

**PRONÓSTICO DE LA DEMANDA DE PASAJEROS DE  
TRANSMILENIO EN BOGOTÁ MEDIANTE MODELOS DE SERIES  
DE TIEMPO**

**FORECAST OF TRANSMILENIO PASSENGER DEMAND IN  
BOGOTÁ USING A TIME SERIES MODEL**

**Diana Marcela Rivera Bernal, [dmriverab@ulibertadores.edu.co](mailto:dmriverab@ulibertadores.edu.co)**

**Erica Constanza Quintero Rosas, [ecquinteror@ulibertadores.edu.co](mailto:ecquinteror@ulibertadores.edu.co)**

**RESUMEN**

En este trabajo, se presenta una estimación de la demanda del Sistema Masivo de Transporte de Bogotá tanto en su componente troncal como zonal, utilizando la metodología de series de tiempo tanto univariados como multivariados. De manera que el documento se desarrolla así: i) introducción, ii) justificación, iii) problema, iv) objetivo, v) marco teórico, vi) revisión de literatura, vii) metodología, viii) resultados y ix) conclusiones. De acuerdo con métricas como el RMSE y el MAPE, los resultados arrojaron el mejor modelo univariado  $SARIMA(2,1,2)X(0,1,1)_{12}$  y por su lado en lo multivariado es aquel donde las validaciones son explicadas a partir de la tarifa y el IPC de Bogotá.

***Palabras clave:*** Transporte Masivo, Demanda, Series de tiempo, Transmilenio.

## **ABSTRACT**

Our text presents the demand estimation of the Bogota mass transit system, both in the trunk and zonal components, using the univariate and multivariate time series methodology. The document includes: i) introduction; ii) justification; iii) problem; iv) objective; v) theoretical framework; vi) literature review; vii) methodology; viii) results; and ix) conclusions. According to the RMSE and MAPE metrics, the results projected the best univariate model SARIMA  $SARIMA(2,1,2)X(0,1,1)_{12}$ ; in relation to the multivariate model, the variations are explained from the rate and Bogota consumer price index.

**Keywords:** Mass Transport, Demand, Time Series, Transmilenio.

### **1. INTRODUCCIÓN**

El crecimiento exponencial de la población y la urbanización en las grandes ciudades a nivel mundial ha generado grandes retos tanto a los gobiernos centrales como a los locales, en términos de política pública, planeación, financiación y desarrollo de los servicios públicos en lo que tiene que ver con cobertura, calidad y accesibilidad, entre los cuales se encuentra el transporte público de pasajeros.

En el caso particular de Colombia, ciudades como Bogotá, Barranquilla, Bucaramanga, Pereira, Medellín, Cali y Cartagena cuentan con un Sistema Integrado de Transporte Masivo (SITM), donde Bogotá fue el pionero en la estructuración tarifaria y operativa, el diseño de estructura y funcionamiento de las estaciones troncales y los paraderos en el componente zonal, así como también los contratos de recaudo y de operación a través de las concesiones privadas.

En términos generales, TransMilenio está compuesto por su infraestructura, el sistema de operación de buses, el sistema de operación de los equipos de recaudo y el sistema de gestión apoyado en un centro de control. El Distrito es responsable de la construcción y mantenimiento de la infraestructura y del suministro y operación de los equipos del centro de control. En lo que respecta al sector privado, se encarga de suministrar y operar, mediante contratos de concesión, los buses y los equipos de recaudo. Los ingresos por la actividad transportadora deberán cubrir todos los gastos de operación, mantenimiento, reposición de equipos y las utilidades de los actores privados en el sistema. Adicionalmente, el sector privado, a través de una fiducia centralizada, recibe el dinero recaudado y realiza la cascada de pagos entre los agentes del sistema conforme a las reglas contractuales establecidas en los procesos licitatorios.

De hecho, la estructura financiera del Sistema está compuesta por los ingresos del Sistema y los costos de operación de este, cuando los ingresos del sistema son menores a los costos se presenta un déficit. Dicho déficit se cubre con transferencias del Distrito, a través de la Secretaría Distrital de Hacienda (SHD). Estos recursos ingresan al Fondo de Estabilización Tarifaria (FET) del Patrimonio Autónomo en el que se administran los recursos del Sistema (Transmilenio 2019, p.254). Se hace relevante aclarar que los ingresos dependen del comportamiento y dinámica de la demanda de los usuarios al sistema.

En términos de Mankiw (2006), la demanda se refiere a la cantidad de un bien que los compradores quieren y pueden comprar, es decir, la demanda está ligada directamente a la restricción presupuestal que cada uno de los individuos y al precio de un producto o bien. Por ejemplo, en el caso del transporte de Bogotá, el usuario tiene que tomar la decisión de cómo transportarse de un punto de origen a un punto de destino a través de la distinta oferta como es el servicio de taxi, transporte de plataformas digitales, transporte público colectivo,

transporte pirata y el sistema integrado de transporte masivo TransMilenio y toma la decisión final en función de su disponibilidad de dinero al momento de querer realizar un desplazamiento.

En la actualidad TransMilenio transporta en promedio al día tres millones quinientos mil pasajeros y el reto es seguir incrementado la demanda dado que se están haciendo varias acciones, entre ellas se encuentran, la renovación de la flota troncal, el desmonte del Transporte Público Colectivo (TPC) y la llegada de buses zonales con tecnologías más amigables con el medio ambiente para el cubrimiento de zonas de Bogotá como Fontibón, Suba y Usme.

Al ser la estadística una herramienta que permite hacer análisis e inferencias de variables cuantitativas y cualitativas en todos los campos del conocimiento, en el artículo de proyecto de aplicación se pretende aterrizar un caso práctico a partir de los conceptos teóricos que se ven en los módulos de la especialización de estadística aplicada. En procura de dar finalidad a lo anterior, se tendrá como objetivo estimar los factores tanto propios del sistema como los externos que pueden estar afectando el comportamiento de la demanda del Sistema Integrado de Transporte Público en Bogotá a lo largo del tiempo. Como punto de partida, se revisará la data existente de la demanda, las variables propias del sistema tales como la tarifa y los kilómetros recorridos y, las variables externas como por ejemplo el PIB de Bogotá, precipitaciones en la ciudad de Bogotá, y troncales y la tarifa del Sistema.

## **2. JUSTIFICACIÓN**

En general, Transmilenio está compuesto por su infraestructura, el sistema de operación de buses, el sistema de operación de los equipos de recaudo y el sistema de gestión apoyado en un centro de control. El Distrito es responsable de la construcción y mantenimiento de la

infraestructura y del suministro y operación de los equipos del centro de control. En lo que respecta al sector privado, se encarga de suministrar y operar, mediante contratos de concesión, los buses y los equipos de recaudo. Los ingresos por la actividad transportadora deberán cubrir todos los gastos de operación, mantenimiento, reposición de equipos y las utilidades de los actores privados en el sistema. Adicionalmente, el sector privado, a través de una fiducia centralizada, recibe el dinero recaudado y realiza la cascada de pagos entre los agentes del sistema conforme a las reglas contractuales establecidas en los procesos licitatorios. De hecho, la estructura financiera del Sistema está compuesta por los ingresos del Sistema y los costos de operación de este, cuando los ingresos del sistema son menores a los costos se presenta un déficit. Dicho déficit se cubre con transferencias del Distrito, a través de la Secretaría Distrital de Hacienda (SHD). Estos recursos ingresan al Fondo de Estabilización Tarifaria (FET) del Patrimonio Autónomo en el que se administran los recursos del Sistema (Transmilenio 2019, p.254). Se hace relevante aclarar que los ingresos dependen del comportamiento y dinámica de la demanda de los usuarios al sistema.

Por lo anterior, es relevante para el Sistema de Transporte Masivo – Transmilenio conocer la demanda de pasajeros en futuros meses, lo cual permitiría lograr una sostenibilidad y planeación financiera y, por tanto, reducir los recursos del FET.

### **3. PROBLEMA**

¿Qué variables exógenas y endógenas permiten el pronóstico de la demanda de pasajeros del sistema de transporte masivo Transmilenio en la ciudad de Bogotá?

### **4. OBJETIVO**

Pronosticar la demanda de pasajeros del sistema de transporte masivo Transmilenio en la ciudad de Bogotá en función de las variables exógenas y endógenas.

## 5. MARCO TEÓRICO

### 5.1. Conceptos Básicos

Teniendo presente el problema y el objetivo del presente documento, se hace vital la definición de conceptos básicos tales como:

Transporte: Es un sistema organizacional y tecnológico que apunta a trasladar personas y mercancías de un lugar a otro para balancear el desfase espacial y temporal entre los centros de oferta y demanda. Lo anterior plantea el problema de realizar este traslado en forma eficiente y sustentable (Garrido, 2001)

Sistema de Transporte: Se refiere al trazado y características de la red vial y colectora, así como también la definición del sistema de transporte urbano y organización de rutas de este (Camargo, 2010). Por su parte Tolley (1995) afirma que los sistemas de transporte son la respuesta a las crecientes necesidades de comunicación entre individuos como entre sociedades para la movilidad de mercancías como parte de las economías regionales y mundial.

Transporte Masivo de Pasajeros: se define a través del Decreto 1079 de 2015, como el servicio que se presta a través de una combinación organizada de infraestructura y equipos, en un sistema que cubre un alto volumen de pasajeros y da respuesta a un porcentaje significativo de necesidades de movilización.

### 5.2. Demanda de Transporte

En general la demanda se define en la siguiente ecuación:

$$Q(p) = a - bp$$

Donde:

- a y b: Parámetros constantes

- p: precio del producto o servicio

De hecho, esta ecuación se puede aplicar al transporte dado que el valor de p para el caso de Transmilenio sería el precio del tiquete, para el taxi el valor de la carrera, o el TPC el costo del trayecto. Por lo tanto, el usuario o consumidor del sistema de transporte es quien elige entre las opciones cual tomar dependiendo sus necesidades y su capacidad adquisitiva. Es así como la demanda de transporte se traduce en la cantidad de viajes requeridos durante un periodo determinado en función de un conjunto de variables explicativas o factores determinantes.

En cuanto a los factores que tienen un efecto en la demanda, según expertos como Stuart Cole, son las características físicas, el precio, los precios relativos de los diferentes modos de transporte o de servicios de transporte similares, ingreso del pasajero, velocidad del servicio, calidad del servicio y seguridad.

### **5.3 Series de Tiempo**

Una serie de tiempo hace referencia al conjunto de mediciones que describen la evolución de un fenómeno o variable a lo largo del tiempo (Pepió, 2001). A este tipo de series se les puede hacer análisis de tipo descriptivo tanto de tipo numérico como gráfico, permitiendo contextualizar al investigador de la evolución de las observaciones a través del tiempo.

De hecho, con el fin de realizar el pronóstico de la demanda de pasajeros del sistema masivo de transporte de Bogotá, se utilizará series de tiempo como marco de referencia, dado que tiene múltiples aplicaciones en áreas como la economía, en las ciencias naturales y las exactas. Es así, como existen dos tipos de enfoques, por un lado, el enfoque de la importancia del tiempo, es decir, como los periodos pasados de las variables están afectando el hoy. De otro lado, el enfoque de frecuencia que consiste en darle importancia a los ciclos.

Es relevante mencionar que dentro de las series de tiempo existe multiplicidad de modelos para realizar la predicción. Sin embargo, se va a hacer referencia a algunos de ellos con una breve explicación y síntesis a partir de los expuesto por Gujarati (2010).

### 5.3.1 ARMA

Estos modelos cuentan con un proceso autorregresivo (AR) y de medias móviles (MA), de esta manera se constituye el orden p,q. En general, se definen en general bajo las siguientes ecuaciones:

$$\text{AR}(p): (Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \dots + \alpha_p(Y_{t-p} - \delta) + \mu_t$$

Donde:

- $\delta$  es la media de Y y  $\mu_t$  es un término de error aleatorio no correlacionado con media cero y varianza constante (ruido blanco)

$$\text{MA}(q): Y_t = \mu + \beta_0\mu_t + \beta_1\mu_{t-1} + \beta_2\mu_{t-2} + \dots + \beta_q\mu_{t-q}$$

- $\mu$  es una constante y es el término de error estocástico de ruido blanco, es así como el proceso MA(q) es una combinación lineal de términos de error de ruido blanco.

Por lo tanto, un proceso ARMA de orden p=1 y q=1 se denota de la siguiente manera:

$$Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 \mu_t + \beta_1 \mu_{t-1}$$

- Donde  $\theta$  representa un término constante.

### 5.3.2 ARIMA

Dado que la mayoría de las series de tiempo no son estacionarias, se hace necesario diferenciar la serie para hacerla estacionaria<sup>1</sup>. De hecho, a partir de la publicación hecha por

---

<sup>1</sup> Si se va ha pronosticar bajo Box-Jenkins, debe suponerse que sus características son constantes a través del tiempo y, en particular, en periodos futuros. Así, la sencilla razón para requerir datos estacionarios es que

G.P.E Box y G.M. Jenkins titulada Times Series Analysis: Forecasting an Control, marcó la nueva generación para la realización de pronósticos. Dicha metodología es la conocida con el nombre ARIMA (Metodología de Box-Jenkins), donde el análisis se centra en las propiedades estocásticas de las series de tiempo.

Consecuentemente, en los procesos ARIMA se tiene un orden p,d,q, donde d es el número de diferencias que se le aplica a la serie.

### 5.3.3. SARIMA

Dado que en ocasiones la serie de tiempo ( $Y_t$ ) no es posible transformar la variable para quitar la estacionalidad, se hace indispensable considerar un modelo ARIMA con componente estacional, es decir, SARIMA. De esta manera, el proceso puede ser simplificado así:

$$\phi(B)\Phi(B^s)[(1 - B)^d(1 - B^s)^D - \mu]y_t = \theta(B)\Theta(B^s)u_t$$

Lo anterior indica lo siguiente:

$$\text{ARIMA } \underbrace{(p, d, q)}_{\text{Componente no estacional}} \times \underbrace{(P, D, Q)_s}_{\text{Componente estacional}}$$

### 5.3.4. SARIMAX

Están compuestos por un componente autoregresivo, medias móviles y un componente de integración. En este tipo de modelo se incluye inputs o variables exógenas en el modelo. En general

$$\Phi_p(B^s)\phi_p(B)\nabla_s^D\nabla^d x_t = \alpha + \theta_q(B^s)\theta_q(B)w_t + \sum_{k=1}^r B y_{t-k}$$

se denotan así:

### 5.3.5. VAR (Vector Autoregresivo)

Los modelos VAR tienen como una de sus principales utilidades la de estudiar la respuesta dinámica de una serie cuando se produce un shock imprevisto de la otra serie. Asimismo, se

---

todo modelo que se infiera a partir de estos datos pueda interpretarse como estacionario o estable en si mismo, y proporcione, por consiguiente, una base válida para pronosticar (Pokorny, 1987,p.343).

utilizan cuando se requiere encontrar relaciones simultaneas en un grupo de variables, generando una estimación de un sistema de ecuaciones. Al tener una serie multivariada ( $z_t$ ), un modelo VAR de orden p esta dado por lo siguiente:

$$z_t = \phi_0 + \sum_{i=1}^p \phi_i z_{t-i} + a_t$$

Donde:

- $\phi_0$ : es una constante dimensión k.
- $\phi_i$ : matrices de dimensión k x k para todo  $i > 0$ ,  $\phi_p \neq 0$
- $a_t$ : sucesión de vectores aleatorios independientes e idénticamente distribuidos con media cero y matriz de covarianza definida positiva.

### **5.3.6. VEC (Vector de corrección de errores vectoriales)**

Es un modelo que depende de la prueba de cointegración, es decir, que las variables guardan una relación de equilibrio de largo plazo entre ellas.

## **6. REVISIÓN DE LITERATURA**

### **6.3. Estudios de Demanda de Sistema de Transporte de Pasajeros en Colombia y en el mundo**

Para realizar una contextualización adecuada, se hizo necesario realizar una revisión inicial de los elementos que pueden afectar el comportamiento de la demanda de Transmilenio y cuyas variaciones tiene diversas implicaciones en términos financieros, económicos, sociales y técnicos. Es así, como se encontró un estudio de los factores que indican en la demanda del sistema TransMilenio en Bogotá, en donde afirma que la demanda está afectada por la competencia del transporte público colectivo, PIB, Tarifa y el crecimiento poblacional

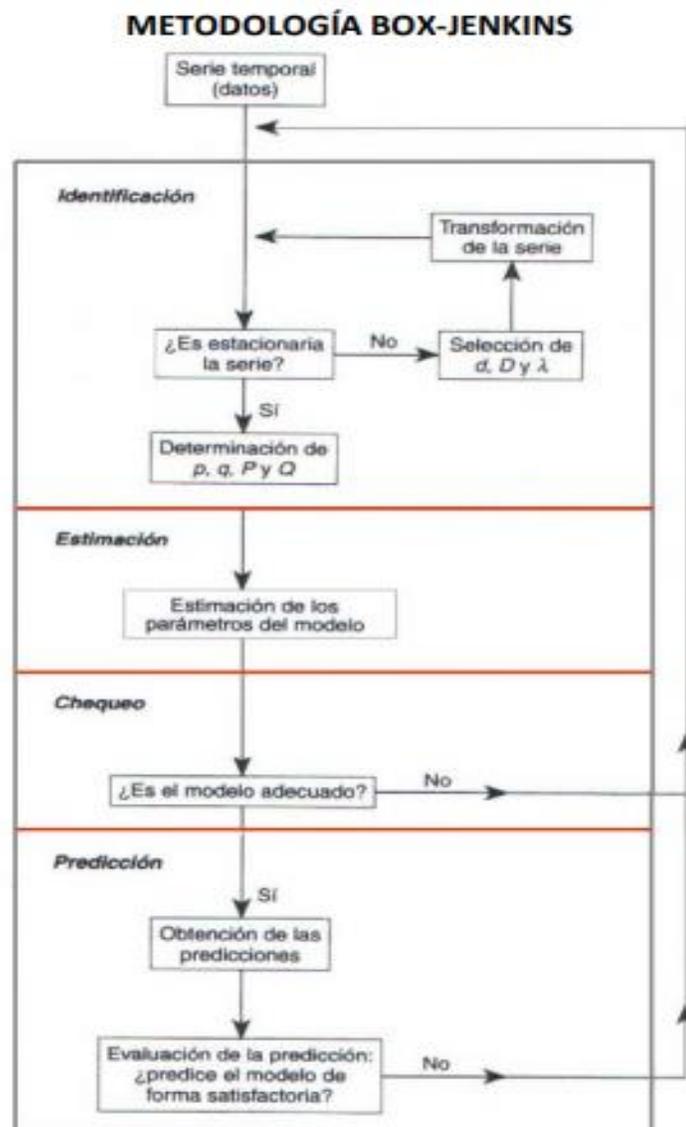
(Velásquez, 2009). Asimismo, el autor Rojas (2005) efectuó una comparación entre los sistemas BRT (autobús de transporte rápido) de Curitiba y Bogotá, donde encontró que afecta directamente la demanda de pasajeros en el caso de Transmilenio la guerra del centavo y que el sector público influya directamente en la prestación del servicio.

Por otro lado, se efectuó una búsqueda de artículos académicos y estudios para predecir la demanda de pasajeros de transporte en diferentes países del mundo. A continuación, se hace una breve compilación de los más relevantes:

- Atlanta, Georgia: Se realizó un análisis basado en series temporales univariadas para la estimación de tiempos de viajes utilizando un modelo SARIMA (3,0,2)(0,1,1), donde obtiene un porcentaje de MAPE del 8,8% (Guin, 2006)
- Madrid, España: En el estudio realizaron varios modelos de regresión múltiple para determinar la demanda potencial de viajeros en el metro de Madrid. De esta forma, las variables que explicaron mejor a la demanda fueron Empleo Total, la población total, la cantidad de accesos que dispone la estación y la cantidad de líneas que pasan por la estación (Cardozo, Garcia & Gutierrez, 2008).
- Santiago de Chile, Chile: El objetivo se centró en poder probar la factibilidad de pronosticar la afluencia de pasajeros con técnicas de minería de Datos con un error aceptable (inferior al 10%). Para este fin tomaron información de 4 estaciones (Universidad de Chile, Pedro de Valdivia, Escuela Militar y San Pablo), y cuyos resultados demostraron que el mejor modelo para pronosticar la afluencia de pasajeros en el corto plazo (mensual) es el de Promedios Móviles, con un MAPE bajo 3.3% al predecir la afluencia de pasajeros en todas las estaciones seleccionadas para los experimentos (Garnica, 2011).

## 7. METODOLOGÍA

Para realizar cada uno de los modelos a través de series de tiempo, se utiliza el software RStudio que permite hacer manejo de datos, análisis estadísticos, simulaciones, entre otros. Es así como se utilizaron librerías tales como: vars, devtools, tseries, urca, highcharter, dplyr, ggfortify, libplotly, TSstudio, forecast, metrics y mlmetrics, utilizando la metodología de Box- Jenkins, cuyas etapas se sintetizan en el siguiente esquema:



Asimismo, es vital comprender lo que se hace en cada una de las etapas para poder llevar a cabo la predicción:

- a. Identificación: Se debe determinar si la serie es estacionaria a través de las pruebas de Dickey Fuller y Phillips Perron. En caso de que arrojase la no estacionariedad, se procede a realizar una diferencia ordinaria o estacionaria.
- b. Estimación: Con los gráficos de las funciones ACF y PACF se prosigue a determinar el valor de  $p$  y  $q$ (componente regular),  $d$ (diferencia),  $P$  y  $Q$  (componente estacional).
- c. Chequeo: Se evalúan los resultados de las pruebas de Ljun Box (Autocorrelación), Jarque Bera(Normalidad), Runs Test (Aleatoriedad) y BIC.
- d. Predicción: Al tener elegido el mejor modelo, se procede a realizar la predicción y se evalúa que tan cercano esta a la realidad de la variable estudiada, a través de dos métricas tales como el RMSE y MAPE.

Por otro lado, las variables que constituyen la base de datos a utilizar son las siguientes:

- a. IPC Bogotá: Se refiere a la variación anual del índice de precios al consumidor, es decir, un valor que refleja las variaciones que experimenta los precios en un momento dado. Se cuenta con el histórico mensual a partir del reporte del DANE desde el año 2001 a abril de 2022 para la ciudad de Bogotá.
- b. Kilómetros Recorridos: Describe el total de kilómetros recorridos por los buses tanto del componente troncal como zonal. Dicha variable fue tomada directamente desde datos oficiales de Transmilenio desde el año 2001 a 2022 por mes.
- c. Precipitaciones: Son el volumen de lluvias mensuales obtenidas en la red de estaciones hidrológicas de la Corporación Autónoma Regional de Cundinamarca - CAR.

- d. Tarifa Promedio Ponderado Usuario: Hace referencia al monto de dinero que los usuarios del Transporte Masivo de Bogotá pagan por la utilización del sistema. Es relevante mencionar que la tarifa tiene inmersa la ponderación tanto de lo troncal como lo zonal. Es así como se cuenta con los datos desde inicio de operación de Transmilenio (2001) a la actualidad.
- e. Validaciones: Número de registros (entradas) que reporta el sistema cada vez que los usuarios colocan la tarjeta en las BCA's para pagar los servicios del SITP.

## 8. RESULTADOS

Aplicando distintos modelos como se describieron en el marco teórico, a continuación, se exhibe los principales resultados.

Para comenzar, se tomó la base de datos en bruto para analizar mediante el summary las medidas de tendencia central y de posición para cada una de las variables:

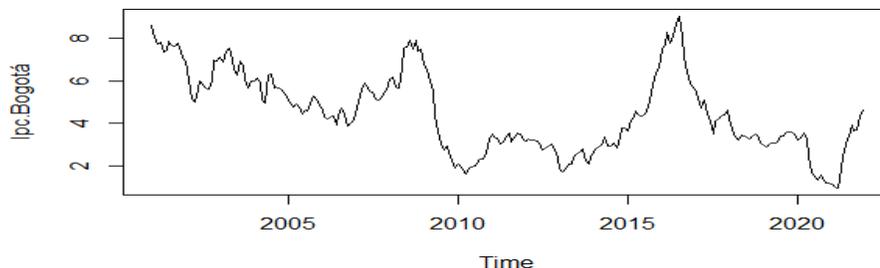
**Tabla 1**

IPC Bogotá	Precipitaciones	Km Recorridos	Tarifa Usuario
Min. :0.970	Min. : 0.00	Min. : 212094	Min. : 800
1st Qu.:3.075	1st Qu.: 38.40	1st Qu.: 5257215	1st Qu.:1200
Median :4.320	Median : 65.90	Median : 7941182	Median :1500
Mean :4.513	Mean : 76.54	Mean :20263801	Mean :1482
3rd Qu.:5.860	3rd Qu.:101.50	3rd Qu.:40004395	3rd Qu.:1725
Max. :9.030	Max. :330.00	Max. :57782119	Max. :2037
	NA's :36		
Validaciones			
Min. : 8770278			
1st Qu.: 79625902			
Median :117007020			
Mean :127954015			
3rd Qu.:188363262			
Max. :246282935			

Se puede decir que, en el IPC de Bogotá, los kilómetros recorridos y las validaciones, se presenta una distancia significativa entre el valor mínimo y máximo, lo cual guarda coherencia con la realidad dado que el sistema con el transcurrir de los años ha ido creciendo en cobertura, infraestructura y precio.

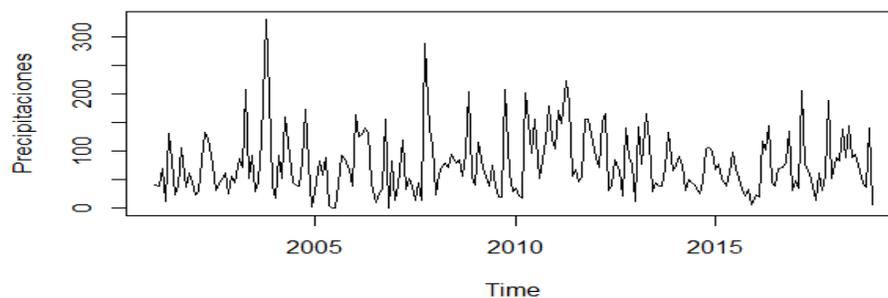
Asimismo, es vital realizar la visualización a través de gráficos del comportamiento de las variables que se quieren analizar así:

**Gráfica 1**



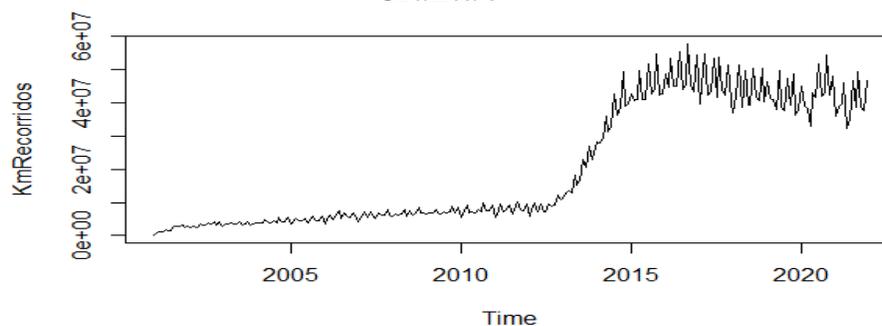
El IPC de Bogotá se puede observar que a primera vista los valores no están rondando en una media y varianza constante, a su vez presenta valores extremos, es decir, o muy altos en los años 2009 y 2016.

**Gráfica 2**



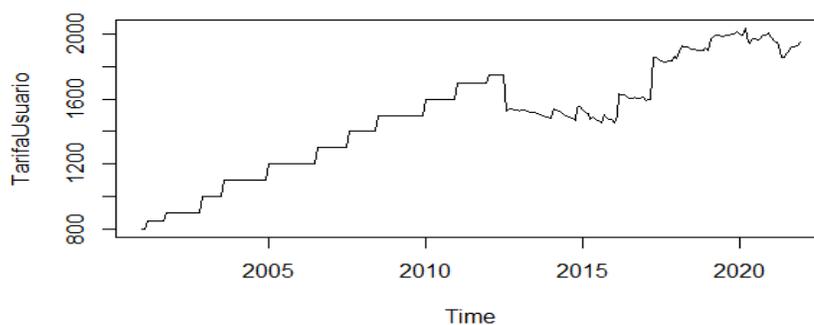
Dado que es una variable climática, en su comportamiento, se visualiza la presencia de ciclos y tal vez que se cuente con una media y varianza constante.

**Gráfica 3**



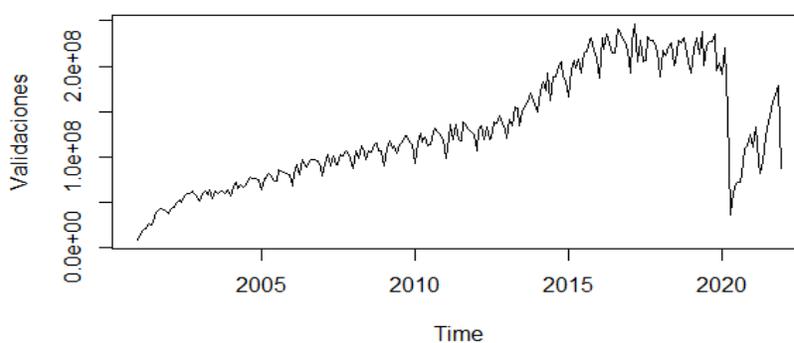
Con respecto a los kilómetros recorridos, se puede decir que la variable experimenta una tendencia al alza a partir del año 2012 dado que ingresó el sistema zonal a movilizar a los bogotanos al interior de las localidades de la ciudad. Asimismo, la serie se puede dividir en dos periodos: 2001 a 2011 y 2012 en adelante, ya que los valores fluctúan alrededor de una media y varianza constante.

**Gráfica 4**

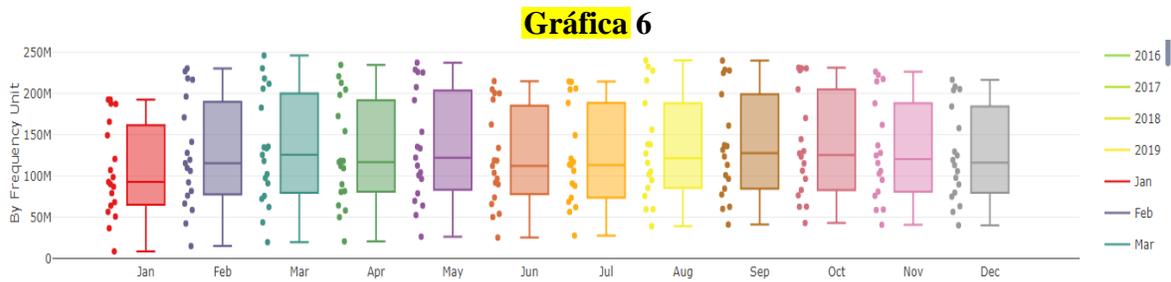


En la tarifa usuario se puede observar un comportamiento alcista en casi toda la serie, sin embargo, para el periodo comprendido entre los años 2012 al 2015 se ve una caída dada las distintas decisiones de política pública en Bogotá.

**Gráfica 5**



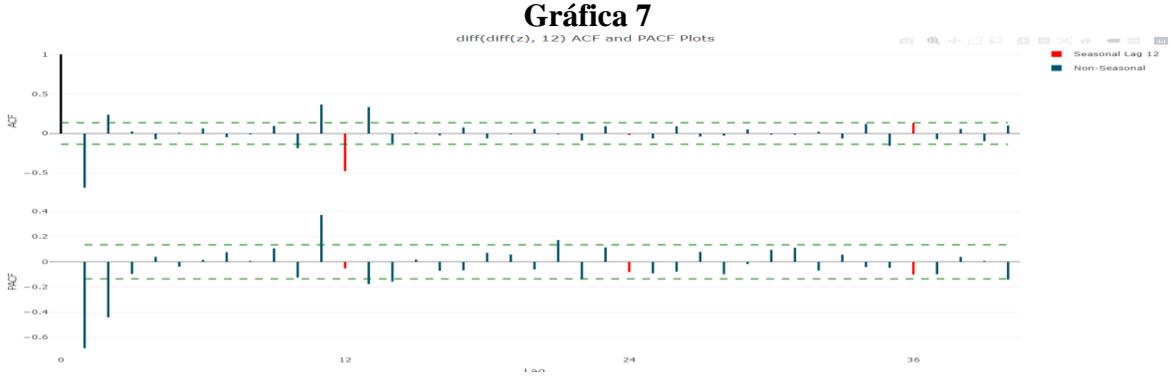
Por último, la variable central de este documento, es decir, las validaciones (demanda) a lo largo del periodo de análisis presentó un comportamiento ascendente con apariencia de ciclos dadas las vacaciones y festividades, por ejemplo, en los meses de enero, junio y diciembre se visualiza en el gráfico 6 como la mediana es menor respecto a los otros meses del año.



Con la declaración de la pandemia en marzo del 2020 y las restricciones de movilidad sobre todo el distanciamiento y aislamiento, generó una disminución en los usuarios que frecuentaban las estaciones y paraderos del SITP. Adicionalmente, para el año 2021 se ve una leve recuperación de la demanda, pero sin lograr alcanzar los niveles anteriores a la pandemia.

De esta manera, se tiene una primera aproximación de las variables con que se cuenta para hacer la predicción. Teniendo en cuenta el marco teórico se procede a mostrar los resultados de los diferentes modelos que se propusieron, con el fin de estimar la demanda mensual de pasajeros del sistema masivo de transporte de Bogotá en el corto plazo. Para ejecutar y estimar los diferentes modelos univariados, se procedió a tomar los datos mensuales desde enero de 2001 hasta junio de 2019 para quitar el impacto en la predicción del periodo de pandemia.

A continuación, se exhibe los gráficos de las funciones ACF y PACF, con el fin de poder determinar los posibles modelos así:



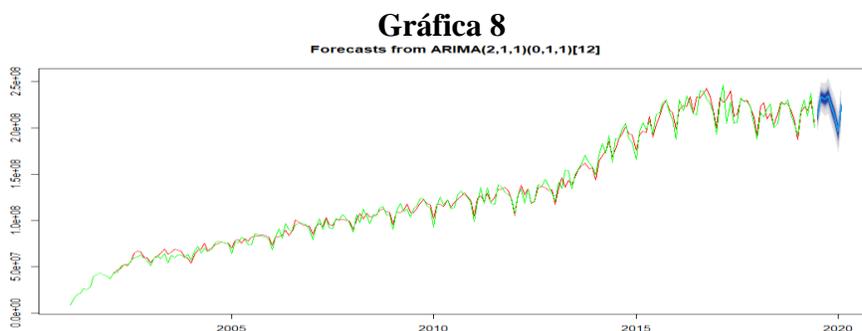
Es así como en la tabla 2, se muestra un resumen con los modelos SARIMA para la variable de interés, es decir, validaciones:

**Tabla 2**

MODELO	BIC	INDEPENDENCIA (LB)	NORMALIDAD (JB)	ALEATORIEDAD (RT)	RMSE	MAPE
SARIMA: (1,1,0) X (0,1,1) <sub>12</sub>	7183.934	4.44E-11	3.66E-15	0.587	11,553,726	4.068%
SARIMA: (1,1,1) X (0,1,1) <sub>12</sub>	7162.147	8.84E-05	4.50E-09	0.007934	11,372,501	4.053%
SARIMA: (2,1,1) X (0,1,1) <sub>12</sub>	7158.574	0.009813	1.11E-07	0.008245	11,108,876	3.925%
SARIMA: (11,0,11) X (0,1,1) <sub>12</sub>	7206.837	0.9969	0.006557	0.2818	11,961,616	4.332%
SARIMA: (1,1,2) X (0,1,1) <sub>12</sub>	7160.911	0.007888	6.29E-07	0.008245	11,290,128	4.037%
SARIMA: (2,1,2) X (0,1,1) <sub>12</sub>	7163.742	0.01311	1.94E-07	0.008245	11,109,707	3.923%

Es relevante indicar que las métricas de RMSE y MAPE, se calcularon con 8 rezagos y su vez se compararon con 8 datos reales. Con los resultados obtenidos, se tomaría el  $SARIMA(2,1,2)X(0,1,1)_{12}$  bajo el criterio de MAPE mas pequeño, mientras si se tiene el de menor BIC se utilizaría el modelo  $SARIMA(2,1,1)X(0,1,1)_{12}$ . Al evaluar los supuestos básicos, el mejor modelo es el  $SARIMA(2,1,2)X(0,1,1)_{12}$  dado que los datos se distribuyen de manera independiente y los residuales exhiben un comportamiento de aleatoriedad. En lo relacionado con el RMSE el margen de error mensual es aproximadamente de 11 millones, lo cual es un valor relativo pequeño dado que la demanda de pasajeros está registrando alrededor de 246 millones en meses típicos antes de pandemia.

El pronóstico del modelo  $SARIMA(2,1,2)X(0,1,1)_{12}$  se ajusta bastante a los datos reales tal como lo presenta la siguiente gráfica:



Teniendo presente los resultados de los modelos univariados y la base de datos con que se cuenta, se decidió realizar modelos multivariados que nos permitieran analizar el

comportamiento de la variable de interés (validaciones) con variables endógenas y exógenas al Sistema Integrado de Transporte de Bogotá, donde se incluyó el periodo de pandemia al análisis. Es vital indicar que las métricas de RMSE y MAPE, se calcularon con 6 rezagos y su vez se compararon con 6 datos reales. De hecho, en la tabla 3, se observan los modelos VAR con los respectivos resultados de RMSE y MAPE:

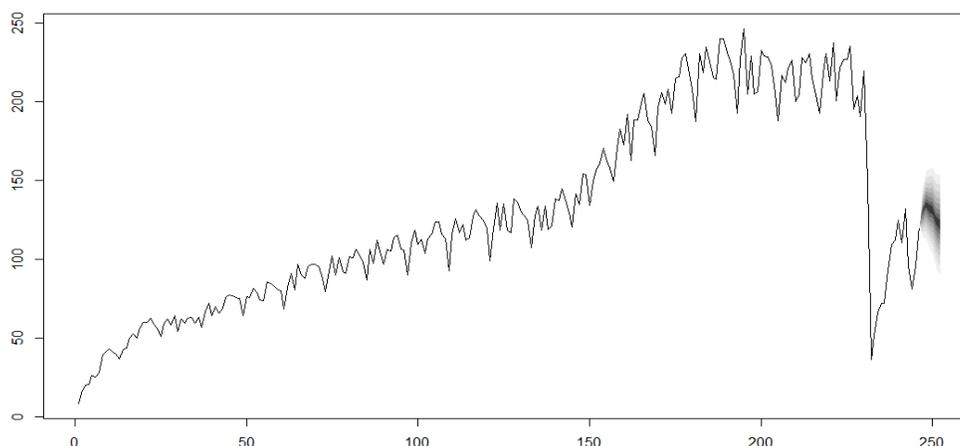
**Tabla 3**

Modelo VAR	Variables Tomadas	RMSE	MAPE
valid_todas	validaciones	39.18033	24.23%
	tarifa	36.03157	1.66%
	ipcb	0.839785	16.11%
	km	7.952432	14.83%
valid_tarifa_ipcb	validaciones	33.66964	21.44%
	tarifa	27.86359	9.18%
	ipcb	0.5268	16.87%
valid_km_tarifa	validaciones	35.61045	23.64%
	tarifa	83.98716	4.01%
	km	9.185667	16.87%
valid_km_ipcb	validaciones	36.50955	23.74%
	ipcb	0.4493429	9.18%
	km	6.641274	12.03%

Al efectuar la revisión, el mejor modelo (criterio de menor MAPE) arrojado es aquel donde se encuentran las variables validaciones, tarifa e IPC de Bogotá y también porque cumple los supuestos de ausencia de correlación serial y que los residuales sean homocedásticos. Es así, como se cuenta con un buen pronóstico para las tres variables al compararlas con los datos reales y cuyos resultados se observa en el gráfico 9 así:

**Gráfica 9**

Fanchart for variable valid



## 9. **CONCLUSIONES**

Las validaciones como variable que explica la demanda del Sistema Integrado de Transporte en Bogotá presentan ciclos y tendencias, por lo cual se hizo necesario realizar diferencias ordinarias y estacionarias para poder realizar los modelos a nivel univariado. Efectivamente, los modelos SARIMA en general tiene una buena métrica en cuanto al MAPE. En particular, en el mes de enero de 2020, en el periodo de predicción con el modelo  $SARIMA(2,1,2)X(0,1,1)_{12}$ , las validaciones (demanda) disminuyen en relación con el segundo semestre de 2019, lo cual es coherente con el comportamiento que presentó la variable en los meses de enero, julio y diciembre durante el periodo de análisis.

Por otro lado, el mejor modelo multivariado para explicar el comportamiento de la demanda a corto plazo se dio tanto en función de variables tanto exógenas como endógenas tales como la tarifa y el IPC de Bogotá, dado que cumplió los supuestos de ausencia de correlación serial y homocedasticidad, así como se logró comprobar la existencia de una ecuación de cointegración. Por tanto, el modelo hallado podría ser usado, para toma de decisiones en cuenta a la maximización de la flota y kilómetros del Sistema Transmilenio, y, por ende, una revisión del FET y de los costos asociados a la operación.

Como recomendación para estudios posteriores para determinar la demanda de pasajeros que se movilizan en Transmilenio, es indispensable incluir otra serie de variables exógenas tales como el número de viajes en bicicleta y el número de compra de motos. Adicionalmente, se podría proponer modelos en función del tipo de día, es decir, días hábiles, sábado, domingo y festivos.

## REFERENCIAS BIBLIOGRÁFICAS

- Camargo, M. (2010). Operacionalización del Plan de Desarrollo Urbano Local (PDUL): Fundamentos de la planificación urbanística y estratégica. *Revista Geográfica Venezolana*, 51 (1), 145-156. ISSN: 1012-1617. Disponible en: <https://www.redalyc.org/articulo.oa?id=3477/347730384009>.
- Cardozo, O. Garcia, J. Gutiérrez, J. (2008). Modelos de demanda potencial de viajeros en redes de transporte público: aplicaciones en el metro de Madrid. IV Seminario de Ordenamiento Territorial. [https://bdigital.uncu.edu.ar/objetos\\_digitales/3242/gutierrezproyeccion4.pdf](https://bdigital.uncu.edu.ar/objetos_digitales/3242/gutierrezproyeccion4.pdf)
- Cryer, J. D., & Chan, K. S. (2008). *Time Series Analysis with Applications in R* (2nd ed.). New York: Springer. <http://dx.doi.org/10.1007/978-0-387-75959-3>
- Escobar L.M., Parra K., Marín C.A., Bolaños R. I. & Martínez D. A (2020). Elasticidad de la demanda en el subsistema de transporte masivo del Área Metropolitana Centro Occidente en función del incremento de la tarifa. *Entre Ciencia e Ingeniería*, Volumen (14), 63-72. <https://doi.org/10.31908/19098367.1780>
- Decreto Único Reglamentario 1079 de 2015 Nivel Nacional. Diario Oficial No.49523 del 26 de mayo de 2015. Disponible en <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=62514>.
- Garnica, D. (2011). Pronóstico a corto plazo de afluencia de pasajeros utilizando técnicas de data mining: Metro S.A. Tesis para optar el grado de magister en gestión de operaciones. Universidad de Chile, facultad de ciencias físicas y matemáticas, departamento de ingeniería industrial.
- Garrido, R. (2001). Modelación de sistemas de distribución de carga. Universidad de Chile.
- Guin, A. (2006). Travel time prediction using a seasonal autoregressive integrated moving average time series model. 2006 IEEE Intelligent Transportation Systems Conference, pages 493–498.
- Guerrero, V. M. (2012). Análisis estadístico de series de tiempo económicas generadas con datos oficiales. *Realidad, datos y espacio. Revista Internacional de Estadística y Geografía*, 3(3).
- Gujarati, D. N., & Dawn, C. P. (2010). *Econometria* (V ed.). Mexico: McGraw-Hill.
- Hidalgo, D. (2001). TransMilenio: el sistema de transporte masivo de Bogotá. *Planeación y Desarrollo*, Volumen (XXXII), 173-186. [https://colaboracion.dnp.gov.co/CDT/RevistaPD/2001/pd\\_vXXXII\\_n2\\_2001\\_art.1.pdf](https://colaboracion.dnp.gov.co/CDT/RevistaPD/2001/pd_vXXXII_n2_2001_art.1.pdf)
- Mankiw, G. (2006). *Macroeconomía*, Sexta edición. New York: Worth Publishers.
- Pepió, M. (2001). *Series Temporales*. Universidad Politécnica de Cataluña. Aula Politécnica/ETSEIT.
- Pokorny, M. (1987). *An introduction to Econometrics*, Basil Blackwell. Nueva York.
- Rojas, F.(2005). El transporte público colectivo en Curitiba y Bogotá. *Revista de Ingeniería*.
- Tolley, R. & Turton, B. (1995). *Transport systems, policy and planning. A geographical approach*. Addison Wesley Longman Ltd. England.

TRANSMILENIO (2019). Informe de Gestión 2019. Obtenido de [https://www.google.com.co/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiE44y\\_j6\\_2AhU0RzABHefxDy8QFnoECAUQAQ&url=https%3A%2F%2Fwww.transmilenio.gov.co%2Fpublicaciones%2F151126%2Frendicion-de-cuentas-de-transmilenio-sa%2Fdescargar.php%3FidFile%3D4073&usg=AOvVaw0ms--2\\_tcH5m59fDZ4\\_jCD](https://www.google.com.co/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiE44y_j6_2AhU0RzABHefxDy8QFnoECAUQAQ&url=https%3A%2F%2Fwww.transmilenio.gov.co%2Fpublicaciones%2F151126%2Frendicion-de-cuentas-de-transmilenio-sa%2Fdescargar.php%3FidFile%3D4073&usg=AOvVaw0ms--2_tcH5m59fDZ4_jCD)

Velásquez, J. (2009). Análisis de factores que inciden en la demanda del sistema TransMilenio en Bogotá, Colombia. Revista de Ingeniería. Universidad de los Andes.

Wei, W. W. S. (1990). Time series analysis: Univariate and multivariate methods. Redwood City, Calif: Addison-Wesley Pub.