246.41	249.75	-3.34
307	270.21	268.83	1.37	353	281.04	260.09	20.96	399	274.64	248.98	25.65
308	283.08	270.86	12.22	354	267.90	268.33	-0.44	400	256.73	258.74	-2.02
309	275.92	270.57	5.35	355	281.04	266.73	14.31	401	238.98	258.27	-19.29
310	280.10	272.92	7.18	356	281.04	271.04	10.00	402	264.92	248.79	16.13
311	250.07	278.60	-28.53	357	253.22	274.83	-21.61	403	246.41	253.48	-7.07
312	277.28	267.27	10.01	358	265.00	268.28	-3.28	404	274.64	253.66	20.98
313	277.28	269.37	7.91	359	265.86	265.48	0.38	405	256.73	263.30	-6.57
314	240.77	275.20	-34.43	360	241.37	266.19	-24.82	406	238.98	259.54	-20.56
315	277.28	262.33	14.95	361	244.19	258.77	-14.58	407	264.92	249.97	14.94
316	240.77	264.69	-23.93	362	265.25	250.73	14.52	408	261.41	255.21	6.20
317	239.59	256.19	-16.60	363	259.12	255.48	3.64	409	239.15	259.62	-20.47
318	239.84	251.02	-11.18	364	273.78	260.32	13.46	410	241.29	253.53	-12.24
319	237.45	244.79	-7.34	365	258.01	264.48	-6.48	411	269.35	244.60	24.75
320	251.18	242.19	8.99	366	258.01	259.80	-1.80	412	259.71	254.39	5.32
321	236.00	247.86	-11.86	367	296.48	258.67	37.81	413	264.92	261.02	3.90
322	276.78	241.44	35.34	368	299.38	274.50	24.88	414	261.41	260.79	0.61

No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t
415	239.15	258.23	-19.09	461	255.78	257.49	-1.71	507	358.58	324.82	33.75
416	241.29	251.33	-10.04	462	252.71	254.84	-2.13	508	354.56	336.61	17.95
417	269.35	247.76	21.58	463	276.70	255.42	21.27	509	331.44	345.06	-13.62
418	247.34	256.67	-9.34	464	251.43	263.57	-12.13	510	323.86	339.61	-15.75
419	247.34	256.10	-8.76	465	287.01	259.08	27.92	511	333.25	329.27	3.98
420	277.20	250.23	26.97	466	282.49	268.05	14.44	512	338.44	331.90	6.54
421	269.96	258.21	11.74	467	262.69	275.15	-12.47	513	338.44	338.97	-0.54
422	241.29	266.19	-24.90	468	262.69	270.68	-8.00	514	335.90	338.83	-2.93
423	277.20	256.83	20.37	469	260.04	265.68	-5.63	515	310.63	335.31	-24.68
424	269.96	260.25	9.71	470	245.97	262.86	-16.89	516	304.75	324.60	-19.84
425	233.52	267.37	-33.85	471	295.53	259.65	35.88	517	306.12	316.18	-10.06
426	235.31	256.43	-21.12	472	254.00	272.83	-18.84	518	306.12	313.07	-6.95
427	225.33	243.32	-17.99	473	277.20	267.05	10.15	519	306.12	312.18	-6.06
428	212.46	236.08	-23.62	474	254.00	268.61	-14.61	520	303.14	309.76	-6.62
429	214.41	230.88	-16.47	475	277.20	261.49	15.71	521	303.39	305.31	-1.93
430	209.64	223.07	-13.43	476	310.97	268.16	42.81	522	295.28	304.02	-8.74
431	206.58	216.43	-9.85	477	272.85	288.39	-15.54	523	312.25	300.73	11.52
432	210.67	213.12	-2.44	478	274.05	282.24	-8.19	524	298.52	305.09	-6.58
433	243.33	211.23	32.10	479	270.46	276.01	-5.55	525	296.73	303.50	-6.77
434	233.44	223.75	9.69	480	302.53	271.27	31.25	526	302.69	300.01	2.68
435	233.44	229.42	4.02	481	303.56	287.81	15.74	527	303.81	300.18	3.62
436	236.78	229.04	7.73	482	299.71	297.64	2.07	528	320.44	302.47	17.97
437	229.95	230.05	-0.10	483	299.71	295.38	4.33	529	307.99	310.65	-2.67
438	243.41	231.35	12.06	484	312.17	294.81	17.35	530	311.33	309.07	2.25
439	239.76	237.81	1.95	485	316.87	301.75	15.12	531	331.19	308.76	22.43
440	243.94	238.56	5.38	486	305.70	310.78	-5.08	532	335.71	317.48	18.23
441	240.68	240.64	0.05	487	307.48	308.70	-1.22	533	352.36	326.66	25.70
442	235.14	240.24	-5.10	488	317.98	306.45	11.53	534	365.65	337.36	28.29
443	237.36	237.83	-0.47	489	302.97	311.21	-8.24	535	330.08	347.20	-17.13
444	234.30	238.04	-3.74	490	302.02	310.01	-7.99	536	308.76	340.69	-31.93
445	241.79	236.82	4.98	491	298.96	305.91	-6.96	537	305.34	326.12	-20.78
446	241.37	239.20	2.17	492	296.14	302.08	-5.93	538	312.08	317.28	-5.20
447	241.46	240.12	1.34	493	295.62	300.78	-5.17	539	299.71	319.18	-19.46
448	240.77	240.28	0.48	494	297.68	299.31	-1.64	540	323.69	313.88	9.81
449	239.06	240.12	-1.05	495	301.33	297.91	3.42	541	323.00	314.69	8.31
450	239.76	239.54	0.21	496	305.87	299.49	6.38	542	300.07	317.63	-17.56
451	227.30	240.05	-12.75	497	313.11	301.96	11.15	543	295.70	310.62	-14.92
452	243.85	235.08	8.77	498	302.44	306.19	-3.75	544	289.74	302.33	-12.59
453	247.25	238.04	9.21	499	317.79	304.68	13.11	545	306.20	297.12	9.08
454	235.31	242.76	-7.45	500	337.43	309.03	28.39	546	301.94	303.65	-1.72
455	292.13	240.24	51.89	501	336.32	321.17	15.14	547	293.33	304.02	-10.69
456	258.95	259.04	-0.09	502	328.04	328.83	-0.79	548	298.35	297.96	0.39
457	266.87	261.24	5.63	503	309.35	327.32	-17.97	549	286.76	296.37	-9.61
458	260.74	262.23	-1.49	504	331.19	317.90	13.29	550	291.10	293.04	-1.94
459	255.36	259.28	-3.92	505	319.68	323.72	-4.03	551	285.89	293.19	-7.29
460	247.86	259.05	-11.18	506	323.86	325.54	-1.68	552	283.61	290.13	-6.53

No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t
553	282.91	287.40	-4.48	599	338.79	338.61	0.18
554	286.15	285.59	0.55	600	342.11	338.07	4.04
555	288.62	285.45	3.18	601	310.88	339.76	-28.88
556	296.14	287.29	8.85	602	381.17	329.30	51.87
557	299.96	290.53	9.43	603	310.88	348.61	-37.73
558	305.61	294.22	11.39	604	345.79	337.24	8.55
559	291.88	298.56	-6.68	605	338.27	337.39	0.88
560	290.60	295.82	-5.22	606	338.79	336.67	2.12
561	291.19	292.88	-1.69	607	338.79	339.34	-0.55
562	299.88	292.66	7.22	608	342.11	341.16	0.95
563	297.34	296.80	0.54	609	310.88	337.62	-26.74
564	302.02	297.79	4.23	610	381.17	329.38	51.79
565	291.88	298.15	-6.27	611	320.71	347.83	-27.12
566	302.02	295.30	6.72	612	316.51	341.20	-24.69
567	296.98	297.62	-0.64	613	343.31	328.60	14.71
568	283.86	298.39	-14.53	614	316.51	331.16	-14.65
569	290.85	292.63	-1.78	615	343.31	329.26	14.05
570	301.16	290.96	10.20	616	324.20	337.27	-13.08
571	297.07	295.45	1.61	617	331.53	328.30	3.22
572	299.46	297.93	1.53	618	339.72	329.62	10.09
573	307.82	297.16	10.66	619	352.44	334.31	18.14
574	306.98	300.12	6.85	620	352.44	341.43	11.02
575	296.31	304.02	-7.71	621	379.13	347.37	31.76
576	297.76	301.18	-3.43	622	333.50	357.91	-24.40
577	311.07	298.50	12.57	623	345.79	349.24	-3.46
578	311.07	304.14	6.94	624	339.72	345.52	-5.81
579	311.07	308.78	2.30	625	352.44	344.12	8.33
580	329.99	309.10	20.89	626	352.44	350.03	2.42
581	324.70	315.83	8.87	627	379.13	352.39	26.74
582	324.70	320.50	4.20	628	333.50	360.15	-26.65
583	324.70	322.29	2.41				
584	327.26	322.14	5.13				
585	330.69	324.67	6.02				
586	327.26	328.41	-1.14				
587	330.69	327.70	2.98				
588	334.45	328.46	5.99				
589	339.30	330.89	8.40				
590	345.79	334.75	11.03				
591	327.26	339.26	-11.99				
592	330.69	334.24	-3.55				
593	327.26	331.66	-4.40				
594	330.69	330.43	0.25				
595	334.45	331.87	2.57				
596	339.30	333.22	6.08				
597	345.79	334.75	11.03				
598	338.27	339.26	-0.99				

Anexo 3

Ajuste ARIMA (5,1,0) a 100 datos de los incrementos de esfuerzos D4-I4

No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t
1	345.79	-	-	47	345.26	358.79	-13.53	93	384.01	384.92	-0.92
2	383.90	-	-	48	345.26	357.00	-11.74	94	404.99	384.71	20.27
3	353.45	-	-	49	392.53	353.95	38.59	95	389.11	393.30	-4.19
4	315.25	-	-	50	393.20	367.11	26.09	96	389.11	392.64	-3.53
5	374.01	-	-	51	393.20	378.01	15.20	97	389.11	389.57	-0.46
6	383.90	-	-	52	398.50	382.83	15.67	98	387.16	388.42	-1.26
7	353.45	369.53	-16.08	53	388.69	385.91	2.78	99	373.76	389.13	-15.37
8	315.25	366.90	-51.65	54	393.20	387.60	5.61	100	373.76	384.19	-10.43
9	377.43	338.37	39.06	55	393.20	392.34	0.86				
10	372.73	350.06	22.67	56	398.50	392.86	5.63				
11	345.79	369.19	-23.40	57	388.69	395.67	-6.98				
12	345.79	360.95	-15.17	58	388.69	393.14	-4.45				
13	366.35	347.26	19.08	59	388.69	390.12	-1.43				
14	371.03	353.82	17.21	60	381.28	389.76	-8.49				
15	371.03	367.57	3.46	61	381.36	387.08	-5.72				
16	371.03	368.11	2.92	62	384.93	384.71	0.22				
17	371.03	366.00	5.03	63	382.20	384.35	-2.15				
18	375.38	367.84	7.53	64	385.96	384.16	1.80				
19	383.57	372.24	11.33	65	384.17	384.58	-0.41				
20	382.54	377.27	5.27	66	399.19	383.95	15.24				
21	382.54	379.66	2.88	67	399.19	390.03	9.15				
22	377.60	380.25	-2.64	68	388.19	394.44	-6.25				
23	382.54	378.71	3.82	69	376.49	391.52	-15.03				
24	377.60	380.41	-2.81	70	396.46	384.24	12.22				
25	378.19	380.20	-2.02	71	396.02	388.80	7.22				
26	342.03	379.28	-37.26	72	384.26	394.70	-10.45				
27	343.81	364.27	-20.46	73	375.21	390.69	-15.48				
28	319.94	354.35	-34.41	74	380.25	381.81	-1.57				
29	311.83	341.97	-30.14	75	387.66	380.46	7.20				
30	330.44	330.82	-0.38	76	387.66	385.91	1.75				
31	342.53	330.53	12.00	77	400.81	387.32	13.48				
32	347.74	335.64	12.10	78	390.14	391.30	-1.16				
33	358.07	341.03	17.04	79	398.24	390.44	7.80				
34	393.73	344.87	48.86	80	387.74	393.05	-5.31				
35	332.56	363.47	-30.92	81	391.08	391.25	-0.17				
36	389.55	353.67	35.88	82	391.08	391.25	-0.17				
37	393.73	364.09	29.64	83	391.08	391.88	-0.80				
38	392.95	378.41	14.54	84	385.01	391.53	-6.51				
39	385.37	387.90	-2.53	85	389.11	389.17	-0.06				
40	388.94	384.87	4.08	86	387.66	388.17	-0.51				
41	388.94	382.23	6.71	87	385.79	388.56	-2.77				
42	402.51	389.15	13.36	88	385.79	387.68	-1.89				
43	402.51	395.14	7.37	89	388.86	386.55	2.30				
44	373.00	398.40	-25.40	90	388.86	387.24	1.61				
45	345.26	387.47	-42.20	91	384.01	388.42	-4.42				
46	345.26	368.55	-23.29	92	384.01	386.48	-2.47				

Anexo 4

Nuevo ajuste ARIMA (5,1,0) a 100 datos de los incrementos de esfuerzos D4-I4

No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t	No.	Y_t (kg/cm ²)	Ajuste ARIMA	Residual a_t
1	345.79	-	-	47	345.26	355.56	-13.53	93	384.01	386.23	-0.92
2	383.90	-	-	48	345.26	342.30	-11.74	94	404.99	384.02	20.27
3	353.45	-	-	49	392.53	344.13	38.59	95	389.11	397.38	-4.19
4	315.25	-	-	50	393.20	381.14	26.09	96	389.11	387.66	-3.53
5	374.01	-	-	51	393.20	376.80	15.20	97	389.11	392.68	-0.46
6	383.90	-	-	52	398.50	388.76	15.67	98	387.16	394.39	-1.26
7	353.45	353.27	-16.08	53	388.69	405.45	2.78	99	373.76	382.03	-15.37
8	315.25	351.64	-51.65	54	393.20	383.76	5.61	100	373.76	380.86	-10.43
9	377.43	353.00	39.06	55	393.20	394.56	0.86				
10	372.73	367.22	22.67	56	398.50	393.51	5.63				
11	345.79	349.40	-23.40	57	388.69	393.88	-6.98				
12	345.79	347.55	-15.17	58	388.69	392.17	-4.45				
13	366.35	372.04	19.08	59	388.69	390.97	-1.43				
14	371.03	352.91	17.21	60	381.28	390.54	-8.49				
15	371.03	358.16	3.46	61	381.36	381.13	-5.72				
16	371.03	371.23	2.92	62	384.93	385.19	0.22				
17	371.03	374.40	5.03	63	382.20	384.42	-2.15				
18	375.38	369.12	7.53	64	385.96	380.48	1.80				
19	383.57	373.35	11.33	65	384.17	386.39	-0.41				
20	382.54	379.45	5.27	66	399.19	384.35	15.24				
21	382.54	379.67	2.88	67	399.19	393.67	9.15				
22	377.60	382.96	-2.64	68	388.19	395.27	-6.25				
23	382.54	380.19	3.82	69	376.49	389.53	-15.03				
24	377.60	381.35	-2.81	70	396.46	387.01	12.22				
25	378.19	378.07	-2.02	71	396.02	393.03	7.22				
26	342.03	378.34	-37.26	72	384.26	388.34	-10.45				
27	343.81	355.42	-20.46	73	375.21	385.72	-15.48				
28	319.94	354.00	-34.41	74	380.25	387.43	-1.57				
29	311.83	330.96	-30.14	75	387.66	380.01	7.20				
30	330.44	315.70	-0.38	76	387.66	382.26	1.75				
31	342.53	334.59	12.00	77	400.81	384.59	13.48				
32	347.74	328.37	12.10	78	390.14	398.09	-1.16				
33	358.07	342.00	17.04	79	398.24	389.74	7.80				
34	393.73	356.42	48.86	80	387.74	397.13	-5.31				
35	332.56	378.03	-30.92	81	391.08	391.70	-0.17				
36	389.55	338.33	35.88	82	391.08	389.15	-0.17				
37	393.73	390.28	29.64	83	391.08	393.81	-0.80				
38	392.95	383.50	14.54	84	385.01	387.76	-6.51				
39	385.37	370.61	-2.53	85	389.11	388.99	-0.06				
40	388.94	406.44	4.08	86	387.66	389.43	-0.51				
41	388.94	383.55	6.71	87	385.79	387.26	-2.77				
42	402.51	387.69	13.36	88	385.79	385.40	-1.89				
43	402.51	396.55	7.37	89	388.86	388.13	2.30				
44	373.00	399.54	-25.40	90	388.86	387.22	1.61				
45	345.26	380.76	-42.20	91	384.01	387.66	-4.42				
46	345.26	366.74	-23.29	92	384.01	385.54	-2.47				

Referencias

AGUIRRE, Jaime Armando. *Introducción al tratamiento de series temporales*. Madrid, España. Ediciones Díaz de Santos, 1994.

ARNAU Gras, Jaume. *Diseño de series temporales: técnicas de análisis*. 2ª ed. Barcelona, España. Edicions de la Universitat de Barcelona, 2001.

BENNETTS Toledo, Felipe. *Seguimiento de las principales propiedades estructurales de la superestructura del puente San Cristóbal*. México, 2008. Tesis de maestría. UNAM, Instituto de Ingeniería.

BOWERMAN, Bruce L., O'CONNELL, Richard T. y KOEHLER, Anne B. *Pronósticos, series de tiempo y regresión. Un enfoque aplicado*. 4ª ed. México D.F. Cengage learning, 2007.

BOX, George E.P. y JENKINS, Gwilym M. *Time series analysis: forecasting and control*. 3ª ed. New Jersey, E.U.A. Prentice Hall, 1994.

BROWNJOHN, James y CARDEN, E. Peter. *ARMA modeled time-series classification for structural health monitoring of civil infrastructure*. Department of Civil and Structural Engineering, University of Sheffield, UK, 2007.

EVERETT, Adam y RONALD J., Ebert. *Administración de la producción y las operaciones: conceptos, modelos y funcionamiento*. 4ª ed. México. Pearson, 1992.

FARRAR, Charles. CZARNECKI, Jerry y SOHN, Hoon. *Structural Health Monitoring Using Statistical Process Control*.

FUGATE, Michael. SOHN, Hoon y FARRAR, Charles. *Vibration-base damage detection usigng statistical process control*. Los Alamos National Laboratory, 2001.

-
- GLISIC, Branko e INAUDI, Daniele. *Fibre Optic Methods for structural Health Monitoring*. Inglaterra. John Wiley & Sons, 2007.
- GÓMEZ Martínez, Roberto. *Monitoreo del comportamiento estructural de puentes especiales durante su construcción: Chiapas, San Cristóbal y el Baluarte*. Artículo. México. UNAM, Instituto de Ingeniería, 2008.
- GÓMEZ Martínez, Roberto. *et al. Monitoring of the main structural parameters of the structure of the San Cristobal Bridge*. México. UNAM, Instituto de Ingeniería, 2009.
- HANKE, John E. y WICHERN, Dean W. *Pronósticos en los negocios*. 8ª ed. México. Pearson Educación, México 2006.
- HERNÁNDEZ, Cesar. SALCEDO, Octavio y ESCOBAR, Andrés. *An ARIMA model for forecasting Wi-Fi data network traffic values*. Revista de ingeniería e investigación de la Universidad Distrital Francisco José de Caldas, Bogotá, Vol. 29, No. 2, Agosto de 2009. pp 65-69.
- MAKRIDAKIS, Spyros y WHEELWRIGHT, Steven. *Métodos de pronósticos*. México D.F. Limusa, 2000.
- ORDAZ, M. y MONTOYA, C. *Degtra A4 Versión 3.1.4*. Instituto de Ingeniería, UNAM. México, 2002.
- PANKRATZ, Alan. *Forecasting with univariate Box-Jenkins models*. E.U.A. John Wiley & Sons, 1983.
- PEPIÓ Viñals, Montserrat. *Series temporales*. 1ª ed. Barcelona, España. Edicions UPC, Universitat Politècnica de Catalunya, 2001.
- RODRÍGUEZ Morilla, Carmen. *Análisis de series temporales*. Editorial La Muralla. Madrid 2000.
- SEMARNAT. *Manifestación de Impacto Ambiental, Modalidad Particular, "Puente San Cristóbal, Km 24+921 de la Carretera Tuxtla Gutiérrez-San Cristóbal de las Casas"*.
- URIEL Jiménez, Ezequiel. *Análisis de datos: Series temporales y Análisis multivariante*. Madrid, España. Colección Plan Nuevo, Editorial AC, 1995.