

# Calibración de Modelo para Estimación de Variación de Tráfico en Perú

Pedro Henrique Souza de Oliveira<sup>1</sup>, Pedro Henrique de Oliveira Cardoso<sup>1</sup>, Ícaro Ramos Nunes Batista<sup>1</sup>, Eduarda Ribeiro de Melo<sup>1</sup> e Frederico Rodrigues<sup>1</sup>

1: IMTRAFF – Consultoria e Projetos de Engenharia

## 1. Introducción

En 2017 se registraron por la administración de unidades de Peaje, 62 324 216 vehículos en las carreteras de Perú. La variación porcentual del tráfico vehicular de ligeros y vehículos pesados fue de -5,5% y -5,8%, respectivamente, en comparación con el año anterior (MTC, 2017), por ejemplo. Las variaciones de este tipo suceden por varios factores, especialmente, por las variables económicas. Y para Melo (2014), caracterizado por ser una dinámica continua, el tráfico en las zonas urbanas es directamente relacionado con el desarrollo económico.

Por esto, comprender el comportamiento futuro del flujo vehicular es importante para planificar acciones para optimizar las inversiones en infraestructura, en vista del crecimiento económico. O sea, el conocimiento de la evolución del flujo de vehículos es importante para la realización de diversos estudios de tráfico, transporte, concesiones, inversiones públicas, dimensionamiento del pavimento, mantenimiento, entre otros (Schiavon y Rodrigues, 2019).

El Ministerio de Transportes y Comunicaciones del Perú proporciona estadísticas de tráfico en el país. Entre los diversos datos, el tráfico de vehículos se puede usar para diversos análisis. Con el uso de estos datos, agregados a los datos de otras fuentes, como las plazas de peaje, se puede llegar a resultados que presentan series históricas del tráfico vehicular y comparar con variables económicas.

Así, este artículo tiene como objetivo desarrollar un modelo para estimar la variación del tráfico vehicular en Perú a partir de datos del Producto Bruto Interno (PBI) disponible en el sitio web del banco mundial y datos de tráfico vehicular disponibles en el sitio web del Instituto Nacional de Estadística e Informática (INEI).

## 2. Metodología de Trabajo

Para la calibración del modelo, se realizaron las siguientes etapas i) obtención de datos ii) Modelaje iii) Análisis de regresión y validación estadística. Estas son mejor presentadas en los ítems siguientes:

### 2.1. Recolección de datos

Para la calibración del modelo de estimación del crecimiento del tráfico en Perú se utilizaron los datos de tráfico total registrados por el centro de recolección y los puntos de control desde 1998 hasta 2017, disponibles en el del sitio web del INEI.

Los datos citados se presentan en la Tabla 1, con la suma de vehículos ligeros y pesados (todos los tipos de vehículos). La variación en el tráfico es la diferencia porcentual entre los años.

Tabla 1: Tráfico total registrado por el centro de recolección y puntos de control - INEI

Año	Vehículos totales	Variación de tráfico
1998	28.620.985	-
1999	30.071.899	5,07%
2000	30.844.475	2,57%
2001	32.408.444	5,07%
2002	33.372.686	2,98%
2003	32.992.117	-1,14%
2004	34.277.570	3,90%
2005	35.930.657	4,82%
2006	39.418.146	9,71%
2007	45.003.684	14,17%
2008	53.110.313	18,01%
2009	57.924.671	9,06%
2010	64.386.441	11,16%
2011	68.987.668	7,15%
2012	78.065.134	13,16%
2013	82.784.099	6,04%
2014	82.099.661	-0,83%
2015	86.284.468	5,10%
2016	89.668.493	3,92%
2017	89.491.544	-0,20%

Además de los datos que se refieren al tráfico, se requieren datos de indicadores económicos, como ya se mencionó, que se refieren a posibles variables independientes en el proceso de modelaje.

En varios estudios de economía de transporte se identificó que el Producto Bruto Interno (PBI) está directamente relacionado con el índice de movilidad de las personas en sus vehículos, así como los vehículos de carga. Una explicación de este hecho es que cuanto más crece la economía, o cuanto mejor es el estado económico de la región, la inversión en productos y servicios se vuelve más atractiva y, por esto, mayor es el flujo de vehículos para transportar productos bien como para salir y consumir, estudiar, entre otras actividades.

Los datos de variación del PBI se obtuvieron por las series históricas puestas a disposición por el banco mundial, como se muestra en el Gráfico 1. La información se recopiló con datos de la plataforma del sitio web de Google, que facilita el acceso a la información.



Gráfico 1: Serie histórica de Variación del PBI en Perú - datos públicos del Google

Los valores detallados del PBI en miles de millones de USD y su variación se muestran en la Tabla 2:

Año	PBI Bilhões USD	Variación del PBI
1998	R\$ 55,50	-
1999	R\$ 50,19	-9,57%
2000	R\$ 51,74	3,10%
2001	R\$ 52,03	0,55%
2002	R\$ 54,78	5,28%
2003	R\$ 58,73	7,22%
2004	R\$ 66,77	13,69%
2005	R\$ 76,06	13,92%
2006	R\$ 88,64	16,54%
2007	R\$ 102,17	15,26%
2008	R\$ 120,55	17,99%
2009	R\$ 120,82	0,23%
2010	R\$ 147,53	22,10%
2011	R\$ 171,76	16,43%
2012	R\$ 192,65	12,16%
2013	R\$ 201,22	4,45%
2014	R\$ 201,08	-0,07%
2015	R\$ 189,93	-5,55%
2016	R\$ 191,64	0,90%
2017	R\$ 211,39	10,31%

Con estas dos informaciones principales se ha desarrollado los esfuerzos para hacer una relación y, así, calibrar un modelo de previsión de variación de tráfico con variables económicas en Peru.

## 2.2. Modelaje

Para calibración del modelo, se identificó que el impacto de la economía es sentido en el tráfico dos años después de su variación (debido al tiempo para ocurrir el completo impacto económico). Además, se excluyeron las discrepancias del análisis, o sea, en las cuales el PBI y el tráfico no mostraron una relación adecuada por diversos motivos. Los datos seleccionados para obtener el modelo se muestran en la Tabla 3

Tabla 3: Datos seleccionados para calibración del modelo

Intervalo de PBI	variación do PIB	Intervalo del tráfico	variación del tráfico
1999-2000	3,10%	2001-2002	2,98%
2001-2002	5,28%	2003-2004	3,90%
2002-2003	7,22%	2004-2005	4,82%
2003-2004	13,69%	2005-2006	9,71%
2006-2007	15,26%	2008-2009	9,06%
2007-2008	17,99%	2009-2010	11,16%
2009-2010	22,10%	2011-2012	13,16%
2012-2013	4,45%	2014-2015	5,10%
2013-2014	-0,07%	2015-2016	3,92%
2014-2015	-5,55%	2016-2017	-0,20%

## 2.3. Análisis de regresión lineal y validación estadística

En el proceso de calibración, los volúmenes de tráfico son la variable dependiente, mientras que el PBI tiene el papel de la variable independiente.

Para calibración del modelo se utilizó el método tradicional de regresión lineal con el software SPSS en el proceso de cálculo. Los datos obtenidos fueron dispuestos en una hoja de cálculo electrónica, que se sometieron a análisis de datos.

Para validar el modelo, se evaluaron los siguientes parámetros:

- Grado de correlación de Pearson (r): El coeficiente, también llamado correlación lineal o r de Pearson, es un grado de relación entre dos variables cuantitativas y expresa el grado de correlación a través de valores entre -1 y 1. Un coeficiente de correlación cercano a cero indica que no hay relación entre las dos variables, y cuanto más cerca estén de 1 o -1, más fuerte será la relación.;
- Coeficiente de determinación ( $R^2$ ): El coeficiente de determinación es el cuadrado del coeficiente de correlación de la muestra r. En el contexto de un modelo de regresión lineal simple, donde la variable explicativa (o predictor) es "x" y la variable de respuesta (o por predecir) es "y", el coeficiente de determinación  $R^2$  proporciona el porcentaje de variabilidad de los "y" (variable por predecir) que permanece explicado de acuerdo con la variabilidad de "x". Por lo tanto, un valor de  $R^2 \approx 1$  significa que la nube de puntos que se muestra en el diagrama de dispersión está cerca de la línea de regresión, o sea, cuanto mayor sea  $R^2$ , más explicativo es el modelo, mejor se ajustará a la muestra. Cuando  $R^2 \approx 0$ , una estructura lineal ya no es visible;

- Error estándar de estimación: El error estándar de la estimación (Se) o la desviación estándar cuantifica la dispersión de las observaciones de la muestra en relación con la línea de regresión. Cuanto menor sea la dispersión en relación, más precisas serán las estimaciones;
- Pruebas estadísticas: En el análisis estadístico, pueden ocurrir problemas con la precisión, el nivel de detalle, el dominio de regresión e incluso la precisión de los datos en sí, entre otros, y, por lo tanto, es necesario tener dominio para que se puede usar evidencia estadística para probar la precisión de la ecuación de crecimiento, y entonces, si es necesario, ajuste. Para verificar los datos, se realizaron otras tres pruebas: Error Medio Absoluto, Teste T Student y Erro Estándar de Estimación

### 3. Resultados y Discusión

#### 3.1. Calibración del Modelo

Después de inúmeras interacciones en software SPSS se obtuvo un modelo con precisión adecuada. La aplicación de la técnica de regresión permitió la obtención de la ecuación (1) para pronosticar la tasa de crecimiento del tráfico:

$$\text{Tasa de Crecimiento (decimal)} = 0,4734 \times \text{Variación del PBI (decimal)} + 0,0241 \quad (1)$$

La tasa de crecimiento en términos porcentuales se obtiene multiplicando el resultado encontrado en la ecuación anterior por 100. Como se puede ver en el 2, el valor de  $R^2$  encontrado fue 0.9576, lo que indica una buena relación entre las dos variables. El Gráfico 2 presenta los datos utilizados y el comportamiento de la ecuación de estimación:

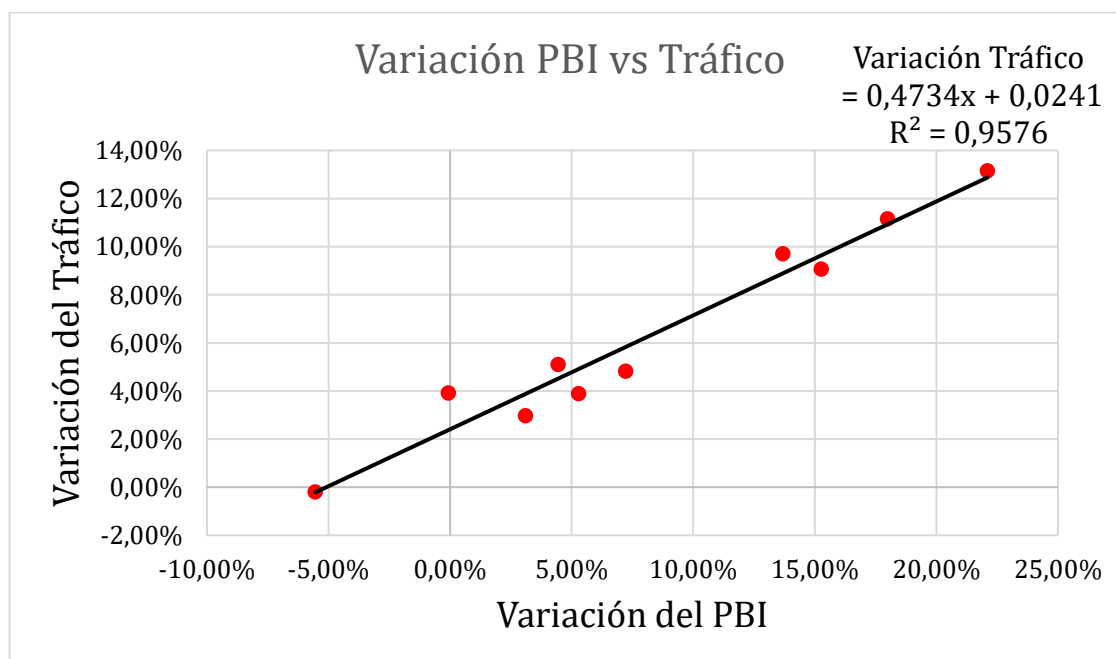


Gráfico 2 : Regresión Linear – PBI y Tráfico

Otra forma de verificar la eficiencia del modelo es aplicarlo a los propios objetos de estudio. La comparación entre lo real y lo estimado se puede ver en el Gráfico 3, que muestra que el modelo estimó variaciones de tráfico más reservadas, donde para valores reales altos, la proyección fue un poco menor. Pero en general, la estimación tiene el mismo comportamiento de los datos reales, lo que indica la aplicabilidad de la ecuación.

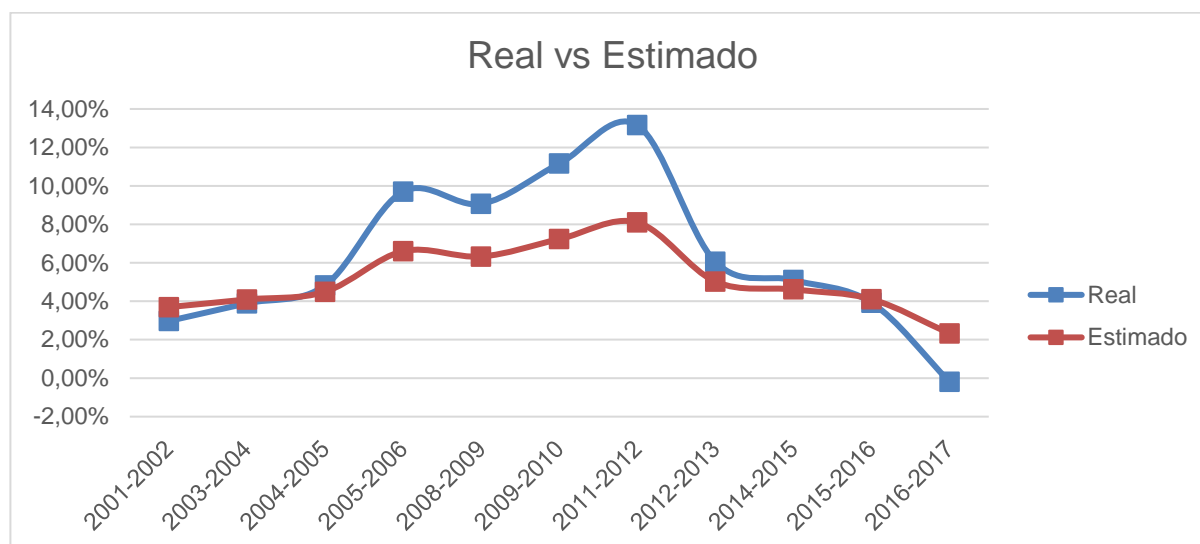


Gráfico 3: Comparación de tráfico real x tráfico estimado

El análisis de datos económicos con variables de flujo de vehículos en las carreteras del Perú, presentó una relación satisfactoria, lo que permitió la calibración de modelos para pronosticar la variación del flujo del tráfico por carreteras en el país.

Adicionalmente al coeficiente de determinación estadística (R<sup>2</sup>), han sido utilizados otros parámetros de verificación: El error medio absoluto (EMA) y el Error estándar de estimación, presentado en los ítems siguientes.

### 3.2. EMA

El EMA, conforme Rodrigues (2017), consiste en calcular el promedio de la diferencia absoluta entre los valores pronosticados por el modelo de regresión y los valores reales conocidos de las variables dependientes. Su ecuación de cálculo se presenta en la ecuación 2:

$$Ema = \sum \frac{\sqrt{(R0-Re)^2}}{n} \tag{2}$$

Donde: Ema es el error medio absoluto;

- R0 es el valor observado;
- Re es el valor estimado por el modelo y;
- n es el número de observaciones.

La

Tabla 4 presenta los valores obtenidos con la diferencia entre lo observado y lo obtenido por los modelos para vehículos ligeros.

Tabla 4: Error absoluto medio

Año	Variación del tráfico	Estimación del modelo	EMA
<b>1999-2000</b>	0,0298	0,0382	0,000843217
<b>2001-2002</b>	0,0390	0,0425	0,000358239
<b>2002-2003</b>	0,0482	0,0469	0,000129607
<b>2003-2004</b>	0,0971	0,0700	0,002701267
<b>2006-2007</b>	0,0906	0,0670	0,002363538
<b>2007-2008</b>	0,1116	0,0769	0,003464471
<b>2009-2010</b>	0,1316	0,0864	0,004519055
<b>2012-2013</b>	0,0510	0,0482	0,0002742
<b>2013-2014</b>	0,0392	0,0427	0,000344706
<b>2014-2015</b>	-0,0020	0,0232	0,002513918
<b>Media =</b>			<b>0,0017512218</b>

Esto demuestra la efectividad de los resultados estimados del modelo, ya que el error obtenido fue pequeño (cerca de 2,75% en relación con los datos reales).

### 3.3. Erro Estándar de Estimación

El error estándar evalúa la precisión del promedio de una comparación de datos. En este caso, se muestra un error estándar para el tráfico y otro para la estimación del tráfico. El cálculo se realiza utilizando la siguiente fórmula:

$$S_x = \frac{S}{\sqrt{N}}$$

Donde,

$S_x$ , es el error estándar

$S$ , es la desviación estándar

$N$ , es o tamaño de muestra

Ha sido obtenido una desviación estándar de 0,041919352 y un error estándar estimado de 0,013256063 para la variación del tráfico. Ya para la estimación del tráfico, ha sido obtenido respectivamente 0,013256063 y 0,00627542. Dado que el intervalo de confianza se centra en la media de la muestra, el error probable máximo que se admite es igual a la mitad de la amplitud del

XXI PANAM 2020 • Universidad del Pacífico, Peru • August 05<sup>th</sup> - 07<sup>th</sup>, 2020  
intervalo. En este caso, la amplitud del intervalo es: 13.36%, por lo que el error está dentro del parámetro estadístico establecido.