

**ANÁLISIS DE LA RELACIÓN ENTRE LA MORBILIDAD POR AFECCIONES  
RESPIRATORIAS Y LA CONTAMINACIÓN DEL AIRE EN BUCARAMANGA  
EN EL 2009 UTILIZANDO UN MODELO DE REGRESIÓN DINÁMICA**

**CARLOS MANUEL SARMIENTO SOTO**



**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE CIENCIAS  
ESCUELA DE MATEMÁTICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA  
BUCARAMANGA  
2013**

**ANÁLISIS DE LA RELACIÓN ENTRE LA MORBILIDAD POR AFECCIONES  
RESPIRATORIAS Y LA CONTAMINACIÓN DEL AIRE EN BUCARAMANGA  
EN EL 2009 UTILIZANDO UN MODELO DE REGRESIÓN DINÁMICA**

**CARLOS MANUEL SARMIENTO SOTO**

**Trabajo de Grado para optar el grado académico de Especialista en  
Estadística**

**Director  
Dr. GABRIEL YÁÑEZ CANAL**



**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE CIENCIAS  
ESCUELA DE MATEMÁTICAS  
ESPECIALIZACIÓN EN ESTADÍSTICA  
BUCARAMANGA  
2013**

## AGRADECIMIENTOS

*Es mi intención agradecer sinceramente a todas aquellas personas e instituciones que hicieron posible llevar a feliz término este trabajo de grado.*

*A mi familia, por su amor y acompañamiento en todas las etapas de mi vida.*

*Al Dr. Gabriel Yáñez Canal, director, por su guía y apoyo incondicional en todo momento.*

*Carlos Manuel Sarmiento Soto*

## TABLA DE CONTENIDO

	Pág.
INTRODUCCIÓN.....	13
1.ANTECEDENTES .....	16
2. LA CONTAMINACIÓN DE PARTÍCULAS .....	18
2.1 TAMAÑO DE LAS PARTÍCULAS .....	18
2.2 LAS PARTÍCULAS Y LA SALUD .....	19
2.3 POBLACIÓN Y RIESGO .....	20
3. MARCO LEGAL Y NORMATIVIDAD .....	22
4. MARCO TEÓRICO .....	26
4.1 CONCEPTOS BÁSICOS DE SERIES DE TIEMPO.....	26
4.1.1 Definición .....	26
4.1.2 Componentes .....	26
4.1.3 Parámetros de una serie .....	28
4.1.4 Clasificación .....	29
4.2 SERIES ESTACIONARIAS .....	29
4.3 RUIDO BLANCO.....	31
4.4 PROCESOS LINEALES ESTACIONARIOS .....	32
4.4.1 Procesos Autorregresivos AR(p).....	32
4.4.2 Proceso de Medias Móviles MA (q).....	34
4.4.3 Proceso Autorregresivo de Medias Móviles ARMA (p,q) .....	34
4.5 PROCESOS LINEALES NO ESTACIONARIOS.....	36
4.5.1 Proceso Autorregresivo Integrado de Media Móvil ARIMA (p, d, q) .....	36
4.5.2 Proceso Estacional Autorregresivo Integrado y de Media Móvil ARIMA (p, d, q)(P, D, Q) <sub>s</sub> .....	37
4.6 METODOLOGÍA DE BOX JENKINS PARA IDENTIFICACIÓN DE MODELOS ARIMA .....	38
4.6.1 Identificación .....	38
4.6.2 Estimación y verificación.....	39
4.6.3 Predicción .....	39

4.7 MODELOS DE FUNCIONES DE TRANSFERENCIA.....	39
4.7.1 La función de correlación cruzada (CCF) .....	40
4.7.2 Relación entre la Función de Correlación Cruzada y la Función de Transferencia .....	42
5. ANÁLISIS DE LA INFORMACIÓN .....	43
5.1 DESCRIPCIÓN DE LOS DATOS.....	43
5.2 ANÁLISIS DE LAS SERIES DE TIEMPO .....	43
5.2.1 Análisis de la serie <i>CONSULTAS</i> .....	44
5.2.2 Análisis de la serie <i>MÁXIMOS_PM10</i> .....	55
5.3 MODELO DE LA FUNCIÓN DE TRANSFERENCIA .....	66
CONCLUSIONES.....	77
BIBLIOGRAFÍA.....	79

## LISTA DE TABLAS

Pág.

Tabla 1. Normas de Calidad del Aire utilizadas en el cálculo del IBUCA.	22
Tabla 2. Descripción epidemiológica del IBUCA. ....	24
Tabla 3. Prueba de Dickey Fuller para la serie CONSULTAS.....	46
Tabla 4. Prueba de Dickey Fuller para la serie SDLCONSULTAS. ....	51
Tabla 5. Estimaciones para el modelo de CONSULTAS. ....	54
Tabla 6. Prueba de Dickey Fuller para la serie MÁXIMOS_PM10.....	57
Tabla 7. Prueba de Dickey Fuller para la serie SDLMÁX_PM10.....	61
Tabla 8. Estimaciones para el modelo de los MÁX_PM10. ....	64

## LISTA DE FIGURAS

Pág.

Figura 1. Gráfico temporal de la serie de tiempo. ....	26
Figura 2. Secuencia del número de consultas en el servicio de urgencias en Bucaramanga por afecciones respiratorias durante el año 2009. ....	44
Figura 3. Tendencia Estimada para la serie CONSULTAS. ....	45
Figura 4. Autocorrelaciones estimadas para la serie CONSULTAS. ....	47
Figura 5. Gráfico de subseries estacionales para CONSULTAS. ....	48
Figura 6. Comparativo del comportamiento de las CONSULTAS en cada día. ....	48
Figura 7. Estadísticos descriptivos para CONSULTAS. ....	49
Figura 8. Transformaciones para estabilizar la serie CONSULTAS. ....	50
Figura 9. Autocorrelaciones estimadas para la serie SDLCONSULTAS. .	51
Figura 10. Estructura de la AC Y ACP para el modelo de CONSULTAS. .	53
Figura 11. Correlograma de los residuos del modelo de CONSULTAS. .	54
Figura 12. Gráfico de los valores actuales, los pronósticos y los residuos para el modelo de consultas. ....	55
Figura 13. Secuencia de los valores máximos de las partículas de PM10 en al aire de Bucaramanga durante el año 2009. ....	56
Figura 14. Tendencia estimada para la serie MÁXIMOS_PM10. ....	57
Figura 15. Autocorrelaciones estimadas para la serie MÁXIMOS_PM10. .	58
Figura 16. Gráfico de subseries estacionales para MÁXIMOS PM <sub>10</sub> . ....	59
Figura 17. Comparativo del comportamiento de los MÁXIMOS DE PM <sub>10</sub> en cada día. ....	59
Figura 18. Estadísticos descriptivos para MÁXIMOS_PM10. ....	60
Figura 19. Transformaciones para estabilizar la serie MÁXIMOS PM10. .	61
Figura 20. Autocorrelaciones estimadas para la serie SDLMÁX_PM10. .	62
Figura 21. Estructura de la AC y ACP para el modelo de MÁX_PM10. .	63
Figura 22. Correlograma de los residuos del modelo de MÁX_PM10. ...	64
Figura 23. Gráfico de los valores actuales, los pronósticos y los residuos del modelo de MÁX_PM10. ....	65

Figura 24. Gráficos comparativos entre la serie de MÁX_PM <sub>10</sub> y CONSULTAS. ....	66
Figura 25. Estimaciones y residuos del modelo 1.....	70
Figura 26. Estimaciones y residuos del modelo 2.....	71
Figura 27. Estimaciones y residuos del modelo 3.....	72
Figura 28. Estimaciones y residuos del modelo 4.....	73
Figura 29. Estimaciones y residuos del modelo 5.....	74
Figura 30. Prueba de estacionariedad de los residuos de la regresión. ....	75

## RESUMEN

**TÍTULO:** ANÁLISIS DE LA RELACIÓN ENTRE LA MORBILIDAD POR AFECCIONES RESPIRATORIAS Y LA CONTAMINACIÓN DEL AIRE EN BUCARAMANGA EN EL 2009 UTILIZANDO UN MODELO DE REGRESIÓN DINÁMICA\*.

**AUTORES:** Carlos Manuel Sarmiento Soto\*\*

**PALABRAS CLAVES** Contaminación del aire  
Afecciones respiratorias  
Regresión dinámica  
Particulado PM10

### DESCRIPCIÓN

La contaminación por partículas es una mezcla de sólidos y microscópicas gotas de líquidos en suspensión en el aire con diferentes concentraciones del particulado ( $PM_{10}$ ). La población residente en la ciudad de Bucaramanga está expuesta a respirar aire contaminado dependiendo del área geográfica donde habitan, reportándose, que los niños que viven en zonas de mayor contaminación presentan mayor incidencia de síntomas de enfermedades respiratorias. El objetivo de este estudio era explorar si la frecuencia de consultas diarias por estas causas en los servicios médicos de urgencias de Bucaramanga se asocia con los niveles de material particulado ( $PM_{10}$ ) cuantificado diariamente en los Centros de Monitoreo de Calidad del Aire por la Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga (CDMB). Se realizó inicialmente un estudio descriptivo, encontrándose que en términos generales la calidad del aire en la ciudad de Bucaramanga es buena en cuanto al particulado  $PM_{10}$  ya que sus niveles se mantienen por debajo del valor establecido por la norma la mayor parte del tiempo. Además al analizar las series de consultas y  $PM_{10}$  por medio de modelos de regresión dinámica se encontró que la estimación proporcionada por el modelo de función de transferencia es demasiado suave, no logrando captar la estructura de la serie.

---

\* Trabajo de grado.

\*\* Facultad de Ciencias. Especialización en Estadística. Dr. Gabriel Yáñez Canal

## ABSTRACT

**TITLE:** ANALYSIS OF THE RELATIONSHIP BETWEEN THE CONDITIONS OF RESPIRATORY DISEASE AND AIR POLLUTION BUCARAMANGA IN 2009 USING A DYNAMIC REGRESSION MODEL \*

**AUTHOR:** Carlos Manuel Sarmiento Soto\*\*

**KEY WORDS:** Air Pollution  
Respiratory conditions  
Dynamic Regression  
PM10

### DESCRIPTION:

Particulate matter contamination is a microscopic mixture of solids and liquid droplets suspended in the air with different concentrations of particles (PM10). The residing population of the city of Bucaramanga is exposed to breathing polluted air, depending on the geographical area where they live, with reports that children living in areas of high pollution have a higher incidence of respiratory symptoms. The aim of this study was to explore whether the frequency of daily queries from these causes in the emergency medical services of Bucaramanga is associated with levels of particulate matter (PM10) quantified daily by the Air Quality Centers Monitoring of the "Corporacion Autonoma Regional para la Defensa de la Meseta de Bucaramanga" (CMDDB). A descriptive study was initially performed, finding that the overall air quality in the city of Bucaramanga is good in terms of PM10 since their levels are kept below the established standard most of the time. In addition after analyzing the series of consultations and PM10 through dynamic regression models found that the estimate given by the transfer function model is too soft, failing to grasp the structure of the series.

---

\* Grade work.

\*\* Faculty of Science. Specialization in Statistics. Dr. Gabriel Yáñez Canal

## INTRODUCCIÓN

“La contaminación del aire es una amenaza aguda, acumulativa y crónica para la salud humana y el ambiente”. Las comunidades en todo el mundo están expuestas a contaminantes del aire que pueden detonar o agravar afecciones respiratorias, cardíacas y otras. Los problemas de salud causados por la contaminación pueden verse influidos por factores tales como: magnitud, alcance y duración de la exposición, edad y susceptibilidad de cada persona entre otros.

La Organización Panamericana de la Salud (OPS), a través de su Centro Panamericano de Ingeniería Sanitaria y Ciencias del Ambiente (CEPIS), han analizado la situación de la gestión de la calidad del aire urbano en exteriores en la Región de los países de América Latina y el Caribe instando a las naciones, con el fin de minimizar el riesgo que representa la contaminación del aire para la salud humana, a establecer estructuras institucionales y técnicas para mejorar las acciones de vigilancia, control y prevención del medio ambiente.

Los materiales particulados (polvo, MP), identificados como PST (partículas suspendidas totales),  $PM_{10}$  (material particulado con un diámetro menor a 10 micrómetro,  $\mu m$ ) y  $PM_{2.5}$  (material particulado con un diámetro menor a 2,5 micrómetro,  $\mu m$ ) son parte de los indicadores criterios para medir la calidad del aire limpio en un área determinada (OMS, 2006). Estas partículas tienen un impacto en la salud, ya que presentan características toxicológicas que dependen de su origen y composición química.

La mayor parte de los problemas de contaminación del aire son el resultado del proceso de desarrollo e industrialización de las ciudades, así como también, del crecimiento del parque automotor. Las emisiones del sector transporte están compuestas por contaminantes entre los que se encuentran el material particulado menor a 10 y a 2,5 micras ( $PM_{10}$  y  $PM_{2.5}$ ), óxidos de azufre ( $SO_x$ ), óxidos de nitrógeno ( $NO_2$ ), metales pesados y óxidos que a su vez son precursores de reacciones

fotoquímicas en la atmósfera de ultra - partículas de ozono (O<sub>3</sub>) entre otros.

En el informe sobre contaminación del aire y salud en América Latina y el Caribe del CEPIS del año 2005 se reportan más de 35.000 muertes que se atribuyen a la contaminación del aire cada año. En este reporte se establece una asociación entre morbilidad y mortalidad con el deterioro de la calidad del aire en los principales centros urbanos. Para Río de Janeiro (Brasil) se ha observado que un aumento de 10 µg/m<sup>3</sup> en la concentración de PM<sub>10</sub> conlleva un incremento de 1,84% en la admisión hospitalaria por causas respiratorias; en Ciudad de México D.F. (México) el mismo incremento en el material particulado correspondió a un cambio de 1,83% en la mortalidad diaria por todas las causas y todas las edades; para Santiago de Chile (Chile) este cambio es del 0,75%.

Además de un incremento en la morbilidad, también se informa de incrementos en las tasas de mortalidad por enfermedad respiratoria y cardíaca y de efectos sobre la función respiratoria con reducción del Volumen Espiratorio Forzado (VEF1), Capacidad Vital Forzada (CVF), Flujo Espiratorio Forzado (FEF25-75%) y Flujo Espiratorio Pico (FEP), siendo la población adulta mayor y la infantil, los que presentan aumento en las admisiones hospitalarias por EPOC y asma e incremento en la prevalencia de condiciones respiratorias como tos y expectoración crónica, dificultad respiratoria diurna y nocturna y disnea de esfuerzo, asociados con mayores exposiciones en las concentraciones promedio anuales de dichos contaminantes.

Este estudio está organizado en seis capítulos cuyos contenidos se resumen a continuación:

En el primer capítulo, ANTECEDENTES, se exponen de manera limitada algunos trabajos e investigaciones relacionados con la contaminación del aire y los efectos sobre la salud de la población.

En el segundo capítulo, LA CONTAMINACIÓN DE LAS PARTÍCULAS, se presenta una breve descripción de los contaminantes del aire, sus características y riesgos para la salud de las comunidades.

En el tercer capítulo, MARCO LEGAL Y NORMATIVIDAD, se relacionan las principales disposiciones legales que existen en Colombia para la calidad del aire.

En el cuarto capítulo, MARCO TEÓRICO, se extractan las principales ideas de la teoría de series de tiempo utilizada en este estudio.

En el quinto capítulo, ANÁLISIS DE LA INFORMACIÓN, se inicia con un análisis de las series de tiempo de las consultas de urgencias en el municipio de Bucaramanga y los registros diarios de los niveles de  $PM_{10}$  realizados en la estación automática del centro de Bucaramanga para el año 2009. Posteriormente el análisis se centra en las series cronológicas de las variables de interés: *consultas y máximo  $PM_{10}$* .

En el sexto capítulo, ANÁLISIS DE RESULTADOS, partiendo de los modelos encontrados en el capítulo anterior, el interés en este capítulo se centra en los modelos de transferencia entre las series cronológicas de las variables de interés: *consultas y máximo  $PM_{10}$*  y los resultados obtenidos con este procedimiento.

En el séptimo capítulo, CONCLUSIONES, se precisan algunos resultados importantes observados en el transcurso de este estudio que pueden generar interés en la continuación de futuras investigaciones en el tema.

## 1. ANTECEDENTES

Hoy en día una de las grandes preocupaciones de la salud pública en varios países, es la contaminación del aire por concentraciones de partículas y contaminantes.

En el mundo son varias las investigaciones realizadas que describen la contaminación del aire con efectos adversos sobre la salud de la población. Los primeros estudios epidemiológicos (O'Nelly et al, 2004; American Heart Association (AHA), 2004) realizados en algunos países, señalan una asociación estadísticamente significativa entre la mortalidad por causas respiratorias y cardiovasculares y los días en que se incrementan en el aire las partículas de diez, y menos micrómetros de diámetro ( $PM_{10}$ ,  $PM_{2.5}$  y UF o  $PM < 0.1$ ). Investigaciones realizadas a nivel mundial reportan asociaciones entre concentraciones de contaminantes y afectaciones en la salud, especialmente en centros urbanos densamente poblados, siendo las variaciones temporales de la concentración de  $PM_{10}$  en el ambiente las de mayor relación con un aumento de la mortalidad diaria.

A raíz de estos, y muchos otros estudios, en los Estados Unidos y países Europeos hay suficiente evidencia que muestra que la contaminación del aire tiene incidencia en el incremento en la prevalencia de asma en niños y adultos y en las exacerbaciones de esta enfermedad. Se ha identificado que el incremento en los niveles de  $PM_{10}$ ,  $SO_2$  y  $NO_2$  se asocia a un incremento en las admisiones hospitalarias por bronquitis y asma, empeoramiento del asma, incremento en los síntomas respiratorios y en el uso de medicamentos.

Estudios como los realizados por (Romero, 1988) en La Habana (Cuba), (Guillen, 1996) en Cartagena (España), (Aristízabal, 1997) en Bogotá (Colombia) entre otros, determinan que la exposición a corto plazo a material particulado suspendido y al dióxido de azufre ( $SO_2$ ), se asocia con el aumento de la morbilidad por problemas respiratorios.

Recientemente en Bucaramanga algunos investigadores han estudiado la prevalencia e incidencia de síntomas respiratorios indicativos de

asma en la población preescolar residente en zonas con diferentes niveles de contaminación atmosférica por material particulado ( $PM_{10}$ ) (Rodríguez et al, 2010 a,b).

Los resultados del primer estudio (Rodríguez et al, 2010 a) no muestran asociación entre la contaminación atmosférica y la presencia de síntomas respiratorios indicativos de asma. La dificultad para analizar esta asociación (asma y contaminación atmosférica) ha sido descrita por otros investigadores (Rosas et al, 1998), reportándose que de los elementos criterios puede ser el ozono, y no las partículas, el causante del incremento de asma y ataques de asma de la población (EPA, 2010).

El segundo estudio de cohorte indica que la incidencia de síntomas respiratorios entre niños preescolares bumanguenses expuestos a mayores concentraciones de  $PM_{10}$  (Md anual  $55,5\mu g/m^3$ ) comparados con los expuestos a menores concentraciones (Md.  $32,5\mu g/m^3$ ) varía según los niveles de exposición. En la zona de mayor contaminación, un incremento de  $10\mu g/m^3$  en la concentración atmosférica de  $PM_{10}$  incrementa en 1,3 el número de síntomas diarios reportados, siendo el estornudo y la tos con flemas los síntomas más frecuentes (Rodríguez et al, 2010b). La sumatoria de los diferentes síntomas respiratorios presentados por los preescolares de Bucaramanga señala además, que los residentes en la zona de mayor contaminación presentan 13% más síntomas respiratorios que los de las otras zonas y que son ellos los que acuden con mayor frecuencia a los servicios de atención médica siendo en mayor número diagnosticados a utilizar dispositivos de inhalación.

En el mismo estudio a pesar de lo observado en el análisis de incidencia, el análisis multivariado de los datos muestra que la contaminación atmosférica no tiene un efecto adverso contundente sobre la presencia de síntomas respiratorios en preescolares, sin embargo, destacan que el alto flujo vehicular frente a la vivienda (variable considerada como de exposición) en niños con historia de asma, se asocia con mayor riesgo de desarrollar síntomas respiratorios.

## 2. LA CONTAMINACIÓN DE PARTÍCULAS

La contaminación por partículas es una mezcla de sólidos y microscópicas gotas de líquidos en suspensión en el aire. Esta contaminación, también conocida como la materia particulada, se produce por varios componentes, incluyendo ácidos (como nitratos y sulfatos), productos químicos orgánicos, partículas de metales, del suelo o el polvo, y alergen<sup>1</sup> (tales como fragmentos de polen o esporas de moho).

### 2.1 TAMAÑO DE LAS PARTÍCULAS

El tamaño de las partículas está directamente relacionado con su potencial para causar problemas de salud. Las pequeñas partículas de menos de 10 micrómetros de diámetro presentan los mayores problemas, ya que pueden penetrar profundamente en los pulmones, y algunas incluso entrar en el torrente sanguíneo. En la bibliografía consultada, se identifica que las partículas vienen en una amplia gama de tamaños:

#### □ Partículas (PM<sub>10</sub>)

También conocidas como "partículas", son las menores de 10 micras de diámetro (PM<sub>10</sub>) y tan pequeñas que pueden entrar en los pulmones y causar serios problemas de salud (diez micrómetros es menor que la anchura de un cabello humano).

---

<sup>1</sup> Un alergeno es una sustancia que puede provocar una reacción alérgica. En algunas personas, el sistema inmunitario considera a los alergen<sup>1</sup> "extraños" o "peligrosos". Esto es lo que lleva a que se presenten síntomas de alergias  
<http://www.nlm.nih.gov/medlineplus/spanish/ency/article/002229.htm>

#### □ Partículas finas (PM<sub>2.5</sub>)

Conocidas como "finas", son las que tienen menos de 2,5 micrómetros de diámetro y solo pueden ser detectadas con un microscopio electrónico. Algunas fuentes de partículas finas incluyen todo tipo de combustión como: automóviles, plantas de energía, quema residencial de madera, incendios forestales, quemas agrícolas, y algunos procesos industriales.

#### □ Partículas de polvo grueso

Se conocen como "gruesas", y corresponde a las partículas entre 2,5 y 10 micrómetros de diámetro. Algunas fuentes de partículas gruesas incluyen las operaciones de triturar o moler en las industrias y el polvo provocado por los vehículos que circulan en las carreteras.

## 2.2 LAS PARTÍCULAS Y LA SALUD

La exposición a estas partículas logra afectar tanto a los pulmones como al corazón. Las partículas más grandes son de menor importancia, aunque pueden irritar los ojos, la nariz y la garganta. Las partículas más pequeñas son las que generan más preocupación, incluyen las "partículas finas" (como las que se encuentran en el humo y la neblina), que tienen 2.5 micrómetros de diámetro o menos, y las "partículas gruesas" (como las que se encuentran en el polvo arrastrado por el viento) que tienen diámetros entre 2,5 y 10 micrómetros.

La exposición de partículas puede conducir a una variedad de efectos sobre la salud. Por ejemplo, numerosos estudios han encontrado relación entre los niveles de partículas y el aumento de los ingresos hospitalarios y visitas al servicio médico de urgencias, e incluso, a la muerte por enfermedades del corazón o de los pulmones.

Tanto la exposición de partículas a largo, como la exposición a corto plazo, se han relacionado con problemas de salud.

La exposición a largo plazo, como las que sufren las personas que viven desde hace muchos años en zonas con altos niveles de partículas, se ha asociado con problemas como la reducción de la función pulmonar y el desarrollo de la bronquitis crónica y muerte prematura incluso.

La exposición a corto plazo (horas o días), puede agravar las enfermedades pulmonares, causando ataques de asma y bronquitis aguda, como también aumentar la susceptibilidad a las infecciones respiratorias. En las personas con enfermedades del corazón, las exposiciones a corto plazo se han vinculado a los ataques del corazón y arritmias.

### **2.3 POBLACIÓN Y RIESGO**

La exposición a las partículas genera varios efectos sobre la salud, incluso si una persona se encuentra sana, puede experimentar los síntomas de carácter temporal, como la irritación de los ojos, la nariz y la garganta, tos, flema, opresión en el pecho, y dificultad para respirar.

Las personas con enfermedades del corazón o de los pulmones, como la enfermedad coronaria, insuficiencia cardíaca congestiva, el asma, o la enfermedad pulmonar obstructiva crónica (EPOC), tienen un mayor riesgo, ya que las partículas pueden agravar estas enfermedades. También se tienen en cuenta en este grupo a las personas con diabetes debido a que son más propensas a tener enfermedad cardiovascular subyacente.

Otro grupo de alto riesgo está conformado por los adultos mayores posiblemente debido a que no siempre tienen diagnosticada una enfermedad cardíaca, pulmonar o diabetes. Muchos estudios muestran que cuando los niveles de partículas son altos, los adultos mayores son más propensos a ser hospitalizados.

Los niños menores también hacen parte de este grupo de alto riesgo por varias razones: sus pulmones todavía se están desarrollando, pasan

más tiempo en actividad y son más propensos a tener asma o enfermedades respiratorias agudas, que pueden verse agravadas cuando los niveles de partículas en el aire son altas.

Parece que el riesgo varía durante toda la vida, por lo general es más alto en la primera infancia, menor en los adolescentes y adultos sanos más jóvenes, y aumenta nuevamente en la edad madura a la vejez por la incidencia de enfermedades cardíacas, pulmonares y la diabetes.

Los factores que aumentan el riesgo de ataques al corazón como son la presión arterial alta o los niveles elevados de colesterol, también pueden aumentar el riesgo de las partículas. Además, los científicos están evaluando nuevos estudios que sugieren que la exposición a niveles altos de partículas también puede estar asociada con bajo peso al nacer en los bebés, partos prematuros, muertes fetales y posiblemente muerte infantil.

### 3. MARCO LEGAL Y NORMATIVIDAD

En Colombia, en el año 2010, el Ministerio de Ambiente, Vivienda y Desarrollo Territorial expidió la Norma de referencia de Calidad del Aire para todo el territorio Nacional actualmente vigente (Resolución N° 610), en la cual se establece el máximo nivel permisible de concentración en el aire de cinco contaminantes criterio que se presentan en la Tabla 1.

*Tabla 1. Normas de Calidad del Aire utilizadas en el cálculo del IBUCA.*

CONTAMINANTE	PERIODO	NORMA	UNIDAD
Partículas suspendidas PM <sub>10</sub>	24 horas	134	µg/m <sup>3</sup>
Óxidos de Azufre, SO <sub>x</sub>	24 horas	86	Ppb
Óxidos de Nitrógeno, NO <sub>2</sub>	1 hora	95	Ppb
Monóxido de carbono, CO	1 hora	31	Ppm
Oxidante Fotoquímico, O <sub>3</sub>	1 hora	54	ppb

*Fuente: Informe IBUCA 2009.*

Con respecto a la Normatividad Colombiana el valor máximo permisible por día para PM<sub>10</sub> es de 100 microgramos por metro cúbico (µg/m<sup>3</sup>) y el promedio de concentración anual establecida entre el 1 de enero y el 31 de diciembre es de 50 µg/m<sup>3</sup>.

Por otro lado, el Decreto 979 del mismo año establece los niveles de prevención, alerta y emergencia por contaminación del aire con base en las mediciones de calidad del aire obtenidas diariamente para que en consecuencia se tomen las medidas de control que sean necesarias para proteger la salud de la población.

En la página web de la Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga (CDMB)<sup>2</sup>, en el marco conceptual del Sistema de Información Ambiental (SIA) declara que tiene como objetivo primordial:

“conservar un medio ambiente sano, para lo cual es necesario que los recursos naturales receptores de los vertimientos, residuos y emisiones de las actividades humanas, a saber, el agua, el suelo y el aire, conserven unas condiciones de cantidad y calidad, que les permitan recibir determinado nivel de contaminantes y tener la capacidad de autodepurarlos de tal forma que se conserve un medio ambiente adecuado para la vida tanto de las personas como de las diferentes especies de fauna y flora que conforman nuestros ecosistemas”.

Para tal fin la CDMB cuenta con unas redes de monitoreo del agua, del aire, del clima y el análisis multitemporal de las coberturas vegetales mediante el procesamiento de imágenes satelitales.

El Sistema de Vigilancia de la Calidad del Aire (SVCA) actualmente está conformado por cinco (5) estaciones automáticas, tres (3) estaciones manuales y cuatro (4) estaciones meteorológicas ubicadas estratégicamente en el área metropolitana de Bucaramanga<sup>3</sup>.

Las estaciones de monitoreo Automáticas entregan un reporte diario de ciertos contaminantes criterio, exigidos por la normatividad colombiana que son generados por fuentes contaminantes fijas y móviles. Además, estas estaciones también ofrecen reportes meteorológicos que miden variables como presión, temperatura, humedad relativa, dirección del viento, brillo solar, etc.

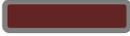
---

<sup>2</sup> <http://www.cdmb.gov.co/web/index.php/sistema-de-informacion-ambiental.html>

<sup>3</sup> Estación Centro, Estación Ciudadela, Estación Florida, Estación Cabecera, Estación Norte, Estación La Concordia, Estación Tránsito, Estación La Joya.

Los resultados de los muestreos son informados a la comunidad utilizando el índice IBUCA que permite establecer como se encuentra la calidad del aire con respecto a los límites locales (ver Tabla 2). El indicador es adimensional y posee una escala de 0 a 10 que depende del grado de contaminación y está relacionado con la afectación que tiene la contaminación del aire sobre la salud humana.

*Tabla 2. Descripción epidemiológica del IBUCA.*

<b>IBUCA</b>	<b>DESCRIPTOR</b>	<b>CALIFICACIÓN EPIDEMIOLÓGICA</b>	<b>COLOR</b>
<b>0 – 1,25</b>	<b>BUENO</b>	<i>La calidad de aire es considerada como satisfactoria u la afectación en la contaminación del aire es pequeña y no evidencia ningún efecto en la salud humana.</i>	
<b>1,26 – 2,50</b>	<b>MODERADO</b>	<i>La calidad de aire es aceptable y no tiene ningún efecto sobre la población en general.</i>	
<b>2,51 – 7,50</b>	<b>REGULAR</b>	<i>Aumento de molestias en personas con padecimientos respiratorios y cardiovasculares, aparición de ligeras molestias en la población en general.</i>	
<b>7,51 – 10,0</b>	<b>MALO</b>	<i>Agravamiento significativo de la salud en personas con enfermedades cardíacas o respiratorias. Afectación de la población sana.</i>	
<b>➤ 10,0</b>	<b>PELIGROSO</b>	<i>Alto riesgo para la salud de la población. Aparición de efectos al nivel de daño</i>	

*Fuente: Informe Anual del Estado de los recursos naturales, CDMB, 2011.*

En el informe Anual del estado de los recursos naturales del año 2009, se lee:

“la calidad del aire durante los tres últimos años se ha enfocado principalmente en el material particulado respirable, realizando monitoreos en 8 sitios diferentes. Este contaminante y el ozono troposférico son los que deterioran en mayor medida el estado de la calidad del

aire que respiramos. En su orden, los sectores de mayor contaminación por  $PM_{10}$  fueron la carrera 27, la carrera 17 y el centro. La causa principal de generación de  $PM_{10}$  en el aire es el gran número de vehículos de diesel que circulan por vías tan angostas y congestionadas del Área Metropolitana de Bucaramanga y debido a la ejecución de las obras civiles del proyecto Metrolínea” (Informe Calidad del aire 2009, CDMB, p, 56).

## 4. MARCO TEÓRICO

A continuación se resumen los aspectos más importantes de la teoría de series de tiempo según la bibliografía consultada.

### 4.1 CONCEPTOS BÁSICOS DE SERIES DE TIEMPO

#### 4.1.1 Definición

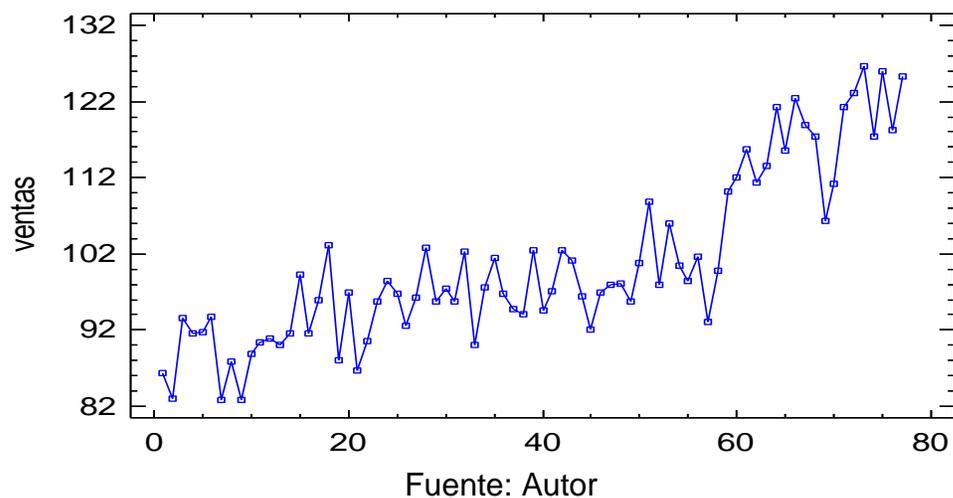
Una serie de tiempo es un conjunto de mediciones de cierto fenómeno o experimento (variables aleatorias) registradas secuencialmente en un período a intervalos iguales de tiempo. Se suelen representar por medio de una ecuación matemática que describe los valores de la variable observada como una función del tiempo, es decir  $X=f(t)$ . (Ver Figura 1)

#### 4.1.2 Componentes

La teoría clásica considera una serie de tiempo formada por cuatro componentes teóricas: tendencia, variaciones estacionales, variaciones cíclicas y variaciones aleatorias.

*Figura 1. Gráfico temporal de la serie de tiempo.*

Gráfico de Serie de Tiempo para ventas



□ *Componente TENDENCIA ( $T_t$ )*

La tendencia viene dada por el movimiento general de la serie a largo plazo o el cambio a largo plazo de la media.

□ *Componente ESTACIONAL ( $S_t$ )*

Las variaciones estacionales son oscilaciones que se producen en un período igual o inferior a un año y que se reproducen de manera reconocible en los diferentes períodos de tiempo (semestral, mensual, etc).

□ *Componente cíclico*

Las variaciones cíclicas son oscilaciones que se producen en un período superior a un año y que se deben, principalmente, a la alternancia de etapas largas (ciclos) en las que se repite el comportamiento de la serie. Es más difícil de detectar ya que estas variaciones tienen un período no fácilmente identificable y en muchos casos es variable, siendo frecuente la existencia de ciclos que se superponen.

□ *Componente ALEATORIA ( $R_t$ )*

Las variaciones aleatorias o también llamadas variaciones residuales son las que no responden a ningún patrón de comportamiento, son el resultado de los factores aleatorios que inciden aisladamente en una serie de tiempo.

Por lo anterior una serie de tiempo se puede denotar como:

$$X_t = T_t + S_t + R_t$$

$T_t$  : *tendencia – ciclo*

$S_t$  : *estacionalidad*

$R_t$  : *aleatoria*

El modelo clásico de series de tiempo supone que la serie se puede expresar como la suma (modelo aditivo) o el producto (modelo multiplicativo) de los tres componentes primordiales: tendencia, estacionalidad y aleatoria.

Para poder llegar a la fase de identificación de la serie, es necesaria la representación gráfica de la serie y la eliminación de la tendencia. La primera es de utilidad para definir sobre la estacionariedad y estacionalidad de la serie.

Pero estas dos características de la serie también pueden detectarse a través de las funciones de autocorrelación (FAC) y autocorrelación parcial (FACP) que miden la relación estadística entre las observaciones de una serie temporal.

La función de autocorrelación de una serie proporciona la estructura de dependencia lineal de la misma y la función de autocorrelación parcial proporciona la relación directa que existe entre observaciones separadas por  $k$  retardos.

La representación gráfica del estimador muestral de la FAC denotado por la expresión  $\gamma_k$  se denomina correlograma muestral y constituye un instrumento de gran interés práctico en el análisis de series.

#### 4.1.3 Parámetros de una serie

- La media: analiza la tendencia

$$\mu_t = E(X_t)$$

- La varianza: analiza la variación

$$\sigma_t^2 = E(X_t - \mu_t)^2$$

- La covarianza: analiza la variabilidad en dos tiempos diferentes, por lo tanto interesa la covarianza entre lo que ocurre en el tiempo  $t$  y en el  $t+k$ :

$$\gamma_{t,t+k} = COV(X_{t,t+k}) = E [(X_t - \mu_t)(X_{t+k} - \mu_{t+k})]$$

- La autocorrelación: da la estructura de autocorrelación de la serie.

$$\rho_{X_t, X_{t+k}} = Corr(X_t, X_{t+k}) = \frac{Cov(X_t, X_{t+k})}{\sqrt{V(X_t)V(X_{t+k})}}$$

#### 4.1.4 Clasificación

Las series temporales de pueden clasificar en:

- Series estacionarias. Son las series que se mantienen estables a lo largo del tiempo, es decir, cuando tanto la media, como la varianza, son constantes en el tiempo.
- Series No estacionarias. Son las series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo.

#### 4.2 SERIES ESTACIONARIAS

Una serie de tiempo es estacionaria si su media, su varianza y su covarianza permanecen iguales sin importar el momento en el cual se midan, es decir, son invariantes respecto al tiempo.

Una serie estacionaria varía alrededor de una media, luego tienen la misma distribución y la varianza es la misma:

- *La media*: no depende del tiempo

$$\mu_t = E(X_t)$$

□ *La varianza: no depende del tiempo*

$$\sigma_t^2 = VAR(X_t) = E(X_t - \mu_t)^2$$

□ *La*

*covarianza: no depende del tiempo sino de la separación. Es decir la covarianza entre  $X_t$  y  $X_{t+k}$  va a depender solo de la separación ( $k$ ).*

$$\gamma_k = COV(X_{t,t+k}) = E[(X_t - \mu_t)(X_{t+k} - \mu_{t+k})]$$

□ *La autocorrelación: se denomina autocorrelación de rezago  $k$ , caracteriza la estructura de autocorrelación de la serie. (autocorrelación entre  $X_t$  y  $X_{t+k}$ ).*

El valor de la autocorrelación para sucesivos valores del retardo  $k$  proporciona la llamada función de autocorrelación que refleja la memoria del proceso: el número de períodos durante los cuales una variable continua sigue teniendo influencia en la evolución del proceso.

$$\rho_{X_t, X_{t+k}} = Corr(X_t, X_{t+k}) = \frac{Cov(X_t, X_{t+k})}{\sqrt{V(X_t)V(X_{t+k})}}$$

$$\rho_k = \frac{\gamma_k}{\sigma^2}$$

Entre las propiedades de la autocorrelación de rezago  $k$  se encuentran:

- Es un coeficiente de correlación, por lo tanto es acotado (está entre -1 y 1) y tiene todas las propiedades de los coeficientes de correlación (si se cambia la escala de datos, no cambia).

$$-1 \leq \rho_k \leq 1$$

$\rho_k = 1$       *correlación perfecta*

$\rho_k = -1$      *correlación negativa*

$\rho_k = 0$       *no hay ninguna relación entre lo que ocurre  
en un tiempo  $t$  y un tiempo  $t+k$*

○  $\rho_0 = 1$

○  $\rho_{-k} = \rho_k$

$\rho_{-k}$  es la correlación entre  $-X_{t-k}$  y  $X_t$  es decir da lo mismo mirar la serie hacia atrás que hacia delante.

○  $\lim_{k \rightarrow \infty} \rho_k = 0$  es decir en la medida que se distancien las medidas ( $k$  aumente) la correlación entre ellas es cada vez menor.

### 4.3 RUIDO BLANCO

Un ruido blanco es un caso simple de los procesos estocásticos, donde los valores son independientes y distribuidos idénticamente a lo largo del tiempo con media cero e igual varianza.

Supongamos una variable  $a_t$  que tiene:

➤  $E(a_t) = 0$

➤  $VAR(a_t) = \sigma^2$ ,

➤  $COV(a_t, a_{t'}^*) = 0 \quad \forall t \neq t^*$

Sea la serie  $X_t$  tal que :

$$\begin{aligned} X_t = a_t &\Rightarrow \mu_x = 0 \\ &\sigma_x^2 = \sigma^2 \\ &\rho_1 = 0 \end{aligned}$$

$$\rho_1 = \frac{\gamma_1}{\sigma^2} \quad y \quad \gamma_1 = COV(a_t = a_{t+1}) = 0 \Rightarrow \rho_1 = 0, \rho_2 = 0, \rho_k = 0$$

$$\rho_k = \begin{cases} 1 & \rightarrow k = 0 \\ 0 & \rightarrow k > 0 \end{cases}$$

A esta serie se le llama *ruido blanco* porque es totalmente aleatoria, lo que ocurre en un punto no está correlacionado con otro, no hay autocorrelación de los residuos.

#### 4.4 PROCESOS LINEALES ESTACIONARIOS

##### 4.4.1 Procesos Autorregresivos AR(p)

Los modelos Autorregresivos se basan en la idea que el valor actual de la serie,  $X_t$ , puede explicarse en función de p valores pasados

$$X_{t-1}, X_{t-2}, \dots, X_{t-p}$$

donde  $p$  determina el número de rezagos necesarios para pronosticar un valor actual.

*Ejemplo 1: Proceso Autorregresivo de orden 1: AR(1)*

En los procesos AR(1) la variable  $X_t$  está determinada únicamente por el valor pasado, esto es  $X_{t-1}$

$$X_t = \theta_0 + \phi_1 X_{t-1} + a_t \quad (1)$$

□ *La media*

*La media no es  $\theta_0$  como en los modelos MA*

$$\mu_X = \frac{\theta_0}{1 - \phi_1}, \quad \phi_1 \neq 1 \quad \Rightarrow \quad \theta_0 = \mu_X (1 - \phi_1)$$

□ *La varianza*

$$X_t - \mu_X = \phi_1 (X_{t-1} - \mu_X) + a_t \quad (2)$$

La estructura de (2) es la misma estructura de (1). El modelo (2) es para las desviaciones de la media.

$$\sigma^2_X = \frac{\sigma^2}{1 - \phi_1^2} \quad |\phi_1| < 1$$

□ *La covarianza*

$$\gamma_1 = \phi_1 \sigma^2_X$$

$$\gamma_k = \phi_1^k \sigma^2_X$$

□ *La función de autocorrelación*

$$\rho_k = \phi_1^k$$

*Esto es exponencialmente decreciente*

#### 4.4.2 Proceso de Medias Móviles MA (q)

El modelo está determinado por una fuente externa. Estos modelos suponen linealidad, el valor actual de la serie,  $X_t$ , está influenciado por los valores de la fuente externa

*Ejemplo 2: Proceso de Media Móvil de orden 1: MA(1)*

La primera autocorrelación no es cero pero las demás sí.

$$X_t = \theta_0 + \theta_1 a_{t-1} + a_t$$

□ *La media*

$$\mu_X = \theta_0$$

□ *La varianza*

$$\sigma_X^2 = \sigma^2 (1 + \theta_1^2)$$

□ *La covarianza*

$$\gamma_1 = \theta_1 \sigma^2, \quad \gamma_2 = 0, \quad \dots, \gamma_k = 0 \quad \text{para } k > 1$$

□ *La función de autocorrelación*

$$\rho_0 = 1, \quad \rho_1 = \frac{\theta_1}{1 + \theta_1^2}, \quad \rho_k = 0, \quad k > 1$$

#### 4.4.3 Proceso Autorregresivo de Medias Móviles ARMA (p,q)

Es muy probable que una serie de tiempo  $X_t$ , tenga características de AR y de MA a la vez, en estos casos se dice que sigue un modelo ARMA. De

esta forma si  $X_t$  sigue un proceso  $ARMA(p, q)$ , en este proceso habrá  $p$  términos Autorregresivos y  $q$  términos de media móvil, esto es,  $ARMA(p,q)=AR(p)+MA(q)$ .

*Ejemplo 3: Proceso Autorregresivo de Medias Móviles ARMA (1,1)*

El modelo es:  $X_t = \theta_0 + \phi X_{t-1} + a_t - \theta_1 a_{t-1}$

Los valores de los parámetros básicos de este modelo se describen a continuación:

□ *La media:*

$$\mu_X = \theta_0 + \phi \mu_X$$

$$\mu_X = \frac{\theta_0}{1 - \phi} \quad |\phi| < 1; \quad |\theta| < 1$$

Estacionario

Invertible

□ *La varianza:*

$$\sigma^2_X = \phi^2 \sigma^2_X + \sigma^2 + \theta^2 \sigma^2 - 2\phi\theta\sigma^2$$

...

$$\sigma^2_X = \frac{\sigma^2(1 + \theta^2 - 2\phi\theta)}{1 - \phi^2}$$

□ *La Autocovarianza:*

$$\gamma_1 = \phi \sigma^2_X - \theta \sigma^2$$

$$\gamma_1 = \sigma^2 \left( \frac{\phi + \phi\theta^2 - \phi^2\theta - \theta}{1 - \phi^2} \right)$$

$$\gamma_1 = \sigma^2 \frac{(\phi - \theta)(1 + \phi\theta)}{1 - \phi^2}$$

$$\gamma_2 = \phi \gamma_1$$

...

$$\gamma_k = \phi \gamma_{k-1}$$

$$\gamma_k = \phi^{k-1} \sigma^2 \frac{(\phi - \theta)(1 + \phi\theta)}{1 - \phi^2}$$

□ *La autocorrelación:*

$$\rho_k = \begin{cases} 1 & k = 0 \\ \frac{(\phi - \theta)(1 + \phi\theta)}{1 + \theta^2 - 2\phi\theta} & k = 1 \\ \phi^{k-1} \frac{(\phi - \theta)(1 + \phi\theta)}{1 + \theta^2 - 2\phi\theta} & k > 1 \end{cases}$$

## 4.5 PROCESOS LINEALES NO ESTACIONARIOS

### 4.5.1 Proceso Autorregresivo Integrado de Media Móvil ARIMA (p, d, q)

Muchas series de tiempo no son estacionarias porque cambian de nivel o porque la varianza no es constante en el tiempo. Una forma de estabilizar la media y la varianza es a través de diferencias, procedimiento que conduce a los modelos ARIMA.

El operador de retraso ( $B$ ), retrasa una posición la observación.

Cuando una serie tiene tendencia es preciso quitarla mediante una diferencia regular. Una diferencia regular le resta a cada elemento de la serie la observación anterior:

Por ejemplo en la serie  $Y_t$  :

$$Y_t = X_t - X_{t-1}$$

$$Y_t = X_t - BX_{t-1} = (1 - B)X_t$$

*1-B: operador diferencia*

Expresado en forma del polinomio operador de retardos, el modelo ARIMA (p, d, q) es:

$$\phi_p(B)(1-B)^d X_t = \theta_q(B)a_t$$

*p*: orden regresivo (factor AR)

*d*: número de diferencias

*q*: orden factor MA

#### 4.5.2 Proceso Estacional Autorregresivo Integrado y de Media Móvil ARIMA (p, d, q)(P, D, Q)<sub>s</sub>

Con frecuencia cuando las series de tiempo muestran intervalos de observación menores a un año (semestral, mensual, diario, etc) tienen variaciones o patrones sistemáticos cada cierto período que deben ser captadas en los llamados factores estacionales dentro de la estructura del modelo a construirse.

Estas series estacionales pueden ser de dos tipos: aditivas o multiplicativas y al mismo tiempo, cada una de estas series puede ser estacionaria o no estacionaria. Con mayor frecuencia se presentan los modelos multiplicativos, de tal forma que se combinan términos ordinarios del proceso ARMA y términos estacionales, así como diferencias regulares y diferencias estacionales para transformarlas en series estacionarias.

Las series estacionales tienen la forma:  $1-B^s$  siendo *s* la longitud de la estación. Por ejemplo, si la serie tiene estacionalidad semanal, la diferencia estacional tiene la forma  $1-B^7$  que al aplicar a la serie  $X_t$  se traduce en:

$$(1-B^7)X_t = X_t - X_{t-7}$$

Este proceso presenta las siguientes características:

- Contiene una componente ARIMA (p, d, q) que modela la componente regular asociada a observaciones consecutivas.
- Contiene una componente ARIMA (P, D, Q) que modela la componente estacional asociada con las observaciones separadas por  $S$  períodos.

## 4.6 METODOLOGÍA DE BOX JENKINS PARA IDENTIFICACIÓN DE MODELOS ARIMA

### 4.6.1 Identificación

Representar gráficamente la serie temporal, la Función de Autocorrelación Simple (FAC) y la Función de Autocorrelación Parcial (FACP).

La exploración visual del gráfico indicará si la serie es estacionaria o no. Si la serie no es estacionaria se pueden hacer los siguientes correctivos según sea el caso:

- Si tiene tendencia: Se toman diferencias regulares hasta que desaparezcan. (normalmente el orden de la diferencia es 1 y raramente mayor a 3).
- Si la serie tiene estacionalidad: Se toman diferencias estacionales hasta que desaparezca el patrón estacional (en la práctica es muy raro aplicar más de una diferencia estacional).
- Si la serie es heterocedástica: es decir, no tiene varianza constante, habrá que transformar la serie (tomando el logaritmo Ln en muchos casos es suficiente).

Una vez que el gráfico de la serie original **transformada** indica que es estacionaria, se puede intentar deducir la estructura de la serie transformada observando su FAC y FACP.

#### 4.6.2 Estimación y verificación

Al observar la función de autocorrelación simple y parcial de la serie transformada, se pueden obtener los primeros modelos a probar. La forma más común de comprobar analíticamente estos modelos es ajustar varios candidatos y escoger aquel cuyos residuos son ruido blanco y sus valores de AIC (Criterio de Información de Akaike) y BIC (Criterio de Información Bayesiana) son inferiores comparados con los demás.

#### 4.6.3 Predicción

Los modelos ARIMA son buenos para realizar predicciones a corto plazo y esta es una de las razones por la que son tan utilizados.

### 4.7 MODELOS DE FUNCIONES DE TRANSFERENCIA

Si se supone que se tienen dos series temporales,  $X_t$  e  $Y_t$ , que interactúan entre ellas, la relación entre ambas se puede modelar de la siguiente forma:

$$Y_t = v(B)X_t + n_t \quad (1)$$

Siendo  $v(B)$  una serie en el operador de retraso  $B$ , definido por la expresión siguiente:

$$v(B) = \sum_{j=-\infty}^{\infty} v_j B^j \quad (2)$$

y  $n_t$  es ruido blanco independiente de  $X_t$ .

La expresión (2) expresa los valores de  $Y_t$  en función de los valores presente, pasado y futuro de la serie  $X_t$ . Como el interés se cifra en la explicación que pueden dar los valores pasados, se asume el modelo llamado *causal* para el cual los valores  $v_j$  son cero si  $j < 0$  (Wei, 2006).

Los valores  $v_j$  se denominan *los impulsos de la función de respuesta*.

De esta manera el modelo (1) toma la forma:

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + \eta_t \quad (3)$$

con  $v(B) = \sum_{j=0}^{\infty} |v_j| < \infty$  para que el modelo sea *estable*. Se asume, además, que  $X_t$  y  $Y_t$  son independientes.

Para evitar tener que estimar infinitos coeficientes  $v_j$  contando con un número finito de datos (los valores de  $X_t$  y  $Y_t$ ) se escribe (3) de la siguiente forma:

$$v(B) = \frac{w(B)B^b}{\delta(B)} \quad (4)$$

Donde  $w(B) = w_0 - w_1 B - \dots - w_s B^s$  y  $\delta(B) = 1 - \delta_1 B - \dots - \delta_r B^r$  y  $B^b$  es un parámetro de retardo que representa el retardo de tiempo real que transcurre antes de que el impulso de la variable de entrada ( $X_t$ ) produzca un efecto sobre la variable de salida ( $Y_t$ ).

#### 4.7.1 La función de correlación cruzada (CCF)

Así como la función de correlación (FAC) es la herramienta clave para determinar el tipo de modelo ARMA asociado a una serie temporal, la Función de Correlación Cruzada (CCF, por su sigla en inglés) es una medida útil de la fuerza y la dirección de la correlación entre series temporales.

Dados dos procesos estocásticos ( $X_t$ ) y ( $Y_t$ ) para  $t = 0, \pm 1, \pm 2, \dots$  se dice que  $X_t$  y  $Y_t$  son *estacionarias conjuntamente* si  $X_t$  y  $Y_s$  son ambas procesos estacionarios univariados y la función de covarianza cruzada

entre  $X_t$  y  $Y_s$ ,  $\text{cov}(X_t, Y_s)$  es una función solamente de la diferencia de tiempo  $s-t$ .

$$\gamma_{xy}(k) = E[(X_t - \mu_X)(Y_{t+k} - \mu_Y)]$$

Para  $k = 0, \pm 1, \pm 2, \dots$   $\mu_X = E(X_t)$ ,  $\mu_Y = E(Y_t)$

Estandarizando se obtiene:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y}$$

Por ejemplo, considerando el proceso AR(1)

$(1 - \phi B)Z_t = a_t$  donde  $|\phi| < 1$ , siendo  $a_t$  ruido blanco con media cero y varianza constante  $\sigma_a^2$ .

Haciendo los cálculos se obtiene que

$$\gamma_{az}(k) = \begin{cases} \phi^k \sigma_a^2 & k \geq 0 \\ 0 & k < 0 \end{cases} \quad y$$

$$\rho_{az}(k) = \begin{cases} \phi^k \sqrt{1 - \phi^2} & k \geq 0 \\ 0 & k < 0 \end{cases}$$

Recordando que un modelo ARMA de la forma  $\phi_p(B)Z_t = \theta_q(B)a_t$  puede ser escrito en la forma  $Z_t = \frac{\theta_q(B)}{\phi_p(B)}a_t$  se concluye que el modelo ARMA univariado puede ser visto como un caso especial de un modelo de función de transferencia sin ruido blanco.

En este caso, la serie  $Z_t$  es la salida, el ruido blanco  $a_t$  es la entrada y la función de transferencia es la función  $v(B) = \frac{\theta_q(B)}{\phi_p(B)}$ .

#### 4.7.2 Relación entre la Función de Correlación Cruzada y la Función de Transferencia

La relación entre CCF y el impulso de la función de respuesta  $v_j$  está dada por la expresión:

$$\rho_{xy}(k) = \frac{\sigma_x}{\sigma_y} [v_0\rho_x(k) + v_1\rho_x(k-1) + v_2\rho_x(k-2) + \dots] \quad (5)$$

Obtener de la ecuación anterior los valores de  $v_j$  en función de  $\rho_{xy}(k)$  y  $\rho_x(k)$  es sumamente complicado. Ahora bien, si se asume que la serie de entrada  $X_t$  es ruido blanco, es decir,  $\rho_x(k) = 0$  para  $k \neq 0$ , la expresión (5) toma la forma:

$$v_k = \frac{\sigma_x}{\sigma_y} \rho_{xy}(k) \quad (6)$$

De donde se desprende que el impulso de la función de respuesta  $v_k$  es directamente proporcional a la función de correlación cruzada. La idea es estimar los valores de  $v_k$  a partir de las estimaciones de los  $\rho_{xy}(k)$ . Los detalles técnicos de los procesos de estimación y de los procesos de blanqueo de las series de entrada y salida se pueden consultar en Wei, 2006).

El objetivo final del proceso es poder escribir los valores de la serie  $Y_t$  en función de la serie  $X_t$  y de un ruido blanco, de la forma siguiente:

$$y_t = \frac{w(B)}{\delta(B)} x_{t-b} + \frac{\theta(B)}{\phi(B)} a_t \quad (7)$$

## 5. ANÁLISIS DE LA INFORMACIÓN

### 5.1 DESCRIPCIÓN DE LOS DATOS

Para este estudio se cuenta con las series de CONSULTAS y MÁXIMOS PM10 que se describen brevemente a continuación:

- Los datos de *CONSULTAS* están constituidos por 9.952 observaciones que corresponden al número de consultas diarias atendidas en el servicio médico de urgencias, urgencias en observación y hospitalización, por enfermedades del sistema respiratorio en el municipio de Bucaramanga durante el año 2009.

Los datos de salud se obtienen de la compilación de RIPS<sup>4</sup> que son enviados anualmente por las EAPB<sup>5</sup> al Observatorio de Salud Pública de Santander donde son consolidados.

- Los datos de *MÁXIMOS PM<sub>10</sub>* están conformados por los valores máximos diarios en microgramos por metro cúbico ( $\mu\text{g}/\text{m}^3$ ) del particulado PM<sub>10</sub> registrados en la Estación Automática Centro de monitoreo de calidad del aire de Bucaramanga durante el mismo período de tiempo.

Los datos ambientales se obtienen de los registros horarios diarios disponibles en la Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga- CDMB.

### 5.2 ANÁLISIS DE LAS SERIES DE TIEMPO

El proceso de análisis que se sigue a continuación es el propuesto por Box-Jenkins (1976) que ofrece una estrategia para la elaboración de

---

<sup>4</sup> Registro Individual de Prestación de Servicio

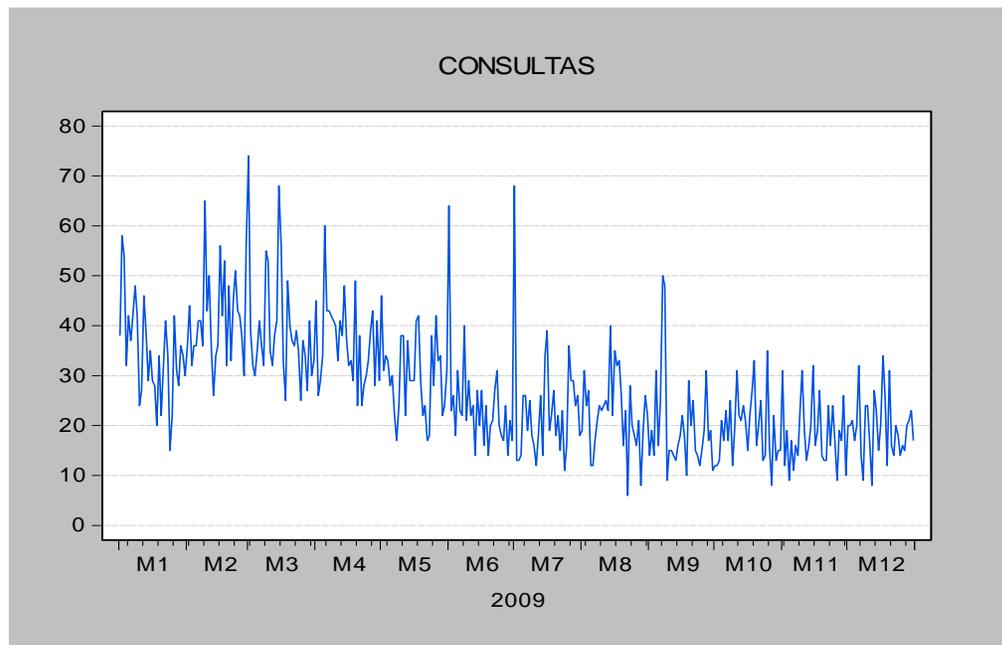
<sup>5</sup> Entidad Administrativa de Planes de Beneficiarios

modelos de series de tiempo. En esta metodología se tiene en cuenta la dependencia entre los datos y se busca sacar el mayor provecho de la información disponible para realizar posteriores pronósticos. Como se mencionó en el capítulo anterior, se destacan cuatro etapas a seguir que en su orden son: identificación, estimación, diagnóstico y pronóstico. A continuación se presenta el análisis de la serie *CONSULTAS* y la serie *MÁXIMOS PM10*.

### 5.2.1 Análisis de la serie *CONSULTAS*

El análisis de esta serie temporal univariante comienza con la representación gráfica de la variable a lo largo del tiempo y la determinación de las autocorrelaciones entre las observaciones.

*Figura 2. Secuencia del número de consultas en el servicio de urgencias en Bucaramanga por afecciones respiratorias durante el año 2009.*



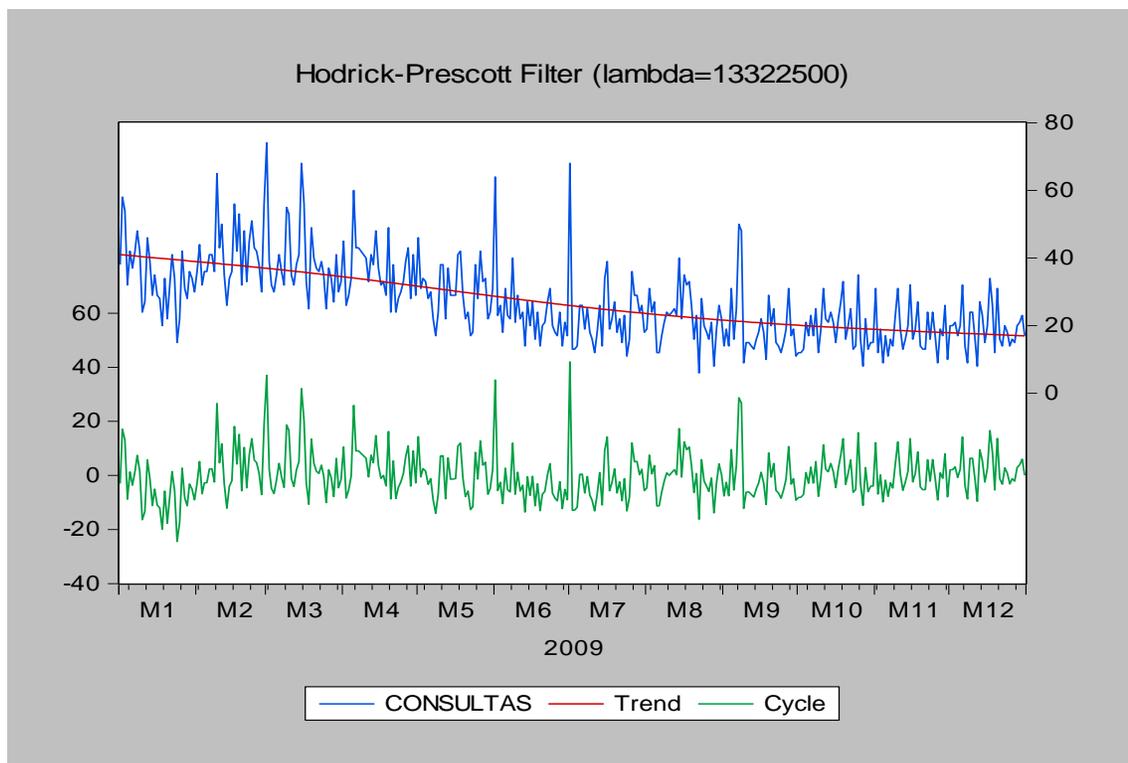
Fuente: Autor

La Figura 2 permite observar el comportamiento diario de las consultas en el servicio de urgencias de Bucaramanga en el año de estudio. Los

datos varían con el tiempo, por lo que la media y la varianza de estos datos están cambiando a medida que pasa el tiempo, es decir, las consultas diarias tienen una media y varianza distinta.

Al utilizar el procedimiento de FILTRO HOUDRICK-PRESCOTT (1997) que ofrece Eviews 7 se puede observar la tendencia de la serie.

*Figura 3. Tendencia Estimada para la serie CONSULTAS.*



Fuente: Autor

En la Figura 3 se advierte que la tendencia estimada de la serie es decreciente (o negativa) a lo largo del período de estudio.

Con el fin de determinar las propiedades de estacionariedad de las series se pueden utilizar distintos procedimientos, en esta ocasión se utilizará el test de Dickey Fuller (DF).

Esta prueba verifica si la serie de tiempo de las consultas de urgencias es o no estacionaria ( $H_0$ : No es estacionaria la serie). Los resultados de esta prueba se presentan en la Tabla 3.

**Tabla 3. Prueba de Dickey Fuller para la serie CONSULTAS.**

Null Hypothesis: CONSULTAS has a unit root		
Exogenous: Constant		
Lag Length: 6 (Automatic - based on SIC, maxlag=16)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.479203	0.1214
Test critical values:		
1% level	-3.448414	
5% level	-2.869396	
10% level	-2.571023	

\*Mackinnon (1996) one-sided p-values.

Fuente: Autor

El resultado de la probabilidad fue de 0,1214, por lo que al comparar este valor con el nivel de significancia de 0,05 se puede decir que la hipótesis nula no se rechaza, es decir, que la serie CONSULTAS no es estacionaria.

Una manera de corroborar si la serie de tiempo es estacionaria es observando el correlograma de la serie que se presenta en la Figura 4.

**Figura 4. Autocorrelaciones estimadas para la serie CONSULTAS.**

Sample: 1/01/2009 12/31/2009  
Included observations: 365

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.531	0.531	103.83	0.000
		2	0.447	0.229	177.41	0.000
		3	0.418	0.167	242.01	0.000
		4	0.492	0.268	331.98	0.000
		5	0.442	0.097	404.57	0.000
		6	0.466	0.162	485.62	0.000
		7	0.523	0.231	588.11	0.000
		8	0.449	0.020	663.59	0.000
		9	0.421	0.047	730.27	0.000
		10	0.385	-0.012	786.23	0.000
		11	0.387	-0.023	843.04	0.000
		12	0.390	0.030	900.84	0.000
		13	0.444	0.101	975.93	0.000
		14	0.517	0.197	1077.8	0.000
		15	0.468	0.084	1161.6	0.000
		16	0.343	-0.114	1206.9	0.000
		17	0.393	0.073	1266.4	0.000
		18	0.373	-0.051	1320.2	0.000
		19	0.437	0.092	1394.1	0.000
		20	0.443	0.089	1470.3	0.000
		21	0.500	0.100	1567.8	0.000
		22	0.427	0.011	1639.0	0.000
		23	0.379	-0.020	1695.3	0.000
		24	0.340	-0.078	1740.8	0.000
		25	0.374	0.024	1796.0	0.000
		26	0.428	0.069	1868.3	0.000
		27	0.411	-0.009	1935.2	0.000
		28	0.442	0.043	2012.7	0.000
		29	0.394	-0.021	2074.6	0.000
		30	0.403	0.082	2139.5	0.000
		31	0.363	0.024	2192.4	0.000
		32	0.328	-0.103	2235.7	0.000
		33	0.359	-0.002	2287.8	0.000
		34	0.423	0.026	2360.3	0.000
		35	0.452	0.068	2443.2	0.000
		36	0.381	-0.004	2502.5	0.000

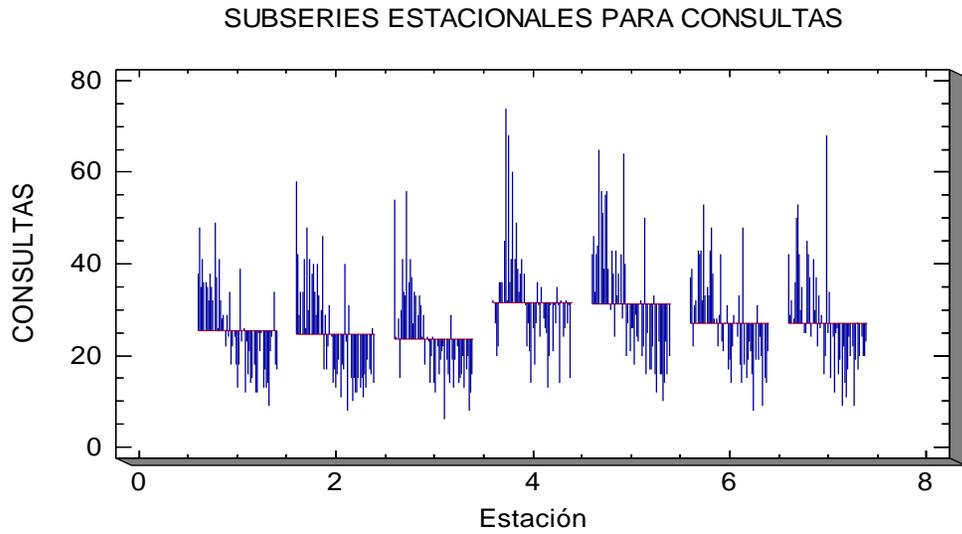
Fuente: Autor

Se puede observar que el autocorrelograma decrece muy lentamente, no hay decrecimiento exponencial, aspecto muy común en las series con tendencia y evidencia falta de estacionariedad.

Se percibe que la serie es estacional, hay un pico en 7, en 14 y 21, el número de consultas está correlacionado con las de hace 7 días.

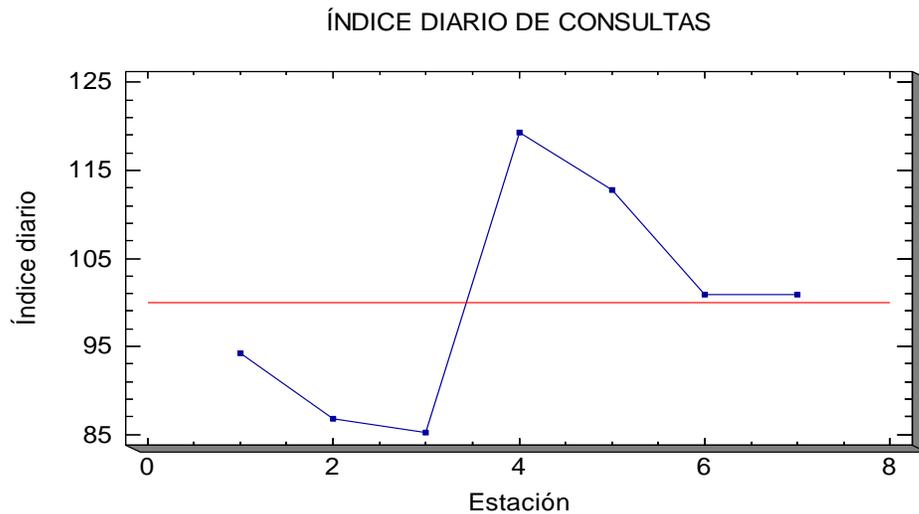
El gráfico de subseries estacionales que se exhibe en la Figura 5 permite observar el movimiento del número de consultas en cada día en el año. En todos los días el comportamiento es muy similar: mayor número de consultas en el primer semestre del año respecto al segundo semestre. El promedio de consultas es muy similar en los primeros días y en los dos últimos, en tanto que en los días intermedios (domingo y lunes) el promedio de consultas es superior a los demás.

**Figura 5. Gráfico de subseries estacionales para CONSULTAS.**



Otra gráfica que da cuenta del comportamiento de la serie y su variación se advierte en la Figura 6 donde se muestran los índices estacionales para cada día a escala, de modo que un día normal es igual a 100.

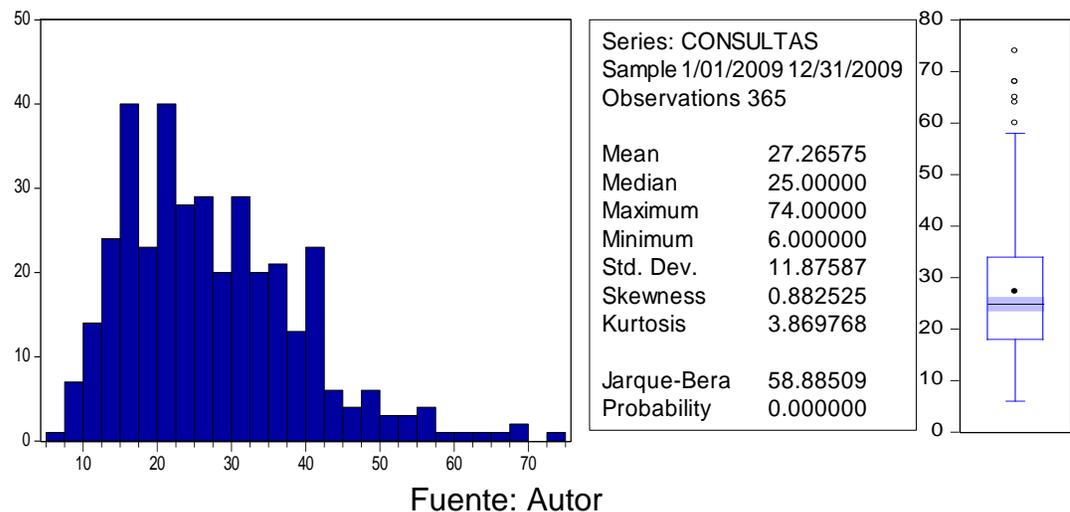
**Figura 6. Comparativo del comportamiento de las CONSULTAS en cada día.**



Se observa que en los días 1,2, 3 (jueves, viernes y sábado) el número de consultas es bajo, aumentan en los días 4 y 5 (domingo y lunes) y vuelve a decrecer en los días 6 y 7 (martes y miércoles) pero con valores superiores al valor de referencia. Se podría interpretar este comportamiento como el efecto fin de semana: desde el jueves disminuyen las consultas para luego aumentar el domingo de guayabo y el lunes de trabajo. El comportamiento aunque estacional no es regular porque no todos los días son iguales y hay una gran variación dentro de ellos.

El histograma de frecuencias y el diagrama de caja de las consultas diarias de urgencias se muestran en la Figura 7. Estas consultas presentan una distribución asimétrica a la derecha con mediana de 25, media de 27,26, un mínimo de 6 y un máximo muy alto de 74 consultas en el día.

**Figura 7. Estadísticos descriptivos para CONSULTAS.**

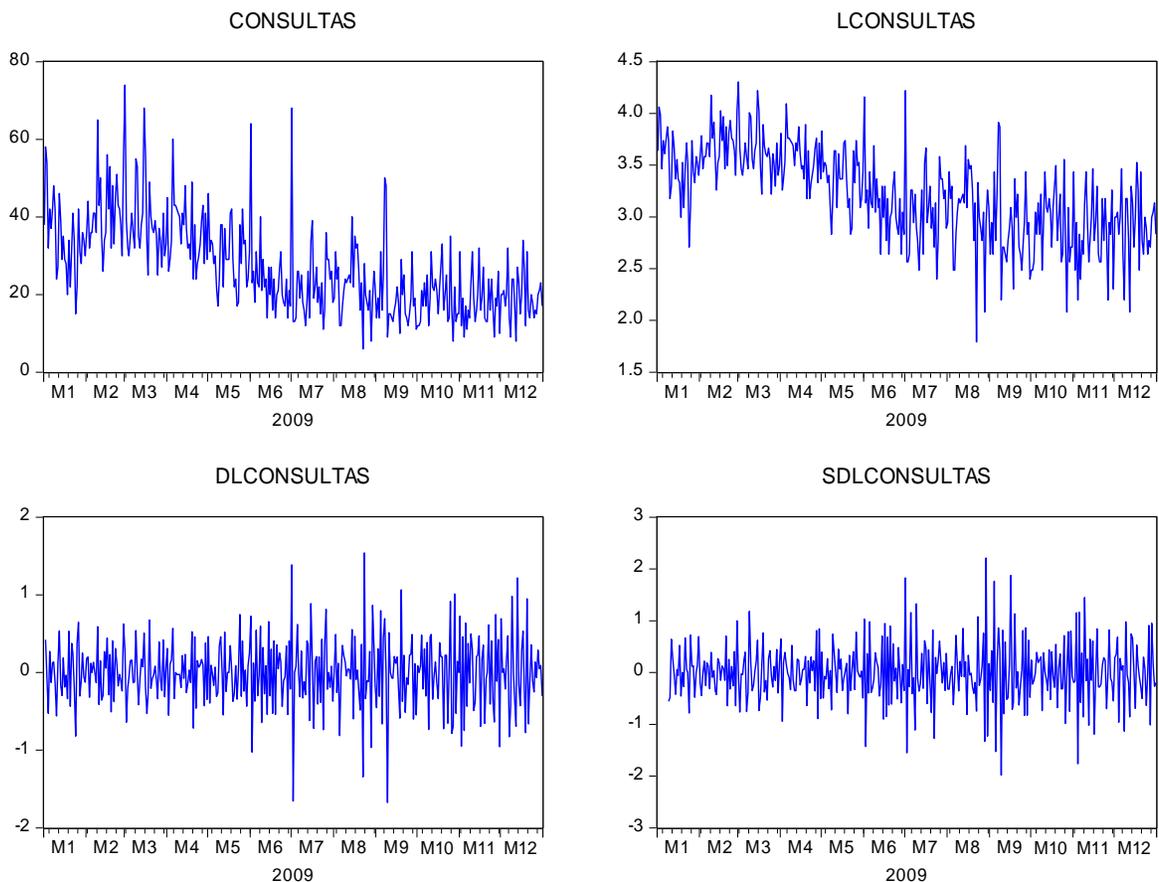


Es necesario transformar la serie para convertirla en estacionaria (media y varianza constante). Para tal fin, buscando estabilizar la varianza se hace una primera transformación de la serie con el logaritmo (*LCONSULTAS*). Para suprimir la tendencia se realiza la primera diferencia de la serie transformada (*DLCONSULTAS*) y para eliminar la

estacionalidad, la primera diferencia estacional a la serie diferenciada en su parte regular (*SDLCONSULTAS*).

En la Figura 8 se pueden observar los diferentes cambios de la serie al realizar estas transformaciones.

**Figura 8. Transformaciones para estabilizar la serie CONSULTAS.**



Fuente: Autor

Con la última transformación ya se observa estacionariedad en la serie a pesar de algunos comportamientos atípicos en ciertos períodos de tiempo que hacen que no sea completamente estacionaria, pero para verificar esta supuesta condición se hará de nuevo el test de Dickey Fuller (DF) a la serie *SDLCONSULTAS* (Tabla 4).

**Tabla 4. Prueba de Dickey Fuller para la serie SDLCONSULTAS.**

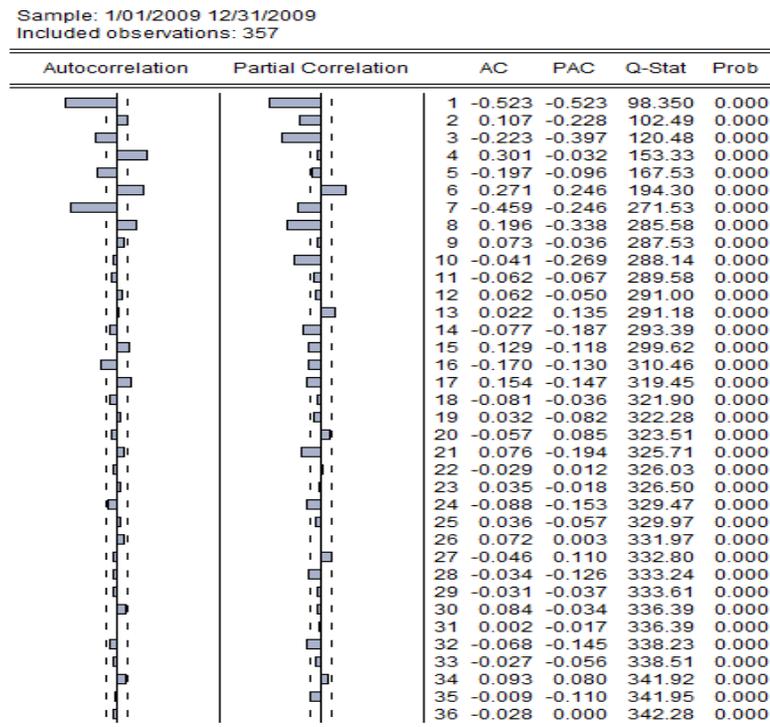
Null Hypothesis: SDLP has a unit root		
Exogenous: Constant		
Lag Length: 16 (Automatic - based on SIC, maxlag=16)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-9.409512	0.0000
Test critical values:		
1% level	-3.449389	
5% level	-2.869825	
10% level	-2.571253	

\*Mackinnon (1996) one-sided p-values.

Fuente: Autor

El valor\_p=0,000 al compararlo con el nivel de significancia de 0,05 hace que se rechace la hipótesis nula, es decir, que la serie transformada SDLCONSULTAS no tiene raíz unitaria y por lo tanto es estacionaria en tendencia y variación. La Figura 9 muestra el correlograma para la serie de consultas (SDLCONSULTAS).

**Figura 9. Autocorrelaciones estimadas para la serie SDLCONSULTAS.**



Fuente: Autor

Con este correlograma es el que nos vamos a quedar para definir los posibles modelos que ajusten a los datos. El mejor modelo debe tener todos los coeficientes significativos, los residuos no deben estar autocorrelacionados; además, exhibir el menor AIC y el mayor R cuadrado.

#### □ Definición del modelo para la serie SDLCONSULTAS

Siguiendo con la metodología de Box y Jenkins, para la definición del modelo se tiene en cuenta en el correlograma de la Figura 9; la columna de autocorrelación (AC) que se relaciona con la parte MA, los coeficientes que se tomarán para el modelo serán los rezagos que se salen o tocan los límites (en la estación). La parte AR se corresponde con la columna de correlación parcial (PAC) en los primeros rezagos y para escoger los coeficientes se seguirá el mismo paso de la parte MA.

Según estas observaciones los posibles modelos a analizar sin constante serían:

*Modelo 1: SDLCONSULTAS AR(1) SMA(7)*

*Modelo 2: SDLCONSULTAS AR(4) MA(1) SMA(7)*

*Modelo 3: SDLCONSULTAS AR(1) MA(1) SMA(7)*

Y entre ellos, el que ofrece el mejor ajuste, menor error y cumplimiento de los supuestos del modelo, es el modelo 2, por consiguiente sobre este modelo que en adelante es el *MODELO DE CONSULTAS*, es que se presentan las estimaciones y respectivos gráficos.

El modelo final según la salida de Eviews se escribe:

*Estimation Equation:*

=====

*SDLCONSULTAS = 0 +*

*[AR(4)=C(1),MA(1)=C(2),SMA(7)=C(3),BACKCAST=1/13/2009,ESTSMPL="1/13/2009 11/30/2009"]*

*Substituted Coefficients:*

=====

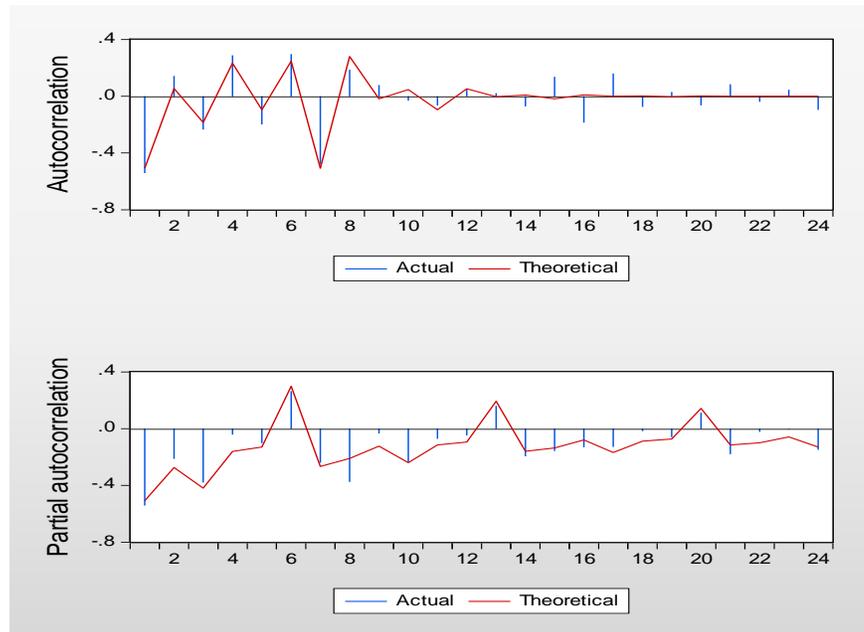
*SDLCONSULTAS = 0 + [AR(4)=0.188220774,MA(1)=-*

*0.934675886178,SMA(7)=-*

*0.934952291805,BACKCAST=1/13/2009,ESTSMPL="1/13/2009 11/30/2009"]*

La Figura 10 permite observar la estructura del modelo de consultas en su parte autorregresiva y de medias móviles, así como su acercamiento a los modelos teóricos.

**Figura 10. Estructura de la AC Y ACP para el modelo de CONSULTAS.**



Fuente: Autor

En la Tabla 5 se aprecian las estimaciones del modelo, lo primero que se verifica en la salida del programa, es en la parte final que todas las raíces invertidas sean menores que 1, como de hecho se puede apreciar.

Posteriormente la atención se centra en la parte superior, donde aparecen las variables que se tuvieron en cuenta, así como los coeficientes con su respectivo valor de significancia que para el modelo,

con un valor\_p=0,000, indican que son significativos y que la probabilidad de que sean cero es nula. El valor del criterio de Akaike (AIC) es de 0,482873, el de Schwarz es de 0,518040 y el  $R^2_{adj}=72,04\%$ .

**Tabla 5. Estimaciones para el modelo de CONSULTAS.**

Sample (adjusted): 1/13/2009 11/30/2009  
 Included observations: 322 after adjustments  
 Convergence achieved after 9 iterations  
 MA Backcast: 1/05/2009 1/12/2009

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(4)	0.188221	0.059524	3.162103	0.0017
MA(1)	-0.934676	0.022245	-42.01810	0.0000
SMA(7)	-0.934952	0.013129	-71.21225	0.0000
R-squared	0.722178	Mean dependent var		-0.001742
Adjusted R-squared	0.720436	S.D. dependent var		0.579914
S.E. of regression	0.306622	Akaike info criterion		0.482873
Sum squared resid	29.99150	Schwarz criterion		0.518040
Log likelihood	-74.74258	Hannan-Quinn criter.		0.496913
Durbin-Watson stat	1.911140			
Inverted AR Roots	.66			
Inverted MA Roots	.99	.93	.62+.77i	.62-.77i
	-.22+.97i	-.22-.97i	-.89-.43i	-.89+.43i

Fuente: Autor

Al mirar Durbin-Watson (1,911140), su valor cercano a dos indica que los residuales son independientes, completamente aleatorios. En la Figura 11 se puede ver el correlograma de los residuos y todos los rezagos están dentro de los límites establecidos.

**Figura 11. Correlograma de los residuos del modelo de CONSULTAS.**

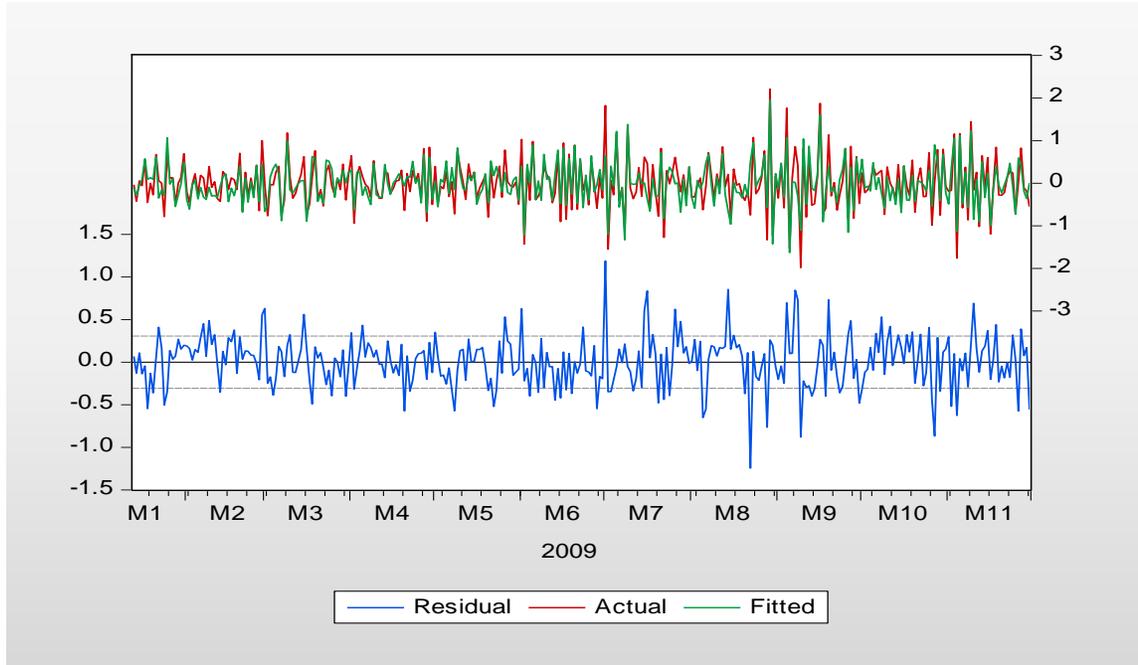
Sample: 1/13/2009 11/30/2009  
 Included observations: 322  
 Q-statistic probabilities adjusted for 3 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.038	0.038	0.4762	
		2 0.080	0.079	2.5833	
		3 -0.013	-0.019	2.6361	
		4 0.006	0.000	2.6466	0.104
		5 -0.074	-0.072	4.4534	0.108
		6 -0.016	-0.012	4.5392	0.209
		7 -0.037	-0.024	4.9882	0.299
		8 -0.044	-0.042	5.6400	0.343
		9 0.013	0.022	5.6980	0.458
		10 -0.049	-0.050	6.4872	0.484
		11 -0.037	-0.039	6.9427	0.543
		12 0.024	0.031	7.1351	0.623
		13 -0.020	-0.025	7.2695	0.700
		14 -0.020	-0.022	7.3999	0.766
		15 0.098	0.097	10.680	0.557
		16 -0.111	-0.128	14.916	0.313
		17 0.059	0.058	16.091	0.308
		18 -0.033	-0.027	16.456	0.352
		19 0.047	0.033	17.229	0.371
		20 -0.006	0.014	17.240	0.438
		21 0.037	0.005	17.702	0.475
		22 0.057	0.071	18.844	0.467
		23 -0.024	-0.037	19.044	0.519
		24 -0.038	-0.057	19.555	0.550
		25 -0.039	-0.008	20.077	0.578
		26 0.065	0.073	21.571	0.546
		27 -0.015	-0.017	21.649	0.600
		28 -0.019	-0.015	21.771	0.649
		29 -0.038	-0.042	22.272	0.674
		30 0.018	0.018	22.385	0.718
		31 -0.018	0.010	22.506	0.757
		32 -0.063	-0.094	23.911	0.733
		33 0.007	0.046	23.928	0.775
		34 0.049	0.030	24.793	0.777
		35 -0.009	-0.011	24.821	0.813
		36 -0.007	-0.016	24.837	0.846

Fuente: Autor

Se puede ver en la Figura 12 la estimación del modelo (en color verde) y la serie original (en color rojo).

*Figura 12. Gráfico de los valores actuales, los pronósticos y los residuos para el modelo de consultas.*



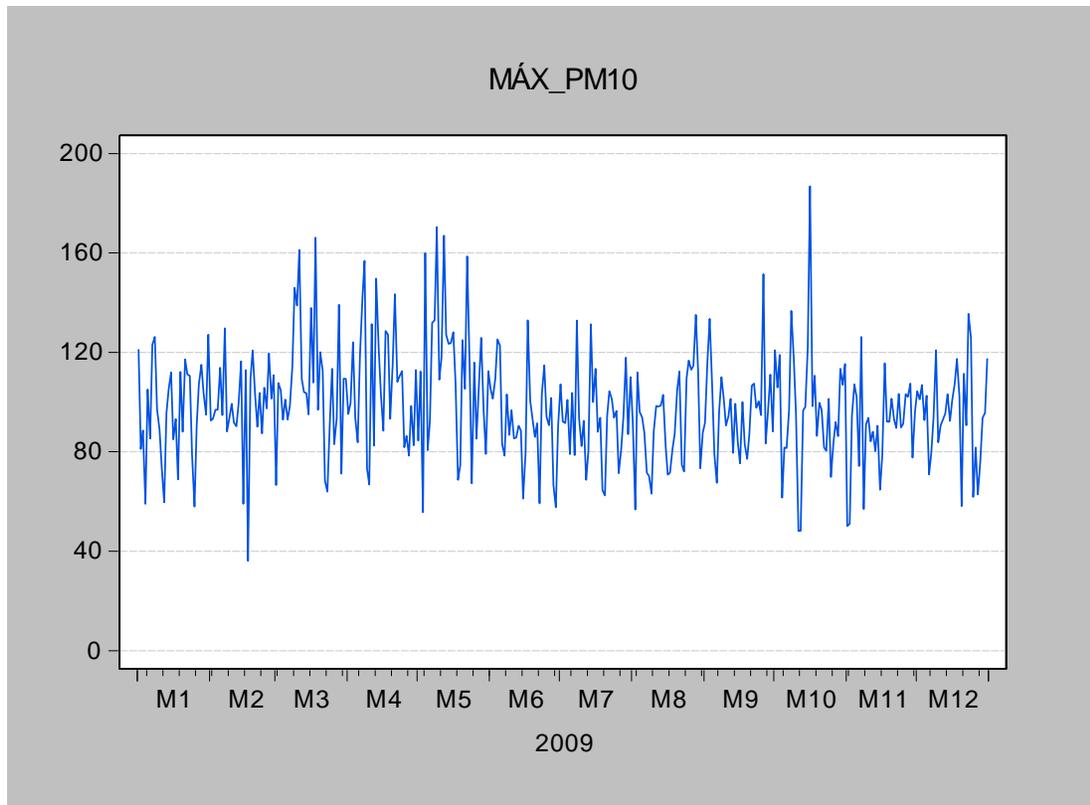
Fuente: Autor

En la gráfica se ve que el modelo presenta un relativo buen ajuste con los datos, aunque se distinguen algunas discrepancias, así como un comportamiento más regular en el primer semestre del año. En la parte inferior del gráfico está la serie de los residuos (en color azul) y se distinguen unos valores muy altos y muy bajos que pueden generar ruido y sesgar los resultados.

### 5.2.2 Análisis de la serie *MÁXIMOS\_PM10*

El gráfico de la serie cronológica de la variable Máximos\_PM10 se presenta en la Figura 13.

**Figura 13. Secuencia de los valores máximos de las partículas de PM10 en al aire de Bucaramanga durante el año 2009.**

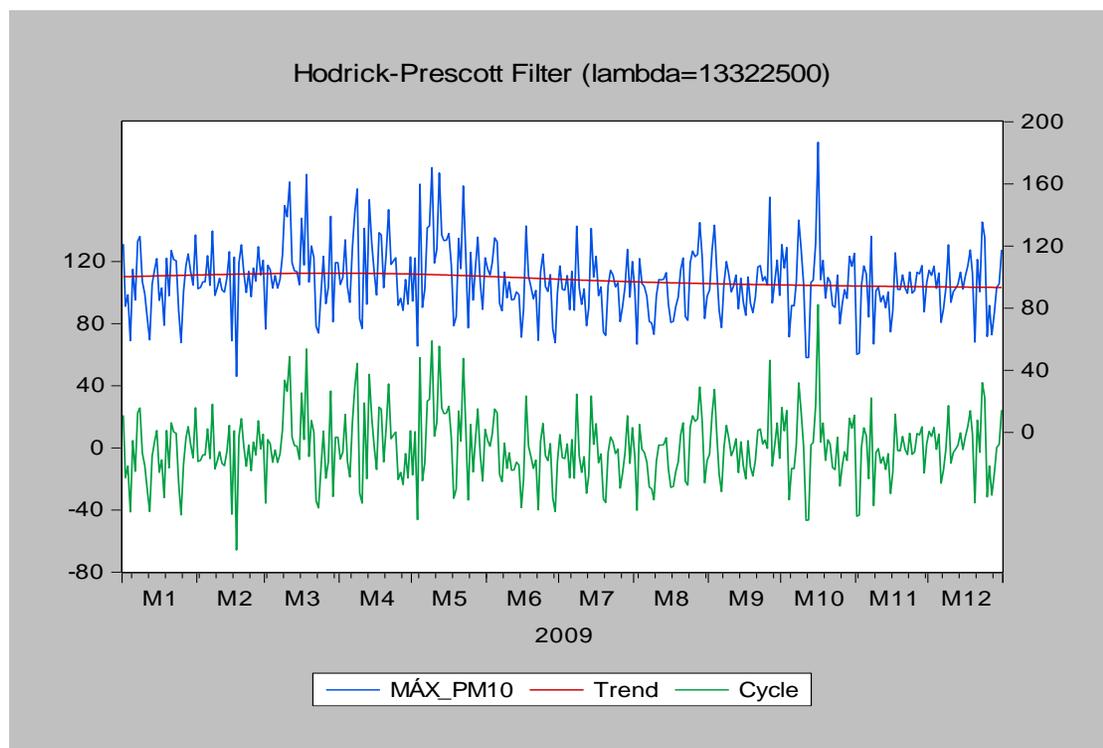


Fuente: Autor

En esta gráfica se observa que la serie pareciera ser estacionaria en media y varianza aunque se identifican unos puntos muy altos o muy bajos en algunos momentos de la serie.

Para obtener una aproximación del componente tendencial se aplica el procedimiento de filtro de Houdrick-Prescott a la serie. La curva de tendencia se muestra en la Figura 14.

**Figura 14. Tendencia estimada para la serie MÁXIMOS\_PM10.**



Fuente: Autor

Se puede considerar que la serie del particulado PM10 es estacionaria en tendencia, apreciación que es confirmada por los resultados de la prueba de Dickey Fuller (DF) que se muestran en la Tabla 6.

**Tabla 6. Prueba de Dickey Fuller para la serie MÁXIMOS\_PM10.**

Null Hypothesis: MAX\_PM10 has a unit root  
 Exogenous: Constant  
 Lag Length: 0 (Automatic - based on SIC, maxlag=16)

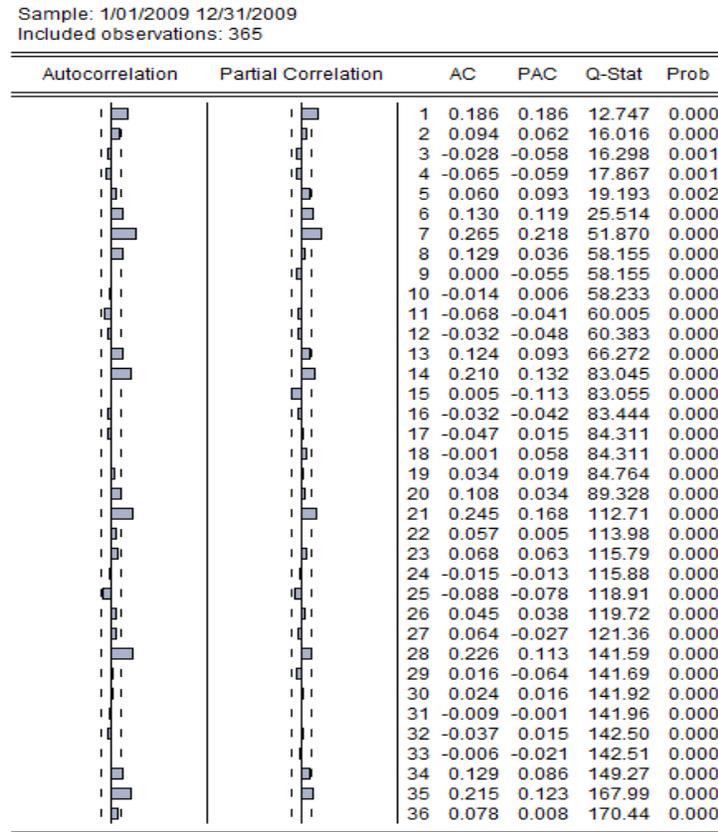
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-15.76153	0.0000
Test critical values:		
1% level	-3.448111	
5% level	-2.869263	
10% level	-2.570952	

\*MacKinnon (1996) one-sided p-values.

Fuente: Autor

El valor\_p de la prueba es 0,000 por lo tanto se rechaza la hipótesis nula y se concluye que la serie de MÁXIMOS\_PM10 es estacionaria en tendencia y por consiguiente no necesita diferenciarse.

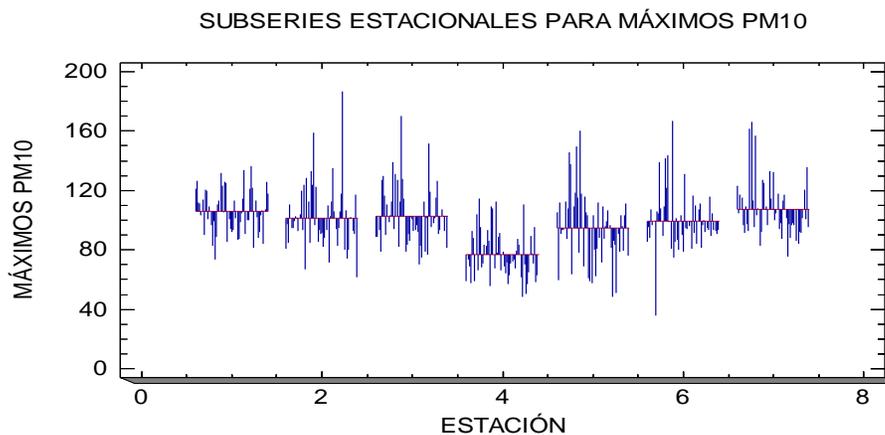
**Figura 15. Autocorrelaciones estimadas para la serie MÁXIMOS\_PM10.**



Fuente: Autor

En la Figura 15 se muestran las funciones FAC y FACP para la serie del particulado. Se corrobora que no hay tendencia, pero se descubre que la serie es estacional: hay un pico en 7, 14, 21, se puede decir que el máximo de PM<sub>10</sub> en un día está correlacionado con el de hace 7 días.

**Figura 16. Gráfico de subseries estacionales para MÁXIMOS  $PM_{10}$ .**

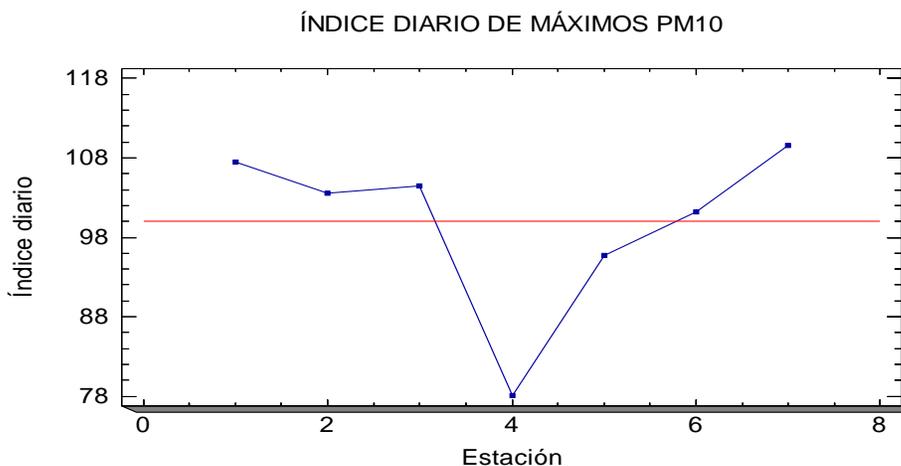


Fuente: Autor

En la Figura 16, las líneas horizontales (en rojo) son el promedio de cada día en todo el período considerado y las verticales (azules) cuánto se desvía del promedio en cada día. En los días 1, 2, 3 (jueves, viernes y sábado) los valores medios son muy parecidos, el día 4 (domingo) se advierten los valores más bajos y vuelve a crecer en los días 5, 6, 7 (lunes, martes y miércoles). Existe variabilidad, relativamente constante, en todo el período considerado.

En la Figura 17 se observa otra gráfica que muestra los valores de cada una de las observaciones para el período de 7 días.

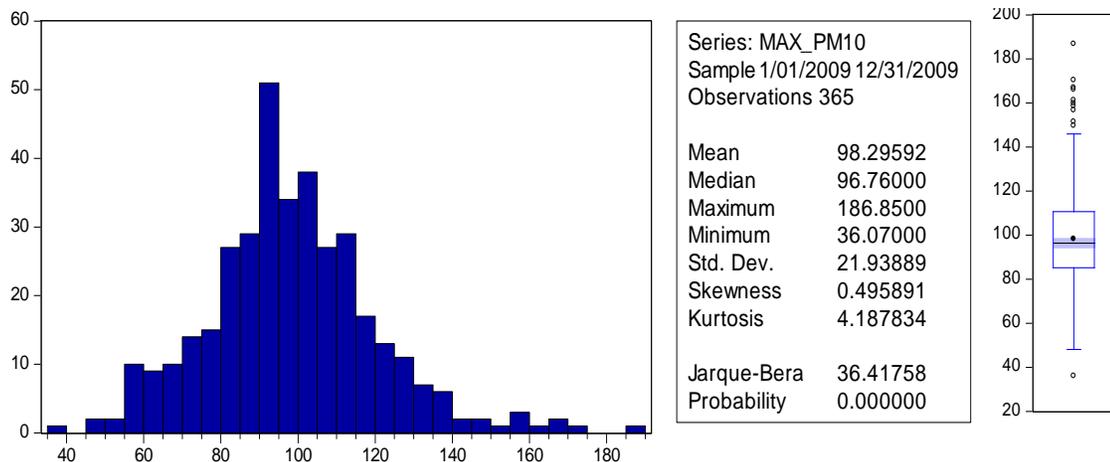
**Figura 17. Comparativo del comportamiento de los MÁXIMOS DE  $PM_{10}$  en cada día.**



Fuente: Autor

La línea horizontal (en rojo) muestra el comportamiento medio en el período de referencia, lo que permite realizar comparaciones entre los diferentes días, es evidente que en los días 4 y 5 (sábado y domingo) los índices son inferiores al promedio (78,06 y 95,71 respectivamente) en tanto que en los demás días el comportamiento es bastante similar y por encima de la media de los valores de PM<sub>10</sub>.

**Figura 18. Estadísticos descriptivos para MÁXIMOS\_PM10.**

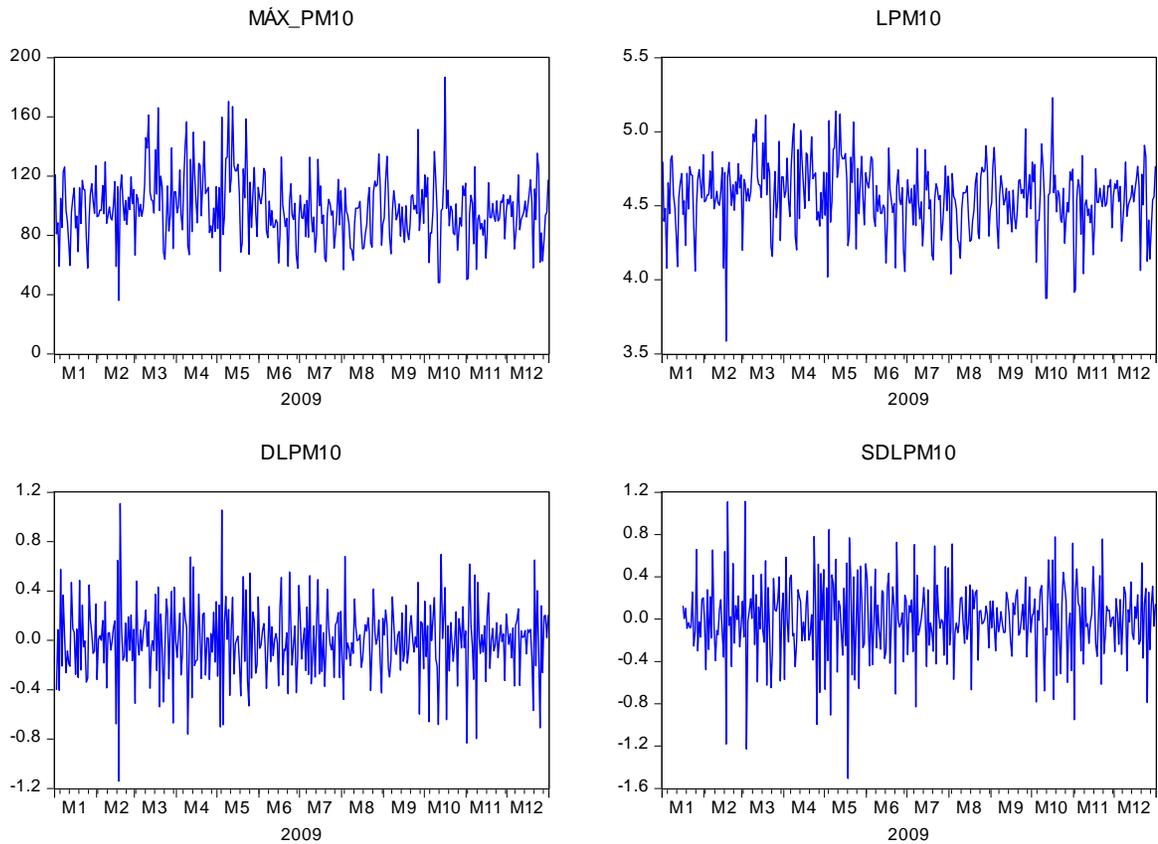


Fuente: Autor

El histograma de frecuencias y el diagrama de caja de los valores máximos en el aire del particulado PM10 se muestran en la Figura 18. Estos valores presentan una distribución bastante simétrica (simetría=0,496), con una media de 98,29, mediana igual a 96,76, un valor mínimo de 36,07 y un máximo de 186,85.

Con el fin de estabilizar la serie se presentan en la Figura 20 los gráficos de la serie con las posibles transformaciones. De izquierda a derecha y de arriba abajo, el primer gráfico corresponde a la serie original, el segundo al logaritmo de la serie, el siguiente con la primera diferencia estacional al logaritmo de la serie y el último con la segunda diferencia estacional al logaritmo de la serie.

**Figura 19. Transformaciones para estabilizar la serie MÁXIMOS PM10.**



Fuente: Autor

La última transformación ofrece la mejor opción y para corroborar la apreciación visual de estacionariedad, se realiza la prueba de Dickey Fuller a esta serie, los resultados obtenidos se muestran en la Tabla 7.

**Tabla 7. Prueba de Dickey Fuller para la serie SDLMÁX\_PM10.**

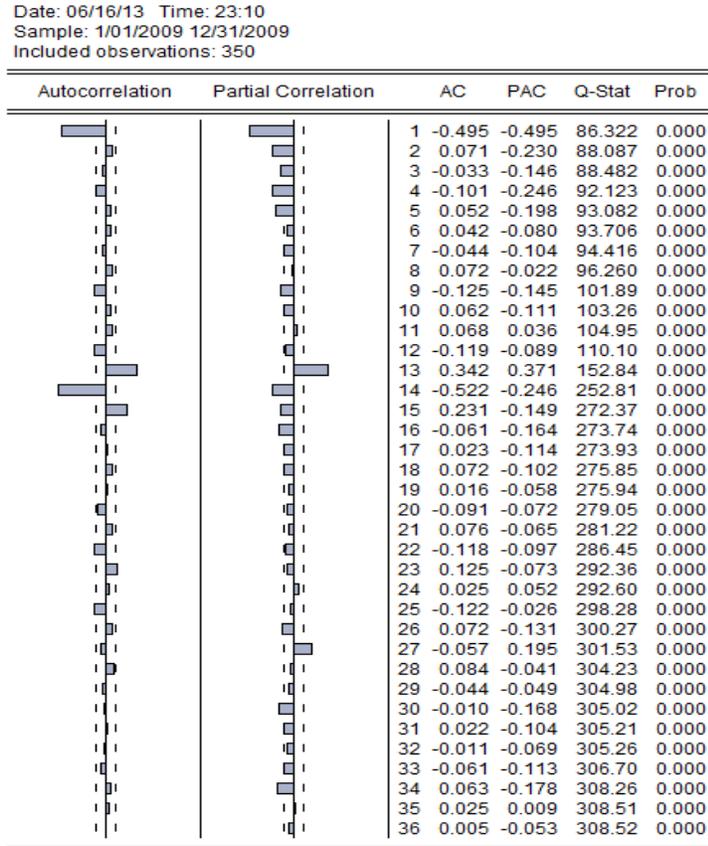
Null Hypothesis: SDLPM10 has a unit root		
Exogenous: Constant		
Lag Length: 15 (Automatic - based on SIC, maxlag=16)		
	t-Statistic	Prob.*
<b>Augmented Dickey-Fuller test statistic</b>	<b>-7.477342</b>	<b>0.0000</b>
Test critical values:		
1% level	-3.449738	
5% level	-2.869978	
10% level	-2.571335	

\*MacKinnon (1996) one-sided p-values.

Fuente: Autor

Los resultados de la prueba concluyen que la serie SDLPM\_10 es estacionaria. En la Figura 20 se muestran las respectivas funciones FAC y FACP para esta serie.

**Figura 20. Autocorrelaciones estimadas para la serie SDLMÁX\_PM10.**



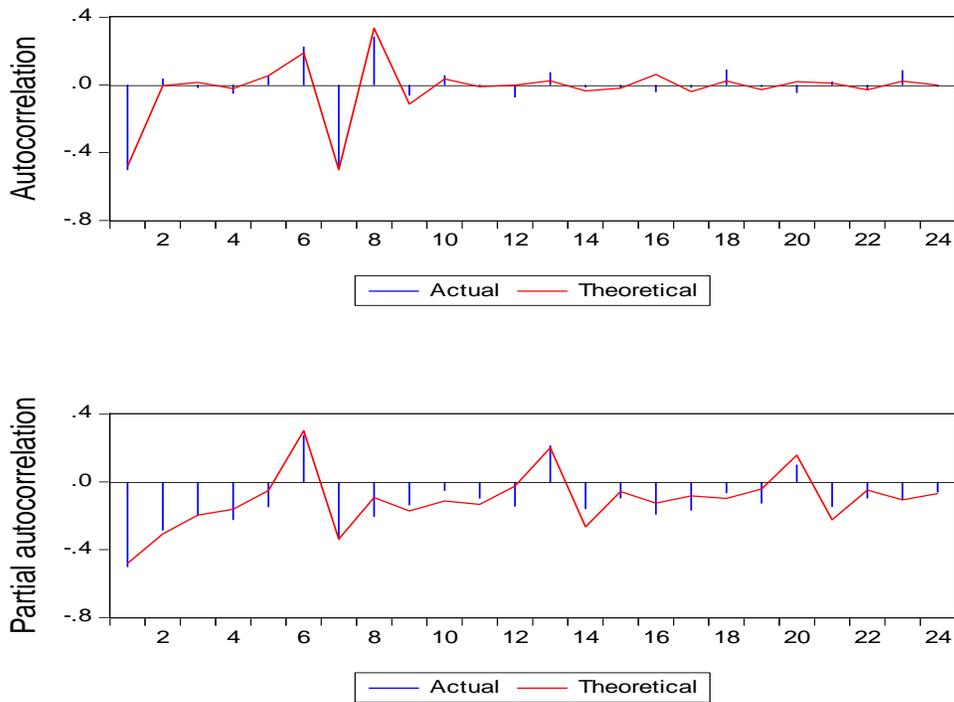
Fuente: Autor

Con este correlograma se definen los posibles modelos que ajusten a los datos. Después de analizar varios modelos, el que mejor ajuste ofrece, con el menor error y cumplimiento de los supuestos es el modelo 3 que se expresa a continuación como:

*Modelo 3: SDLMÁX\_PM10 AR(1) AR(7) AR(14) AR(21) MA(1) MA(2) MA(7)*

Seguidamente, en la Figura 21, se exhibe la estructura gráfica de la serie que ofrece este modelo, indicando en el correlograma con líneas continuas el modelo teórico que subyace a las funciones de autocorrelación y autocorrelación parcial.

**Figura 21. Estructura de la AC y ACP para el modelo de MÁX\_PM10.**



Fuente: Autor

Las estimaciones del modelo se pueden apreciar en la Tabla 8, se cumple que todas las raíces invertidas son menores que uno, además, todos los coeficientes con significativos en el modelo (valor\_p=0,000).

El valor del criterio de Akaike (AIC) es de -0,128210, el de Schwarz es de -0,048687 y el  $R^2=64,35\%$ . Al mirar Durbin-Watson (2,14), su valor cercano a dos indica que los residuales deben ser independientes, completamente aleatorios, pero para verificarlo se presentan a continuación las pruebas a los residuos.

**Tabla 8. Estimaciones para el modelo de los MÁX\_PM10.**

Sample (adjusted): 1/30/2009 12/31/2009  
 Included observations: 336 after adjustments  
 Convergence achieved after 27 iterations  
 MA Backcast: 1/23/2009 1/29/2009

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.278416	0.069987	-3.978113	0.0001
AR(7)	-0.586393	0.071015	-8.257324	0.0000
AR(14)	-0.434231	0.064623	-6.719464	0.0000
AR(21)	-0.222887	0.053730	-4.148256	0.0000
MA(1)	-0.441701	0.087006	-5.076690	0.0000
MA(2)	-0.315644	0.061534	-5.129613	0.0000
MA(7)	-0.242629	0.052711	-4.603041	0.0000
R-squared	0.649961	Mean dependent var		1.42E-05
Adjusted R-squared	0.643578	S.D. dependent var		0.376239
S.E. of regression	0.224619	Akaike info criterion		-0.128210
Sum squared resid	16.59925	Schwarz criterion		-0.048687
Log likelihood	28.53926	Hannan-Quinn criter.		-0.096510
Durbin-Watson stat	2.140033			
Inverted AR Roots	.90+.20i .73+.59i .19+.89i -.42-.83i -.75+.60i -.94	.90-.20i .73-.59i .19-.89i -.42+.83i -.75-.60i	.81-.40i .39+.84i -.01+.95i -.59-.71i -.93+.19i	.81+.40i .39-.84i -.01-.95i -.59+.71i -.93-.19i
Inverted MA Roots	1.00 -.15-.75i	.59-.56i -.72-.33i	.59+.56i -.72+.33i	-.15+.75i

Fuente: Autor

En la Figura 22 se muestra el correlograma de los residuos del modelo. Se puede ver que los rezagos tanto en la autocorrelación, como en la autocorrelación parcial se mantienen dentro de los límites.

**Figura 22. Correlograma de los residuos del modelo de MÁX\_PM10.**

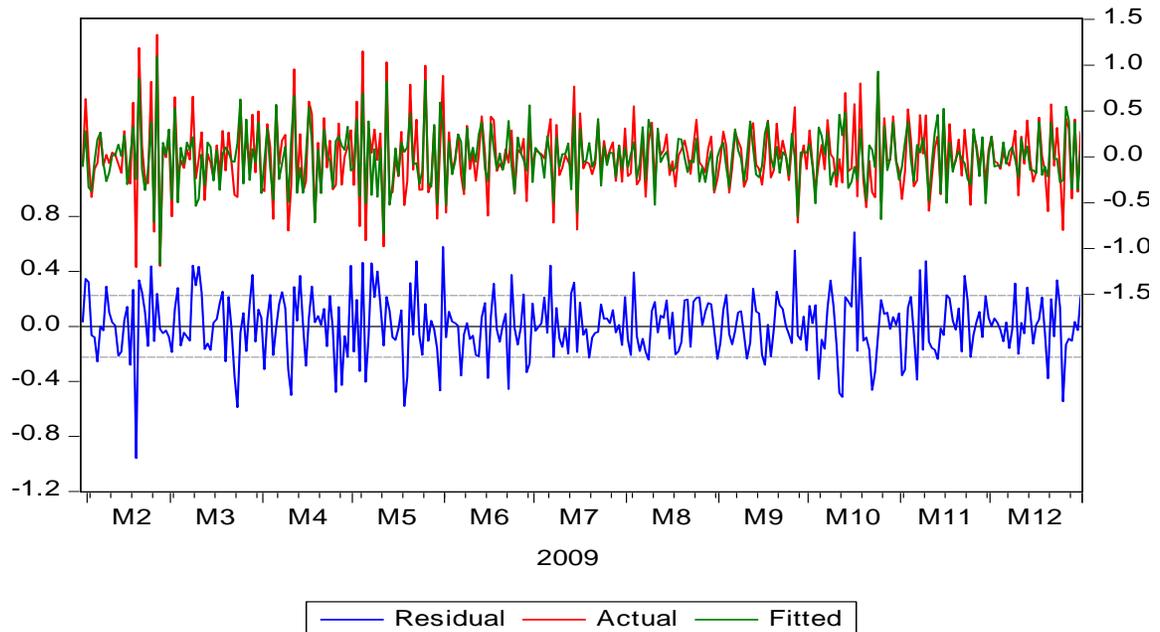
Sample: 1/30/2009 12/31/2009  
 Included observations: 336  
 Q-statistic probabilities adjusted for 7 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.072	-0.072	1.7682	
		2 -0.021	-0.026	1.9175	
		3 -0.104	-0.108	5.6156	
		4 -0.115	-0.134	10.146	
		5 -0.072	-0.102	11.912	
		6 0.010	-0.027	11.949	
		7 -0.041	-0.081	12.532	
		8 -0.060	-0.116	13.793	0.000
		9 0.064	0.015	15.215	0.000
		10 0.047	0.023	15.984	0.001
		11 0.023	-0.005	16.167	0.003
		12 -0.077	-0.101	18.241	0.003
		13 0.085	0.078	20.777	0.002
		14 -0.074	-0.052	22.733	0.002
		15 -0.086	-0.122	25.334	0.001
		16 -0.078	-0.119	27.499	0.001
		17 0.048	0.024	28.324	0.002
		18 0.082	0.059	30.699	0.001
		19 -0.005	-0.075	30.707	0.002
		20 0.044	-0.003	31.410	0.003
		21 -0.071	-0.047	33.250	0.003
		22 0.010	-0.008	33.285	0.004
		23 0.074	0.049	35.268	0.004
		24 0.048	0.049	36.095	0.004
		25 -0.021	0.034	36.262	0.007
		26 -0.016	-0.010	36.356	0.010
		27 0.028	0.047	36.646	0.013
		28 -0.146	-0.136	44.516	0.002
		29 -0.018	-0.045	44.643	0.003
		30 -0.029	-0.051	44.956	0.004
		31 0.032	-0.014	45.345	0.005
		32 0.001	-0.031	45.345	0.008
		33 -0.045	-0.118	46.109	0.009
		34 0.105	0.106	50.260	0.004
		35 0.025	0.023	50.502	0.006
		36 -0.020	-0.078	50.657	0.008

Fuente: Autor

A continuación en la Figura 23 se exhibe el gráfico de la serie con los valores actuales, los pronosticados por el modelo y en la parte inferior, la serie de los residuos del modelo.

*Figura 23. Gráfico de los valores actuales, los pronósticos y los residuos del modelo de MÁX\_PM10.*



Fuente: Autor

Se puede ver que el modelo ajusta relativamente bien a los datos, aunque se distinguen algunas discrepancias. La serie de los residuos (en la parte inferior) permite ver regularidad en la mayor del período de estudio aunque sobresalen algunos valores muy grandes.

El modelo para la serie de MÁX\_PM10 en Eviews se define:

```
SDLMÁX_PM10 = [AR(1)=-0.278416295686,AR(7) =-
0.586393174199,AR(14)=-0.434231141885,AR(21) =-
0.222887234236,MA(1)=-0.441700587727,MA(2) =-
0.315643975135,MA(7) =-0.242628641287,BACKCAST=
1/30/2009, ESTSMPL="1/30/2009 12/31/2009"]
```

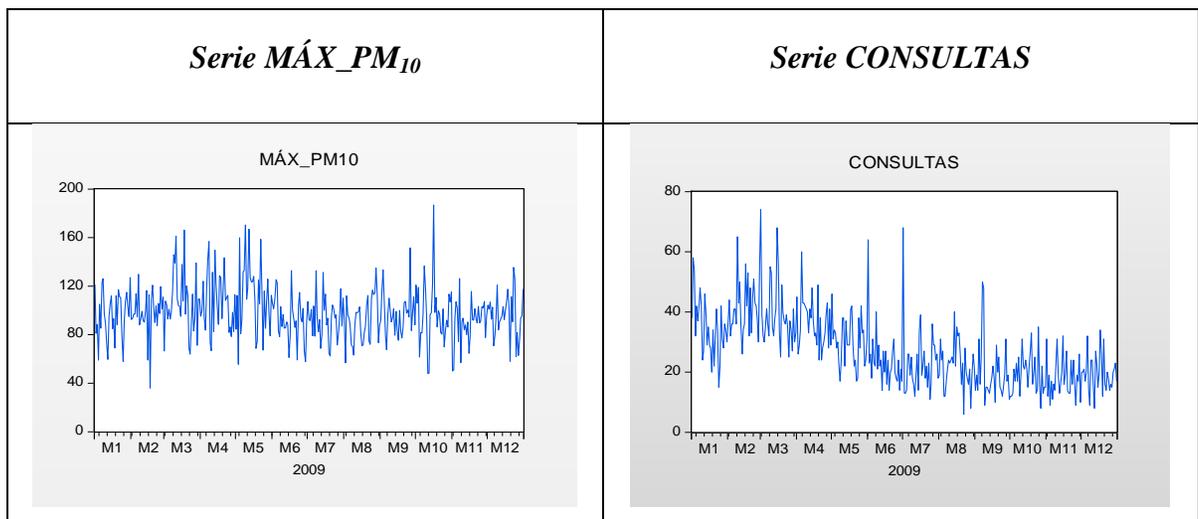
### 5.3 MODELO DE LA FUNCIÓN DE TRANSFERENCIA

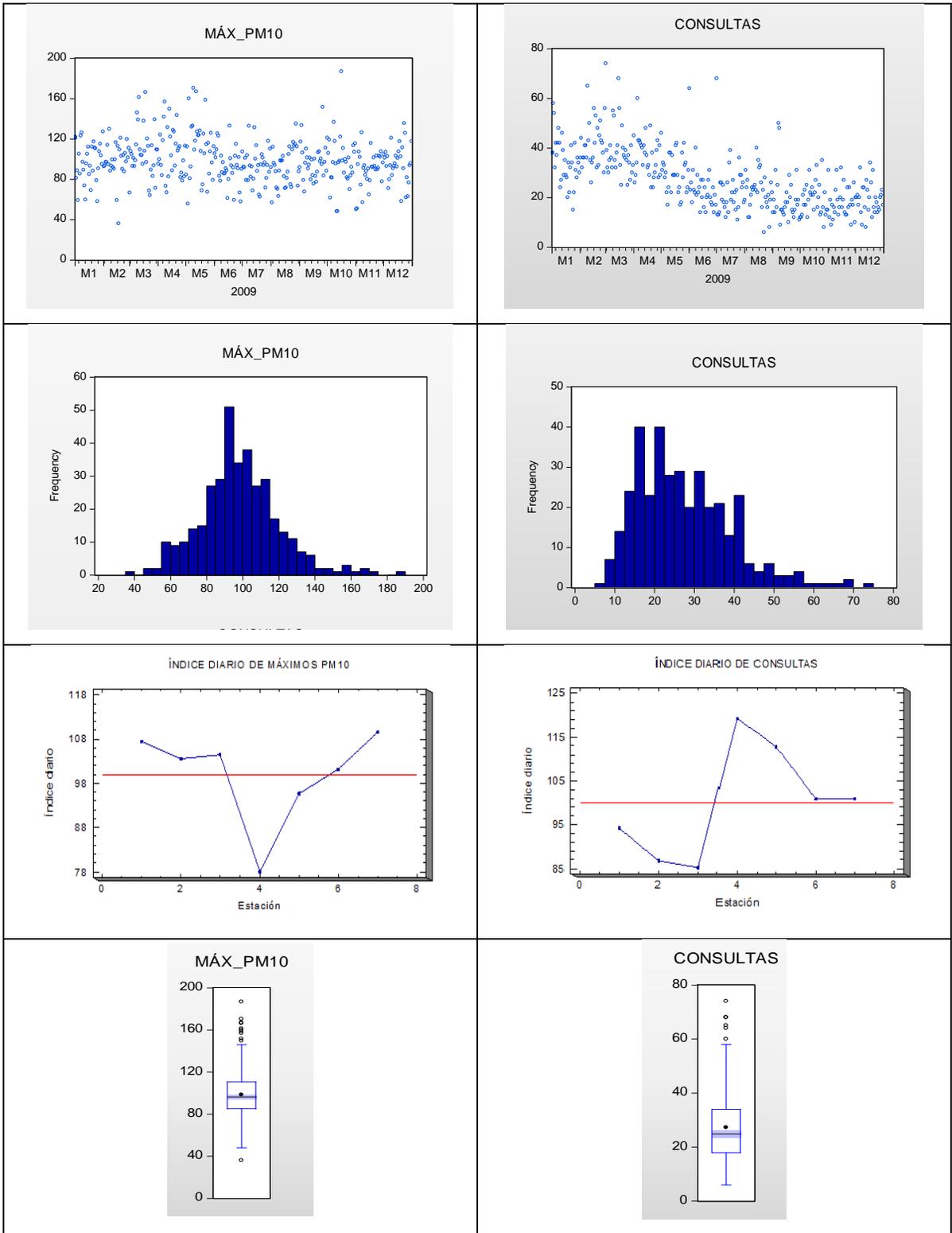
Analizadas por aparte las series del número de consultas en los servicios de urgencias en Bucaramanga y los niveles máximos de  $PM_{10}$  en el aire en el 2009, se aborda la pregunta que motivó esta investigación: ¿Se relacionan los servicios de urgencias con los niveles de  $PM_{10}$ ? O, en palabras más dicientes, ¿los niveles de  $PM_{10}$  podrían contarse entre las causas de las consultas de urgencia, o, mejor dicho, de su crecimiento en ciertos momentos?

Los valores máximos de  $PM_{10}$  pueden ser un indicador del número de consultas en el servicio de urgencias por lo que se pretende construir un modelo de función de transferencia que explique el número de consultas ( $Y_t$ ) en función de las lecturas máximas de  $PM_{10}(X_t)$ , considerando que existe una relación causal entre ellos.

En la Figura 25 se resumen los gráficos temporales de las series, los histogramas de frecuencia, los diagramas de punto y los diagramas de cajas para las dos series en estudio, aspectos que ya fueron previamente analizados en este trabajo.

**Figura 24. Gráficos comparativos entre la serie de MÁX\_ $PM_{10}$  y CONSULTAS.**





Fuente: Autor

Se puede ver, en una primera comparación a priori de las series que hay algunas diferencias entre las dos: mientras la serie de *MÁXIMOS PM<sub>10</sub>* tiene un comportamiento relativamente estable, la serie de *CONSULTAS* tiene una marcada tendencia negativa; además, los índices estacionales de ambas series tienen comportamientos contrarios.

Entre las similitudes, se encuentra el hecho de que ambas son estacionales con estación 7 días y valores atípicos muy grandes en ciertos períodos del año.

Como se entiende que el efecto del aumento de las partículas PM<sub>10</sub> en términos del número de consultas por enfermedades respiratorias se debe producir un tiempo después de observar el valor de PM<sub>10</sub>, el desfase en los gráficos de los índices (últimos gráficos en la Figura 24) puede ser indicativo de este hecho. Es así que los valores relativamente altos de los días miércoles, jueves y viernes se podrían explicar los altos valores de las consultas del domingo y el lunes. A su vez, los bajos valores de PM<sub>10</sub> el día jueves se podrían reflejar en el descenso en el número de consultas de los días martes y miércoles.

Para construir el modelo, se aborda el enfoque cuantitativo, es decir, se seguirá con el proceso de ajuste de la función de transferencia que en resumen es el siguiente:

- a. Estimar la regresión para un valor alto k

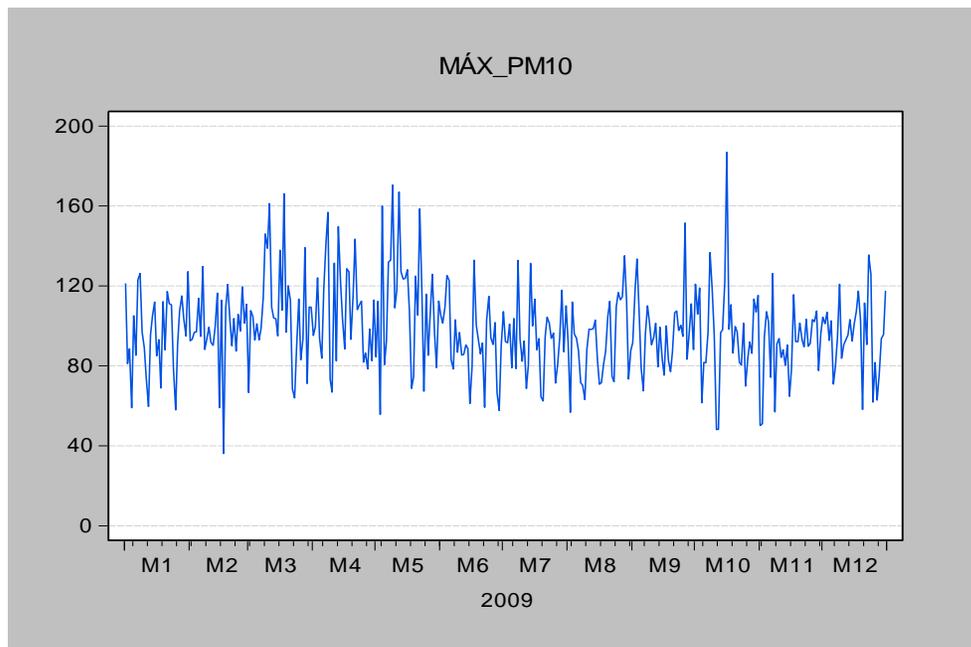
$$y_t = v_0x_t + v_1x_{t-1} + \dots + v_kx_{t-k} + n_t$$

- b. Comprobar si el ruido es estacionario o no
  - Si es estacionario identificar un modelo para el ruido y
  - Estimar los parámetros conjuntamente.
- c. Si no lo es, se toman diferencias en la relación (a) se estima el modelo de regresión y se vuelve a comprobar si el ruido es estacionario.

- d. Se repite el proceso hasta que el ruido sea estacionario
- e. Se ajusta el modelo lineal al ruido y se estima conjuntamente la función de transferencia y los parámetros del modelo de ruido.

Para este procedimiento se seguirá paso a paso el procedimiento en el programa Eviews 7, por facilidad en la interpretación, para generar el modelo se renombra la serie MÁXIMOS\_PM10 por *particulado*.

Se trae de regreso la Figura 13 que muestra la serie temporal de los MÁXIMOS\_PM<sub>10</sub>



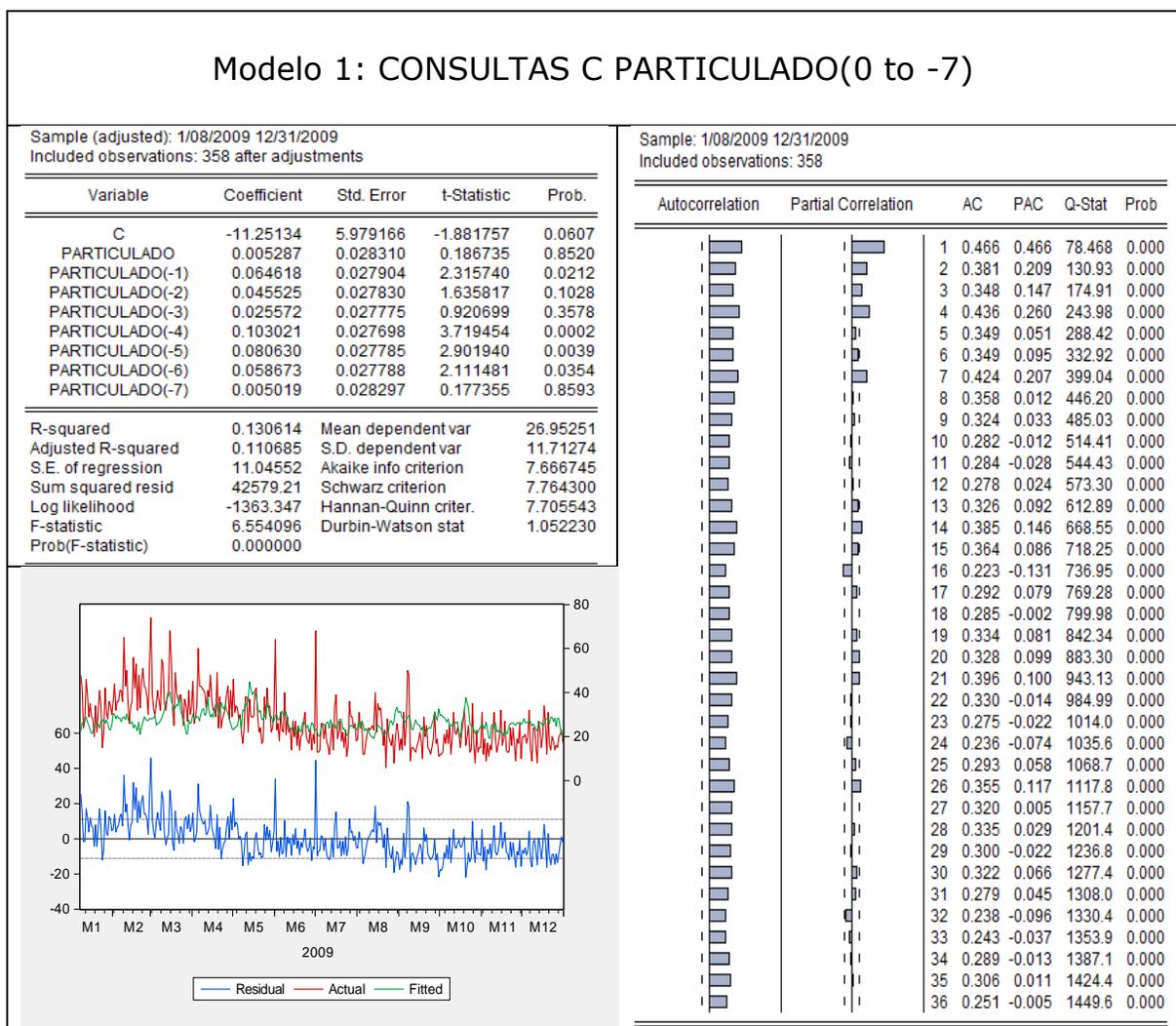
Fuente: Autor

Como se mencionó en el análisis anterior, la serie de *particulado* es estacionaria en tendencia y tiene cambios en la estación lo que implica que para volverla una serie estacionaria en la estación se debe tomar como mínimo una diferencia regular. Esto significa que en la estructura de los residuos se debe incluir algún término estacional. Iniciamos con que asume 7 retardos para la variable particulados.

$$CONSULTAS = C + PARTICULADO(0-7)$$

Para facilidad en la interpretación de los resultados que se van obteniendo, se organizan las salidas del programa correspondientes a las estimaciones del modelo, el gráfico de los residuos, y el correlograma de los mismos, en única tabla. En la parte superior de la tabla se escribe el modelo como se introduce en el programa.

**Figura 25. Estimaciones y residuos del modelo 1**



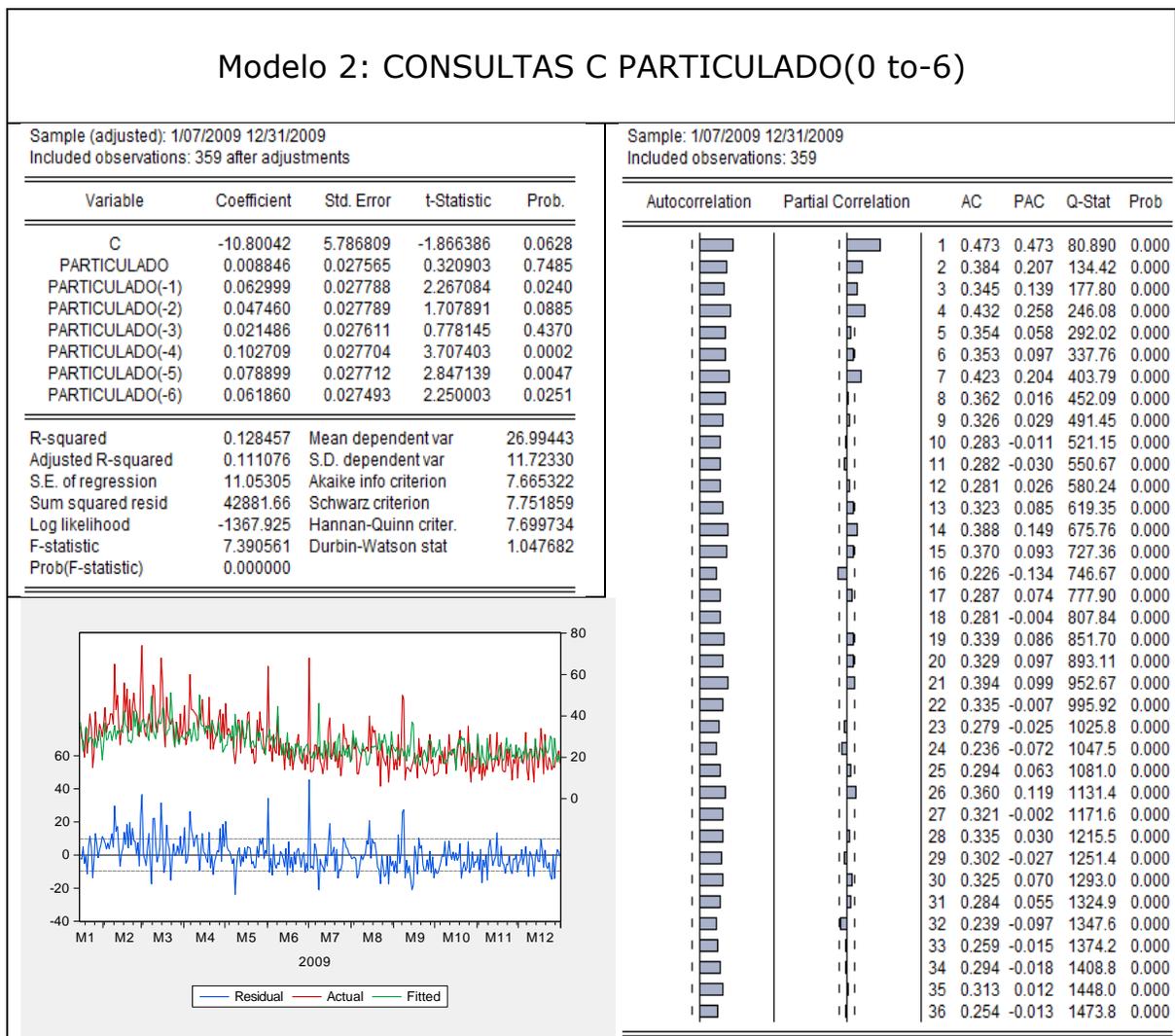
Fuente: Autor

Las estimaciones del modelo dejan ver los coeficientes en: (-1), (-4), (-5) y (-6) que son significativos, por el contrario, la constante y los demás términos no lo son. Se aprecia que el modelo no tiene un buen

ajuste y los residuos no parecieran ser estacionarios, muestran la estacionalidad de consultas, valoración que se corrobora con el correlograma.

Para el segundo modelo se retira el coeficiente en el retardo 7 que no es significativa

**Figura 26. Estimaciones y residuos del modelo 2.**

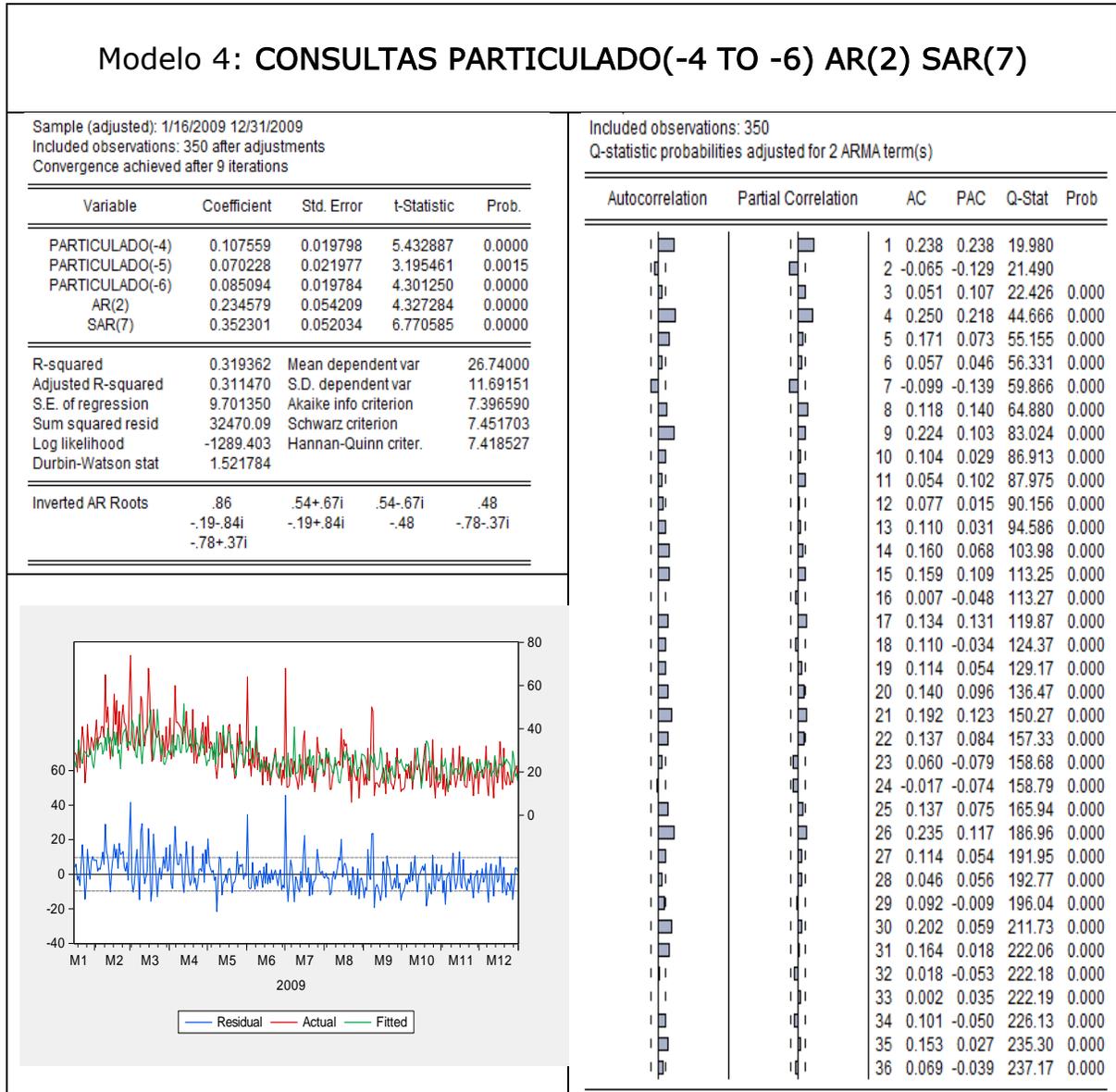


Fuente: Autor

Aunque el modelo presenta un mejor ajuste, no son significativos la constante y los términos en los retardos (0) y (-3). Para el tercer



Figura 28. Estimaciones y residuos del modelo 4.

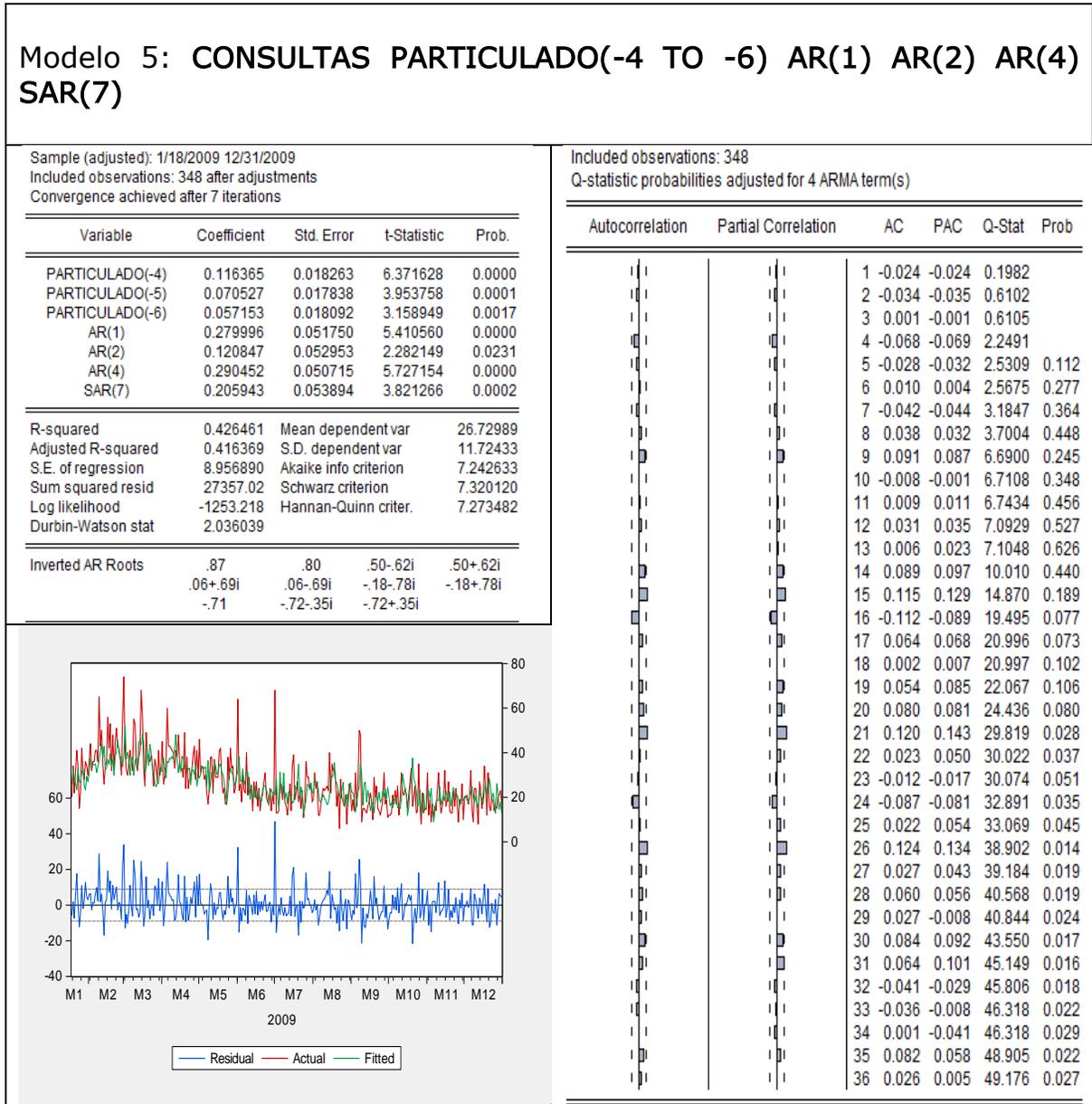


Fuente: Autor

En el gráfico se puede ver que los residuos ganan estacionariedad aunque el modelo no presenta un buen ajuste a los datos. El correlograma de los residuos deja ver una estructura compleja, se puede ver en la autocorrelación parcial un modelo AR(1) y AR(4)

Se va a intentar con el siguiente modelo:

Figura 29. Estimaciones y residuos del modelo 5.



Fuente: Autor

Se observa que el modelo presenta un mejor ajuste, la serie de los residuos es más estable y el correlograma muestra que aparentemente se tiene ruido blanco. Los términos del ruido son todos significativos y las estimaciones de la función de transferencia también lo son. Para corroborar la estacionariedad de los residuos, en la Figura 30 se presenta la prueba de Dickey Fuller que con un valor\_p igual a cero

rechaza la hipótesis nula de no estacionariedad, y se concluye que los residuos de la regresión son estacionarios.

**Figura 30. Prueba de estacionariedad de los residuos de la regresión.**

Null Hypothesis: ERR5 has a unit root  
 Exogenous: Constant  
 Lag Length: 0 (Automatic - based on SIC, maxlag=16)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-19.03037	0.0000
Test critical values:		
1% level	-3.448998	
5% level	-2.869653	
10% level	-2.571161	

\*Mackinnon (1996) one-sided p-values.

Fuente: Autor

En la FAC y en la FACP se observan muy pocos valores significativamente diferentes de cero y el valor del  $R^2_{adj}$  es de 41,63% que indica que el modelo de regresión explica casi el 42% de la variación de la variable consultas.

El modelo finalmente estimado es:

$$(Consultas)_t = (w_3)B_{t-4} + (w_4)B_{t-5} + (w_5)B_{t-6} \frac{(1 - \theta B^7)}{(1 - \phi_1 B - \phi_2 B^2 - \phi_4 B^4)} a_t$$

Siendo :

$$w_3 = 0,116365; w_4 = 0,070527; w_5 = 0,057153$$

$$\phi_1 = 0,279996; \phi_2 = 0,120847; \phi_4 = 0,290452$$

Que se podría interpretar como:

El número de consultas de urgencias por afecciones respiratorias en Bucaramanga depende del 0,116365 de la lectura máxima del particulado de hace 4 días más el 0,070527 de la lectura máxima del particulado de hace 5 días, más el 0,057153 de la lectura máxima del particulado de hace 6 días más

## CONCLUSIONES

- En promedio en la ciudad de Bucaramanga se atendieron diariamente 27 pacientes (con desviación de 12) en el 2009, por afecciones respiratorias en el servicio médico de urgencias que incluye la consulta de urgencias, las urgencias en observación y la hospitalización.
- La variación estacional de las consultas médicas en el servicio de urgencias en la ciudad de Bucaramanga durante el año 2009 oscila entre el 85,2% del valor medio en los días sábados y un 119,3% en los días miércoles.
- En los días domingo y lunes se registra la mayor atención en el servicio de urgencias. Por el contrario, en los días viernes y jueves (más el viernes que el jueves) se presenta la menor demanda en este servicio médico.
- Las lecturas diarias máximas en microgramos por metro cúbico ( $\mu\text{g}/\text{m}^3$ ) del particulado  $\text{PM}_{10}$  registrados en la Estación Automática Centro de monitoreo de calidad del aire de Bucaramanga presentaron un valor promedio de  $96,76 \mu\text{g}/\text{m}^3$  con desviación de  $21,93 \mu\text{g}/\text{m}^3$ , valor inferior al valor de la norma que es de  $134 \mu\text{g}/\text{m}^3$ . Al considerar la desviación estándar y asumir un comportamiento relativamente normal, se puede afirmar la existencia de valores superiores a los de la norma.
- En términos generales se puede decir que la calidad del aire en la ciudad de Bucaramanga es buena en cuanto al particulado  $\text{PM}_{10}$  ya que sus niveles se mantienen por debajo del valor establecido por la norma la mayor parte del tiempo.
- La variación estacional los máximos de  $\text{PM}_{10}$  en Bucaramanga durante el año 2009 oscila entre el 78,06% de la media los días sábados y 109,57% de la media, los días miércoles.

- Los índices para los días sábados son los más bajos de la semana, sus valores son bastante más pequeños que los del promedio semanal. Para el resto de la semana estos valores se mantienen por encima del promedio.
- La estimación proporcionada por el modelo de función de transferencia, es demasiado suave no logrando captar la estructura de la serie y proporcionando valores muy alejados de los originales.
- La serie de valores máximos del particulado  $PM_{10}$  no es un buen predictor del número de consultas en el servicio de urgencias por afecciones respiratorias. El modelo presenta un moderado ajuste ( $R^2_{adj} = 41,63\%$ ).
- Se podría pensar en modelar los valores atípicos (intervención) que se observan en los residuos para tratar de ganar un mejor ajuste del modelo.
- Otra opción que se puede contemplar es analizar los meses de marzo, abril y mayo donde se observan las lecturas más altas del particulado, inclusive por encima de la norma ( $134 \mu\text{g}/\text{m}^3$ ) para poder detectar si en este espacio de tiempo la relación entre las dos series es más contundente.

## BIBLIOGRAFÍA

Cardona J. Contaminación ambiental y enfermedad respiratoria. Revista de Neumología. <http://www.encolombia.com/medicina/neumologia/>

Castillejos M. Caracterización de las circunstancias clínicas y epidemiológicas que rodean las muertes en días de alta y baja concentración de partícula (PM<sub>10</sub> y PM<sub>2.5</sub>) y otros contaminantes atmosféricos en la zona suroeste del área metropolitana de la Ciudad de México. Consejo de Estudios para la Restauración y Valoración Ambiental (CONSEVA), CGF 1998.

CDMB (Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga). Red de Monitoreo de Calidad del Aire del AMB (Área Metropolitana de Bucaramanga). Informe 2004

CDMB (Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga). Red de Monitoreo de Calidad del Aire del AMB (Área Metropolitana de Bucaramanga). Informe Enero – Marzo 2010

Dockery D. y Pope A. Epidemiology of acute health effects: Summary of time-series studies. En *Particles in our air: concentration and health effects* (J.D. Spangler & R. Wilson Eds) Harvard University Press. 1996 <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1480527/>>

DTB (Dirección de Tránsito y Transporte de Bucaramanga) - Vehículos matriculados [en línea] <[w.w.w.transitobucaramanga.gov.co](http://w.w.w.transitobucaramanga.gov.co)>

García A, Icaza G, Angeles F, Blanco S, Gómez J, Delgadillo M, Mandujano F, Fentanes O, Segovia E, Salgado P, García H, Sepúlveda J, León E. Proyecto de monitoreo ambiental, microambiental y de exposición personal a partículas menores a diez micrómetros (PM<sub>10</sub>) y menores a 2.5 micrómetros (PM<sub>2.5</sub>) en la Zona Metropolitana de la

Ciudad de México (ZMCM). Consejo de Estudios para la Restauración y Valoración Ambiental (CONSERVA), GDF 2000

Manquían A., Tovar LR., Yáñez G. Aluminio en el tracto respiratorio bajo de residentes de la Ciudad de México. *Rev. Salud Pública* 2008, vol.10, n.3 pp. 452-462

Peña D., *Análisis de series de temporales*. Alianza Editorial. Madrid. 2005.

Ramírez M, Rojas RI, Rosa I, García R. Influencia de los Contaminantes Atmosféricos en las Infecciones Respiratorias Agudas en Mexicali-Baja California, México. *Inf. tecnol.* [En línea]. 2009, vol.20, n.3, pp. 89-100. <[http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0718-07642009000300011&lng=es&nrm=iso](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-07642009000300011&lng=es&nrm=iso)>.

Rodríguez LA, Rey JJ, Berena A, CastroH, Niederbacher J, Vera LM, Cala L, Bolívar F. Prevalencia de síntomas respiratorios indicativos de asma y asociación con contaminación atmosférica en preescolares de Bucaramanga, Colombia. *Biomédica* 2010; 30:0-0a

Rodríguez LA, Berena A, Castro H, Niederbacher J, Vera LM. Incidencia de síntomas respiratorios y su asociación con contaminación atmosférica en preescolares: un análisis multinivel. *Cad. Saúde Pública*, Rio de Janeiro, 26(7):1411-1418, 2010b

Rodríguez LA, Berena A, Castro H. Niveles de contaminación atmosférica y síntomas respiratorios en población con morbilidad en Bucaramanga. Colombia. VIII Congreso SOLANEP, 2010 c

Solarte I, Caicedo M, Restrepo S. Contaminación atmosférica y enfermedad respiratoria en niños menores de 14 años en Santafé de Bogotá. Pontificia Universidad Javeriana 1999 Bogotá

Téllez M, Romieu I, Polo M, Ruiz S, Meneses F, Hernández M. Efecto de la contaminación ambiental sobre las consultas por infecciones respiratorias en niños de la Ciudad de México. *Salud Pública México* 1997, 39(6): 513 - 21

Valencia A. Consultor del Ministerio de Ambiente. Entrevista Radio Caracol, [En línea] Julio 17 de 2009,  
<<http://www.caracol.com.co/nota.aspx?Id=846956>>

Wei, W.W.S.. Time Series Analysis, Pearson/Addison-Wesley. 2006.