

MODELO DE INTEGRACIÓN ENTRE EXPECTATIVAS DE EMPRESARIOS Y MÉTODO ESTADÍSTICO PARA PRONOSTICAR EL CRECIMIENTO DE LAS VENTAS DEL SECTOR INDUSTRIAL EN COLOMBIA, UNA APLICACIÓN A LA PLANEACIÓN DE LAS EMPRESAS.

Trabajo de Grado

PAOLA ANDREA VARGAS VIGOYA

MILTON MIER SANTANDER



Pontificia Universidad
JAVERIANA
Colombia

**MODELO DE INTEGRACIÓN ENTRE EXPECTATIVAS DE EMPRESARIOS Y MÉTODO
ESTADÍSTICO PARA PRONOSTICAR EL CRECIMIENTO DE LAS VENTAS DEL SECTOR
INDUSTRIAL EN COLOMBIA, UNA APLICACIÓN A LA PLANEACIÓN DE LAS EMPRESAS.**

PAOLA ANDREA VARGAS VIGOYA

MILTON MIER SANTANDER

TRABAJO DE GRADO

DIRECTOR:

ING. JORGE ANDRÉS ALVARADO VALENCIA

PONTIFICIA UNIVERSIDAD JAVERIANA

BOGOTÁ D.C.

2012

Bogotá, Abril 23 de 2012

Señores

COMITÉ DE CARRERA

INGENIERÍA INDUSTRIAL

PONTIFICIA UNIVERSIDAD JAVERIANA

Señores Comité de Carrera:

La presente comunicación con el fin de manifestar mi conocimiento y aprobación del trabajo de grado titulado “*Modelo de Integración entre Expectativas de Empresarios y Método Estadístico para Pronosticar el Crecimiento de las Ventas del Sector Industrial en Colombia, Una Aplicación a la Planeación de las Empresas*”, elaborada por los estudiantes **PAOLA ANDEA VARGAS VIGOYA** , C.C. 1020 738 144 Y **MILTON MIER SANTANDER** , C.C 1032 427 097, en mi calidad de Director.

Declaro conocer y aceptar el reglamento y disposiciones de los trabajos de grado en la Carrera de Ingeniería Industrial de la Pontificia Universidad Javeriana.

Cordialmente,

ING. JORGE ANDRÉS ALVARADO VALENCIA

TABLA DE CONTENIDO

RESUMEN EJECUTIVO.....	10
1. BASE DE DATOS	13
1.1. DESCRIPCIÓN DE DATOS	13
1.1.1. Expectativas de Empresarios del Sector Industrial en Colombia	13
1.1.2. Histórico de las Ventas del Sector Industrial en Colombia	14
1.2. PREPARACIÓN DE DATOS	16
1.2.1. Crecimiento de las Ventas Reales	16
1.2.2. Expectativas de Empresarios.....	17
1.2.3. Ventas Reales Desagregadas.....	18
1.2.4. Datos de Entrenamiento y Datos de Prueba.....	19
1.2.5. Resumen Base de Datos.....	19
2. MODELO BOX-JENKINS.....	21
2.1. SELECCIÓN DE MODELOS ARMA(p,q)	21
2.2. ESTIMACIÓN DE PARÁMETROS DE MODELOS SELECCIONADOS	25
2.3. PRONÓSTICOS DE MODELOS ARMA SELECCIONADOS	27
2.4. EVALUACIÓN DE PRONÓSTICOS.....	28
3. MODELO FILTRO DE KALMAN.....	31
3.1. MODELO TEÓRICO.....	31
3.1.1. Representación Estado-Espacio	32
3.1.2. Filtro de Kalman	33
3.2. DEFINICIÓN DEL MODELO	35
3.2.1. Capacidad Predictiva de las Expectativas de Empresarios.....	35
3.2.2. Identificación de Variables Exógenas.....	35
3.2.2. Representación Estado-Espacio	38
3.3. ESTIMACIÓN DEL MODELO	39
3.3.1. Estimación de Parámetros Iniciales.....	39
3.3.2. Filtro de Kalman	41
3.3.3. Evaluación del Pronóstico	42
4. RESULTADOS.....	44
4.1. PRONÓSTICO EN LA ETAPA DE PRUEBA	44

4.2.	EVALUACIÓN Y COMPARACIÓN	47
5.	DESAGREGACIÓN.....	48
5.1.	METODOLOGÍAS DE DESAGREGACIÓN	48
5.2.	ESTIMACIÓN DE PROPORCIONES PARA SUBSECTORES	49
5.3.	DESAGREGACIÓN DEL MEJOR PRONÓSTICO Y EVALUACIÓN	52
5.4.	APLICACIÓN EN LA PLANEACIÓN DE LAS EMPRESAS	56
6.	CONCLUSIONES	59
	BIBLIOGRAFÍA.....	62
	ANEXOS	64
	ANEXO I. Cuestionario EMEE Septiembre de 2011.....	64
	ANEXO II. CIU Rev. 3 A.C.	66
	ANEXO III. Correlativa Entre MMM y EAM	67
	Anexo IV. Gráficas pronósticos de los subsectores	70

ÍNDICE DE TABLAS

<i>Tabla 1. Correlaciones EMEE con Ventas reales.....</i>	<i>17</i>
<i>Tabla 2. Resumen Base de Datos</i>	<i>20</i>
<i>Tabla 3. Modelos ARMA Seleccionados Para la Estimación.....</i>	<i>24</i>
<i>Tabla 4. Estimación de Modelos Arma Seleccionados</i>	<i>26</i>
<i>Tabla 5. Evaluación de los Modelos ARMA Seleccionados</i>	<i>29</i>
<i>Tabla 6. Estimación de los modelos para la Ecuación de Transición.....</i>	<i>40</i>
<i>Tabla 7. Estimación de los modelos para la Ecuación de Medición</i>	<i>40</i>
<i>Tabla 8. Modelos a estimar por Filtro de Kalman</i>	<i>41</i>
<i>Tabla 9. Evaluación de los Modelos de Filtro de Kalman</i>	<i>43</i>
<i>Tabla 10. Evaluación de los Modelos ARMA(3,1) y FK3</i>	<i>47</i>
<i>Tabla 11. Promedio de las proporciones históricas y Proporción del promedio histórico.....</i>	<i>50</i>
<i>Tabla 12 . Evaluación de Pronósticos desagregados.....</i>	<i>53</i>
<i>Tabla 13. Participación y establecimientos de los subsectores de mejor ajuste</i>	<i>54</i>
<i>Tabla 14. Objetivos de las empresas</i>	<i>56</i>

ÍNDICE DE GRÁFICAS

<i>Gráfica 1. Índices de Ventas Reales y Nominales del Sector Industrial</i>	15
<i>Gráfica 2. Crecimiento de las Ventas Reales</i>	17
<i>Gráfica 3. Expectativas de Empresarios</i>	18
<i>Gráfica 4. Pronóstico ARMA(3,0)</i>	27
<i>Gráfica 5. Pronóstico ARMA(3,1)</i>	27
<i>Gráfica 6. Pronóstico ARMA(2,3)</i>	27
<i>Gráfica 7. Pronóstico ARMA(1,3)</i>	28
<i>Gráfica 8. Pronóstico ARMA(1,4)</i>	28
<i>Gráfica 9. Pronóstico ARMA(7,3)</i>	28
<i>Gráfica 10. Pronóstico ARMA(0,5)</i>	28
<i>Gráfica 11. Pronóstico ARMA(7,0)</i>	28
<i>Gráfica 12. Pronóstico FK 1</i>	42
<i>Gráfica 13. Pronóstico FK 3</i>	42
<i>Gráfica 14. Pronóstico FK 2</i>	42
<i>Gráfica 15. Pronóstico FK 4</i>	42
<i>Gráfica 16. Pronósticos del Modelo ARMA(3,1) y FK 3 en comparación con la Serie</i>	44
<i>Gráfica 17. Pronóstico del Modelo ARMA(3,1) en comparación con la Serie</i>	45
<i>Gráfica 18. Pronóstico del Modelo FK 3 en comparación con la Serie.</i>	46
<i>Gráfica 19. Proporciones Subsectores</i>	51
<i>Gráfica 20. Subsector 1580</i>	55
<i>Gráfica 21. Subsector 2510</i>	55
<i>Gráfica 22. Subsector 1740</i>	55

<i>Gráfica 23. Subsector 2030</i>	55
<i>Gráfica 24. Subsector 2800</i>	55

ÍNDICE DE ILUSTRACIONES

<i>Ilustración 1. Salida de ACF y PACF del Crecimiento de las Ventas</i>	<i>22</i>
<i>Ilustración 2. Salida de la Identificación del Modelo: MIMIC – ESACF - SCAN</i>	<i>24</i>
<i>Ilustración 3. Diagrama de Flujo del Filtro de Kalman</i>	<i>34</i>
<i>Ilustración 4. ACF y PACF de las Expectativas de empresarios</i>	<i>36</i>
<i>Ilustración 5. Salida selección de variables exógenas ecuación de medición.....</i>	<i>37</i>
<i>Ilustración 6. Salida selección de variables exógenas ecuación de transición</i>	<i>38</i>
<i>Ilustración 7. Modelo de Planeación Estratégica de las empresas</i>	<i>57</i>

RESUMEN EJECUTIVO

El Propósito de esta investigación es evaluar la metodología Box Jenkins y Filtro de Kalman para pronosticar el Crecimiento de las Ventas del Sector Industrial en Colombia, utilizando la serie de tiempo y las expectativas de empresarios, con el fin de presentar a las empresas información precisa del comportamiento futuro de su entorno de tal forma que sea útil para implementar dentro de su planeación estratégica

Las empresas tienen como objetivo principal maximizar sus utilidades, por lo tanto deben planear la cantidad de producción, materia prima a adquirir y factores productivos a contratar restringido a los pronósticos de demanda. Para realizar este pronóstico las grandes empresas utilizan modelos estadísticos sencillos, empleando la información histórica como un pronóstico inicial (Goodwin, Lawrence, Fildes, & Nikolopoulos, 2009). Adicionalmente, (Armstrong, 2006) presenta una evaluación de modelos de pronóstico para identificar las características que comparten los modelos exitosos concluyendo las siguientes:

- Diversifican el riesgo utilizando diferentes métodos.
- Usan representaciones realistas de la situación.
- Incluyen mayor información.
- Utilizan el conocimiento previo.
- Aplican métodos estructurados.

Estas características no están presentes en los métodos de pronósticos de las empresas en la actualidad, si únicamente utilizan simples modelos estadísticos. La principal problemática de estos modelos es que son incapaces de predecir cambios y puntos de inflexión, o de considerar nuevas situaciones (Makridakis & Taleb, 2009), por lo tanto, si las empresas involucraran mayor información como las tendencias del mercado, las tendencias de la economía o las tendencias del sector económico en sus métodos de pronósticos cumplirían con un mayor número de características al incluir mayor información, realizar representaciones más realistas de la situación, utilizar el conocimiento previo y aplicar métodos estructurados, lo que conduciría a mayor precisión de su pronóstico mejorando la planeación de su cadena productiva.

Para el contexto empresarial Colombiano un tipo de información que se encuentra disponible acerca del entorno de las empresas es la serie histórica de las ventas del sector Industrial publicada por el DANE a través de la Muestra Mensual Manufacturera y las Expectativas de empresarios presentada por el Banco de la Republica a través de la Encuesta Mensual de Expectativas Económicas. El propósito es utilizar esta información disponible para presentar a los empresarios un pronóstico del comportamiento de su entorno de tal forma que este pueda

ser implementado en su pronóstico de demanda y al mismo tiempo incluir y evaluar su planeación.

Las expectativas de los empresarios constituyen en si un pronóstico que logra identificar cambios próximos en el mercado o tendencias inesperadas causadas por el entorno macroeconómico. Sin embargo, utilizar las expectativas como única fuente de predicción, conlleva a pronósticos imprecisos (Öller L.-E. , 1990). Los juicios son inconsistentes, puesto que los humanos pueden procesar pequeñas cantidades de información y presentar sesgos cognitivos, en contraste, las series de tiempo modeladas a través de modelos estadísticos son consistentes y pueden capturar mayor volumen de información, pero son rígidos a cambios imprevistos de otras variables causales. En efecto integrar modelos estadísticos y expectativas de expertos conducirá a un mejor pronóstico (Goodwin P. , 2002), donde el modelo estadístico involucre los datos históricos, y el juicio otras variables causales.

(Rahiala & Teräsvirta, 1993), (Öller & Tallbom, 1993) y (Dahllöf, 2003) desarrollan un modelo de Filtro de Kalman para integrar las expectativas de empresarios y la serie de tiempo con el objetivo de pronosticar el crecimiento de la producción de Suecia y Finlandia. La metodología consiste en utilizar el pronóstico de la serie como regresor de las expectativas, ajustarlo adicionando la diferencia entre la expectativa observada y la expectativa estimada y finalmente multiplicándolo por *“la ganancia de Kalman”*. La diferencia entre los autores radica en la periodicidad de las series utilizadas. (Rahiala & Teräsvirta, 1993) y (Öller & Tallbom, 1993) quienes utilizan series trimestrales obtienen resultados satisfactorios ya que el pronóstico presenta un buen ajuste especialmente en los puntos de inflexión. (Dahllöf, 2003) Utiliza la misma metodología para series mensuales, sin embargo sus resultados no presentan un buen ajuste aunque permite identificar cambios anticipados en la tendencia.

El Objetivo de esta investigación es pronosticar el crecimiento de las ventas del Sector Industrial en Colombia mediante la evaluación del modelo de integración entre expectativas de empresarios y modelos estadísticos como se desarrolla en (Dahllöf, 2003), teniendo como marco de referencia un modelo estadístico de series de tiempo. Adicionalmente, dada la clasificación en subsectores del Sector Industrial que presenta la Muestra Mensual Manufacturera es posible realizar una desagregación del pronóstico del entorno que complementariamente otorgue al empresario una información más específica del subsector.

Para el desarrollo de los objetivos se inicia con la construcción de un modelo Box-Jenkins utilizando únicamente la serie de tiempo con el fin de tener un marco de referencia. A continuación se desarrolla el modelo de Filtro de Kalman donde se implementa la integración de las expectativas de empresarios y la serie del crecimiento de las ventas. El desempeño del pronóstico de los modelos se evalúa sobre una muestra de prueba que no fue incluida en la estimación de los parámetros, donde los resultados indican que la información de las expectativas de empresarios tiene un impacto significativo en la precisión del pronóstico. La serie de pronósticos del mejor modelo, es finalmente desagregada sobre cada uno de los subsectores del Sector Industrial de tal forma que los empresarios adicionalmente tengan disponible información precisa del comportamiento del entorno económico para ser implementada dentro de su planeación estratégica.

El documento se compone de seis Capítulos. En el primer capítulo se presenta una descripción detallada de los datos, para en el segundo y tercer capítulo construir el modelo Box-Jenkins y el modelo de Filtro de Kalman. En el cuarto capítulo se presentan la evaluación de cada uno los modelos establecidos sobre la muestra de prueba, posteriormente en el quinto se realiza la desagregación por subsectores, y finalmente las conclusiones son presentadas en la último capítulo.

1. BASE DE DATOS

1.1. DESCRIPCIÓN DE DATOS

Con el propósito de pronosticar el crecimiento de las ventas del sector industrial en Colombia por medio de la integración de la expectativa de expertos y la serie de tiempo, como herramienta para la planeación estratégica de las empresas, la información disponible para el contexto colombiano es:

1. Expectativas de los empresarios del sector industrial en Colombia, resultados Encuesta Mensual de Expectativas Económicas, Banco de la Republica de Colombia.
2. Histórico de las ventas del sector industrial en Colombia, resultados Muestra Mensual Manufacturera, Departamento Administrativo Nacional de Estadística, DANE

A continuación se describe con exactitud cada una de las series, las cuales se utilizaran a lo largo de este trabajo.

1.1.1. Expectativas de Empresarios del Sector Industrial en Colombia

La Subgerencia de Estudios Económicos del Banco de la República realiza la Encuesta Mensual de Expectativas Económicas (EMEE), con el objetivo de identificar la percepción que tienen los empresarios sobre algunas variables de la actividad económica Colombiana. El contenido de la EMEE se puede clasificar en cuatro categorías:

- Actividad de los negocios
- Capacidad de producción
- Salarios
- Precios

Con cobertura temática en Agricultura, Industria, Construcción, Comercio, Transporte y Comunicaciones, Servicios Financieros y Actividades Inmobiliarias. La cobertura geográfica es Eje Cafetero, Costa Atlántica, Bogotá, Nor-occidente, Sur-occidente, Central Oriental, Antioquia y Llanos Orientales.

La EMEE es una encuesta de tipo cualitativo cuyo cuestionario se compone de once preguntas, con opciones de respuesta ordinal clasificadas en tres categorías las cuales son: “aumenta”, “disminuye” o “permanece igual”. El cuestionario completo se presenta en el Anexo 1.

Los Resultados de la EMEE son publicados en la página Web del Banco de la República el mes posterior a la aplicación de la encuesta a los empresarios. La serie histórica se presenta desde Octubre de 2005 clasificada por Regiones y Actividad económica. Para cada publicación se

presenta el resultado desagregado por cada una de las once preguntas clasificado en cuatro grupos, porcentaje de empresarios que respondieron Mayor, Menor y El mismo, y Balance que representa la resta entre el porcentaje mayor menos el porcentaje menor.

Dentro de la categoría que indaga acerca de la actividad de los negocios se encuentra la expectativa del crecimiento de las ventas, específicamente la pregunta No. 2

“El crecimiento del volumen de ventas (cantidades) en los próximos 12 meses, comparado con el crecimiento del volumen de ventas (cantidades) en los pasados 12 meses, se espera sea: Mayor, Menor, o El mismo”

Esta pregunta pretende medir la expectativa de los empresarios acerca del crecimiento de las ventas basadas en la información de los 12 meses anteriores. Los resultados de esta pregunta para el sector industrial, constituye la serie de expectativas de empresarios.

1.1.2. Histórico de las Ventas del Sector Industrial en Colombia

El Departamento Administrativo Nacional de Estadística DANE con el ánimo de obtener información de la evolución de las principales variables económicas del sector fabril en el corto plazo realiza la Muestra Mensual Manufacturera (MMM). Mediante esta investigación el DANE recolecta la información entregada por los industriales acerca de la producción, ventas, personal ocupado, empleo, sueldos y salarios, prestaciones sociales y horas trabajadas, y genera índices y variaciones con año base 2001 para la presentación de dicha información.

La MMM se realiza bajo una muestra probabilística estratificada de 1344 establecimientos (con 10 o más empleados) representativa de la industria manufacturera nacional, desagregado en 48 grupos o clases de actividades económicas diferentes de acuerdo a la Clasificación Industrial Internacional Uniforme CIIU adaptada a Colombia (CIIU Rev. 3 A.C.), presentada en el Anexo 2.

La serie histórica de los resultados de la MMM agregados para el total nacional, y desagregados de acuerdo al CIIU adaptada a Colombia, son publicados en el boletín de estadística de la página Web del DANE a partir de Enero de 2001 y empalme de los índices para el total nacional de la serie anterior desde 1980 a 2001. Los índices son actualizados con base en la Encuesta Anual Manufacturera, como marco de referencia.

Específicamente la MMM incluye dos variables referentes a la evolución de las ventas, que se definen a continuación:

- Ventas nominales: valor de las ventas totales de los productos y subproductos elaborados por el establecimiento, además de las transferencias recibidas de otros establecimientos de la misma empresa.
- Ventas reales: valor nominal de las ventas deflactadas por el índice de precios al productor según clase industrial.

Como ya se ha afirmado la MMM publica su información por medio de índices con año base 2001, cuya construcción se define a continuación:

$$\hat{I}_{(i)y} = \frac{t_{(i)y}}{t_{(i0)y}}$$

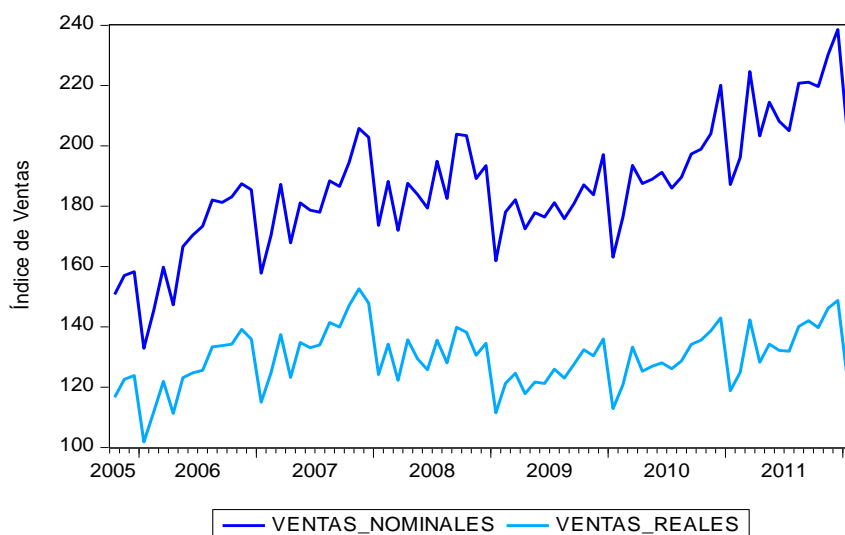
Dónde:

$\hat{I}_{(i)y}$: Índice nacional total de la variable y en el sector industrial, estimado para el mes i .

$t_{(i)y}$: Total nacional de la variable y en el sector industrial, estimado para el mes i .

$t_{(i0)y}$: Total nacional de la variable y en el sector industrial, estimado para el mes $i0$ (periodo base).

A continuación se presenta la gráfica de las series de Ventas Agregadas Nominales y Reales para el periodo Octubre de 2005 a Enero 2012:



Gráfica 1. Índices de Ventas Reales y Nominales del Sector Industrial

Los índices representan el comportamiento en el tiempo de las ventas mensuales del sector industrial. Su cálculo respecto al año base solo representa un cambio de escala que busca mantener la confidencialidad de los datos, por lo tanto no afecta la calidad ni la precisión con la que los índices representan el comportamiento de las ventas. En consecuencia, estos índices reflejan la realidad nacional y son utilizados por las entidades encargadas de hacer proyecciones de las ventas del sector industrial como el DANE, la ANDI, Planeación Nacional y Banco de la República.

Para el Desarrollo del presente trabajo, las ventas reales constituyen la serie de tiempo histórica por las siguientes razones:

- El interés de este estudio es pronosticar cambios en la demanda de productos industriales sin incluir cambios por efectos inflacionarios.

- La serie de ventas reales se encuentra deflactada por el índice de precios al productor, es decir que representa únicamente el cambio en las cantidades vendidas mas no cambio en los precios.
- La naturaleza de la serie de Expectativas de Empresarios, esta expresada en términos del cambio en las cantidades vendidas.

1.2. PREPARACIÓN DE DATOS

1.2.1. Crecimiento de las Ventas Reales

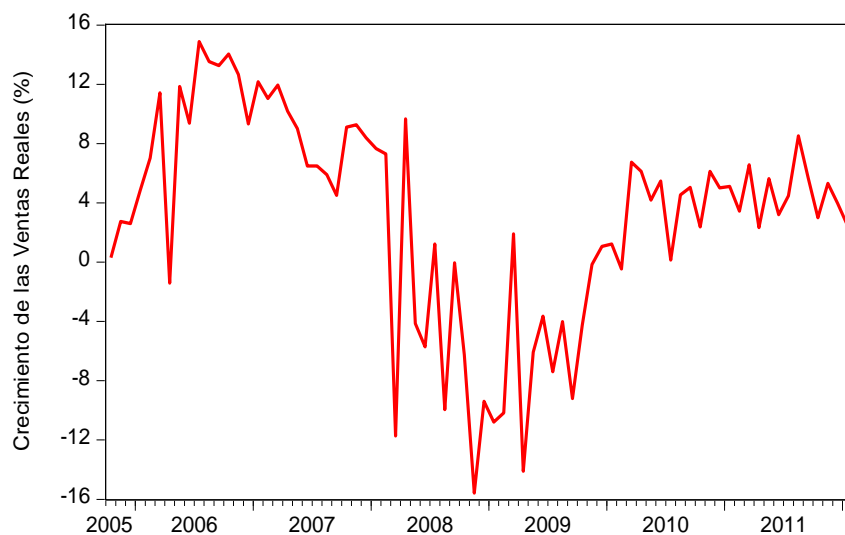
La naturaleza de la serie de expectativas de empresarios está determinada por la pregunta No.2 del cuestionario de la EMEE, la cual se expresa en términos del crecimiento que esperan los empresarios de las ventas para el sector industrial dentro del próximo año. Es por esta razón que es necesario calcular el crecimiento anual para el histórico de las ventas, con fin de hacer comparable la naturaleza de las dos series.

Sea v_t el índice de ventas reales del sector industrial en el momento t , el cálculo de x_t crecimiento porcentual anual de las ventas mensuales del sector industrial en el momento t será:

$$x_t = \ln v_t - \ln v_{t-1}$$

Donde la muestra de v_t se compone del periodo Octubre de 2004 a Enero de 2012, con el fin de unificar el tamaño de la base de datos con la de las expectativas.

La gráfica 2 Muestra el comportamiento de la serie calculada para el Crecimiento de las Ventas Reales.



Gráfica 2. Crecimiento de las Ventas Reales

1.2.2. Expectativas de Empresarios

La EMEE presenta cuatro series para la respuesta a la pregunta No 2.

1. Porcentaje de empresarios que respondieron que el crecimiento de las ventas crecería: *Mayor*
2. Porcentaje de empresarios que respondieron que el crecimiento de las ventas descendería: *Menor*
3. Porcentaje de empresarios que respondieron que el crecimiento de las ventas permanecería contante: *Igual*
4. Diferencia entre el porcentaje *Mayor* y *Menor* : *Balance*

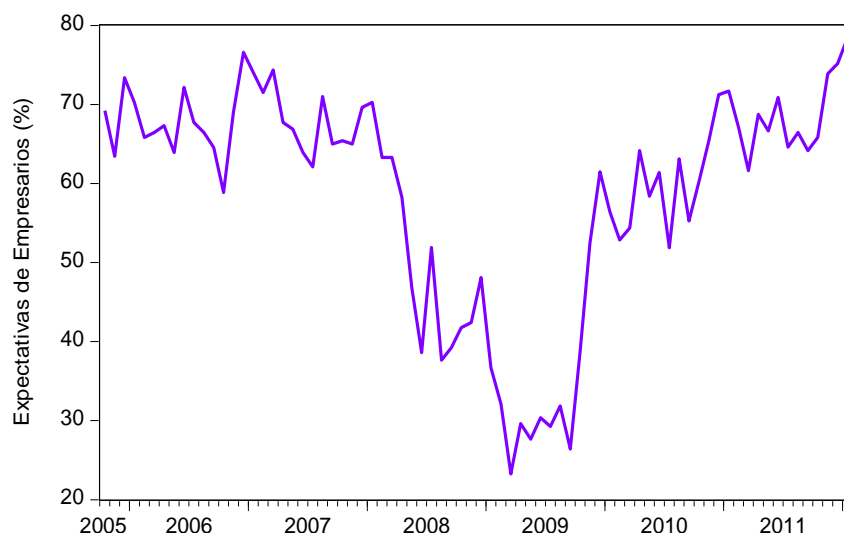
Con el fin de identificar la serie de Expectativas de empresarios, se realizó una tabla de correlación entre cada una de las cuatro categorías de respuesta y la serie de ventas reales presentada a continuación:

	<i>Mayor</i>	<i>Menor</i>	<i>El mismo</i>	<i>Balance</i>	x_t
<i>Mayor</i>	1				
<i>Menor</i>	-0.89268636	1			
<i>El mismo</i>	-0.72809294	0.34102829	1		
<i>Balance</i>	0.98075815	-0.96349387	-0.58025943	1	
x_t	0.72434692	-0.68744646	-0.46528695	0.72809685	1

Tabla 1. Correlaciones EMEE con Ventas reales

La serie que presenta mayor correlación con la serie del crecimiento de las Ventas Reales es el Balance, que adicionalmente presenta altas correlaciones con las demás series, Mayor, Menor y El mismo. Por lo tanto la serie Expectativas de empresarios estará representada por el Balance correspondiente a la pregunta No. 2 de la EMEE. El gráfico 3 muestra el comportamiento de las expectativas de los empresarios.

A continuación se presenta la gráfica de la serie Expectativas de Empresarios:



Gráfica 3. Expectativas de Empresarios

1.2.3. Ventas Reales Desagregadas

Como se había mencionado anteriormente la MMM presenta los índices para las ventas agregadas y desagregadas de acuerdo a la clasificación CIU adaptada a Colombia. Para conocer la participación de cada uno de los Subsectores dentro del sector industrial, es necesario estimar el valor neto en pesos de la muestra mensual, con base en la Encuesta Anual Manufacturera (EAM).

La EAM¹ es realizada y publicada por el DANE en su página Web, y consiste en una investigación estadística de carácter censal que determina el comportamiento de todas las variables básicas del sector manufacturero. A partir de la EAM es posible identificar el promedio de las ventas para el sector industrial y de cada subsector del año base de la MMM (2001) de acuerdo a la correlativa entre las clases del CIU presentadas en la EAM y en la MMM presentada en el Anexo3, y así despejar el estimador de las ventas reales mensuales para cada subsector, bajo la misma notación del cálculo de los índices, este procedimiento es:

$$t_{(i)j} = \hat{I}_{(i)j} * t_{(i0)j}$$

¹ Tomado y adaptado de (Departamento Administrativo Nacional de Estadística DANE, 2009)

Dónde:

$t_{(i)j}$: Total nacional del subsector j del sector industrial, estimado para el mes i .

$\hat{I}_{(i)j}$: Índice nacional total del subsector j del sector industrial, estimado para el mes i .

$t_{(i0)j}$: Total nacional del subsector j del sector industrial, estimado para el mes $i0$ (periodo base).

1.2.4. Datos de Entrenamiento y Datos de Prueba

Con el fin probar la eficiencia de los modelos de pronóstico, la base de datos completa que se compone de 76 observaciones para el periodo Octubre de 2005 a Enero de 2012, es dividida en dos subconjuntos, datos de entrenamiento y datos de prueba.

Los datos entrenamiento se componen de 60 observaciones correspondientes al periodo entre octubre de 2005 y septiembre de 2010, este subconjunto será utilizado para la identificación y estimación de todos los modelos que se realizarán en el presente trabajo.

Los datos de prueba se componen de 16 observaciones correspondientes al periodo Octubre 2010 a enero de 2012, este subconjunto será utilizado para la validación de los modelos y así mismo el ajuste que se presente con estos datos, permitirá concluir y validar las metodologías de pronóstico utilizadas en el presente trabajo, por lo tanto solo serán utilizados hasta la etapa de resultados.

1.2.5. Resumen Base de Datos

La base de datos completa para la realización de este trabajo se compone de tres tipos de series: las expectativas de empresarios, crecimiento de las ventas reales, y las ventas reales desagregadas (en miles de pesos) por cada subsector de acuerdo a la clasificación CIIU. A continuación de se presenta la tabla resumen de las series:

RESUMEN BASE DE DATOS			
Nombre de la serie	Expectativas empresarios	Crecimiento Ventas reales	Ventas Desagregadas por subsector
Publicación	Banco de la Republica	DANE	DANE
Fuente	EMEE	MMM	MMM
Periodo	Oct 2005 – Ene 2012	Oct 2005 – Ene 2012	Oct 2005 – Sept 2011
Variable	Balance	Crecimiento porcentual anual del Índice de Ventas reales	Ventas en pesos
No. de Observaciones	60 datos de entrenamiento y 16 datos de prueba	60 datos de entrenamiento y 16 datos de prueba	69 por cada subsector.
Periodicidad	Mensual	Mensual	Mensual
Notación	y_t	x_t	v_{tj}
			47 Subsectores para la MMM

Tabla 2. Resumen Base de Datos

MODELO BOX-JENKINS

La metodología Box-Jenkins presenta modelos estadísticos de series de tiempo que expresan la serie en función de su propio pasado, donde este provee toda la información relevante del comportamiento futuro del proceso aleatorio que la genera (Montenegro García, 2011).

En este apartado se pretende establecer un marco de referencia para el pronóstico de la serie del crecimiento de las ventas únicamente en función de su propio pasado, por medio de la identificación del mejor modelo a partir de la Metodología Box-Jenkins, bajo un diagnóstico ordinario univariado (Ljung-Box, AIC, BIC) y el criterio del mínimo error relativo al pronóstico ingenuo.

1.3. SELECCIÓN DE MODELOS ARMA(p,q)

En primer lugar para realizar la identificación del modelo, se grafica el coeficiente de autocorrelación (ACF) y autocorrelación parcial (PACF) de la serie, con el fin de identificar el número de rezagos autoregresivos y rezagos de promedio móvil que pueden ser relevantes en el modelo. A través del procedimiento ARIMA del software (SAS® 9.1), se presenta a continuación las gráficas de autocorrelación y autocorrelación parcial en la Ilustración 1.

El estadístico Q **1** calculado es 161.79 y 205.86 para seis y doce rezagos respectivamente, los cuales deben ser comparados respecto a la distribución chi-cuadrado. El valor-p para los estadísticos es menor a 0.001 concluyendo que no hay significancia estadística para afirmar que la serie es ruido blanco.

A partir de la ACF y PACF, se evidencia que no son cero después de varios rezagos, indicando hasta siete **2** y seis **3** rezagos significativos para un modelo AR y MA respectivamente, sin embargo utilizando esta herramienta de identificación no es posible establecer con certeza algún modelo a utilizar.

Para definir los modelos que pueden tener el mejor ajuste a la serie y como complemento al análisis basado en la observación de la ACF y PACF, es posible utilizar los procedimientos MINIC, ESACF y SCAN del software (SAS® 9.1).

The ARIMA Procedure

Name of Variable = x
 Mean of Working Series 0.026695
 Standard Deviation 0.078628
 Number of Observations 60

Lag	Covariance	Correlation	Autocorrelations													Std Error								
			-1	9	8	7	6	5	4	3	2	1	0	1	2		3	4	5	6	7	8	9	1
0	0.0061824	1.00000																						0
1	0.0040075	0.64822																						0.129099
2	0.0045125	0.72989																						0.175136
3	0.0044303	0.71660																						0.220070
4	0.0033990	0.54979																						0.256024
5	0.0040185	0.65000																						0.274998
6	0.0032435	0.52464																						0.299511
7	0.0025038	0.40499																						0.314455
8	0.0029311	0.47410																						0.323031
9	0.0019791	0.32013																						0.334428
10	0.0016681	0.26982																						0.339496
11	0.0013124	0.21228																						0.343052
12	0.00028458	0.04603																						0.345234
13	0.00061320	0.09918																						0.345336
14	0.00003767	0.00609																						0.345811
15	-0.0001768	-0.02860																						0.345813

"," marks two standard errors

Lag	Correlation	Inverse Autocorrelations																						
		-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1		
1	0.07460																							*
2	-0.35528																							*
3	-0.15941																							*
4	0.06449																							*
5	-0.15744																							*
6	-0.01041																							*
7	0.17325																							*
8	-0.11778																							*
9	-0.23853																							*
10	-0.04047																							*
11	0.13539																							*
12	0.21633																							*
13	0.03711																							*
14	-0.02367																							*
15	-0.08979																							*

Lag	Correlation	Partial Autocorrelations																						
		-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1		
1	0.64822																							*
2	0.53415																							*
3	0.36187																							*
4	-0.20209																							*
5	0.10101																							*
6	-0.02614																							*
7	-0.30574																							*
8	0.00916																							*
9	0.02096																							*
10	-0.17120																							*
11	-0.23320																							*
12	-0.18950																							*
13	0.12006																							*
14	0.08131																							*
15	0.15800																							*

Autocorrelation Check for White Noise

To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	161.79	6	<.0001	0.648	0.730	0.717	0.550	0.650	0.525
12	205.86	12	<.0001	0.405	0.474	0.320	0.270	0.212	0.046

Ilustración 1. Salida de ACF y PACF del Crecimiento de las Ventas

A continuación se definen los criterios que utilizan los algoritmos en los procedimientos de (Dickey & Brocklebank, 2003) para la identificación del modelo ARMA(p,q), presentados en la Salida 2:

- **Mínimo Criterio de Información (MINIC):** identifica el orden de un proceso ARIMA estacionario e invertibles por medio del cómputo de diferentes valores de p y q con los cuales identifica los que presentan un mejor criterio de información.
- **Función de Autocorrelación Parcial de Muestra Extendida (ESACF):** identifica el orden de un proceso ARIMA estacionario o no estacionario, basado en la estimación de mínimos cuadrados para diferentes valores de p y q.
- **Mínima Correlación Canónica (SCAN):** identifica el orden de un proceso ARIMA estacionario o no estacionario, basado en el cálculo de los valores propios de la matriz de correlaciones.

Minimum Information Criterion							
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5	MA 6
AR 0	-5.2726	-5.24785	-5.25626	-5.24892	-5.2245	-5.30043	-5.29088
AR 1	-5.76882	-5.93591	-5.9554	-5.90352	-5.88529	-5.94366	-5.93435
AR 2	-6.02767	-6.01891	-5.95205	-5.91746	-5.92904	-5.93799	-5.89052
AR 3	-6.15951	-6.09134	-6.04086	-5.98113	-5.938	-5.90949	-5.86031
AR 4	-6.15691	-6.10783	-6.04057	-6.00975	-5.94302	-5.89115	-5.84451
AR 5	-6.11855	-6.10693	-6.05558	-5.98833	-5.95386	-5.8921	-5.86068
AR 6	-6.05936	-6.05725	-6.03411	-5.98547	-5.92915	-5.92954	-5.86638
AR 7	-6.08431	-6.0747	-6.03293	-5.96471	-5.89932	-5.96705	-5.91372

Error series model: AR(12)
Minimum Table Value: BIC(3,0) = -6.15951

④

Extended Sample Autocorrelation Function							
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5	MA 6
AR 0	0.6482	0.7299	0.7166	0.5498	0.6500	0.5246	0.4050
AR 1	-0.5910	0.1398	0.2642	-0.3410	0.3472	0.0277	-0.2266
AR 2	-0.5103	0.2167	0.2730	-0.2847	0.3594	0.0305	-0.2591
AR 3	0.4467	0.0307	-0.0272	0.1638	0.2464	0.1725	0.1184
AR 4	0.3801	0.1551	0.0054	0.2809	0.1214	-0.0730	-0.0379
AR 5	0.2616	0.0387	0.0115	0.3400	0.2327	-0.0288	-0.0760
AR 6	-0.0871	0.3218	-0.4233	0.1468	-0.0712	0.0661	0.0107
AR 7	0.1010	0.0930	-0.4015	-0.2369	-0.1945	0.0488	0.0416

ESACF Probability Values							
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5	MA 6
AR 0	<.0001	<.0001	0.0011	0.0318	0.0181	0.0798	0.1978
AR 1	<.0001	0.4181	0.1054	0.0511	0.0572	0.8844	0.2480
AR 2	0.0001	0.2475	0.0442	0.0482	0.0538	0.8718	0.1167
AR 3	0.0007	0.8431	0.8623	0.2876	0.1069	0.3544	0.5169
AR 4	0.0045	0.3945	0.9737	0.0925	0.4599	0.6796	0.8039
AR 5	0.0523	0.8129	0.9440	0.0793	0.1487	0.8660	0.6548
AR 6	0.5221	0.0187	0.0054	0.4491	0.7631	0.8005	0.9670
AR 7	0.4623	0.5062	0.0067	0.1620	0.2750	0.8552	0.8283

ARMA(p+d,q) Tentative Order Selection Tests
---ESACF---

p+d	q
3	1
4	1
1	3
7	3
0	5

(5% Significance Level)

⑤

Squared Canonical Correlation Estimates

Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5	MA 6
AR 0	0.4215	0.5349	0.5165	0.3050	0.4288	0.2862	0.1721
AR 1	0.2874	0.0202	0.0541	0.1426	0.1089	0.0004	0.0745
AR 2	0.1303	0.1260	0.1426	0.0726	0.1641	0.0901	0.0872
AR 3	0.0445	0.0004	0.0003	0.0578	0.0591	0.0200	0.0146
AR 4	0.0096	0.0003	<.0001	0.0737	0.0087	0.0026	0.0073
AR 5	0.0008	0.0121	0.0373	0.0716	0.0049	0.0116	0.0225
AR 6	0.1025	0.0827	0.0652	0.0042	0.0154	0.0056	0.0001
AR 7	0.0011	0.0010	0.0327	0.0169	0.0131	0.0052	<.0001

SCAN Chi-Square[1] Probability Values							
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5	MA 6
AR 0	<.0001	<.0001	0.0008	0.0323	0.0184	0.0816	0.2007
AR 1	<.0001	0.4120	0.1819	0.0350	0.0965	0.9168	0.2004
AR 2	0.0044	0.0171	0.0320	0.0633	0.0571	0.1076	0.1175
AR 3	0.1072	0.8964	0.9142	0.1393	0.1228	0.3935	0.5254
AR 4	0.4633	0.9075	0.9685	0.0696	0.5881	0.7636	0.5955
AR 5	0.8341	0.4217	0.2455	0.1238	0.7146	0.5552	0.4149
AR 6	0.0157	0.0800	0.1835	0.7132	0.4519	0.7320	0.9624
AR 7	0.8094	0.8252	0.2582	0.4844	0.6095	0.7001	0.9725

ARMA(p+d,q) Tentative Order Selection Tests

----SCAN----

p+d	q
3	1
2	3
1	4
0	5
7	0

(5% Significance Level)

Ilustración 2. Salida de la Identificación del Modelo: MIMIC – ESACF - SCAN

Dados los resultados de los algoritmos presentados en la Ilustración 2 en ④, ⑤ y ⑥, los Modelos a estimar son:

	MINIC	ESACF	SCAN
ARMA(3,0)		ARMA(3,1)	ARMA(3,1)
		ARMA(4,1)	ARMA(2,3)
		ARMA(1,3)	ARMA(1,4)
		ARMA(7,3)	ARMA(0,5)
		ARMA(0,5)	ARMA(7,0)

Tabla 3. Modelos ARMA Seleccionados Para la Estimación

1.4. ESTIMACIÓN DE PARÁMETROS DE MODELOS SELECCIONADOS

Los modelos ARMA(p, q) conforman un proceso aleatorio donde el valor contemporáneo de la variable se encuentra en función de una constante, sus propios rezagos y los rezagos del término de error, representado en la siguiente ecuación:

$$x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_p x_{t-p} + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \dots + \alpha_q e_{t-q} + e_t$$

Para la estimación de los parámetros de cada uno de los modelos seleccionados se utilizó el software (E-views® 7), los resultados se presentan en la Tabla 4, adicionalmente se presenta la desviación estándar ubicada entre paréntesis, el coeficiente de determinación (R^2) y el coeficiente de determinación ajustado (R^2 Ajustado).

Todos los modelos seleccionados presentan adecuada bondad de ajuste superando en todos los casos 60% de la variación explicada por el modelo, sin embargo los de mejor bondad de ajuste son ARMA (7,3), ARMA (0,5) y ARMA (7,0). Adicionalmente se evidencia que los parámetros de promedio móvil tienden a ser significativos en la mayoría de los modelos respecto a los parámetros autoregresivos.

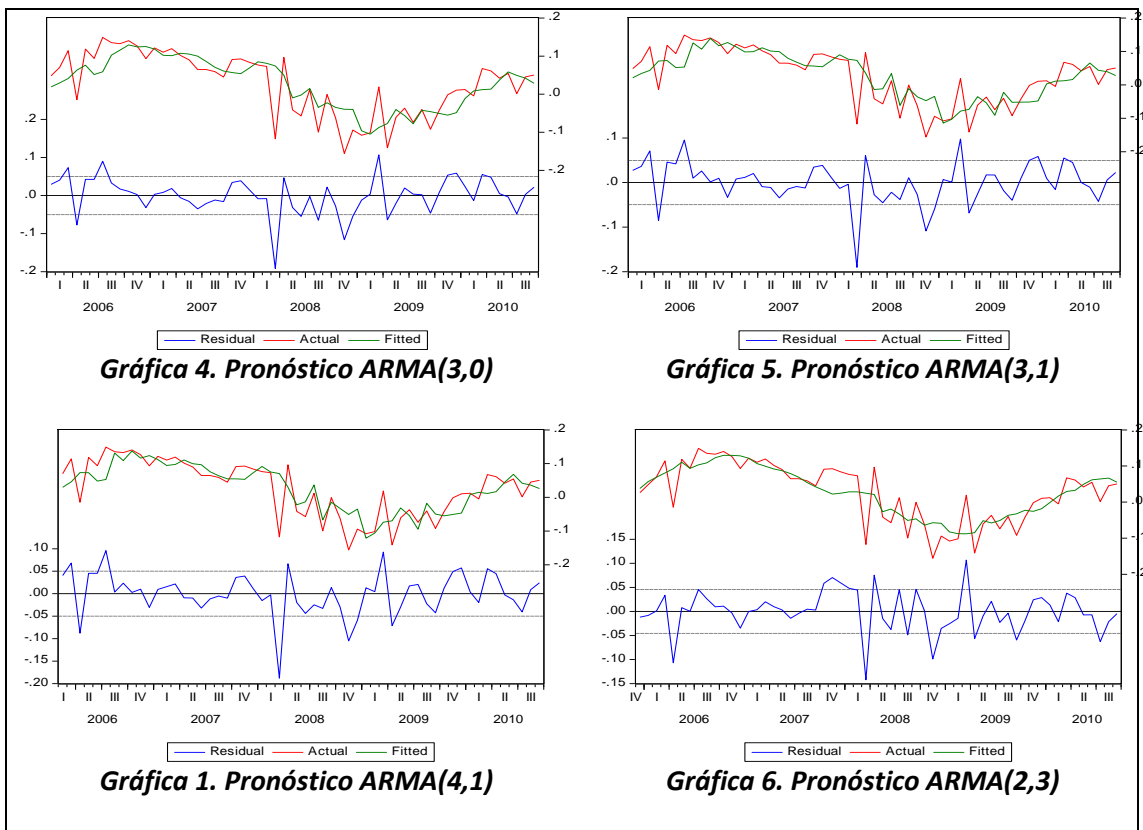
MODELO	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	α_1	α_2	α_3	α_4	α_5	R^2	R^2 Ajustado
ARMA(3,0)	0.0034 (0.0071)	0.1063 (0.1284)	0.4271 (0.1149)	0.3635 (0.1287)										0.6421	0.6219
ARMA(3,1)	0.0053 (0.0100)	-0.2645 (0.2335)	0.5309 (0.1251)	0.5738 (0.1384)					0.4373 (0.2730)					0.6578	0.6315
ARMA(4,1)	0.0042 (0.0091)	-0.0630 (0.5559)	0.5449 (0.1297)	0.4963 (0.2525)	-0.1282 (0.2639)				0.2592 (0.5582)					0.6615	0.6276
ARMA(2,3)	0.0002 (0.0001)	1.9247 (0.0487)	-0.9409 (0.0500)						-1.9340 (0.1479)	1.1737 (0.2595)	-0.2397 (0.1582)			0.7078	0.6797
ARMA(1,3)	0.0024 (0.0041)	0.9289 (0.0646)							-0.7973 (0.1506)	0.3866 (0.1673)	-0.0308 (0.1488)			0.6306	0.6032
ARMA(1,4)	0.0022 (0.0041)	0.9307 (0.0635)							-0.9411 (0.1349)	0.7367 (0.1622)	-0.5756 (0.1614)	0.3579 (0.1326)		0.6378	0.6036
ARMA(7,3)	0.0015 (0.0174)	-0.1089 (0.3048)	0.0153 (0.1765)	0.1082 (0.1560)	0.3501 (0.1195)	0.6089 (0.1139)	0.0736 (0.2054)	-0.3581 (0.1467)	0.3459 (0.3189)	0.7690 (0.1064)	0.6399 (0.3007)			0.7792	0.7267
ARMA(0,5)	0.0297 (0.0222)								0.3905 (0.0967)	0.7285 (0.1015)	0.9221 (0.0724)	0.3585 (0.0868)	0.7741 (0.0882)	0.7113	0.6846
ARMA(7,0)	0.0031 (0.0068)	0.2165 (0.1365)	0.5026 (0.1355)	0.2833 (0.1499)	-0.1126 (0.1551)	0.2638 (0.1496)	0.0140 (0.1367)	-0.3137 (0.1337)						0.7256	0.6829

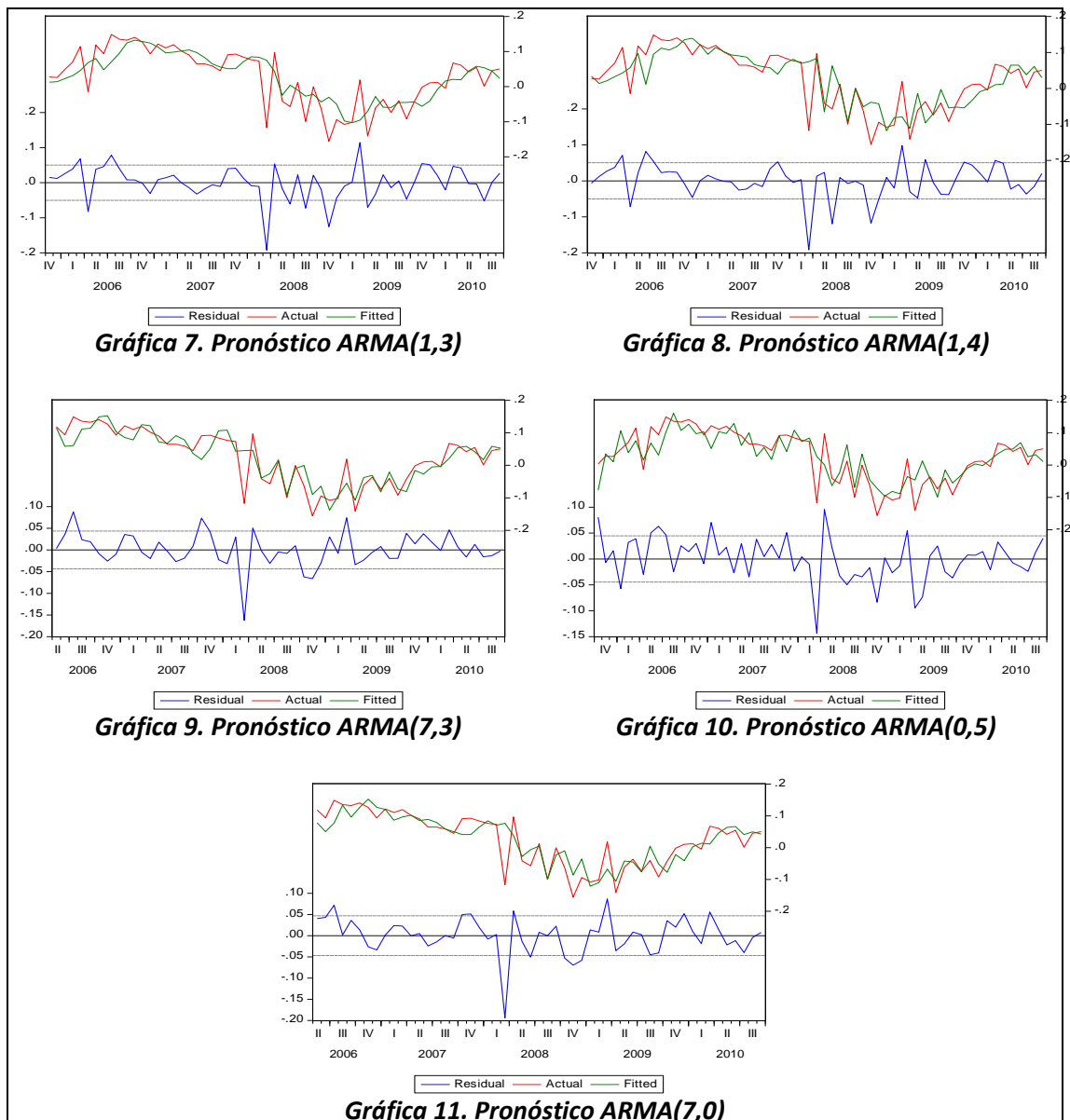
Tabla 4. Estimación de Modelos Arma Seleccionados

1.5. PRONÓSTICOS DE MODELOS ARMA SELECCIONADOS

Así como la estimación de los parámetros de los modelos seleccionados, los pronósticos de los mismos se calcularon utilizando el comando “*forecast - Static forecast*” en el software (E-views® 7). La gráfica de la serie pronosticada comparada con la serie real y la gráfica de los residuos se presentan a partir del Gráfico 4 hasta el Gráfico 12 para todos los modelos seleccionados.

Como ya se había mencionado en la sección de estimación, el ajuste de los modelos es muy bueno, sin embargo se evidencia en la gráficas de residuales que para todos los casos hay puntos donde superan significativamente las dos desviaciones estándar, basado únicamente en la observación de las gráficas, esto ocurre cuando hay cambios drásticos en la serie. Por ejemplo para marzo de 2006 y el periodo Diciembre de 2007 y Junio de 2009, para el primer caso se experimenta un punto de inflexión y para el segundo caso se experimenta alta volatilidad generando que la inercia de la serie no sea suficiente para mantener el residuo dentro de las dos desviaciones estándar. Esta problemática, genera la necesidad de incluir variables exógenas al modelo, de tal forma que sea posible pronosticar puntos de inflexión, garantizando un mejor ajuste.





1.6. EVALUACIÓN DE PRONÓSTICOS

Con el fin de seleccionar el mejor modelo para la estimación del crecimiento de las ventas a partir de los rezagos de la serie, se utilizan los Criterios basados en errores relativos para la evaluación del pronóstico de los modelos seleccionados siguiendo los resultados presentados en (Hyndman & Koehler, 2006), debido a que la serie al representar un cambio porcentual tiene valores alrededor del cero.

El Criterio de errores relativos, se calcularon respecto al error del pronóstico ingenuo F_t (paseo aleatorio donde F_t es igual a la última observación de la serie) así:

$$r_t = \frac{e_t}{e_t^f} * 100$$

Dónde r_t el error relativo, e_t es el error de pronóstico del modelo, y e_t^f es el error del pronóstico ingenuo.

Se define:

- Media del Error Absoluto Relativo:
 $MRAE = E(|r_t|)$
- Mediana del Error Absoluto Relativo:
 $MDRAP = Mediana(|r_t|)$
- Media Geométrica del Error Absoluto Relativo:
 $GMRAE = Media Geométrica(|r_t|)$

Cada criterio representa la relación del pronóstico obtenido a través del modelo evaluado respecto al pronóstico ingenuo de forma conjunta para toda la muestra con el fin de evaluar la precisión del modelo. Para cada criterio superior a 100, este indicara que el pronóstico del modelo tiene una precisión menor que el pronóstico ingenuo, en caso contrario si es menor que 100 indica que el pronóstico del modelo supera la precisión del pronóstico ingenuo. Se busca el modelo que tenga menor error relativo respecto al pronóstico ingenuo.

El ranking de los modelos seleccionados de acuerdo a estos criterios se presenta en la tabla 5 continuación:

	MRAE		MDRAP		GMRAE	
1	ARMA(2,3)	1.301.834	ARMA(3,1)	74.580	ARMA(3,0)	71.077
2	ARMA(4,1)	3.436.518	ARMA(7,3)	80.366	ARMA(1,4)	74.200
3	ARMA(7,3)	4.212.420	ARMA(4,1)	83.474	ARMA(0,5)	87.405
4	ARMA(3,1)	4.269.309	ARMA(1,4)	84.460	ARMA(2,3)	88.457
5	ARMA(7,0)	4.514.680	ARMA(0,5)	88.267	ARMA(7,3)	89.024
6	ARMA(1,3)	5.235.486	ARMA(2,3)	89.397	ARMA(3,1)	94.970
7	ARMA(3,0)	6.101.034	ARMA(7,0)	94.912	ARMA(4,1)	100.601
8	ARMA(1,4)	6.485.813	ARMA(1,3)	97.307	ARMA(1,3)	102.866
9	ARMA(0,5)	10.942.650	ARMA(3,0)	107.099	ARMA(7,0)	112.373

Tabla 5. Evaluación de los Modelos ARMA Seleccionados

Antes de escoger el modelo a implementarse para pronosticar el crecimiento de las ventas del sector industrial en Colombia, es importante aclarar los resultados presentados en la Tabla 5. El criterio MRAE presenta resultados que indican un gran error en los modelos propuestos respecto al pronóstico ingenuo, sin embargo para este caso específico el criterio puede no ser un buen indicador debido a que por ejemplo en el mes de julio de 2007 el crecimiento de las ventas permaneció casi igual que el crecimiento del periodo exactamente anterior provocando un error absoluto relativo al pronóstico ingenuo muy alto, lo que ocurre varias veces en el periodo de referencia. En consecuencia para este caso es recomendable prestar mayor

atención en los criterios MDRAP y GMRAE que le dan una importancia menor a los valores extremos.

Los modelos que tienen el mejor comportamiento respecto los Criterios de Errores Relativos son AR(3), ARMA(2,3) , ARMA(3,1). Comparando este resultado con la Tabla, se evidencia que el modelo ARMA(3,1) representa el mejor modelo sugerido por los procedimientos ESCAF y SCAN, por lo tanto es posible concluir que este modelo será el seleccionado como marco de referencia para pronosticar la serie del crecimiento de las ventas utilizando únicamente la información del pasado de la misma.

El modelo estadístico bajo la metodología Box-Jenkins implementado para pronosticar el crecimiento de las ventas del sector industrial en Colombia en función del pasado de la serie es **ARMA (3,1)** cuya ecuación general y ecuación estimada se representan continuación:

$$x_t = \mu + \hat{\beta}_1 x_{t-1} + \hat{\beta}_2 x_{t-2} + \hat{\beta}_3 x_{t-3} + \varepsilon_t + \gamma_1 \varepsilon_{t-1}$$

$$x_t = 0.0053 - 0.2646x_{t-1} + 0.5309x_{t-2} + 0.5738x_{t-3} + \varepsilon_t + 0.4373\varepsilon_{t-1}$$

MODELO FILTRO DE KALMAN

La implementación de modelos estadísticos que integran la información histórica con las expectativas de los empresarios conduce a pronósticos más precisos en el corto plazo (Goodwin P. , 2002). El Filtro de Kalman es implementado en la presente investigación como una metodología que integra las expectativas de los empresarios con el histórico de las ventas reales, con el fin de pronosticar el crecimiento de las ventas del sector industrial en Colombia. (Rahiala & Teräsvirta, 1993), (Öller & Tallbom, 1993) y (Dahllöf, 2003) implementan esta misma metodología de integración con el objetivo de pronosticar el crecimiento de la producción de Suecia y Finlandia, partiendo del supuesto de que la expectativas de los empresarios son capaces de incorporar toda la información que explica la variable de interés que la serie histórica es incapaz de proveer.

El Filtro de Kalman es un método de estimación cuyos parámetros se corrigen en cada iteración dependiendo del error de predicción que se haya cometido (Montenegro, 2005). Este método matemático opera por medio de un mecanismo de predicción y corrección, en el cual se hace la estimación del nuevo estado en función de su predicción previa y la inclusión de un término de corrección proporcional al error de predicción, el cual se minimiza estadísticamente; el método incorpora los supuestos de normalidad sobre los vectores de Estado Inicial y Perturbaciones del sistema lo cuales hacen posible calcular la Función de Máxima Verosimilitud sobre el error de predicción, permitiendo la estimación de los parámetros no conocidos dentro del sistema (Solera, 2003).

En consecuencia, el presente apartado pretende estimar el crecimiento de las ventas del sector industrial en Colombia, involucrando la información que ofrece la serie de Ventas Reales, y las Expectativas de los Empresarios, por medio del método del Filtro de Kalman. En primer lugar se hará una revisión teórica de la metodología para proceder a la definición del modelo a implementarse y su respectivo desarrollo.

1.7. MODELO TEÓRICO

Siguiendo el marco teórico y notación de (Harvey, 2003), a continuación se explica el Filtro de Kalman, partiendo desde la notación estado-espacio hasta la explicación detallada del Filtro de Kalman

1.7.1. Representación Estado-Espacio

La notación en estado-espacio es una fuerte herramienta para el manejo de un extenso rango de modelos de series de tiempo, y fundamental para el desarrollo de un Filtro de Kalman. La forma general del estado-espacio se describe a continuación:

Ecuación de Medición

Para una Serie de Tiempo Multivariada la ecuación de Medición se define como:

$$\mathbf{y}_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T$$

Donde \mathbf{y}_t es el vector de variables observables (Contiene N elementos), $\boldsymbol{\alpha}_t$ es el vector de variables inobservables que se relacionan con \mathbf{y}_t (Denominado Vector de Estado), \mathbf{d}_t es el conjunto de otras variables que explican a \mathbf{y}_t , y $\boldsymbol{\varepsilon}_t$ es el vector de perturbaciones serialmente no correlacionadas que tiene $E[\boldsymbol{\varepsilon}_t] = 0$ y $Var[\boldsymbol{\varepsilon}_t] = \mathbf{H}_t$ (matriz de covarianzas).

En el caso de una Serie de Tiempo Univariada:

$$y_t = \mathbf{z}'_t \boldsymbol{\alpha}_t + d_t + \varepsilon_t, \quad t = 1, \dots, T$$

Donde y_t es la variable observable (Contiene N=1 Elementos), $\boldsymbol{\alpha}_t$ es el vector de variables inobservables que se relacionan con y_t (Denominado Vector de Estado), d_t es el conjunto de otras variables que explican a y_t , y ε_t es el vector perturbaciones serialmente no correlacionadas que tiene $E[\varepsilon_t] = 0$ y $Var[\varepsilon_t] = h_t$ (matriz de covarianzas).

Ecuación de Transición.

Volviendo al caso Multivariado, el Vector de Estado $\boldsymbol{\alpha}_t$ (comúnmente de variables inobservables) es generado por un proceso Markoviano de primer orden:

$$\boldsymbol{\alpha}_t = \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t \boldsymbol{\eta}_t, \quad t = 1, \dots, T$$

Donde $\boldsymbol{\alpha}_{t-1}$ es el primer rezago de $\boldsymbol{\alpha}_t$, \mathbf{c}_t es el conjunto de otras variables que explican a $\boldsymbol{\alpha}_t$, y $\boldsymbol{\eta}_t$ es el vector de perturbaciones serialmente no correlacionadas que tiene $E[\boldsymbol{\eta}_t] = 0$, y $Var[\boldsymbol{\eta}_t] = \mathbf{Q}_t$ (Matriz de covarianzas).

Finalmente la especificación del Sistema estado-espacio se completa con los siguientes supuestos:

1. El Vector de Estado Inicial $\boldsymbol{\alpha}_0$ tiene una media $E[\boldsymbol{\alpha}_0] = \boldsymbol{\alpha}_0$ y $Var[\boldsymbol{\alpha}_0] = \mathbf{P}_0$ (Matriz de covarianzas).
2. Los términos de perturbación $\boldsymbol{\varepsilon}_t$ y $\boldsymbol{\eta}_t$ no están correlacionados en todos los periodos y no están correlacionados con el Vector de Estado Inicial $\boldsymbol{\alpha}_0$, es decir:

$$E[\boldsymbol{\varepsilon}_t \boldsymbol{\eta}'_s] = 0, \quad \forall s, t = 1, \dots, T$$

$$E[\boldsymbol{\varepsilon}_t \boldsymbol{\alpha}'_0] = 0, \quad E[\boldsymbol{\eta}'_s \boldsymbol{\alpha}'_0] = 0, \quad \forall t = 1, \dots, T$$

El ánimo de la Formulación estado-espacio es definir α_t de tal manera que contenga toda la información relevante del sistema en el momento t con el menor número de elementos posible (parsimonia).

A las matrices Z_t, d_t, H_t (de la Ecuación de Medición), $T_t, c_t, R_t, y Q_t$ (de la Ecuación de Transición) se definen como las Matrices del Sistema. En el caso donde el sistema de matrices no cambia en el tiempo, se dice que el modelo es Invariante en el Tiempo u Homogéneo en el Tiempo. Igualmente a las matrices $Z_t, H_t, T_t, R_t, y Q_t$, se le denomina Hiperparámetros ya que determinan las propiedades estocásticas del modelo, y a d_t y c_t se les denomina parámetros y determinan el valor esperado del Estado y de las observaciones en forma determinística.

1.7.2. Filtro de Kalman

El Filtro de Kalman es un procedimiento que busca el estimador óptimo del vector de estado en el tiempo t , basándose en la información del momento exactamente anterior $t - 1$ (Predicción), o en toda la información disponible hasta dicho momento T (Suavizado). El Filtro de Kalman asume que el Sistema de Matrices y los parámetros a_0 y P_0 son conocidos en todos los periodos.

La derivación del Filtro de Kalman, establece el supuesto de normalidad del Vector de Estado Inicial y de las Perturbaciones, por lo tanto la distribución para el Vector de Estado será normal (combinación lineal de variables normalmente distribuidas). Dado este supuesto es posible calcular la función de verosimilitud sobre el error e predicción y llevar a cabo la estimación de los parámetros no conocidos del sistema.

Forma General

Bajo la Formulación de estado-espacio, considere a_{t-1} como el estimador óptimo de α_{t-1} basado en la información incluida en y_{t-1} , y $P_{t-1} = E[(\alpha_{t-1} - a_{t-1})(\alpha_{t-1} - a_{t-1})']$ como la matriz de covarianza del error de estimación. El estimador óptimo de α_t y la matriz de covarianza del error de estimación están dados por las ecuaciones de predicción:

$$a_{t/t-1} = T_t a_{t-1} + c_t, \quad t = 1, \dots, T$$

$$P_{t/t-1} = T_t P_{t-1} T_t' + R_t Q_t R_t', \quad t = 1, \dots, T$$

Una vez que la nueva observación de y_t está disponible el estimador de α_t , $a_{t/t-1}$ debe actualizarse, de tal manera las Ecuaciones de Actualización se definen como:

$$a_t = a_{t/t-1} + P_{t/t-1} Z_t' F_t^{-1} (y_t - Z_t a_{t/t-1} - d_t)$$

$$P_t = P_{t/t-1} - P_{t/t-1} Z_t' F_t^{-1} Z_t P_{t/t-1}$$

Dónde:

$$F_t = Z_t P_{t/t-1} Z_t' + H_t, \quad t = 1, \dots, T$$

En forma de resumen, la Ilustración 3 disponible en (Misas A., Arango A., & López E., 2005) muestra el funcionamiento del Filtro de Kalman:

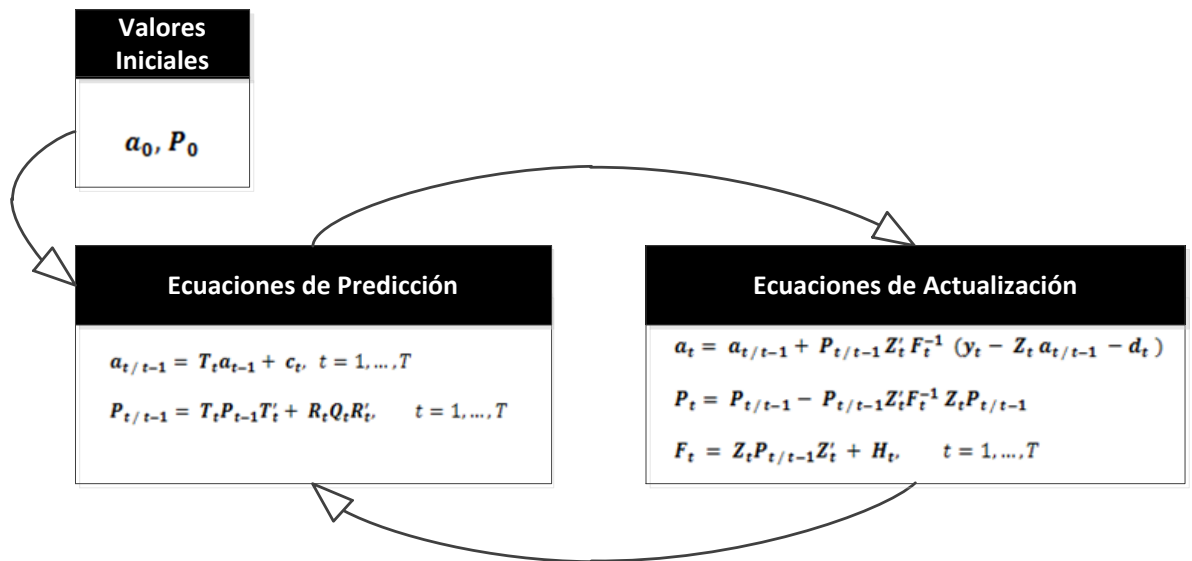


Ilustración 3. Diagrama de Flujo del Filtro de Kalman

De manera alternativa el Filtro de Kalman puede describirse como:

$$a_{t+1|t} = (T_{t+1} - K_t Z_t) a_{t/t-1} + K_t y_t + (c_{t+1} - K_t d_t)$$

Donde

$$K_t = T_{t+1} P_{t/t-1} Z_t' F_t^{-1}, \quad t = 1, \dots, T$$

$$P_{t+1/t} = T_{t+1} (P_{t/t-1} - P_{t/t-1} Z_t' F_t^{-1} Z_t P_{t/t-1}) T_{t+1}' + R_{t+1} Q_{t+1} R_{t+1}'$$

Para el caso de un Modelo Univariado:

$$f_t = z_t' P_{t/t-1} z_t + h_t$$

En este caso la varianza de las perturbaciones son proporcionales a un escalar positivo σ_*^2 , es decir que $Var[\varepsilon_t] = \sigma_*^2 h_t$ y $Var[\eta_t] = \sigma_*^2 Q_t$. Igualmente los valores iniciales también dependerán de σ_*^2 , por lo tanto:

$$P_{t+1/t} = \sigma_*^2 P_{t+1/t}^*$$

$$f_t = \sigma_*^2 f_t^*$$

1.8. DEFINICIÓN DEL MODELO

La evaluación de la metodología de Filtro de Kalman para pronosticar el crecimiento de las ventas a partir de las expectativas de empresarios y la serie de tiempo, compone un Modelo Univariado el cual se expondrá a continuación.

1.8.1. Capacidad Predictiva de las Expectativas de Empresarios

La pregunta No. 2 de la EMEE indaga sobre las expectativas de los empresarios de la siguiente manera:

“El crecimiento del volumen de ventas (cantidades) en los próximos 12 meses, comparado con el crecimiento del volumen de ventas (cantidades) en los pasados 12 meses, se espera sea: Mayor, Menor, o El mismo”

Específicamente la pregunta indaga sobre la expectativa que tiene el empresario acerca del crecimiento del volumen de ventas en los próximos doce meses, dado que conoce sus doce meses pasados, por lo tanto representa un pronóstico del crecimiento de las cantidades de las ventas doce meses adelante. Por ejemplo, los resultados de la pregunta No. 2 para el mes de septiembre del año 2005, tiene una predictibilidad de doce meses ubicándola en el mes de septiembre del año 2006.

La naturaleza de la pregunta ubica las Expectativas de los Empresarios doce meses adelante del momento de la generación de la misma, situación que ocasiona una pérdida 12 periodos en la base de datos de prueba.

1.8.2. Identificación de Variables Exógenas

Las ecuaciones de medición y transición del modelo se componen de un vector de variables exógenas que afectan al Vector de Estado c_t y al Vector de variables observadas d_t respectivamente. Para esta investigación cada uno de estos vectores se compone de rezagos de las expectativas de empresarios y de la serie de Crecimiento de las ventas Reales.

Con el fin de seleccionar las variables explicativas se utilizó el procedimiento REG, MODEL SELECTION del software (SAS® 9.1), al cual se declara todas las posibles variables a introducir en el modelo, los niveles de significancia para la prueba F y los criterios de ordenamiento de los modelos seleccionados.

Los criterios de ordenamiento de los modelos seleccionados son:

1. Mínimo criterio de información de Mallows Cp.
2. Mínimo criterio de información bayesiano BIC.
3. Mayor coeficiente de determinación R^2 y R^2 ajustado.

Variables exógenas para la ecuación de medición

Para realizar la modelación de las expectativas inicialmente se realiza el Correlograma de la serie con el fin de identificar si deben incluirse rezagos de la propia variable. El Correlograma es presentado en la ilustración 4, y evidencia que la ACF ① tiene una caída lenta y en la PACF ② el primer rezago es significativo indicado un modelo autoregresivo AR(1).

```

The SAS System      22:24 Wednesday, April 15, 2012  1

The ARIMA Procedure
Name of Variable = y

Mean of Working Series  0.561575
Standard Deviation     0.149929
Number of Observations 57

Autocorrelations

Lag  Covariance  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1  Std Error

0    0.022479    1.00000    |          | ***** | ①    0
1    0.020301    0.90312    |          | ***** |    0.132453
2    0.018806    0.83660    |          | ***** |    0.214855
3    0.017118    0.76154    |          | ***** |    0.265933
4    0.015780    0.70202    |          | ***** |    0.301777
5    0.014473    0.64384    |          | ***** |    0.329183
6    0.012976    0.57728    |          | ***** |    0.350580
7    0.011131    0.49519    |          | ***** |    0.366878
8    0.0096331   0.42855    |          | ***** |    0.378423
9    0.0080307   0.35726    |          | ***** |    0.386843
10   0.0066495   0.29581    |          | ***** |    0.392589
11   0.0060373   0.26858    |          | ***** |    0.396480
12   0.0045801   0.20375    |          | ***** |    0.399659
13   0.0029562   0.13151    |          | ***** |    0.401478
14   0.0010381   0.04618    |          | ***** |    0.402233

"." marks two standard errors

Partial Autocorrelations

Lag  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

1    0.90312    |          | ***** | ②
2    0.11375    |          | **      |
3   -0.05937    |          | *       |
4    0.03334    |          | *       |
5   -0.00453    |          |         |
6   -0.08078    |          | **      |
7   -0.13663    |          | ***     |
8    0.01190    |          |         |
9   -0.05380    |          | *       |
10   -0.02053    |          |         |
11    0.16826    |          | ***     |
12   -0.18101    |          | ****    |
13   -0.15527    |          | ***     |
14   -0.12666    |          | ***     |
    
```

Ilustración 4. ACF y PACF de las Expectativas de empresarios

Para la realización del procedimiento REG, la variable dependiente se declaró como las Expectativas de Empresarios y como variables independientes el primer rezago de las expectativas de Empresarios (Y1) y doce rezagos de la serie Crecimiento de las Ventas Reales (X1, X2, X3, ... , X12) . Los resultados del procedimiento se presentan en la Ilustración 5. El primer y segundo modelo ③ presentado en la lista presenta un buen coeficiente de determinación y un buen Cp, por lo tanto estos dos modelos serán probados en la etapa de estimación para la ecuación de medición.

The SAS System 19:29 Wednesday, April 14, 2012 5

The REG Procedure
Model: MODEL1
Dependent Variable: Y

C(p) Selection Method

Number in Model	C(p)	R-Square	BIC	Root MSE	Variables in Model	
4	-0.8544	0.8981	-264.1995	0.05436	Y1 X2 X4 X10	③
3	-0.4632	0.8914	-264.3510	0.05548	Y1 X2 X4	
2	-0.0372	0.8846	-264.3012	0.05654	Y1 X9	
5	0.1785	0.9009	-262.4901	0.05428	Y1 X2 X4 X10 X11	
4	0.3027	0.8949	-263.0767	0.05523	Y1 X2 X4 X11	
5	0.3282	0.9004	-262.3509	0.05440	Y1 X2 X4 X9 X10	
4	0.3658	0.8947	-263.0163	0.05527	Y1 X2 X4 X9	
5	0.3697	0.9003	-262.3124	0.05443	Y1 X2 X4 X5 X10	
4	0.8226	0.8934	-262.5820	0.05561	Y1 X4 X5 X10	
5	0.8395	0.8990	-261.8790	0.05479	Y1 X2 X3 X4 X10	
3	0.8408	0.8877	-263.0743	0.05641	Y1 X9 X10	
4	0.8570	0.8933	-262.5494	0.05563	Y1 X5 X9 X10	
4	0.9282	0.8931	-262.4821	0.05569	Y1 X2 X4 X7	
4	0.9688	0.8930	-262.4438	0.05572	Y1 X2 X4 X12	
5	0.9958	0.8986	-261.7357	0.05491	Y1 X2 X4 X6 X10	

Ilustración 5. Salida selección de variables exógenas ecuación de medición

Variables exógenas para la ecuación de transición

La variable dependiente se declaró como el crecimiento de las ventas Reales y como variables independientes se declararon los rezagos del uno hasta el doce con base en el correlograma presentado en la Ilustración 1. Los resultados del ordenamiento de los mejores modelos se presentan en la Ilustración 6. El primer y segundo modelo ④ presentado en la lista presentan los mejores coeficiente de determinación y los mejores criterios de información Cp, por lo tanto estos dos modelos serán probados en la etapa de estimación para la ecuación de transición.

The REG Procedure
 Model: MODEL1
 Dependent Variable: Y

C(p) Selection Method

Number of Observations Read 60
 Number of Observations Used 60

Number in Model	C(p)	R-Square	BIC	Root MSE	Variables in Model
4	2.9022	0.7116	-366.4535	0.04411	X2 X3 X9 X12
3	3.8473	0.6955	-365.9248	0.04491	X2 X3 X12
5	3.9340	0.7168	-365.0519	0.04410	X2 X3 X7 X9 X12
5	4.0313	0.7163	-364.9642	0.04414	X2 X3 X5 X9 X12
6	4.1917	0.7263	-364.2554	0.04376	X2 X3 X5 X7 X9 X12
4	4.2318	0.7043	-365.2288	0.04466	X2 X3 X5 X12
5	4.3619	0.7145	-364.6672	0.04428	X2 X3 X9 X11 X12
5	4.4534	0.7140	-364.5852	0.04432	X2 X3 X9 X10 X12
5	4.4674	0.7139	-364.5727	0.04433	X2 X3 X8 X9 X12
5	4.6221	0.7131	-364.4345	0.04439	X2 X3 X4 X9 X12
5	4.7233	0.7125	-364.3443	0.04444	X1 X2 X3 X9 X12
4	4.7257	0.7016	-364.7807	0.04486	X2 X3 X6 X12
5	4.7437	0.7124	-364.3261	0.04445	X2 X3 X6 X9 X12
4	4.8449	0.7010	-364.6731	0.04491	X2 X3 X8 X12
4	4.9590	0.7003	-364.5702	0.04496	X2 X3 X10 X12

Ilustración 6. Salida selección de variables exógenas ecuación de transición

3.2.2. Representación Estado-Espacio

Inicialmente, para la representación estado-espacio es necesaria la identificación de las variables observadas y las variables no observadas. Para el modelo particular de esta investigación, la variable observada es las Expectativas de Empresarios denotada y_t y la variable no observada es el pronóstico del Crecimiento de las Ventas Reales denotada $x_{t|t-1}$, que conforma el Vector de Estado α_t . La representación estado-espacio está identificada por un modelo univariado de parámetros desconocidos e invariantes en el tiempo.

Ecuación de medición

La relación existente entre las Expectativas de Empresarios con el pronóstico del Crecimiento Ventas Reales y las variables exógenas conforman la ecuación de medición:

$$y_t = \rho\alpha_t + \Omega D_t + \varepsilon_t, \quad t = 1, \dots, 48$$

$$[y_t]_{1 \times 1} = [\theta_1]_{1 \times 1} [\alpha_t]_{1 \times 1} + [\theta_2 \quad \theta_3 \quad \theta_4 \quad \theta_5 \quad \theta_6]_{1 \times 5} \begin{bmatrix} 1 \\ y_{t-1} \\ x_{t-2} \\ x_{t-4} \\ x_{t-10} \end{bmatrix}_{5 \times 1} + [\varepsilon_t]_{1 \times 1}$$

Dónde y_t es el vector 1×1 de variables observables que contiene las Expectativas de Empresarios, α_t es el Vector 1×1 de Estado que contiene el pronóstico del Crecimiento de las Ventas ($x_{t|t-1}$), D_t es el vector $\times 1$ de otras variables explicativas que contiene todos los rezagos encontrados en el paso anterior, y ε_t es el vector 1×1 de perturbaciones no correlacionadas serialmente, con $E[\varepsilon_t] = 0$ y $Var[\varepsilon_t] = \sigma_\varepsilon^2 = \theta_{13}$.

Ecuación de transición

La relación existente entre las expectativas de los empresarios con el pronóstico del crecimiento ventas y otras variables existentes define la ecuación de medición:

$$\alpha_t = \tau\alpha_{t-1} + \Phi C_t + \eta_t, \quad t = 1, \dots, 48$$

$$[\alpha_t]_{1 \times 1} = [\theta_7]_{1 \times 1} [\alpha_{t-1}]_{1 \times 1} + [\theta_8 \quad \theta_9 \quad \theta_{10} \quad \theta_{11} \quad \theta_{12}]_{1 \times 5} \begin{bmatrix} 1 \\ x_{t-2} \\ x_{t-3} \\ x_{t-9} \\ x_{t-12} \end{bmatrix}_{5 \times 1} + [\eta_t]_{1 \times 1}$$

Dónde α_t es el Vector 1×1 de Estado que contiene el pronóstico del crecimiento anual de las ventas ($x_{t|t-1}$), α_{t-1} es su rezago, C_t es el vector 5×1 de otras variables explicativas que contiene todos los rezagos encontrados en el paso anterior, y η_t es el vector 1×1 de perturbaciones no correlacionadas serialmente, con $E[\eta_t] = 0$ y $Var[\eta_t] = \sigma_\eta^2 = \theta_{14}$.

1.9. ESTIMACIÓN DEL MODELO

1.9.1. Estimación de Parámetros Iniciales

A continuación, la tabla 6 y la tabla 7 presentan las estimaciones por medio del software (E-views® 7) de los parámetros iniciales tanto para la ecuación de medición como para la ecuación de transición, que entrarán al Filtro de Kalman.

Se definió como α_t^0 como el pronóstico del modelo ARMA(3,1) dado que representa el mejor modelo Box-Jenkins deducido en el capítulo anterior.

MODELO	θ_1	θ_2	θ_3	θ_3	θ_5	θ_6	θ_{13}	R^2	R^2 Ajustado
	Ω_1	-0.1745 (0.2914)	0.0813 (0.0421)	0.8258 (0.0767)	-0.2495 (0.2448)	0.4198 (0.1480)	0.2719 (0.1581)	0.0030	0.8996
Ω_2	-0.1445 (0.2976)	0.0325 (0.0318)	0.9198 (0.0551)	-0.1876 (0.2477)	0.4586 (0.1496)	NA	0.0031	0.8924	0.8822

Tabla 6. Estimación de los modelos para la Ecuación de Transición

MODELO	θ_7	θ_8	θ_9	θ_{10}	θ_{11}	θ_{12}	θ_{14}	R^2	R^2 Ajustado
	Φ_1	-0.2351 (0.0584)	0.0047 (0.0012)	0.5219 (0.0247)	0.5512 (0.0350)	0.0270 (0.0229)	-0.0394 (0.0195)	0.0001	0.9880
Φ_2	-0.2142 (0.0558)	0.0046 (0.0012)	0.5177 (0.0245)	0.5484 (0.0351)	-0.0245 (0.0149)	NA	0.0001	0.9876	0.9865

Tabla 7. Estimación de los modelos para la Ecuación de Medición

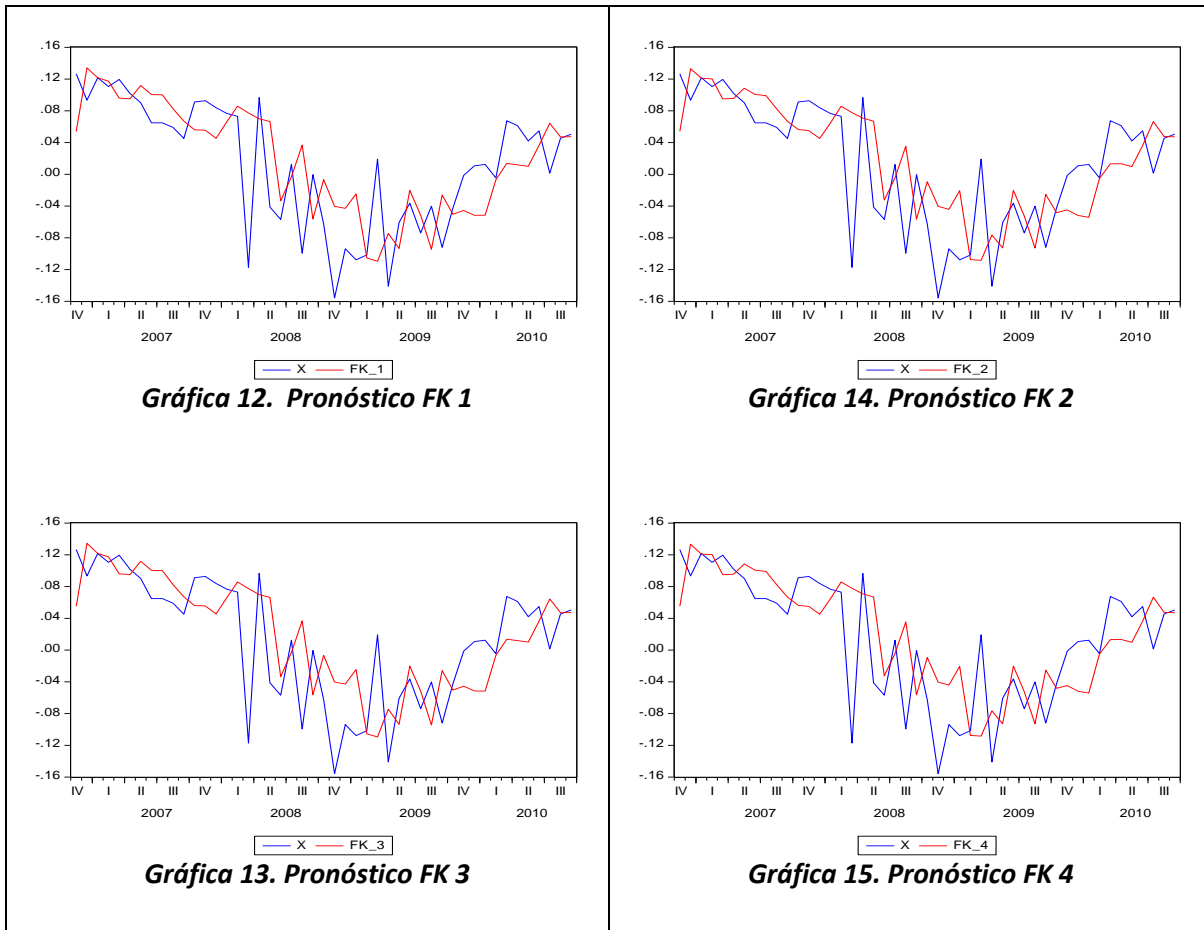
Como se aprecia en las tablas anteriores existen dos posibles modelos para la ecuación de medición denominados Ω_1 y Ω_2 , y dos posibles modelos para la ecuación de transición denominados Φ_1 y Φ_2 , de forma que existen cuatro diferentes combinaciones a evaluar en el Filtro de Kalman y que se explican a continuación en la tabla 8.

MODELO	ECUACIÓN DE MEDICIÓN		ECUACIÓN DE TRANCISIÓN	
	Modelo	Parámetros	Modelo	Parámetros
FK 1	Ω_1	$\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6$ y θ_{13}	Φ_1	$\theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}$ y θ_{14}
FK 2	Ω_1	$\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6$ y θ_{13}	Φ_2	$\theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}$ y θ_{14}
FK 3	Ω_2	$\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$ y θ_{13}	Φ_1	$\theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}$ y θ_{14}
FK 4	Ω_2	$\theta_1, \theta_2, \theta_3, \theta_4, \theta_5$ y θ_{13}	Φ_2	$\theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}$ y θ_{14}

Tabla 8. Modelos a estimar por Filtro de Kalman

1.9.2. Filtro de Kalman

Para la estimación del crecimiento de las ventas por medio de Filtro de Kalman se han establecido cuatro posibles formas de modelación con el fin de comparar su ajuste y escoger el modelo más apropiado de estimación. Por medio de la función KALCVF del software (SAS® 9.1) se realizan los pronósticos de los cuatro modelos que pueden verse a continuación en el gráfico 12 al gráfico 15. Cabe resaltar que la función KALCVF del software (SAS® 9.1) realiza el procedimiento que se expone en la Ilustración 3.



Los gráficos anteriores muestran que el comportamiento que tienen los cuatro modelos propuestos de la estimación del Filtro de Kalman tiene un comportamiento muy similar. El gráfico 10 al gráfico 13 de igual manera muestra el comportamiento que tienen los pronósticos del Vector de Estado α_t o de la variable inobservada en cada periodo, que ante cada movimiento de la variable real trata de acomodarse y de seguir su tendencia, aunque la gran volatilidad no permite hacerlo de una manera muy precisa o por contar con un bajo número de observaciones en la muestra de entrenamiento.

1.9.3. Evaluación del Pronóstico

Posterior a la estimación de los modelos por medio de la función KALCVF del software (SAS® 9.1) se evaluarán estos modelos bajo los criterios de errores relativos (Hyndman & Koehler, 2006), que también fueron implementados en el primer capítulo. A continuación la tabla 9 muestra los resultados para cada una de las cuatro combinaciones.

	FK 1	FK 2	FK 3	FK 4
MRAE	12325.3122	11990.0449	12306.8200	11971.8583
MDRAP	117.9908	119.5936	118.0281	120.5107
GMRAE	120.1863	125.5820	119.3171	125.5898

Tabla 9. Evaluación de los Modelos de Filtro de Kalman

Al igual que en el primer capítulo el criterio MRAE puede no ser un buen indicador para la elección del modelo debido a que se afecta significativamente ante la presencia de valores extremos muy altos por la gran precisión del pronóstico ingenuo en algunos periodos, provocando que este indicador sea muy alto. De esta manera el criterio de decisión más oportuno debe ser el GMRAE ya que en su cálculo le da una ponderación menor a esos valores extremos, y no los omite como ocurre en el criterio MDRAP.

Con estas consideraciones y como se ve en la tabla 9 el modelo que tiene un error relativo menor es FK3 con un valor de 119.31; resultado que aunque es el mejor de los modelos que se están probando se encuentra por debajo del error del pronóstico ingenuo en aproximadamente un 20% razón que nos haría pensar que la implementación de la metodología del Filtro de Kalman es inútil para pronosticar el Crecimiento de las Ventas Reales. Sin embargo es importante recordar que los resultados de ajuste obtenidos hasta el momento corresponden a la evaluación sobre la muestra de entrenamiento, por lo tanto no es posible realizar una conclusión certera sobre la invalidez del método hasta que el mismo sea evaluado sobre la muestra de prueba. Dicha validación se tratará en el capítulo siguiente donde se presentan los primeros resultados de la investigación.

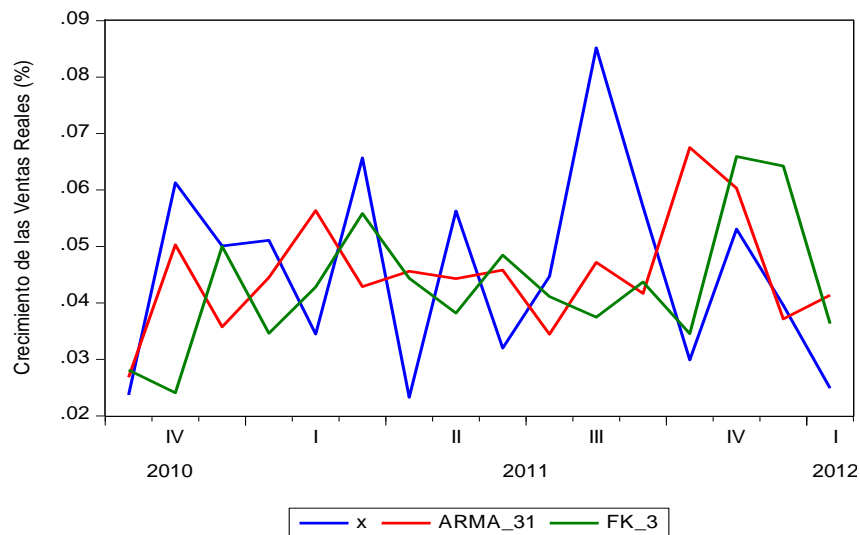
RESULTADOS

El objetivo del presente capítulo es evaluar el pronóstico del Crecimiento de las Ventas Reales de los mejores modelos seleccionados en el capítulo 2 y 3 sobre la muestra de prueba, que abarca el periodo entre octubre de 2010 y enero de 2012.

1.10. PRONÓSTICO EN LA ETAPA DE PRUEBA

En primer lugar, el capítulo 2 concluyó que el mejor modelo bajo la metodología de Box-Jenkins para pronosticar el Crecimiento de las Ventas Reales es el ARMA(3,1). Para el pronóstico de este modelo se utilizó el software (E-views® 7) sobre la muestra de prueba, y los resultados se presentan en el gráfico 17.

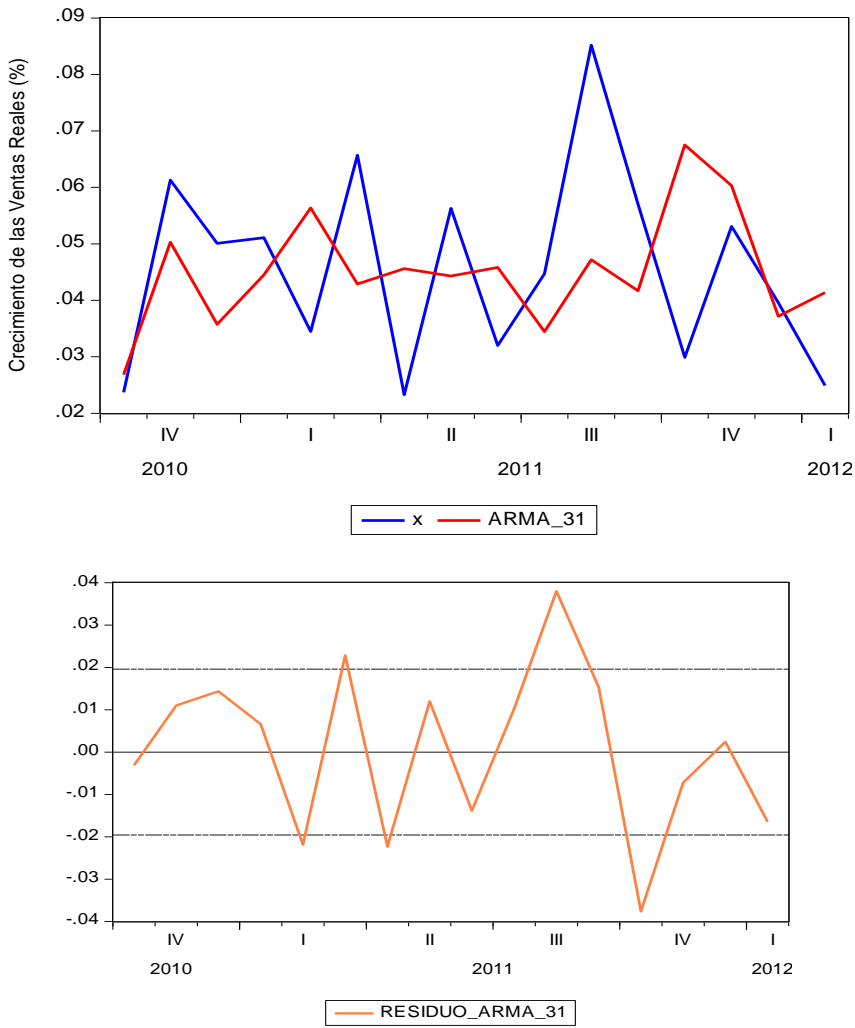
Por su parte el modelo FK 3 fue seleccionado en el capítulo 2 como mejor modelo para implementar la metodología de Filtro de Kalman. La estimación de este modelo se realizó con el software (SAS® 9.1) y los resultados se muestran en el gráfico 16.



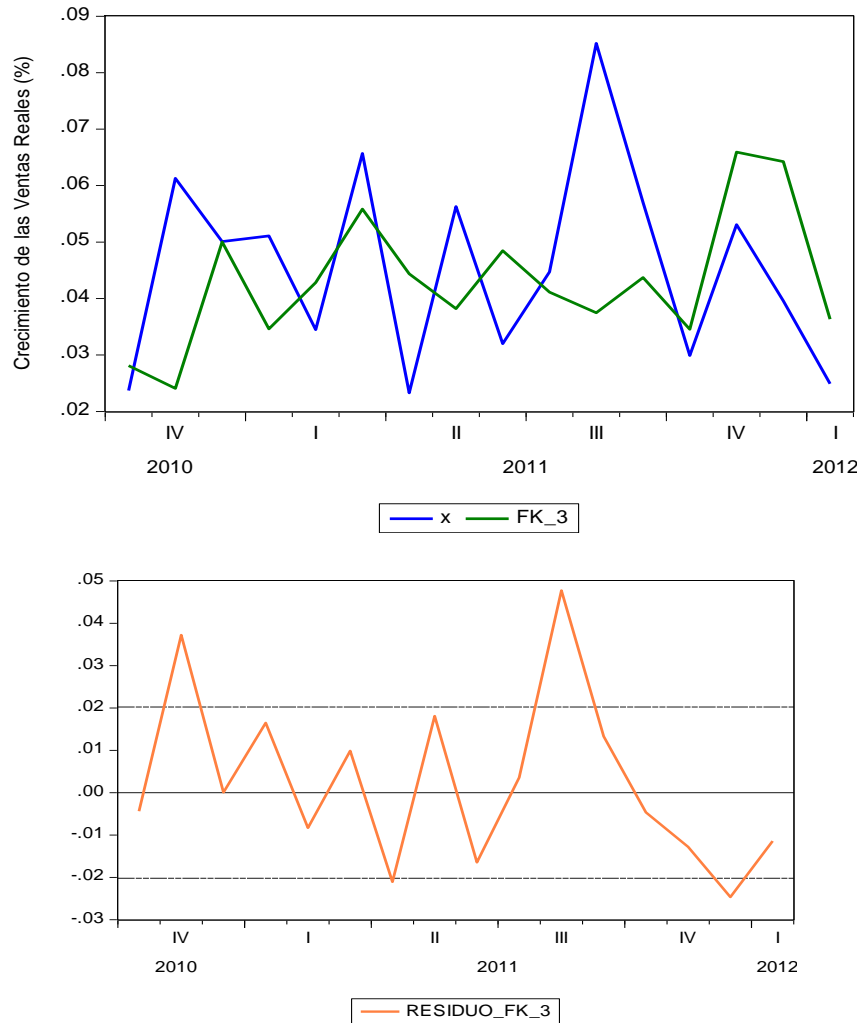
Gráfica 16. Pronósticos del Modelo ARMA(3,1) y FK 3 en comparación con la Serie

La Grafica 16, evidencia que los modelos ARMA(3,1) y FK 3, presentan gran similitud entre sí, presentándose un comportamiento suave en la mayoría de la muestra con excepción del último trimestre de 2011. Ninguno de los dos modelos logra capturar perfectamente los cambios bruscos de la serie como se esperaba al incluir las expectativas. Sin embargo, si se presenta un mejor ajuste en algunos puntos para el modelo FK 3 ya que logra capturar el cambio casi al mismo tiempo mientras que el ARMA(3,1) lo hace al menos un periodo después o antes, como se evidencia en el primer y último trimestre de 2011.

A continuación el gráfico 17 y el gráfico 18 muestran los pronósticos obtenidos por la metodología Box-Jenkins y el Filtro de Kalman individualmente así como el residuo para cada uno.



Gráfica 17. Pronóstico del Modelo ARMA(3,1) en comparación con la Serie



Gráfica 18. Pronóstico del Modelo FK 3 en comparación con la Serie.

Visualmente el gráfico 17 y el Gráfico 18 demuestran que el ajuste presente con los dos métodos de pronóstico no es bueno debido a la gran volatilidad de la serie, y a la poca disponibilidad de datos. Por su parte los gráficos de los residuos muestran como esta gran volatilidad de la serie produce errores grandes, donde para el modelo ARMA(3,1) salen de la primera desviación estándar 5 picos y para el FK 3 salen 3 picos, esto se presenta en los puntos de inflexión más grandes dentro de la serie corroborando que los modelos no logran capturar cambios bruscos o altas volatilidades de la series de forma anticipada. Al mismo tiempo se evidencia que el movimiento de la serie de residuos para el modelo ARMA(3,1) dentro de la región de la primera desviación estándar percibe mayores distancias de los errores respecto a cero, mientras que el FK 3 evidencia distancias más pequeñas. Para hacer un análisis más exhaustivo de los errores presentados, la siguiente sección evaluará los modelos por medio de los criterios de errores relativos que han venido manejándose a lo largo de la investigación.

1.11. EVALUACIÓN Y COMPARACIÓN

Posterior a la revisión de las gráficas de los pronósticos este apartado los evaluará bajo los criterios de errores relativos (Hyndman & Koehler, 2006). A continuación la tabla 10 muestra los resultados obtenidos por los pronósticos:

	ARMA(3,1)	FK 3
MRAE	105.3175	155.3624
MDRAP	65.0096	52.4611
GMRAE	65.1185	48.1110

Tabla 10. Evaluación de los Modelos ARMA(3,1) y FK3

Como se explicó en el capítulo anterior, el criterio MRAE no es un indicador fiable debido a los casos extremos que se presentan por lo que el indicador idóneo para este problema sería el GMRAE debido a que le da un menor peso a los valores extremos, y no los excluye como lo hace el MDARP. Como se ve en la tabla 10 los dos modelos tienen un muy buen comportamiento respecto al pronóstico ingenuo superándolo en aproximadamente un 35% el ARMA(3,1) y en aproximadamente 52% el FK 3, según el GMRAE. En consecuencia los resultados demuestran que la metodología del Filtro de Kalman mejora el pronóstico aproximadamente en un 52% respecto al pronóstico ingenuo y en aproximadamente un 28% respecto al ARMA(3,1), lo que lo hace el modelo más efectivo para pronosticar las Ventas Reales del Sector Industrial reafirmando las conclusiones de (Armstrong, 2006) en las que se afirma que modelos más sofisticados y que involucran una mayor cantidad de información son más exitosos.

DESAGREGACIÓN

La MMM presenta sus resultados clasificados en subsectores del sector industrial de acuerdo al CIU Rev. 3 A.C. como se ve en el anexo 3. A partir de esta clasificación es posible realizar una desagregación del pronóstico realizado por medio de la metodología de Filtro de Kalman para el crecimiento de las Ventas Reales del sector industrial en los 46 subsectores con el fin de otorgarle al empresario información del comportamiento del entorno más específica y contrastar la eficiencia de las expectativas de empresarios sobre el comportamiento de las series de ventas desagregadas.

A continuación, el presente capítulo va a realizar la desagregación del mejor pronóstico encontrado en los capítulos anteriores por medio de las metodologías propuestas por (Boylan, 2010), posteriormente se evaluarán dichos pronósticos para cada uno de los subsectores, y finalmente los se hará un análisis sobre cómo este pronóstico puede tener relevancia en las empresas.

1.12. METODOLOGÍAS DE DESAGREGACIÓN

Siguiendo a (Boylan, 2010) las series agregadas permiten tener mejor visión a largo plazo del comportamiento futuro de la serie, y basados en series históricas largas es posible tener modelos de pronóstico más confiables que cuando se parte de información desagregada.

Para realizar la desagregación de pronósticos existen dos métodos que han sido efectivos:

1. *Promedio de las proporciones históricas:*

Esta metodología consiste en calcular la proporción individual de la contribución de serie desagregada respecto a la serie agregada, para cada periodo de la muestra y a continuación calcular el promedio de las proporciones. A partir de esta metodología es posible realizar pronósticos de la proporción de la serie desagregada.

2. *Proporción del promedio histórico:*

Esta metodología consiste en calcular el total de la serie desagregada y de la serie agregada para todos los periodos de la muestra y posteriormente calcular la proporción del total de cada serie desagregada respecto al total de la serie agregada.

1.13. ESTIMACIÓN DE PROPORCIONES PARA SUBSECTORES

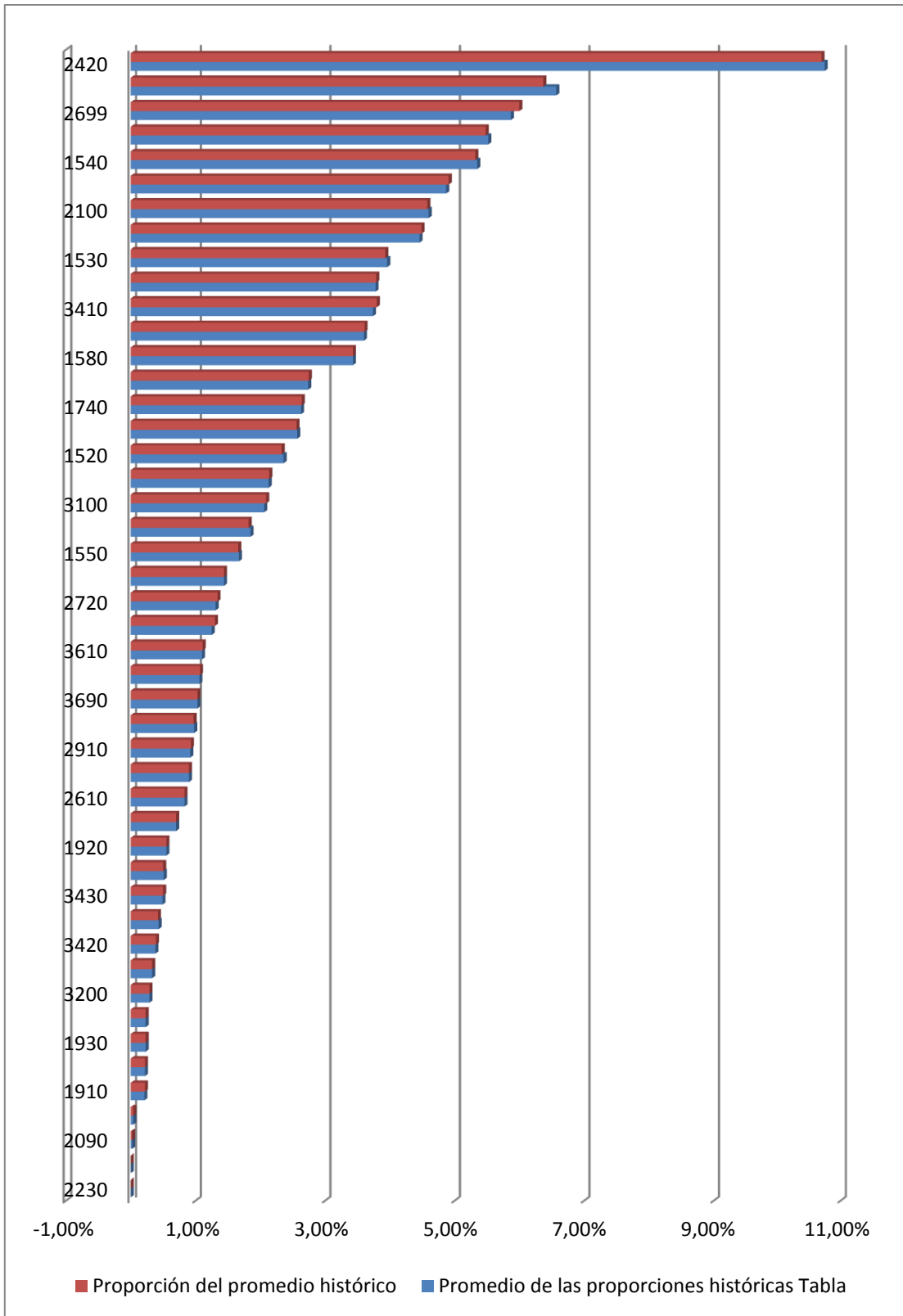
A partir de las series de Ventas Desagregadas por Subsector se calculó la ponderación respecto a la serie de ventas reales utilizando las dos metodologías, *Promedio de las proporciones históricas* y *Proporción del promedio histórico* como se presenta en la Tabla 11.

		Promedio de las proporciones históricas	Proporción del promedio histórico
1500	Total nacional	100.00%	100.00%
1510	Producción, transformación y conservación de carne y pescado	3.78%	3.79%
1520	Aceites y grasas; elaboración de frutas, legumbres, hortalizas	2.36%	2.33%
1530	Productos lácteos	3.96%	3.93%
1540	Productos de molinería y almidones	5.35%	5.32%
1550	Productos de panadería	1.67%	1.66%
1561	Trilla de café	1.85%	1.82%
1570	Ingenios, refinerías de azúcar y trapiches	2.74%	2.75%
1580	Otros productos alimenticios	3.43%	3.43%
1590	Bebidas	5.52%	5.48%
1600	Productos de tabaco	0.51%	0.50%
1720	Hilatura, tejedura y acabado de productos textiles	0.43%	0.42%
1740	Otros productos textiles	2.63%	2.64%
1750	Tejidos y artículos de punto y ganchillo	1.06%	1.07%
1800	Prendas de vestir, confecciones	3.60%	3.61%
1910	Curtido y preparado de cueros	0.21%	0.22%
1920	Fabricación de calzado	0.55%	0.55%
1930	Artículos de viaje, bolsos y artículos similares	0.23%	0.23%
2020	Aserrado, de madera, hojas de madera	0.22%	0.22%
2030	Partes y piezas de carpintería para construcciones	0.04%	0.04%
2090	Otros productos de madera; corcho, cestería y espartería	0.02%	0.02%
2100	Papel, cartón y sus productos	4.60%	4.58%
2210	Actividades de edición	1.44%	1.44%
2220	Actividades de impresión	0.98%	0.97%
2230	Actividades de servicios relacionadas con las de impresión; materiales grabados	0.00%	0.00%
2320	Fabricación de productos de la refinación del petróleo	6.57%	6.37%
2410	Sustancias químicas básicas, fabricación de fibras sintéticas y artificiales	4.46%	4.49%

2420	Otros productos químicos	10.75%	10.70%
2510	Productos de caucho	0.90%	0.90%
2520	Productos de plástico	4.87%	4.91%
2610	Vidrio y sus productos	0.83%	0.83%
2691	Fabricación de productos de cerámica no refractaria, para uso no estructural	0.00%	0.00%
2699	Productos minerales no metálicos	5.87%	6.00%
2710	Industrias básicas de hierro y acero; fundición de metales	2.57%	2.56%
2720	Industrias básicas de metales preciosos y de metales no ferrosos	1.31%	1.34%
2800	Fabricación de productos elaborados de metal	2.13%	2.14%
2910	Maquinaria de uso general	0.92%	0.93%
2920	Maquinaria de uso especial	0.33%	0.33%
2930	Aparatos de uso doméstico	0.70%	0.70%
3100	Maquinaria y aparatos eléctricos	2.06%	2.09%
3200	Equipos y aparatos de radio, televisión y comunicaciones	0.29%	0.29%
3300	Aparatos e instrumentos médicos, ópticos y de precisión	0.23%	0.23%
3410	Vehículos automotores y sus motores	3.74%	3.80%
3420	Carrocerías para vehículos automotores	0.38%	0.39%
3430	Partes, piezas y accesorios para vehículos automotores	0.49%	0.50%
3500	Fabricación de otros tipos de equipo de transporte	1.25%	1.30%
3610	Fabricación de muebles	1.10%	1.11%
3690	Otras industrias manufactureras	1.03%	1.03%

Tabla 11. Promedio de las proporciones históricas y Proporción del promedio histórico

Los resultados de las proporciones para cada subsector difieren únicamente en el 0,02% en promedio entre las metodologías. Esto permite concluir que no existen dentro del periodo estudiado cambios significativos de la participación de las ventas de cada subsectores dentro de las ventas agregadas para el Sector Industrial, y al mismo tiempo muestra que el comportamiento de las participaciones a lo largo del tiempo es estable es decir que no presenta altas volatilidades.



Gráfica 19. Proporciones Subsectores

1.14. DESAGREGACIÓN DEL MEJOR PRONÓSTICO Y EVALUACIÓN

Como se concluyó en el capítulo 4 el mejor pronóstico para la estimación del crecimiento de las Ventas Reales del Sector Industrial es el obtenido por la metodología de Filtro de Kalman, específicamente con la modelación FK 3. Este será el pronóstico que a continuación se utilizara para realizar la desagregación en los subsectores del Sector Industrial.

El pronóstico del crecimiento de las ventas reales es inicialmente transformado a su valor en términos nominales (ventas reales en pesos) con el objeto de poder comparar la metodología de desagregación con el valor de las ventas reales de cada subsector. Este procedimiento se explica a continuación:

El pronóstico obtenido se define como el crecimiento porcentual anual de las reales mensuales, y es calculado de la siguiente manera:

$$x_t = \ln v_t - \ln v_{t-1}$$

Donde x_t crecimiento porcentual anual de las ventas mensuales del sector industrial en el momento t , y v_t el índice de ventas reales del sector industrial en el momento t . Por lo tanto para encontrar el índice de ventas reales del sector industrial v_t :

$$v_t = e^{(x_t + \ln v_{t-1})}$$

Por medio del índice de ventas reales del sector industrial es posible encontrar el Total Nacional de las ventas así:

$$r_{(t)} = r_{(t0)} * v_t$$

Donde $r_{(t)}$ es el total nacional de las ventas del sector industrial, estimado para el mes t y $r_{(t0)}$ es el total nacional de las ventas del sector industrial estimado para el mes $t0$ (periodo base). Finalmente $r_{(t)}$ será desagregada para cada uno de los sectores industriales en Colombia.

Para validar la precisión del pronóstico desagregado es necesario utilizar criterios de evaluación comparado frente a otro pronóstico. En este caso, los pronósticos desagregados serán comparados con el pronóstico ingenuo por medio de los criterios de errores relativos enunciados en (Hyndman & Koehler, 2006).

La tabla 12 muestra los resultados obtenidos por la desagregación.

	MRAE	MDRAP	GMRAE		MRAE	MDRAP	GMRAE
1510	329.59	112.57	115.90	2230	8548.86	1313.56	1516.74
1520	627.46	231.17	196.66	2320	1264.60	309.52	304.51
1530	350.06	176.80	159.27	2410	2575.90	167.58	154.31
1540	349.31	107.65	116.51	2420	1256.01	90.47	88.60
1550	603.82	96.25	87.89	2510	159.17	71.85	70.73
1561	662.27	132.59	164.33	2520	332.98	160.69	163.66
1570	210.60	102.54	98.60	2610	327.42	94.87	110.03
1580	239.35	87.21	68.68	2699	1995.96	287.21	311.21
1590	253.29	155.54	117.23	2710	402.14	107.54	111.11
1600	329.14	118.54	128.88	2720	510.84	141.33	119.39
1720	804.16	170.79	215.31	2800	143.75	90.59	77.69
1740	150.70	78.35	72.35	2910	460.38	100.96	92.57
1750	387.36	94.37	83.17	2920	347.15	69.30	88.17
1800	993.02	181.90	179.68	2930	292.36	125.87	105.39
1910	598.90	98.61	127.62	3100	196.10	127.39	86.12
1920	1966.80	148.20	137.67	3200	518.80	88.83	104.95
1930	196.45	110.64	102.28	3300	295.10	182.60	145.69
2020	733.91	160.73	175.96	3410	1625.40	157.95	186.17
2030	569.20	71.27	76.22	3420	270.04	110.44	105.33
2090	334.60	85.34	91.50	3430	382.62	142.38	120.73
2100	678.94	110.03	143.55	3500	2415.28	193.49	234.07
2210	296.58	112.13	112.16	3610	5010.62	132.24	119.05
2220	609.85	123.76	124.62	3690	216.08	104.48	93.98

Tabla 12 . Evaluación de Pronósticos desagregados

Al igual que los capítulos anteriores, el criterio MRAE no es un buen indicador para evaluar los pronósticos debido a la incidencia que tienen los valores extremos, por lo que el criterio GMRAE será el indicado para la evaluación de los pronósticos. Como se ve en la Tabla 12, hay 14 subsectores en los cuales el pronóstico desagregado supera al pronóstico ingenuo, y en los cuales el modelo presenta mejores resultados. De tal manera, las empresas pertenecientes a estos 14 grupos, les puede ser útil la información esta específica de cada subsector.

La siguiente tabla presenta los 14 subsectores con mejor ajuste así como su participación dentro de las ventas del Sector Industrial dado la ponderación estimada a través de la Proporción del promedio histórico y el número de establecimientos definido en la EAM.

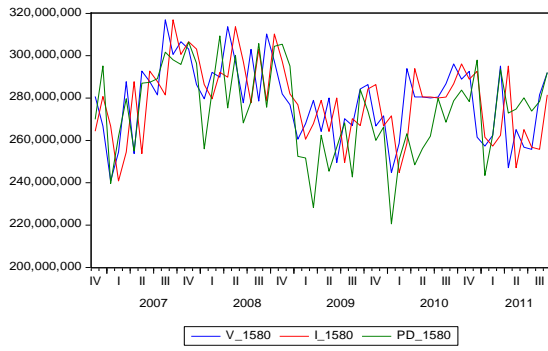
	GMRAE	CLASIFICACIÓN	PARTICIPACIÓN EN EL SECTOR INDUSTRIAL	ESTABLECIMIENTOS
1580	68.68	Otros productos alimenticios	3.43%	135
2510	70.73	Productos de caucho	0.90%	65
1740	72.35	Otros productos textiles	2.64%	49
2030	76.22	Partes y piezas de carpintería para construcciones	0.04%	23
2800	77.69	Fabricación de productos elaborados de metal	2.14%	145
1750	83.17	Tejidos y artículos de punto y ganchillo	1.07%	91
3100	86.12	Maquinaria y aparatos eléctricos	2.09%	39
1550	87.89	Productos de panadería	1.66%	467
2920	88.17	Maquinaria de uso especial	0.33%	155
2420	88.6	Otros productos químicos	10.70%	389
2090	91.5	Otros productos de madera; corcho, cestería y espartería	0.02%	16
2910	92.57	Maquinaria de uso general	0.93%	186
3690	93.98	Otras industrias manufactureras	1.03%	168
1570	98.6	Ingenios, refinerías de azúcar y trapiches	2.75%	28

Tabla 13. Participación y establecimientos de los subsectores de mejor ajuste

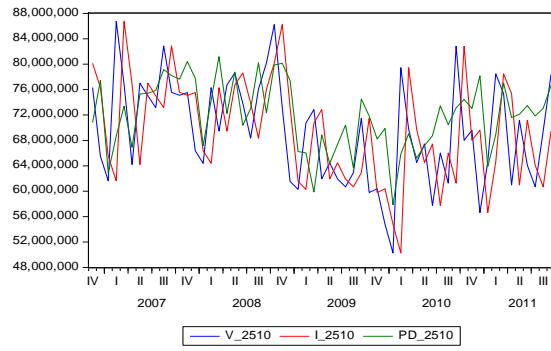
Dentro de los modelos de mejor ajuste se encuentre el subsector de mayor participación del volumen dentro del sector industrial representado por Otros productos químicos con un 10.7% y 389 establecimientos. Así mismo se encuentran subsectores con pocos establecimientos como Otros productos de madera; corcho, cestería y espartería con 16, Partes y piezas de carpintería para construcciones con 23 e Ingenieros, refinerías de azúcar y trapiches con 28 donde el pronósticos del subsector puede ser desagregado para evaluar o generar su propio pronóstico de ventas.

A continuación, el gráfico 19 al gráfico 23 muestra el comportamiento del pronóstico para cada uno de los cinco subsectores que presentan mejor ajuste a través de la desagregación de la integración de expectativas y serie de tiempo por Filtro de Kalman: Otros productos alimenticios, Productos de caucho, Otros productos textiles, Partes y piezas de carpintería para construcciones y Fabricación de productos elaborados de metal.²

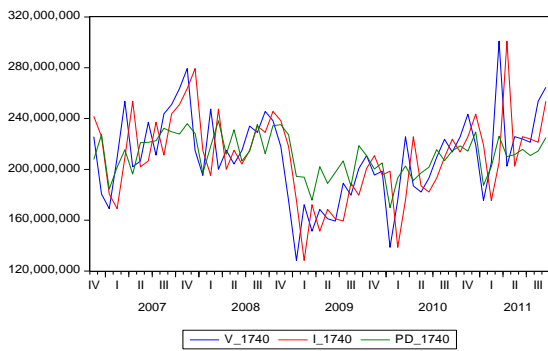
² Las gráficas para los pronósticos de los 9 subsectores restantes se presenta en el Anexo 4.



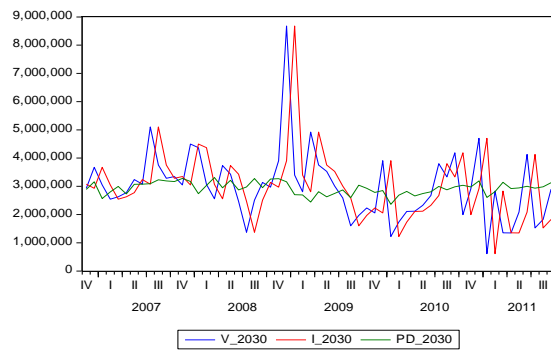
Gráfica 20. Subsector 1580



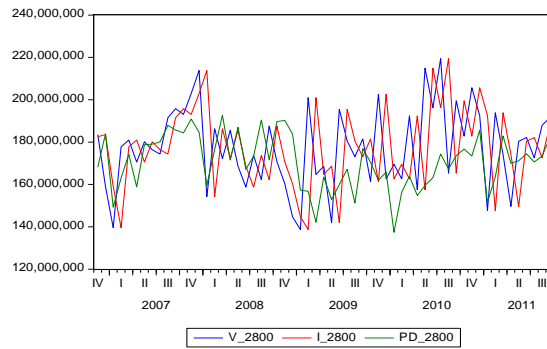
Gráfica 21. Subsector 2510



Gráfica 22. Subsector 1740



Gráfica 23. Subsector 2030



Gráfica 24. Subsector 2800

1.15. APLICACIÓN EN LA PLANEACIÓN DE LAS EMPRESAS

Siguiendo a (Steiner, 2006) los objetivos de las empresas se pueden clasificar en dos grandes grupos, financieros y estratégicos. Los objetivos financieros permiten mantener las utilidades económicas de la empresa y los objetivos estratégicos permiten mantener a la empresa competitiva en el mercado.

La siguiente tabla presenta algunos de los objetivos financieros y de los objetivos estratégicos:

OBJETIVOS FINANCIEROS	OBJETIVOS ESTRATEGICOS
<ul style="list-style-type: none">• Dinamizar el crecimiento de los ingresos de la empresa• Generar dividendos más altos para los inversionistas• Incrementar márgenes de beneficio, rentabilidad en capital invertido y flujos de efectivo.	<ul style="list-style-type: none">• Aumentar la participación del mercado• Mejorar la reputación frente a los clientes• Mayor capacidad para competir en mercados internacionales• Identificar oportunidades de crecimiento

Tabla 14. Objetivos de las empresas

La planeación estratégica son las actividades encaminados a prevenir y anticipar situaciones que pueden afectar el cumplimiento de los objetivos de la empresa, buscando anticiparse a cambios y sucesos futuros para afrontarlos correctamente.

La estructuración de la planeación estratégica sigue el modelo presentado en la ilustración 7.

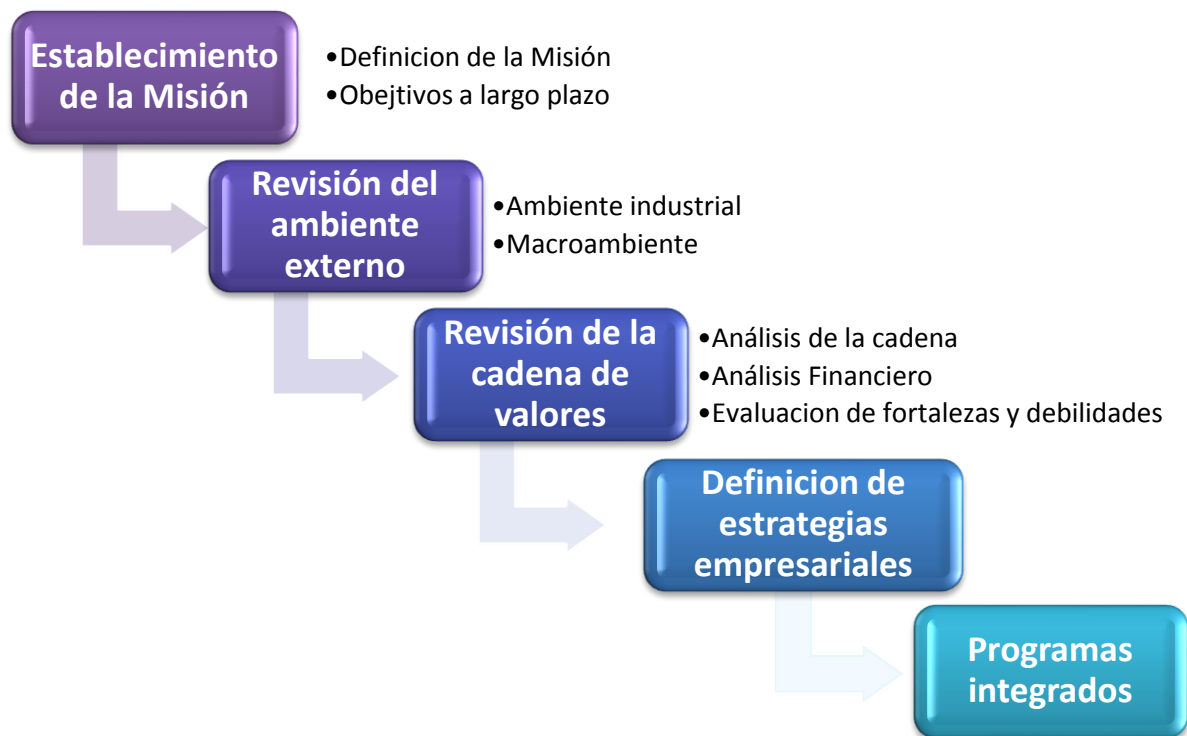


Ilustración 7. Modelo de Planeación Estratégica de las empresas

La cadena de planeación estratégica depende del estudio del mercado y del pronóstico de la demanda, a partir de la cual se estructuran planes para la toma de decisiones en cada eslabón.

Esta investigación contribuye a la planeación de las empresas en tres de los eslabones que conforman la estructuración de la planeación estratégica de las empresas así:

1. **Revisión del Ambiente Externo:**
El pronóstico que se encuentra en esta investigación le presenta a las empresas colombianas el comportamiento de las ventas reales del sector industrial, indicando información sobre el macroambiente, ya que al integrar las expectativas se involucra información sociopolítica y económica.
2. **Revisión de la Cadena de Valores:**
Dentro de la planeación de la cadena de abastecimiento se estructuran planes entorno a la cantidad de producción, materia prima a adquirir y factores productivos a contratar restringido a los pronósticos de demanda. Los empresarios pueden implementar la metodología expuesta en esta investigación incluyendo las expectativas de sus propios expertos con su serie de demanda dado que se demostró que las expectativas de empresarios constituyen información útil para la precisión del pronóstico.
Los empresarios pueden integrar la información del entorno dentro de sus modelos de pronósticos actuales con fin de constituir modelos realistas del comportamiento de la demanda, utilizar el conocimiento previo al involucrar expectativas y utilizar modelos más estructurados, características que permiten mejorar la precisión del pronóstico.
3. **Definición de Estrategias Empresariales:**

Las empresas pueden evaluar sus propios pronósticos de demanda frente a los pronósticos del Sector Industrial e Incluso dentro de su subsector para definir por ejemplo estrategias de ajuste de su pronóstico, estrategias de mercadeo, estrategias de inversión.

CONCLUSIONES

El modelo estado-espacio estimado a través del Filtro de Kalman es una metodología que permite pronosticar una variable económica no observada en función de variables explicativas observadas que influyen en la predicción. Esta metodología es aplicada al caso colombiano con el fin de pronosticar el crecimiento de las ventas del sector industrial mediante la integración de la información contenida en las expectativas de empresarios presentadas por el Banco de la República a través de la EMEE y la serie de tiempo mensual de indicadores presentada por el DANE a través de la MMM, demostrando que las expectativas contienen información útil para mejorar la precisión del pronóstico para el periodo de Octubre de 2006 a Enero de 2012. Comparando con el pronóstico ingenuo la integración de las expectativas y la serie de tiempo mejora el pronóstico en el 52%, mientras que una metodología Box-Jenkins que utiliza únicamente la serie de tiempo mejora el pronóstico en el 34% sobre la base de prueba. Este resultado permite concluir que el juicio de los empresarios logra capturar anticipadamente variables exógenas que afectan el comportamiento futuro de la serie en cada momento del tiempo.

Esta conclusión coincide con los resultados encontrados en (Rahiala & Teräsvirta, 1993) Y (Dahlöf, 2003) para Suecia y Finlandia, lo que conlleva a un nuevo aporte con evidencia empírica que demuestra que el Filtro de Kalman es eficiente para integrar información contenida en encuestas de expectativas y series de tiempo como método de predicción. Pese a las similitudes encontradas en el método de desarrollo y en los resultados encontrados, la presente investigación ofrece las siguientes innovaciones:

- La metodología fue comparada contra el pronóstico ingenuo y con la metodología Box-Jenkins aplicada solo con el pasado de la serie, demostrando su eficiencia y viabilidad en futuras investigaciones.
- Adicional al pronóstico del sector industrial, este es desagregado en subsectores en busca de otorgar mayor información valiosa para las empresas.

Aunque la investigación es concluyente en afirmar que las expectativas de los empresarios contiene información útil que ayuda a mejorar la precisión de los pronósticos, estas no son acertadas en presencia de puntos de inflexión o momentos de alta volatilidad, en los que se percibe que el Filtro de Kalman al integrar las expectativas con la serie de tiempo no logra anticipar los cambios bruscos de tendencia.

Durante el desarrollo de la presente investigación, se identificaron algunos puntos de gran relevancia y que deben ser considerados en futuras investigaciones que empleen la metodología, los cuales son:

- El modelo es sensible a las variables exógenas, por lo tanto es necesario hacer una exhaustiva revisión de las variables para identificar correctamente sus rezagos, teniendo en cuenta especialmente criterios de información y criterios de ajuste.
- El modelo es sensible especialmente a los parámetros iniciales asignados a la matriz de transición (coeficientes que relacionan al vector de estado con el vector de estado en el momento anterior) y a la matriz de coeficientes que relaciona la variable observada con el vector de estado.

Dentro de la etapa de desagregación se encontró que de los cuarenta y siete subsectores en que se encuentra clasificada la MMM, para diecisiete de dichas desagregaciones se supera el pronóstico ingenuo. Las empresas que hacen parte de estos diecisiete subsectores tienen a su disposición información adicional que ofrece la desagregación del pronóstico integrador de las expectativas de empresarios y la serie de tiempo. Esta información del comportamiento de las ventas, puede ser incorporada dentro de sus métodos de pronósticos de demanda y/o dentro de su planificación permitiéndoles conocer de manera anticipada el comportamiento de su subsector para apoyar la planeación estratégica. Para los subsectores restantes es recomendable el estudio de otros métodos de desagregación específicos con el fin de tener disponibilidad de dicha información, debido a que la metodología implementada en la investigación no presenta buenos resultados.

Esta investigación contribuye a la planeación de las empresas desde tres puntos de vista:

- Las expectativas de empresarios constituyen información útil para la precisión del pronóstico, es por esto que los empresarios pueden implementar esta metodología de integración incluyendo las expectativas de sus propios expertos con su serie de demanda.
- Las empresas pueden evaluar sus propios pronósticos de demanda frente a los pronósticos del Sector Industrial e Incluso dentro de su subsector para definir por ejemplo estrategias de ajuste de su pronóstico, estrategias de mercadeo, estrategias de inversión.
- La información del entorno puede ser integrada en los modelos de pronósticos de demanda propios de las empresas con el fin de mejorar su precisión.

Finalmente, futuras investigación sobre este tema pueden estar dirigidas a:

- Involucrar en el modelo otras variables macroeconómicas explicativas de las ventas del sector industrial como el PIB, la inflación, la tasa de cambio y la tasa de interés, ampliamente utilizadas en modelos de predicción macroeconómica.
- Aplicar la integración de expectativas y la serie para predecir las ventas totales de la economía colombiana con el uso de los resultados totales de pregunta 2 de la EMEE.

- Aplicar la integración entre las expectativas y la serie de tiempo a otros sectores de la Economía de los que se dispone información de la EMEE como construcción, transporte, agricultura y comercio, además del uso de las metodologías de desagregación del pronóstico.
- Utilizar una metodología de integración más sencilla que el Filtro de Kalman para integrar las expectativas de empresarios y la serie de ventas que podría tener mayor efectividad.

BIBLIOGRAFÍA

- Armstrong, J. S. (2006). Findings from evidence-based forecasting: Methods for reducing forecast error. *International Journal of Forecasting*, 22, 583– 598.
- Banco de la Republica de Colombia. (2005). *Ficha Técnica Encuesta Mensual de Expectativas Economicas*.
- Boylan, J. (2010). Choosing Levels of Aggregation for Supply Chain Forecasts. *The International Journal of Applied Forecasting is the property of International*, 9-13.
- Dahllöf, K.-J. (2003). Monthly leading indicators using the leading information in the monthly Business Tendency Survey. *Mathematical Statistics Stockholm University Research Report*.
- Departamento Administrativo Nacional de Estadística DANE. (2008). *Ficha Metodológica Muestra Mensual Manufacturera MMM*.
- Departamento Administrativo Nacional de Estadística DANE. (2009). *Ficha Metodológica Encuesta Anual Manufacturera EAN*.
- Dickey, D. A., & Brocklebank, J. C. (2003). *SAS® for Forecasting Time Series*. Cary, NC: SAS Institute and Wiley.
- E-views® 7. (s.f.).
- Goodwin, P. (2002). Integrating management judgment and statistical methods to improve short-term forecasts. *The International Journal of Management Science*, 30, 127 – 135.
- Goodwin, P., Lawrence, M., Fildes, R., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, 25, 3–23.
- Granada, S. (2008). Caracterización y contextualización de la dinámica del desplazamiento forzado interno en Colombia 1996 - 2006. *Documentos de CERAC No. 12*.
- Harvey, A. (2003). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* 22, 679-688.
- Makridakis, S., & Taleb, N. (2009). Decision making and planning under low levels of predictability. *International Journal of Forecasting*, 25, 716–733.
- Misas A., M., Arango A., C. A., & López E., E. (2005). Economía Subterránea en Colombia 1976-2003: Una Medición a Partir de La Demanda de Efectivo. *Borradores de Economía No.* 335.
- Montenegro García, Á. (2011). *Análisis de Series de Tiempo*. Pontificia Universidad Javeriana.
- Montenegro, Á. (2005). *Introducción al Filtro de Kalman*. Bogotá: Pontificia Universidad Javeriana.
- Öller, L.-E. (1990). Forecasting the business cycle using survey data. *International Journal of Forecasting*, 6, 453-461.
- Öller, L.-E., & Tallbom, C. (1993). Smooth and timely business cycle indicators for noisy Swedish data. *International Journal of Forecasting*, 12, 389-402.
- Rahiala, M., & Teräsvirta, T. (1993). Business Survey Data in Forecasting the Output of Swedish and Finnish Metal and Engineering: A Kalman Approach. *Journal of Forecasting*, 12, 255-271.
- SAS Institute Inc. (2004). SAS/IML 9.1 User's Guide. Cary, North Carolina, USA.
- SAS® 9.1. (s.f.).
- Solera, Á. (2003). El Filtro de Kalman. *Departamento de Investigaciones Económicas, Banco Central de Costa Rica*.
- Steiner, G. A. (2006). *Planeación estratégica : lo que todo director debe saber*. México D.F.: Patria Cultural.
- Zarate, H., Sanchez, K., & Marin, M. (2011). Cuantificación de Encuestas Ordinales y Pruebas de Racionalidad: Una aplicación a la encuesta medular de Expectativas Económicas. *Borradores de Economía*, Núm. 649.

ANEXOS

ANEXO I. Cuestionario EMEE Septiembre de 2011

- 1) El crecimiento del volumen de ventas (cantidades) en el mes de septiembre de 2009 respecto a septiembre de 2008, comparado con el crecimiento del volumen de ventas (cantidades) de septiembre de 2008 respecto a septiembre de 2007 fue:
 - a) Mayor
 - b) Menor
 - c) El mismo

- 2) El crecimiento del volumen de ventas (cantidades) en los próximos 12 meses, comparado con el crecimiento del volumen de ventas (cantidades) en los pasados 12 meses, se espera sea:
 - a) Mayor
 - b) Menor
 - c) El mismo

- 3) El nivel de inversión en maquinaria y equipo en los próximos 12 meses, comparado con el nivel de inversión en maquinaria y equipo en los pasados 12 meses, se espera sea:
 - a) Mayor
 - b) Menor
 - c) El mismo

- 4) El número de trabajadores de tiempo completo que su empresa espera emplear en los próximos 12 meses, comparado con el número actual será:
 - a) Mayor
 - b) Menor
 - c) El mismo

- 5) ¿Enfrenta su empresa “cuellos de botella” por dificultad para conseguir los empleados u obreros necesarios para poder suplir la demanda de sus productos?
 - a) Si
 - b) No

- 6) ¿En las condiciones actuales, cree usted que su empresa enfrente dificultades para suplir un aumento inesperado en la demanda o venta de sus productos?
 - a) No dificultad
 - b) Algo de dificultad
 - c) Mucha dificultad

- 7) En caso de haber elegido la opción b. ó la opción c. en la respuesta de la pregunta anterior, ordene de 1 a 5 según su importancia (1 más importante 5 menos importante) las razones que explican las dificultades para atender incrementos inesperados en la demanda o venta de sus productos:
Limitación en la disponibilidad de:
Maquinaria y Equipo (); Mano de obra calificada (); Materias primas ();
Financiamiento (); Otra (especifique) ()
- 8) Los incrementos salariales en los próximos 12 meses afectarán su rentabilidad:
a) Más que en los últimos 12 meses
b) Menos que en los últimos 12 meses
c) Igual que en los últimos 12 meses
- 9) El incremento en los precios de las materias primas totales (nacionales más importadas) que comprará en los próximos 12 meses, comparado con el incremento de las materias primas totales que compró en los últimos 12 meses, se espera que sea:
a) Mayor
b) Menor
c) El mismo
- 10) El incremento de los precios de los productos que venderá en los próximos doce meses, comparado con el incremento de los costos operacionales en los próximos doce meses, se espera que sea:
a) Mayor
b) Menor
c) El mismo
- 11) El incremento en los precios de los productos que venderá en los próximos 12 meses, comparado con el incremento de los precios de los productos que vendió en los últimos 12 meses, se espera que sea:
a) Mayor
b) Menor
c) El mismo

ANEXO II. CIU Rev. 3 A.C.

Código	Clasificación
1510	Producción, transformación y conservación de carne y pescado
1520	Aceites y grasas; elaboración de frutas, legumbres, hortalizas
1530	Productos lácteos
1540	Productos de molinería y almidones
1550	Productos de panadería
1561	Trilla de café
1570	Ingenios, refinerías de azúcar y trapiches
1580	Otros productos alimenticios
1590	Bebidas
1600	Productos de tabaco
1720	Hilatura, tejedura y acabado de productos textiles
1740	Otros productos textiles
1750	Tejidos y artículos de punto y ganchillo
1800	Prendas de vestir, confecciones
1910	Curtido y preparado de cueros
1920	Fabricación de calzado
1930	Artículos de viaje, bolsos y artículos similares
2020	Aserrado, de madera, hojas de madera

ANEXO III. Correlativa Entre MMM y EAM

COD	MUESTRA MENSUAL MANUFACTURERA	COD	ENCUESTA MENSUAL MANUFACTURERA
1510	Producción, transformación y conservación de carne y pescado	151	Producción, transformación y conservación de carne y pescado
1520	Aceites y grasas; elaboración de frutas, legumbres, hortalizas	152	Elaboración de frutas, legumbres, hortalizas, aceites y grasas
1530	Productos lácteos	153	Elaboración de productos lácteos
1540	Productos de molinería y almidones	154	Elaboración de productos de molinería, de almidones y productos derivados del almidón, y alimentos
1550	Productos de panadería	155	Elaboración de productos de panadería, macarrones, fideos, alcuizcuz y productos farináceos similares
1561	Trilla de café	156	Elaboración de productos de café
1570	Ingenios, refinerías de azúcar y trapiches	157	Ingenios, refinerías de azúcar y trapiches
1580	Otros productos alimenticios	158	Elaboración de otros productos alimenticios
1590	Bebidas	159	Elaboración de bebidas
1600	Productos de tabaco	160	Fabricación de productos de tabaco
1720	Hilatura, tejedura y acabado de productos textiles	171	Preparación e hilatura de fibras textiles
1740	Otros productos textiles	172	Tejedura de productos textiles
		173	Acabado de productos textiles no producidos en la misma unidad de producción
		174	Fabricación de otros productos textiles
1750	Tejidos y artículos de punto y ganchillo	175	Fabricación de tejidos y artículos de punto y ganchillo
1800	Prendas de vestir, confecciones	181	Fabricación de prendas de vestir, excepto prendas de piel
		182	Preparado y teñido de pieles, fabricación de artículos de piel
1910	Curtido y preparado de cueros	191	Curtido y preparado de cueros
1920	Fabricación de calzado	192	Fabricación de calzado
1930	Artículos de viaje, bolsos y artículos similares	193	Fabricación de artículos de viaje, bolsos de mano y artículos similares; fabricación de artículos
2020	Aserrado, de madera, hojas de madera	201	Aserrado, acepillado e impregnación de la madera
		202	Fabricación de hojas de madera para enchapado, fabricación de tableros contrachapados,
2030	Partes y piezas de carpintería para construcciones	203	Fabricación de partes y piezas de carpintería para edificios y construcciones
2090	Otros productos de madera; corcho,	204	Fabricación de recipientes de madera

	cestería y espartería	209	Fabricación de otros productos de madera, fabricación de artículos de corcho, cestería y espartería
2100	Papel, cartón y sus productos	210	Fabricación de papel, cartón y productos de papel y cartón
2210	Actividades de edición	221	Actividades de edición
2220	Actividades de impresión	222	Actividades de impresión
2230	Actividades de servicios relacionadas con las de impresión; materiales grabados	223	Actividades de servicios relacionadas con la impresión
2321	Refinación del petróleo	232	Fabricación de productos de la refinación del petróleo
2322	Derivados del petróleo, fuera de refinería		
2410	Sustancias químicas básicas, fabricación de fibras sintéticas y artificiales	241	Fabricación de sustancias químicas básicas
2420	Otros productos químicos	242	Fabricación de otros productos químicos
2510	Productos de caucho	251	Fabricación de productos de caucho
2520	Productos de plástico	252	Fabricación de productos de plástico
2610	Vidrio y sus productos	261	Fabricación de vidrio y productos de vidrio
2699	Productos minerales no metálicos	269	Fabricación de productos minerales no metálicos ncp
2710	Industrias básicas de hierro y acero; fundición de metales	271	Industrias básicas de hierro y acero
2720	Industrias básicas de metales preciosos y de metales no ferrosos	272	Industrias básicas de metales preciosos y de metales no ferrosos
2800	Fabricación de productos elaborados de metal	281	Fabricación de productos metálicos para uso estructural, tanques, depósitos y generadores de vapor
		289	Fabricación de otros productos elaborados de metal y actividades de servicios relacionados con el trabajo de metales
2910	Maquinaria de uso general	291	Fabricación de maquinaria de uso general
2920	Maquinaria de uso especial	292	Fabricación de maquinaria de uso especial
2930	Aparatos de uso doméstico	293	Fabricación de aparatos de uso doméstico ncp
3100	Maquinaria y aparatos eléctricos	311	Fabricación de motores, generadores y transformadores
		312	Fabricación de aparatos de distribución y control de la energía eléctrica
		313	Fabricación de hilos y cables aislados
		314	Fabricación de acumuladores y de pilas eléctricas
		315	Fabricación de lámparas eléctricas y equipos de iluminación
		319	Fabricación de otros tipos de equipo eléctrico ncp
3200	Equipos y aparatos de radio, televisión y comunicaciones	321	Fabricación de tubos y válvulas electrónicas y de otros componentes electrónicos

		322	Fabricación de transmisores de radio y televisión y de aparatos para telefonía y telegrafía
		323	Fabricación de receptores de radio y televisión, de aparatos de grabación y de reproducción del sonido
3300	Aparatos e instrumentos médicos, ópticos y de precisión	331	Fabricación de aparatos e instrumentos médicos y de aparatos para medir, verificar, ensayar, navegar
		332	Fabricación de instrumentos ópticos y de equipo fotográfico
3410	Vehículos automotores y sus motores	341	Fabricación de vehículos automotores y sus motores
3420	Carrocerías para vehículos automotores	342	Fabricación de carrocerías para vehículos automotores, fabricación de remolques y semirremolques
3430	Partes, piezas y accesorios para vehículos automotores	343	Fabricación de partes, piezas y accesorios (autopartes) para vehículos automotores y para sus motores
3500	Fabricación de otros tipos de equipo de transporte	351	Construcción y reparación de buques y de otras embarcaciones
		353	Fabricación de aeronaves y de naves espaciales
		359	Fabricación de otros tipos de equipo de transporte ncp
3610	Fabricación de muebles	361	Fabricación de muebles
3690	Otras industrias manufactureras	369	Industrias manufactureras ncp

Anexo IV. Gráficas pronósticos de los subsectores



