

MODELIZACION DE LA FRECUENCIA DEL TRANSPORTE PÚBLICO INTERURBANO EN LA PROVINCIA DE CORRIENTES

CARDOZO, Osvaldo Daniel ¹; INSAURRALDE, Juan Ariel ²

¹ Universidad Nacional del Nordeste (UNNE). Laboratorio de Tecnologías de la Información Geográfica (LabTIG) y Grupo de Estudio en Movilidad, Servicios, Infraestructura y Territorio (GEMSIT). Resistencia, Chaco.

² CONICET - Universidad Nacional del Nordeste (UNNE). Grupo de Estudio en Movilidad, Servicios, Infraestructura y Territorio (GEMSIT). Resistencia, Chaco.

odcardo@hum.unne.edu.ar

RESUMEN

Los desequilibrios espaciales en el acceso al transporte público constituyen un importante desafío a superar para los organismos estatales encargados de su planificación. En este sentido, se propone un Modelo de Regresión Lineal Múltiple (MRLM) para predecir el comportamiento de la frecuencia semanal del transporte interurbano de pasajeros en la Provincia de Corrientes, a partir de un conjunto de variables explicativas referidas a características de la población (demográficas, sociales, económicas), y de la red vial (topológicas, geográficas), cuyo tratamiento y análisis se apoyó en el uso de herramientas estadísticas y los Sistemas de Información Geográfica (SIG). Su aplicación muestra resultados alentadores, revelándose como una alternativa válida en el marco de los llamados modelos flexibles de respuesta rápida, por su sencillez y eficacia.

Palabras Claves: Transporte Público – Modelo Predictivo – Análisis de Regresión – Planificación Territorial – Provincia de Corrientes.

MODELLING OF THE FREQUENCY OF THE INTERURBAN PUBLIC TRANSPORT IN THE PROVINCE OF CORRIENTES

ABSTRACT

The spatial imbalances in the access to public transport are a major challenge to overcome for the state agencies responsible for planning. In this regard, we propose a Multiple Linear Regression Model (MRLM) to predict the behavior of weekly frequency of the interurban transport of passengers in the Province of Corrientes, from a set of explanatory variables concerning population characteristics (demographic, social, economic), and of the road network (topological, geographic), whose processing and analysis was based on the use of statistical tools and Geographic Information Systems (GIS). Its implementation shows encouraging results, proving to be a valid alternative in the context of called flexible models of rapid response for its simplicity and effectiveness.

Key words: Public Transport – Predictive Model – Regression Analysis – Territorial Planning – Province of Corrientes.

Introducción

La planificación del transporte, la evaluación de proyectos de inversión en infraestructuras y, en general, la toma de decisiones en política de transportes, continúa siendo uno de los problemas más importantes en el devenir del Siglo XXI (Ortúzar y Willumsen, 2009). En dicho proceso intervienen numerosas actividades, pero sin lugar a dudas, unas de las más relevantes consisten en analizar el comportamiento de la *oferta* y la *demanda*, ya que en función a ellas se determinará la magnitud en la dotación de servicios e infraestructuras necesarias para el funcionamiento eficiente del sistema de transporte.

El destacado papel de la demanda dentro de los estudios de transporte es evidente al observar el extenso cuerpo teórico-metodológico acumulado hasta el momento, y que permitió explorar, ensayar o validar numerosos procedimientos para su estudio a diferentes escalas geográficas, de tiempo, con distintos niveles de agregación o desagregación, de

variada complejidad, etc. A consecuencia de ello también, se ha impuesto como condición muy valorada la pluralidad de enfoques al momento de abordar su análisis, consolidando los enriquecedores aportes provenientes de diversas disciplinas como las ingenierías, Economía, Geografía, la Ciencia Regional, Planificación, Urbanismo o Sociología, entre otras, lo cual ha permitido una mejora sustancial en la amalgama de conocimientos existentes (Seguí Pons y Martínez, 2003; Cardozo, 2007).

Una de las ideas rectoras más enfatizadas en los clásicos manuales de planificación del transporte (Wingo, 1979; Black, 1981; Bruton, 1985; Ortúzar, 1992; Daganzo, 1997; Banister, 2002) es fomentar el uso del transporte público. Una de las posibles medidas a considerar para atender esta preocupación compartida por parte de investigadores, empresarios del sector y organismos de planificación estatal, es el incremento de la oferta. Dicha acción busca generar alternativas para reducir las inequidades espaciales en la distribución del servicio sobre el territorio.

Esta situación fue observada en el Atlas Geográfico de Corrientes, donde la representación cartográfica de la oferta del servicio de transporte público *permite advertir que los mayores registros de frecuencia se producen fundamentalmente en el Oeste de la Provincia, sobre la Ruta Nacional N° 12 desde la ciudad Capital hasta Goya...*(Manoiloff y otros, 2004: 14), corroborado también por el cálculo de una serie de medidas descriptivas referidas a la accesibilidad y centralidad de la red vial de la Provincia a través de Teoría de Grafos y técnicas SIG (Insaurralde y Cardozo, 2010).

En el marco de las posturas antes mencionadas, se propone como objetivo principal del trabajo generar un modelo que permita estimar la frecuencia semanal del servicio de transporte público interurbano en la Provincia de Corrientes, por medio de un Modelo de Regresión Lineal Múltiple (MRLM). Esto permitirá por un lado, medir la relación lineal entre las variables del modelo y realizar extrapolaciones -guardando las precauciones necesarias- para conocer el su comportamiento con datos actuales, y por otra parte, constituir una herramienta metodológica sencilla -básicamente se trata de una ecuación lineal- que permita a los planificadores realizar proyecciones en forma rápida y bajo costo, pero con resultados fiables.

A partir de lo anterior formulamos una hipótesis de trabajo (H_1) con la cual pretendemos contrastar la idea de que algunas de las 24 variables identificadas sobre aspectos demográficos, socioeconómicos, internos de la red y el territorio, presentan coeficientes con valores distintos a 0, es decir:

$$\begin{aligned} H_0: \beta_i &= 0 \\ H_1: \beta_i &\neq 0 \end{aligned}$$

Expresión 1: Contraste de Hipótesis

No rechazar la hipótesis nula (H_0) en palabras de Montgomery, Peck y Vining (2007) implicaría comprobar que no existe relación lineal alguna entre las variables de nuestro modelo (X) y la variable dependiente (γ), es decir, la frecuencia semanal del servicio interurbano de transporte público.

Vinculada a esta hipótesis, se formulan las siguientes preguntas de investigación:

- ¿Cómo es la capacidad predictiva del modelo?

- ¿Existe algún tipo de dependencia espacial en los errores del modelo estimado?

El área de estudio corresponde a la Provincia de Corrientes. Esta se localiza en el noreste del país, entre los 27° y 30° latitud Sur y entre los 28° y 30° longitud Oeste aproximadamente. Se encuentra delimitada por dos importantes cursos fluviales: los ríos Paraná y Uruguay. El primero de ellos separa la Provincia de la República del Paraguay al Norte, y al Oeste de las provincias de Chaco y Santa Fe; el segundo sirve de límite internacional de nuestro país con las Repúblicas del Brasil y Uruguay al Este. Por último hacia el Sur, el río Guayquiraró lo separa de la Provincia de Entre Ríos (Figura 1).



Figura 1: Ubicación geográfica de la Provincia de Corrientes en el contexto nacional. Fuente: SIG-250 (Instituto Geográfica Nacional).

El trabajo se estructura de la siguiente forma: en el apartado introductorio se plantea el problema abordado y los postulados teóricos que consideramos necesarios, se formulan los objetivos e hipótesis de trabajo; en la sección metodológica se discuten brevemente los últimos avances teóricos, se describen las variables seleccionadas y el tratamiento aplicado a las mismas; en el segmento de resultados se comentan las características más sobresalientes del Modelo de Regresión Lineal Múltiple (MRLM) obtenido y algunos parámetros para su validación empírica; para finalizar en las consideraciones finales se desatacan las bondades del modelo ajustado y se plantean futuras líneas de trabajo a seguir.

Materiales y Métodos

Dentro del conjunto de técnicas empleadas para la modelización y planificación de sistemas de transporte, las herramientas matemático-estadísticas ocupan un lugar preferente. Esto responde por un lado, a su tradicional empleo en la ingeniería del tráfico o la economía del transporte, y por otra parte, hasta el momento no se han encontrado alternativas superadoras. Pese a no ser la panacea, suponen una importante disminución -o tal vez control- de los componentes del error, y su carácter objetivo las convierte en instrumentos fundamentales a la hora de tomar decisiones en el ámbito de la planificación (Ortúzar y Willumsen, 2009).

Una tecnología relativamente nueva pero cuyo desarrollo en el ámbito de los transportes muestra un franco crecimiento, son los Sistemas de Información Geográfica (SIG). En este sentido, Seguí Pons y Martínez, 2003 remarcan que uno de los aportes más interesantes y fecundos de los últimos años es la aplicación de tecnología SIG, tanto para el análisis de redes de transporte, como para su gestión y planificación. De hecho, es común encontrar en casi todos los manuales de SIG actuales un capítulo dedicado al análisis de redes donde abundan los ejemplos con transporte. Pero esto no hubiera sido posible sin la evolución del pensamiento geográfico en los años '60, cuando los primeros estudios cuantitativos se introdujeron desde la Geografía de los Transportes con Ullman, Garrison, Marble, o Kansky, entre otros.

La masiva incorporación de los últimos avances tecnológicos, permite a los SIG y otras tecnologías análogas (GPS, Cartografía, Sensores Remotos) ocupar un lugar privilegiado entre los *Intelligent Transportation System* (ITS). Concretamente, los SIG se han incorporado de lleno a tareas de planificación y gestión, gracias al desarrollo de software híbridos para modelización de sistemas de transporte y SIG. Los esfuerzos más serios provienen de productos como *Visum*, *Cube* y *EMME-2/3*, seguidos por otros con prestaciones más específicas como *Accesion*, *Chronomap* o *TRANUS*. Finalmente, entre los software's SIG se desatan algunos por las herramientas específicas para el tratamiento espacial de datos orientados al transporte, como *TransCAD* y *FlowMap*, o los módulos *Traffic Analyst* y *SANET* para *ArcGIS*; sin embargo la mayoría de ellos -tanto comerciales como gratuitos- poseen herramientas de variada complejidad para análisis de redes (*ArcView/ArcGIS*, *Geomedia*, *GE Smallworld*, *TNT*, *SuperMap*, *SuperGIS*, *Manifold*, *GRASS*, *Spring*, *gvSIG*, *El-Shayal GIS*, *TerraView*, etc). Este importante desarrollo tecnológico consolida lo que algunos llaman los Sistemas de Información Geográfica orientados al Transporte o SIG-T (Miller y Shaw, 2001; Seguí Pons y Martínez, 2003; Rodrigue, 2006).

El incremento de la oferta para fomentar el uso del transporte público, una acción básica y sencilla en principio, debe realizarse sobre bases sólidas y científicas. Esto nos lleva al concepto de *modelización flexible* donde la generación de modelos de respuesta rápida es una de las metas más importantes a lograr (Ortúzar y Willumsen, 2009). Por su parte, una característica sobresaliente de este enfoque, es la posibilidad de formular modelos escalables que se ajusten a las necesidades de cada caso, y adoptando una complejidad progresiva sobre algún procedimiento estadístico. En este marco, una de las alternativas más conocidas es el análisis de regresión.

El empleo de regresiones en la modelística del transporte no es algo nuevo, si tenemos en cuenta que se remontan a la aparición de los primeros modelos de cuatro etapas en los años '50. Tampoco lo son en el análisis espacial del transporte, porque desde la década del '60 estas herramientas estadísticas se utilizan para comprobar relaciones de índices elaborados a partir de Teoría de Grafos o con modelos de interacción espacial, y variables socio-

económicas referidas al desarrollo de un territorio. Algunos ejemplos de aplicación en transporte público cercanos en el tiempo de los autores son Cardozo (2007), Gutiérrez Puebla y otros (2008), o Cardozo, Gutiérrez Puebla y García Palomares (2010), aunque todos ellos referidos a la demanda.

Las fuentes consultadas para obtener la información estadística y cartográfica han sido diversas, entre ellos podemos destacar las cartográficas y estadísticas. Por su parte, los datos fueron procesados y analizados con el apoyo de herramientas estadísticas y de SIG, así como para la representación gráfica y cartográfica de los resultados.

Fuentes de Información

La base cartográfica corresponde al SIG-250 del Instituto Geográfico Nacional (límites, red vial y localidades). La mayor parte de las variables referidas a las características demográficas, económicas y sociales fueron obtenidas del Censos Nacional de Población y Vivienda 2001, tanto en su versión impresa como digital, disponibles en la página web del Instituto Nacional de Estadística y Censos (INDEC). [<http://www.indec.mecon.ar>]

Otro conjunto de variables -que podrían denominarse como *internas* de la red-, se obtuvieron en un reciente estudio (Insaurralde y Cardozo, 2010) por medio de la Teoría de Grafos. Se trata de una serie de medidas descriptivas referidas a la accesibilidad y centralidad de la red vial pavimentada provincial, tales como el Número de Shimbel, Número Asociado, Índice Omega, Índice de Accesibilidad Media, e Índice de Centralidad Media.

En total se reunió un conjunto de 24 variables explicativas para intentar explicar el comportamiento -o *variabilidad* en términos estadísticos- de la frecuencia del servicio, las cuales se presentan a continuación:

1. Distancia a la Capital
2. Logaritmo Natural de Distancia a la Capital
3. Población Total
4. Logaritmo Natural del % de Población Total
5. Población con más de 15 años
6. Población sin instrucción y Educación Primaria Incompleta
7. Población con Educación Primaria Completa y Secundaria Incompleta
8. Población con Educación Secundaria Completo y Terciaria/Universitaria Incompleta
9. Población con Educación Terciaria/Universitaria Completa
10. Población con Obra Social
11. Población sin Obra Social
12. Crecimiento Absoluto de la Población (1991-2001)
13. Crecimiento Relativo de la Población (1991-2001)
14. Terminal de Ómnibus
15. Número Asociado
16. Número de Shimbel (topológico)
17. Número de Shimbel (en distancias reales)
18. Accesibilidad Media
19. Índice Omega
20. Índice Omega (en distancias reales)

21. Centralidad Media
22. Población Empleada en Transporte Terrestre
23. Población Empleada en Actividades Complementarias del Transporte
24. Población Empleada en Comercio al por Mayor, en Comisión y al por Menor

Por su parte, la variable dependiente, denominada Frecuencia del Servicio de Transporte Público Interurbano, se obtuvo de la plancha cartográfica N° 12 del Atlas Geográfica de la Provincia de Corrientes (Manoiloff y otros, 2004) que fue elaborada con datos referidos al número de servicios -en días hábiles y fines de semana- que ofrecen las empresas de transporte de pasajeros en el año 2002, según la Dirección de Transporte y Puertos de la Provincia.

Tratamiento de los datos

Las variables espaciales, de gran importancia para este trabajo, fueron calculadas con herramientas informáticas. Las distancias topológicas fueron estimadas a través del software *Grafos 1.2.9* desarrollado por Rodríguez Villalobos y de libre distribución [<http://personales.upv.es/aroDRIGU/grafos/index.htm>]. Por su parte, las distancias geográficas -expresadas en metros para nuestro caso- se estimaron con el software *ArcGIS 9.x*, concretamente a través de una matriz de origen-destino (*O-D cost matrix*) de la extensión *Network Analyst* que permite medir distancias sobre la red entre todas las localidades.

Por otra parte, una de las condiciones básicas para incluir la variable en el modelo de regresión múltiple es que los datos tengan una distribución normal. Por esta razón, se realizó un análisis gráfico de normalidad, y como resultado se detectó que algunas de ellas (población total, distancia a la Capital) no cumplían esta condición (Figura 2), por lo que fue necesario normalizarlas a partir del cálculo de sus *Logaritmos Naturales*.

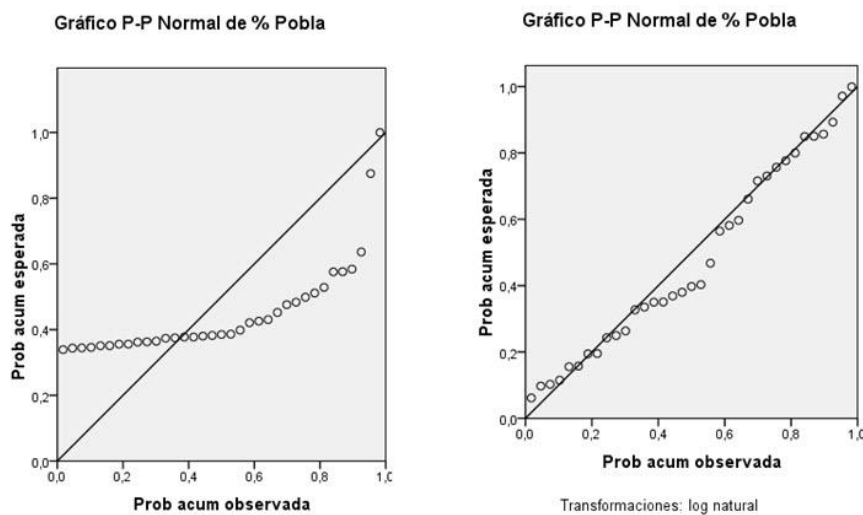


Figura 2: % de Población antes y después de ser transformada por su Logaritmo Natural.

Este tipo de análisis (normalidad en la distribución) otorga una primera aproximación al proceso de selección de las variables: las que no muestran un comportamiento normal,

deberán ser transformadas o excluidas del modelo final. Para avanzar en dicho proceso, también se realizó un análisis de correlación entre todas las variables buscando que, las explicativas estén correlacionadas con la frecuencia semanal, y al mismo tiempo, tengan la menor correlación posible entre ellas.

A tal efecto, se calculó la matriz de correlación bi-variadas utilizando el coeficiente de Pearson y una significación al 95 %. También se realizó el análisis gráfico de la nube de puntos, "par a par" entre cada una de las variables explicativas y la frecuencia semanal, con el objeto de obtener los R^2 y de esta manera, conocer en forma individual el poder explicativo de cada variable. A partir de la matriz de correlaciones y los valores de R^2 , es posible comenzar a definir el conjunto de variables seleccionadas para su inclusión en el modelo de regresión lineal múltiple.

El Modelo de Regresión Múltiple

Por medio del análisis de regresión es posible expresar las complejas relaciones entre variables de manera sencilla y matemáticamente precisa (Estébanez y Bradshaw, 1979). En concreto, el Modelo de Regresión Lineal Múltiple (MRLM) tiene por objeto explicar el comportamiento de una variable dependiente designada γ , utilizando información suministrada por los valores de un conjunto de variables explicativas $X_1, X_2 \dots X_n$. En general los modelos lineales de regresión múltiple se expresan bajo la siguiente forma:

$$\gamma = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots + \beta_n \cdot X_n + \varepsilon$$

Expresión 2: Ecuación de un Modelo de Regresión Lineal Múltiple

Donde los coeficientes $\beta_1 \dots \beta_n$ muestran la magnitud del efecto que las variables explicativas $X_1 \dots X_n$ tienen sobre la variable explicada γ . El coeficiente β_0 es el término constante del modelo, y ε representa al término de perturbación o error existente en el modelo.

La presencia de *multicolinealidad* se verificó por medio del Índice de Tolerancia, el valor F y el Factor de Inflación de la Varianza (FIV), mientras que, la *autocorrelación* se contrastó con el estadístico de Durbin-Watson (DW). Para el primer caso, la Tolerancia es una medida del grado de asociación lineal entre las variables explicativas medida en una escala de 0 (máxima) a 1 (nula), por lo tanto, interesan valores altos de tolerancia, mientras que e FIV es el recíproco del anterior, y sus valores indican la varianza del coeficiente de regresión. Según Kleinbaum (citado en Pérez López, 2005: 98) existe una regla empírica según la cual, valores de FIV mayores a 10 o de Tolerancia menores a 0,1 indican problemas de colinealidad, es decir, correlación entre las variables explicativas del modelo.

Por su parte, el estadístico F representa el incremento producido en la explicación tras la incorporación de una variable, respecto al total alcanzado por las ya existentes. Permite determinar la significación conjunta de los coeficientes -parámetros- estimados, por lo tanto, valores altos son considerados favorables ya que implican alta significatividad conjunta del modelo.

En el segundo caso, para detectar indicios de autocorrelación serial se adoptó contraste de Durbin-Watson, conforme a la siguiente regla general expuesta en Pérez López (2005):

- DW \cong 0 indica autocorrelación perfecta positiva
- DW \cong 2 no hay signos de autocorrelación
- DW \cong 4 indica autocorrelación perfecta negativa

Así mismo, una vez aplicada la ecuación de ajuste y obtenidos los errores en las estimaciones, se efectuó un contraste de la autocorrelación espacial por medio del test *I de Moran*, disponible en las herramientas de modelamiento de patrones en *ArcGIS 9x*.

Resultados

Del conjunto de variables explicativas analizadas -unas 24-, finalmente fueron seleccionadas 3 para el modelo de regresión estimado. Las variables incluidas son: el % Población Total, el Ln de las Distancias a la Capital, y el Índice de Shimbel (en distancias reales).

Las características principales del modelo se muestran en la Tabla 1. En cuanto a la calidad del ajuste puede considerarse como buena, ya que, pese al escaso número de variables su R^2 indica que el modelo logra explicar el 79% de la variabilidad de la frecuencia semanal. Para evitar el problema que supone incluir variables que no sean significativas, se calcula el R^2 Ajustado, que también mide el porcentaje de variación de la variable dependiente pero teniendo en cuenta el número de variables incluidas en el modelo. Es un buen indicador que las diferencias entre los R^2 no sean grandes, como en nuestro caso apenas un 2 %.

R	R^2	R^2 Ajust.	Error típico	Durbin-Watson	F	Sig. (95%)
0,889	0,790	0,770	87,741	2,4	38,962	,000

Tabla 1: Resumen del Modelo Estimado

Los coeficientes o parámetros $b_1, b_2... b_n$ muestran el efecto que las variables explicativas $X_1, X_2... X_n$ tienen sobre la variable frecuencia semanal. Debido a que los coeficientes estimados no son comparables entre ellos, se calculan los *Coefficientes Estandarizados* (β), para conocer la magnitud relativa de la explicación que aporta cada variable al modelo, independientemente del signo (+ o -) que posea; valores altos indican mayor contribución de las variables en el modelo.

Ello cobra importancia cuando se analizan las elasticidades del modelo. En nuestro caso, los coeficientes estandarizados son bastante similares, sin embargo el más alto corresponde al % de Población Total (0,487), lo que representa una explicación de casi el 50 % de las variaciones en la frecuencia de transporte público, y el más bajo, al índice de Shimbel en distancias reales (-0,323).

En cuanto a la significación de las variables incluidas en el modelo, se utiliza el cálculo del *p-value*. Se buscan valores bajos en el test, debido a que, cuanto más pequeño sea mayor

será la evidencia para rechazar la hipótesis nula. En nuestro caso, para un nivel de confianza del 95 % (*p-values* inferiores a 0,05) la significación global del modelo está asegurada, mientras que todas las variables resultan significativas porque ninguna supera el umbral establecido (Tabla N° 2).

Por último, el estadístico *t* contribuye a decidir la inclusión de una variable en el modelo de regresión, al informar si el efecto de esa variable sobre la dependiente es significativo desde el punto de vista inferencial (Bosque y Moreno, 1994). En general, se considera que si son mayores a 2 o menores a -2, la variable ejerce una influencia significativa, por lo que es recomendable incluirla en la ecuación de la regresión; en nuestro modelo la única variable con valores positivos se hallan por encima de 2, y las dos variables con valores negativos también están por debajo del umbral señalado.

	Coeficientes			t	Sig.	Estadísticos de Colinealidad		
	no estandarizados		tipificados			Tolerancia	Índice de Condición	FIV
	b	Error típico	β					
(Constante)	802,249	159,815		5,020	,000		1,000	
Poblac. Total	14,332	4,356	,487	3,290	,003	,309	1,875	3,234
LogNat Dist.	-27,914	12,849	-,328	-2,172	,038	,297	14,032	3,366
Shimbel (dist. Real)	-,025	,007	-,323	-3,786	,001	,929	24,593	1,076

Tabla 2: Coeficientes del Modelo y Estadísticos de Colinealidad

Con los resultados expuestos en la Tabla 2 donde se comprueban que los coeficientes estimados son \neq de 0, existe razón suficiente para rechazar la hipótesis nula (H_0). A partir de los mismos escribimos la siguiente ecuación (Expresión 3) que ajusta al modelo de regresión:

$$\gamma = 802,249 \text{ (Constante)} + 14,332 * \text{(Poblac. Total)} + -27,914 * \text{(Ln Distancia a Capital)} - ,025 * \text{(Índice de Shimbel en dist. real)}$$

Expresión 3: Ecuación Lineal del Modelo de Regresión Múltiple y coeficientes estimados

Validación del Modelo

Una vez construido el modelo y aceptado su nivel de explicación Pérez López (2005) señala que, deben verificarse algunos supuestos teóricos propios del análisis de regresión, a través de la observación de los residuos. Al respecto Seguí Pons y Petrus Bey (1991) justifican su estudio también al afirmar que, puede conducir a valorar causas externas a nuestras variables explicativas, que influyen de manera efectiva en el comportamiento del fenómeno geográfico analizado.

La verificación de su distribución normal se realiza representando en un histograma los residuos tipificados y los valores pronosticados tipificados (Figura 3); en dicho histograma la distribución debe ajustar una curva normal, lo cual constata la ausencia de valores atípicos. También se pueden representar los datos en un gráfico de probabilidades (Figura 3), donde la

detección valores que se alejen mucho de la diagonal es indicio de que pueda existir correlación. Los resultados obtenidos del modelo muestran claras evidencias de la ausencia de este problema, lo cual nos conduce a aceptar la hipótesis de linealidad y homocedasticidad.

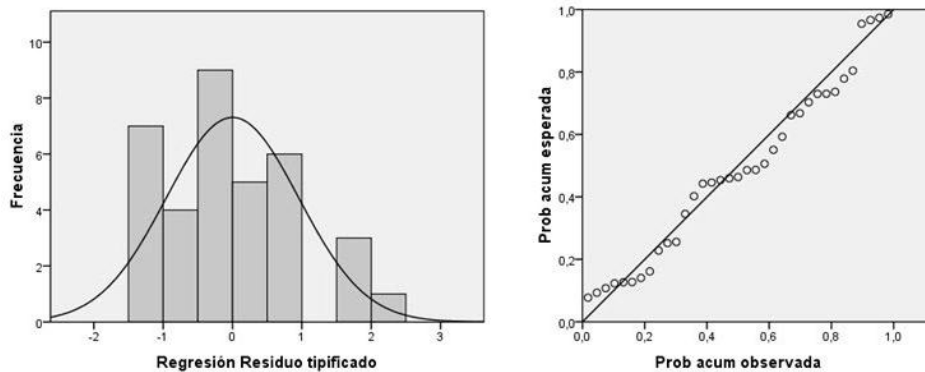


Figura 3: Distribución Normal (Histograma y Diagrama de Dispersión) de los residuos tipificados estimados

Pero a fin de descartar la presencia de autocorrelación, se aplicó el contraste de Durbin-Watson. En nuestro caso, la prueba arroja un valor de 2,4 lo cual permite rechazar la hipótesis de autocorrelación.

Para analizar con efectividad la cuestión anterior, se calcularon los Estadísticos de Colinealidad del modelo (Tabla 2), tales como el Índice de Tolerancia, FIV, valor F y el Índice de Condición. En nuestro caso, todas las variables se hallan dentro de los parámetros establecidos.

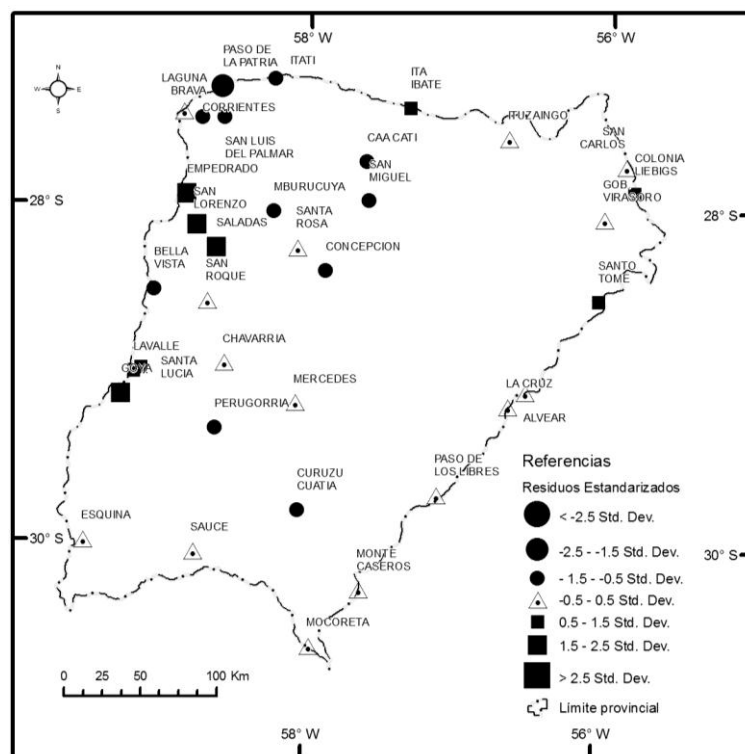


Figura 4: Cartografía de los residuos del Modelo de Regresión Lineal Múltiple. Fuentes: SIG-250 (Instituto Geográfica Nacional) y Censo Nacional de Población y Viviendas 2001 (INDEC).

Para finalizar se elaboró la correspondiente cartografía de residuos (Figura 4), para lo cual, fueron reescalados por sus desvíos estándar para hacerlos comparables, aunque manteniendo los signos: los errores positivos corresponden a sobreestimaciones y los negativos a las subestimaciones. La impresión visual del mapa puede indicarnos la presencia de alguna tendencia en la distribución geográfica de los errores, es por ello que se estimó también, el índice *I de Moran*. El resultado del test arroja un valor de $-0,02$ con lo cual se despejan las dudas sobre la presencia de autocorrelación espacial en los residuos del modelo.

Consideraciones Finales

La propuesta de modelos flexibles puede representar la solución para muchas administraciones públicas que no pueden invertir grandes sumas de dinero en complejos modelos. Con herramientas de este tipo pueden realizarse estimaciones fiables y con la suficiente rapidez al momento de tomar decisiones en el ámbito de la planificación de este sector.

El Modelo de Regresión Lineal Múltiple (MRLM) obtenido es estable desde el punto de vista estadístico, lo cual puede comprobarse con los resultados negativos que arrojan los test de diagnóstico de colinealidad. Por su parte, la prueba de Durbin-Watson indica el rechazo de la hipótesis de autocorrelación serial, y el test *I de Moran* corrobora ausencia de autocorrelación espacial en los errores de las estimaciones.

Se considera como aceptable la respuesta del conjunto de variables explicativas seleccionadas para alimentar el modelo, ya que con la intervención de solo tres variables su nivel de explicación llega al 77 %, aunque se considera oportuna su revisión previa a cualquier aplicación sobre la realidad. Por su parte, los valores de β indican que la variable con mayor contribución explicativa es distancia a la Capital.

De todas formas en estudios posteriores podría explorarse la capacidad explicativa de variables de otro tipo, o de mayor complejidad, de cara a mejorar la capacidad explicativa del modelo de regresión. Sería importante también aplicar métodos de análisis locales sobre los residuos (LISA) o evaluar otros modelos espaciales tipo *Geographically Weighted Regression* de Fotheringham, así como avanzar en el análisis de la variable demanda, al estar íntimamente ligada a la frecuencia del servicio.

Agradecimientos

Queremos expresar nuestra gratitud a los Prof. Erica Gómez y Alejandro Parras del Departamento de Geografía de la Universidad Nacional del Nordeste, por los comentarios y sugerencias aportados.

Referencias

- Bosque Sendra, J. y Moreno Jiménez, A. (1994). *Prácticas de Análisis Exploratorio y Multivariante de Datos*. Oikos-Tau Ediciones, Barcelona.
- Cardozo, O.D. (2007). *Modelos de Demanda Potencial en la Red del Metro de Madrid*. Departamento de Geografía. Universidad de Alcalá (UAH), Alcalá de Henares. [Inédito].
- Cardozo, O.D; Gutiérrez Puebla, J. y García Palomares, J.C. (2010). Influencia de la Morfología Urbana en la Demanda de Transporte Público. Análisis mediante SIG y Modelos de Regresión Múltiple. *GeoFocus*, no 10, 82-102.
- Estébanez Álvarez, J. y Bradshaw, R. (1979). *Técnicas de Cuantificación en Geografía*. Ed. Tebar Flores, Madrid.
- Gutiérrez Puebla, J. y otros. (2008). Modelo de Estimación Directa de la Demanda de Metro a Nivel de Estación: un enfoque alternativo basado en Regresión Múltiple, Funciones de Caída de la Demanda y Herramientas SIG. XVº Congreso Panamericano de Ingeniería de Tránsito y Transporte 2008. Universidad del Norte - Grupo de Investigación TRANVIA. Cartagena de Indias, URL: <http://www.uninorte.edu.co/panam2008/CDPANAM/index.htm> acceso 05/08/10
- Insaurralde, J.A. y Cardozo, O.D. (2010). Análisis de la Red Vial de la Provincia de Corrientes con Teoría de Grafos y Sistemas de Información Geográfica. *Geográfica Digital*, 7, no 13.
- Manoiloff, R.O; Rey, W; Alberto, J.A; Cardozo, O.D; Snaider, P.P; González, C.D. (2004). *Atlas Geográfico de la Provincia de Corrientes: Tomo V - Las Actividades Secundarias y Terciarias*. Instituto de Geografía. Universidad Nacional del Nordeste (UNNE). Resistencia.
- Miller, H.J. y Shaw S.L. (2001). *Geographic Information Systems for Transportation (GIS-T): Principles and Applications*. Oxford University Press, New York.
- Montgomery, D; Peck, E. y Vining, G. (2007). *Introducción al Análisis de Regresión Lineal*. Ed. Patria, México.
- Ortúzar, J.D; Willumsen, L.G. (2009). *Modelos de Transporte*. Ed. Universidad de Cantabria, Madrid.
- Pérez López, C. (2005). *Métodos Estadísticos Avanzados con SPSS*. Ed. Thompson, Madrid.
- Rodrigue, J.P; Comtois, C. y Slack, B. (2009). *The Geography of Transport Systems*, Hofstra University, Department of Global Studies & Geography, URL: <http://people.hofstra.edu/geotrans> acceso 05/08/10
- Seguí Pons, J.M. y Martínez, M.R. (2003). Pluralidad de métodos y renovación conceptual en la Geografía de los Transportes del Siglo XXI. *Scripta Nova*, VII, no 139. URL: <http://www.ub.es/geocrit/sn/sn-139.htm> acceso 05/08/10
- Seguí Pons, J.M y Petrus Bey, J.M. (1991) *Geografía de Redes y Sistemas de Transporte*. Edit. Síntesis, Madrid.