

Evaluación y ajuste de las proyecciones de áreas menores: el caso de la provincia de Buenos Aires. .

María Silvia Tomas, Malena Monteverde y
María Virginia Pérez.

Cita:

María Silvia Tomas, Malena Monteverde y María Virginia Pérez (2015). *Evaluación y ajuste de las proyecciones de áreas menores: el caso de la provincia de Buenos Aires. XIII Jornadas Argentinas de Estudios de Población. Asociación de Estudios de Población de la Argentina, Salta.*

Dirección estable: <https://www.aacademica.org/xiiijornadasaepa/42>

ARK: <https://n2t.net/ark:/13683/e7Bo/pbe>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

**XII Jornadas Argentinas de Estudios de Población,
Salta, 16-18 de septiembre de 2015
Evaluación y ajuste de las proyecciones de áreas menores:
el caso de la provincia de Buenos Aires**

María Silvia Tomás, Dirección Provincial de Estadística de la provincia de Buenos Aires, silviat06@yahoo.com.ar

Malena Monteverde, CIECS-CONICET, UNC, montemale@yahoo.com

María Virginia Pérez, Dirección Provincial de Estadística de la provincia de Buenos Aires, mvirginiaper@gmail.com

Resumen

El presente trabajo propone una metodología para: 1- identificar el grado de ajuste global de las proyecciones de población de las áreas menores (partidos de la provincia de Buenos Aires) y en los partidos cuyas proyecciones de población mediante la aplicación del método de Incrementos Relativos (IR) presenten importantes desvíos respecto al crecimiento esperable en base a sus tendencias demográficas; 2- identificar las posibles causas de los desvíos y 3- proponer ajustes a las estimaciones basadas en IR de forma tal de poder reducir los mismos. Se busca establecer objetivamente si los patrones de crecimiento resultantes del método de IR se corresponden con las tendencias de los componentes demográficos, y en función de los hallazgos establecer hipótesis apropiadas para corregir las tendencias que no están relacionadas a la variación real observada.

El ajuste del IR se evaluó mediante el Análisis de Regresión Lineal, relacionando la tasa de crecimiento intercensal obtenida por IR con una tasa construida con la variación intercensal de los componentes que intervienen en el crecimiento de la población entre dos momentos: la tasa de crecimiento natural y la tasa de variación de los saldos migratorios. En esta evaluación se detectaron divergencias.

Para establecer hipótesis sobre las causas de las divergencias, se utiliza el Análisis de las Componentes Principales (ACP), usando aquellas variables estructurales que dan indicio del potencial de crecimiento o decrecimiento de la población como también otras variables menos directamente asociadas.

Detectadas las posibles causas, se proponen correcciones por fuera del método de IR y se comprueba que éstas mejoren el ajuste de los resultados del modelo.

Esta técnica permitió mejorar las proyecciones, en tanto que sirvió para formular hipótesis acerca de las causas por las cuales el modelo de IR arroja resultados que no responden al comportamiento natural de los componentes demográficos.

I. OBJETIVOS

La población de las áreas menores constituye un insumo fundamental para el diseño, implementación y evaluación de las políticas públicas; en particular cuando se considera a la gestión en su dimensión territorial es un requisito mínimo conocer el número de personas (Cabrera, 2012, pp 11).

El objetivo del presente trabajo es proponer una técnica que permita, por un lado, identificar las áreas menores (partidos en el caso de la provincia de Buenos Aires) para las cuales la tasa de variación intercensal resultante a partir de la aplicación del método denominado de los Incrementos Relativos (IR) no responda a las causas de variación intrínsecas de una población, es decir crecimiento natural y saldo migratorio, y, por otro lado, identificar los partidos cuyas proyecciones presenten importantes desvíos respecto al crecimiento real para luego proponer ajustes.

Se busca establecer objetivamente si los patrones de crecimiento resultantes del método de IR se corresponden con las tendencias de los componentes demográficos y en función de los hallazgos establecer hipótesis adecuadas que permitan establecer las correcciones necesarias en la población proyectada.

Adicionalmente, y a efectos de realizar el Análisis de Regresión Lineal mencionado, se ensaya un modelo lineal dónde la tasa de crecimiento total de la población se obtiene de la suma de las tasas de crecimiento natural y migratorio correspondientes.

II. METODOLOGÍA

II.a Estructura general de la metodología empleada

Para la detección de los desvíos de las proyecciones de población por el método de IR se utilizan herramientas de análisis estadístico tales como el Análisis de Regresión Lineal Simple y Análisis Multivariante, específicamente el Análisis de los Componentes Principales (ACP). Los ajustes se realizan en la medida que las herramientas empleadas muestren divergencias entre lo esperado y lo proyectado.

El Análisis de Regresión Lineal se usa para evaluar el ajuste de las tasas proyectadas con IR a una tasa construida con la variación intercensal de los componentes que intervienen en el crecimiento de la población entre dos momentos: la tasa de crecimiento natural y la tasa de variación de los saldos migratorios.

De surgir divergencias, se emplea el ACP para determinar sus posibles causas. En un primer paso del análisis se hacen intervenir todas las variables estructurales que son representativas del potencial de crecimiento o decrecimiento de la población, y en un segundo paso se introducen otras variables que no se relacionan directamente con el crecimiento poblacional y sin embargo podrían influir en él.

Detectadas las posibles causas, se proponen correcciones por fuera del método de IR y se verifica que éstas mejoren el ajuste de los resultados del modelo a través del Análisis de Regresión Lineal Simple.

Esta técnica permite mejorar las proyecciones, en tanto que resulta útil para formular hipótesis acerca de las causas por las cuales el modelo de IR arroja resultados que no responden al comportamiento natural de los componentes demográficos.

II.b. Tasa de variación anual proyectada mediante el método de IR (TV_IR)

Existen diversas técnicas para la elaboración de proyecciones de población, procedimientos que son cada vez más complejos gracias al desarrollo de la informática que ha posibilitado el procesamiento de grandes volúmenes de datos y facilitado su análisis.

Para las proyecciones globales, es decir, las que solo contemplan los grandes volúmenes sin desagregaciones (geográficas, por sexo y edad, etc.), se usan procedimientos demográficos basados en la ecuación compensadora (en el nivel provincial).

En la ecuación compensadora intervienen todos los componentes demográficos responsables de la variación de la población en un período de tiempo. La misma viene dada por:

$$N^{t+n} = N^t + B^z - D^z + I^z - E^z$$

Donde:

N^{t+n} es la población en t+n,

N^t es la población en t,

B^z son los nacimientos ocurridos durante z,

D^z son las defunciones ocurridas durante z ,

I^z son los inmigrantes que arribaron durante z , y

E^z son los emigrantes que salieron durante z ,

Siendo z el período transcurrido entre t y $t+n$.

En las áreas menores se utilizan modelos semidemográficos que tienen en cuenta la variación intercensal dado que no se dispone de todos los elementos necesarios para realizar proyecciones estrictamente demográficas.

El método de los incrementos relativos o de participación en el crecimiento se fundamenta en la participación del crecimiento absoluto del área mayor (provincia en el caso que se describe) que le ha correspondido a cada área menor (partido en el caso que se describe) en un determinado período de referencia (INDEC (2015 cita de Madeira y Simoes, 1972, p.8-9)). Este método ha permitido obtener resultados razonables en tanto cada partido o municipio mantenga los patrones de crecimiento observados durante el período intercensal anterior. En general es posible sostener esta hipótesis en el corto y mediano plazo, en tanto los componentes del crecimiento intercensal estén asociados a los componentes demográficos, es decir, crecimiento natural y saldo migratorio.

Esta técnica parte de la consideración de un área mayor cuya población estimada en un momento t es N^t . El área mayor se subdivide en " n " áreas menores. La población de una determinada área " i ", en un tiempo " t " se expresa como:

$$N_i^t \text{ donde } : i = 1, 2, 3, \dots, n$$

Luego:

$$N^t = \sum_i^n N_i^t$$

La población del área " i " se descompone en dos términos: $a_i N^t$, que depende del crecimiento de la población del área mayor, y b_i

$$N_i^t = a_i N^t + b_i$$

El coeficiente a_i se denomina coeficiente de proporcionalidad del crecimiento del área menor " i ", en relación al incremento de población del área mayor, y b_i es el coeficiente lineal de correlación. Para la determinación de estos coeficientes se utiliza el

período delimitado por dos censos. Siendo t_0 y t_1 las fechas de los dos censos respectivamente, al reemplazar t_0 y t_1 en la ecuación previa se tiene que:

$$N_i^{t_0} = a_i N_i^{s_0} + b_i$$

$$N_i^{t_1} = a_i N_i^{s_1} + b_i$$

Resolviendo el sistema resulta:

$$a_i = (N_i^{t_1} - N_i^{t_0}) \div (N_i^{s_1} - N_i^{s_0})$$

$$b_i = [(N_i^{t_1} + N_i^{t_0}) + a_i(N_i^{s_1} - N_i^{s_0})] \div 2$$

Para la suma todas las áreas menores:

$$\sum a_i = 1 \quad \sum b_i = 0$$

Así si el área mayor se proyectó siguiendo una tasa media r^t (resultado de la aplicación del método de los componentes demográficos), cada área menor tendrá una tasa

r_i^t , derivada de la aplicación de la técnica de IR.

Esta tasa puede calcularse como:

$$r_i^t = [(\ln r_i^{t+5} \div \ln r_i^t) \div 5]$$

La variación de la población del área menor entre dos censos consecutivos es causada por el comportamiento de los componentes demográficos (vegetativo y migratorio) entonces la tasa derivada a partir de la aplicación de IR, r_i^t , no debería diferir demasiado de la verdadera tasa de crecimiento del área menor.

Los problemas en la tasa derivada de la aplicación del IR se producirían cuando la tasa observada del área menor no guarda relación con el comportamiento de los componentes demográficos de la misma en el pasado, lo que sugiere que existen interferencias de otro tipo, por ejemplo, problemas en las mediciones.

II.c. Tasa de variación anual calculada a partir de las estimaciones de los componentes demográficos (TV_CD)

Partiendo de la expresión de de la ecuación compensadora se obtiene que la tasa de crecimiento total es la suma algebraica de las tasas de crecimiento natural y migratorio correspondientes a un determinado período.

La tasa de crecimiento natural: es el ritmo al que la población aumenta (o disminuye) durante un período dado, debido al balance entre nacimientos y defunciones. Esta tasa no incluye los efectos de la inmigración o la emigración.

La tasa de crecimiento migratorio: es el cociente entre el saldo neto migratorio anual (inmigrantes - emigrantes) ocurrido durante un período determinado, generalmente un año calendario, y la población media del mismo período. Refleja la incidencia del saldo migratorio en la población media de dicho período.

Luego la tasa de crecimiento total viene dada por:

$$r^z = (b^z - d^z) + (i^z - e^z)$$

Donde;

r^z es la tasa anual media de crecimiento del período,

$b_i^z = (B_i^z / N_i) * 1000$ es la tasa de natalidad del período,

$d_i^z = D_i^z / N_i * 1000$ es la tasa de mortalidad del período,

$i_i^z = (I_i^z / N_i^z) * 1000$ es la tasa de inmigración del período, y

$e_i^z = (E_i^z / N_i^z) * 1000$ es la tasa de emigración del período

Considerando que la variación de la población en un período es consecuencia de la acción de las variables demográficas, lo que ya se ha descrito en la ecuación compensadora, la tasa de variación intercensal a nivel municipio debería ser el resultado de la suma de las tasas de crecimiento natural y de crecimiento migratorio a nivel partido, es decir:

$$r_i^{t,z} = (b_i^z - d_i^z) + (i_i^z - e_i^z)$$

Donde:

$r_i^{t,z}$ es la tasa resultante de la aplicación del modelo,

b_i^z es la tasa de natalidad a nivel partido,

d_i^z es la tasa de mortalidad a nivel partido, y

$(i_i^z - e_i^z)$ es la tasa de crecimiento migratorio a nivel partido,

Siendo z el período transcurrido entre t y t+n.

Si bien no se disponen de todos los insumos necesarios para la estimación directa de todos los componentes demográficos a nivel de áreas menores (partidos), los mismos se calculan y aproximan con estadísticas vitales, datos provistos por el censo y

un método de aproximación de los saldos migratorios de movimientos residenciales a nivel de municipios.

Para estimar la tasa de crecimiento vegetativo (cv_i^z) se dispone de tasas brutas de natalidad y mortalidad a nivel partido de fuente oficial, siendo la misma:

$$cv_i^z = b_i^z - d_i^z$$

Donde:

cv_i^z es la tasa de crecimiento vegetativo en z

b_i^z es la tasa de natalidad en z

d_i^z es la tasa de mortalidad en z

Siendo z el período transcurrido entre t y t+n.

En el caso de las tasas de saldo migratorio el problema es más complejo. A nivel municipio, además de encontrar saldos internacionales e internos, existen movimientos entre municipios para los cuales no existen mediciones oficiales. Como variable proxy se utiliza la variación de stocks de migrantes internacionales entre ambos censos, bajo el supuesto de que la tasa de variación intercensal se mantendrá constante en el período, de modo que:

$$i_i = \left[\sqrt[z]{I_t/I_{t_0}} - 1 \right] * 1000$$

i_i es la tasa de variación media anual del stock de inmigrantes internacionales a nivel partido,

I_{t_0} es el stock de inmigrantes internacionales en t0, y

I_t es el stock de inmigrantes internacionales en t.

Para los inmigrantes internos,

$$m_i = \left[\sqrt[z]{M_t/M_{t_0}} - 1 \right] * 1000$$

m_i es la tasa de variación media anual del stock de inmigrantes internos a nivel partido,

M_t es el stock de inmigrantes internos a nivel partido en t0, y

M_{t_0} es el stock de inmigrantes internos a nivel partido en t.

Y donde:

$$s_i = \left[\sqrt[z]{S_t/S_{t_0}} - 1 \right] * 1000$$

s_i es el doble de la tasa de variación media anual del stock de personas que residían en otro municipio 5 años antes a nivel partido,

S_i es el stock de personas que residían en otro municipio 5 años antes a nivel partido en t_0 , y

S_{i0} es el stock de personas que residían en otro municipio 5 años antes a nivel partido en t .

Cabe destacar que la construcción de la tasa de variación media anual de migrantes internos y de la de movimientos residenciales no incluye al stock de población correspondiente que habita en viviendas colectivas dado que esa pregunta no estaba incluida en el cuestionario correspondiente.

Luego la ecuación quedará conformada de la siguiente manera:

$$r_i^z = (b_i^z - d_i^z) + i_i^z + m_i^z + s_i^z$$

Dado que las variaciones de estos stocks pueden ser muy abruptas y sin embargo el volumen poblacional afectado muy pequeño, se procede a ajustar las tasas ponderándolas por el peso relativo de cada componente a nivel partido, de modo que la tasa ponderada resulta:

$$r_i^{z'} = (b_i^z - d_i^z) + (I_i/N_i) * i_i^z + (M_i/N_i) * m_i^z + (S_i/N_i) * s_i^z \quad (1)$$

II.d. Modelo de regresión para evaluar la asociación entre las tasas: TV_IR y TV_CD.

La variación de la población entre dos censos es consecuencia del aporte de los componentes demográficos, luego la tasa del área menor resultante del Modelo IR, r_i , y del área menor por componentes demográficos aproximada según el procedimiento anterior, r_i^z , deberían ser muy similares, siempre y cuando en el pasado no se hayan registrado movimientos poblacionales atípicos que no estén relacionados directamente con las tendencias demográficas

Tomando a r_i obtenida a partir de la aplicación de IR como variable dependiente y a r_i^z la tasa agregada (1) se realiza un análisis de regresión.

Se aplica el modelo a la provincia de Buenos Aires. La Provincia está dividida en forma exhaustiva en 134 partidos que constituyen a la vez áreas de gobierno local o municipios. Es el estado provincial más extenso, teniendo en cuenta el territorio continental, y el más poblado del país. Las 134 jurisdicciones que lo componen presentan una alta heterogeneidad tanto en lo que refiere a su superficie como a su población.

Esta gran disparidad también se refleja en las tendencias demográficas. Para que esta disparidad no afecte el análisis la provincia de Buenos Aires se divide en las dos áreas que se utilizan habitualmente en casi todas las regionalizaciones: 24 Partidos de GBA e Interior. Además, en la elaboración de las proyecciones por departamento, si bien el IR se aplica a los 134 partidos en conjunto, los ajustes se realizan teniendo en cuenta esta subdivisión.

A los fines de este trabajo sólo se realiza el análisis de los partidos de Interior de la Provincia para respetar la extensión requerida.

II.e. Método de Componentes Principales para Evaluar las posibles divergencias entre la TV_IR y la TC_CD

Para analizar las posibles causas de la divergencia entre las tasas de variación de la población por partidos se aplica el método de componentes principales. Este método permite investigar cuáles serían los factores que están interfiriendo en la relación anterior, ya que permite visualizar en el plano ortogonal de los CP partidos cuyo crecimiento / decrecimiento proyectado es consistente con las variables demográficas asociadas a esas tendencias como así también identificar aquellos partidos con crecimiento / decrecimiento inconsistente con las variables demográficas asociadas.

El Análisis de los Componentes Principales (ACP) es una técnica del análisis factorial que tiene por objeto reducir las dimensiones a analizar. Son métodos multivariantes de interdependencia, en el sentido que todas las variables tienen una importancia equivalente ya que ninguna se destaca como principal para el objetivo de la investigación.

En el caso que nos ocupa contamos con una amplia gama de variables observadas $X_1, X_2, X_3, \dots, X_p$ y se trata de encontrar dimensiones comunes o factores que ligan a las aparentemente no relacionadas variables, un conjunto k de factores, con $k < p$ que expliquen a las variables observadas perdiendo el mínimo de información, de modo que sean fácilmente interpretables (principio de la interpretabilidad) y donde k sea pequeño (principio de la parsimonia). Además los factores se extraen de modo tal que son independientes entre sí, es decir ortogonales. Con el análisis factorial es posible conocer la estructura subyacente de los datos a partir del análisis de la interrelación entre ellos (Valederrey Sanz, 2010).

Entre los múltiples usos del ACP, en este trabajo se emplea como técnica exploratoria que permite descubrir interrelaciones entre los datos. Es así como es

posible encontrar patrones ocultos de asociación entre las variables, difíciles de observar de otro modo. Además, la incorporación al análisis de criterios de ordenación permite encontrar indicios sólidos sobre las causas de la asociación.

En una primera etapa del ACP se consideran una gran cantidad de variables tanto relacionadas con el potencial de crecimiento de la población como de otro tipo. El contar con las bases del Censo Nacional de Población, Hogares y Viviendas 2001 y 2010 desagregadas geográficamente a nivel de las áreas menores (partidos) abrió la posibilidad de agregar al ACP la mayor parte de las variables disponibles y seleccionar aquellas que resulten más eficaces para explicar las causas del crecimiento / decrecimiento experimentado en el período.

En un primer procesamiento se ingresaron al análisis un gran número de variables:

- Relación Ancianos / Niños
- % de mujeres en edad fértil
- Relación niños 0-4 años / mujeres en edad fértil
- Índice de Envejecimiento
- % de migrantes 2010 (Inmigrantes internacionales / Población total)
- R/U (Relación rural / urbano)
- % de población urbana
- Promedio de hijos por mujer
- Densidad relativa
- Variación de la densidad en el período
- IDP jóvenes 2010 (Población de 0 a 14 años / Población de 15 a 64 años)
- IDP Ancianos 2010 (Población de 65 años y más / Población de 15 a 64 años)
- TEP (Tasa de Escolarización Primaria)
- TEM (Tasa de Escolarización Media)
- TES (Tasa de Escolarización Superior)
- etc.

El ACP permite seleccionar un subconjunto de variables sobre la base de aquellas que tienen la más alta correlación con el CP. En este proceso es posible desechar a las consideradas redundantes por tener el mismo nivel de correlación y el mismo sentido que otra de similares características. Por ejemplo entre niños de 0-4 años

/ mujeres en edad fértil e IDP jóvenes se optó por este último por tener una correlación más alta con el primer CP.

Interesa estudiar a la variable r_i (resultante de IR) por lo cual no la incluimos en el análisis pero se va a utilizar como variable de ordenamiento, así quedarán determinados 4 grupos:

Grupo 1: Tasa de crecimiento (r_i) menor a 3 por mil, corresponde al primer cuartil, el 25% de los partidos con menores tasas de crecimiento

Grupo 2: Tasa de crecimiento (r_i) de 3 a 6,1 por mil, 25% con crecimiento medio bajo

Grupo 3: Tasa de crecimiento (r_i) de 6,2 a 11,1 por mil, 25% con crecimiento medio alto

Grupo 4: Tasa de crecimiento (r_i) mayor a 11,1 por mil, 25% con crecimiento alto

Tras sucesivos ensayos se seleccionaron las variables que presentan la más alta correlación con el primer y segundo componente principal. Esto permitió acotar el estudio a las variables que describen mejor los patrones de crecimiento esperado de las áreas menores según estos CP.

III. FUENTES

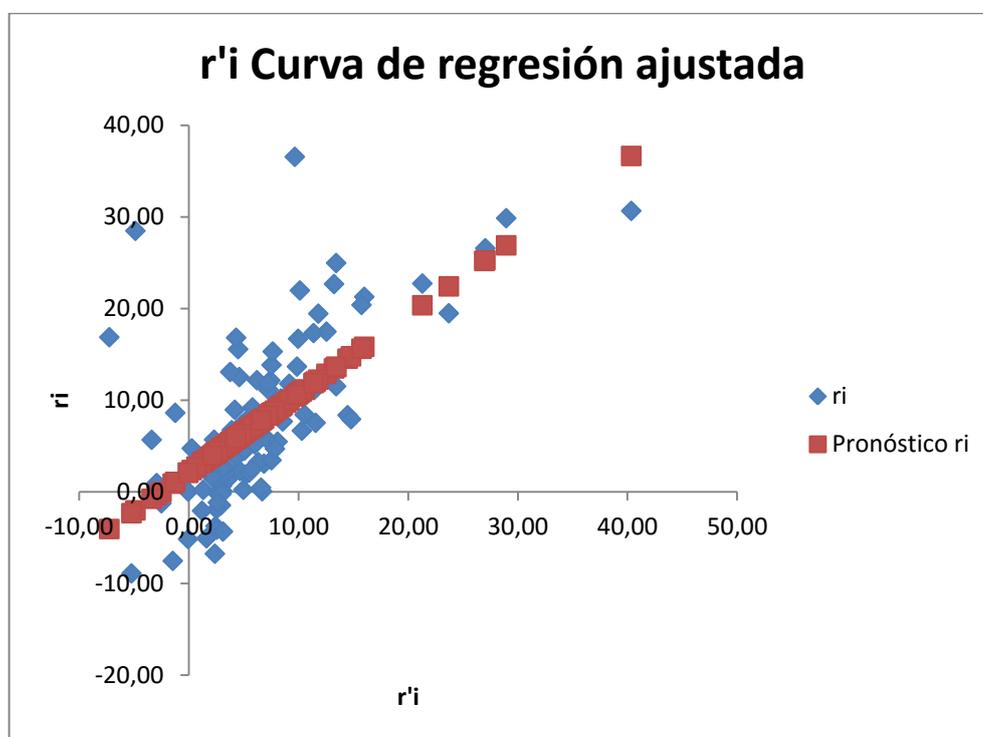
Las fuentes de información utilizadas son:

- a) Los Censos Nacionales de Población, Hogares y Viviendas para los años 2001 y 2010, a partir de los cuales se obtienen las características demográficas básicas de la población en cada momento, así como también se estiman los stocks de grupos poblacionales específicos como inmigrantes, población en viviendas colectivas, u otros de interés.
- b) Las series de nacimientos y defunciones publicados por el Ministerio de Salud de la provincia de Buenos Aires para el mismo período (2001 a 2010), a partir de las que se obtiene el crecimiento vegetativo en el período intercensal en estudio.
- c) Las proyecciones de población por sexo de los municipios de la Provincia, de donde se obtienen las tasas de crecimiento intercensal proyectadas.

IV. RESULTADOS

A continuación se presentan los resultados del modelo de regresión que estima la asociación entre la TV_IR y la TV_CD en el período 2010- 2015 para los 110 partidos del Interior de la Provincia.

Gráfico 1. Modelo de regresión simple entre la Tasa de Variación de la Población por Partidos del método de IR (TV_IR) y la Tasa de Variación del método de Componentes Demográficos (TV_CD) Período 2010-2015. Partidos del Interior.



Fuente: Elaboración propia en base a INDEC (2005, 2013, 2015) MS (s.f)

Cuadro 1.

<i>Estadísticas de la regresión</i>	
Coefficiente de determinación R ²	0,45154557
R ² ajustado	0,44646728
Error típico	6,51510993
Observaciones	110

De los resultados obtenidos, el coeficiente de determinación R² de 0,45 estaría indicando que menos del 50% de la variabilidad de la tasa de crecimiento de la

población resultante obtenida mediante IR es explicada por la suma de las tasas relacionadas con los componentes demográficos.

Para analizar las posibles causas de la divergencia entre las tasas de variación de la población por partidos (de IR y CD) se aplica el método de componentes principales según lo explicado en el apartado de la metodología. Las variables que tienen una correlación más alta con los dos primeros CP (ejes 1 y 2) son:

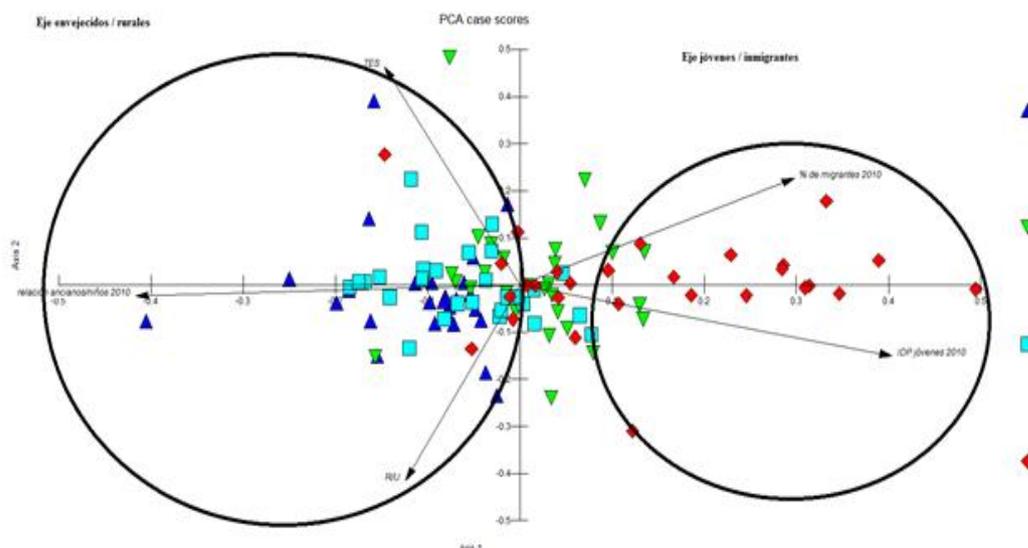
- Relación Ancianos / Niños
- % de migrantes 2010 (Inmigrantes internacionales / Población total)
- R/U (Relación rural / urbano)
- IDP jóvenes 2010 (Población de 0 a 14 años / Población de 15 a 64 años)
- TES (Tasa de Escolarización Superior)

En el gráfico se evidencia que, en general, los partidos con:

- 1- Alto crecimiento de población de acuerdo a la TV_IR tienden a ubicarse sobre el sector derecho (Eje 2) que agrupa a los parámetros que implican alto potencial de crecimiento natural: porcentaje de inmigrantes, e IDP jóvenes. Municipios como Presidente Perón, Villarino, General Rodríguez, Exaltación de la Cruz, entre otros, presentan características compatibles con altos potenciales de crecimiento natural.
- 2- Por encima del Eje 2 están municipios como Pinamar, Villa Gesell, Escobar y Pilar que también se caracterizan por tener un considerable crecimiento migratorio en los últimos períodos censales.
- 3- En cambio, los partidos con bajo crecimiento según la TV_IR se desplazan hacia el sector izquierdo del Eje 2, caracterizado por bajo potencial de crecimiento (alta relación ancianos /niños). Puan, General Arenales, Coronel Dorrego, y Alberti presentan características estructurales envejecidas y se destacan por mostrar un crecimiento lento.
- 4- En el Eje 1, hacia arriba, se agrupan los elementos que presentan valores altos de la TES como Tandil, La Plata, Bahía Blanca, y General Pueyrredón, que además son los partidos más poblados del Interior de la Provincia y los más urbanizados. Hacia abajo del Eje 1

se encuentran los municipios más rurales: Tordillo, General Pinto, Guaminí, etc.

Gráfico 2. Dispersión de los partidos en el plano factorial de los dos primeros CP del análisis con primer set de variables seleccionadas. Período 2010-2015. Partidos del Interior



Fuente: Elaboración propia en base a INDEC (2005, 2013, 2015)

En el cuadro 2 se muestra el grado de asociación entre las variables. El porcentaje de migrantes presenta grados de asociación bastante parejos con las otras variables. Entre el IDP jóvenes y la relación ancianos niños hay obviamente una alta asociación de signo opuesto.

Cuadro 2. Matriz de similitud

	% de migrantes 2010	IDP jóvenes 2010	relación ancianos/niños 2010	R/U	TES
% de migrantes 2010	1				
IDP jóvenes 2010	0,374	1			
relación ancianos/niños 2010	-0,307	-0,803	1		
R/U	0,281	-0,113	0,176	1	
TES	-0,316	-0,414	0,136	-0,227	1

En lo que respecta a los ejes se observa que el Eje 1 acumula el 44,6% de la varianza de las variables y entre los dos primeros se explica más del 70% de la variabilidad del modelo.

Cuadro 3. Autovalores

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
Eigenvalues	2,229	1,377	0,728	0,521	0,146
Percentage	44,572	27,533	14,565	10,418	2,912
Cum. Percentage	44,572	72,105	86,67	97,088	100

En el Cuadro 4 se detallan la correlación de cada variable con todos los ejes.

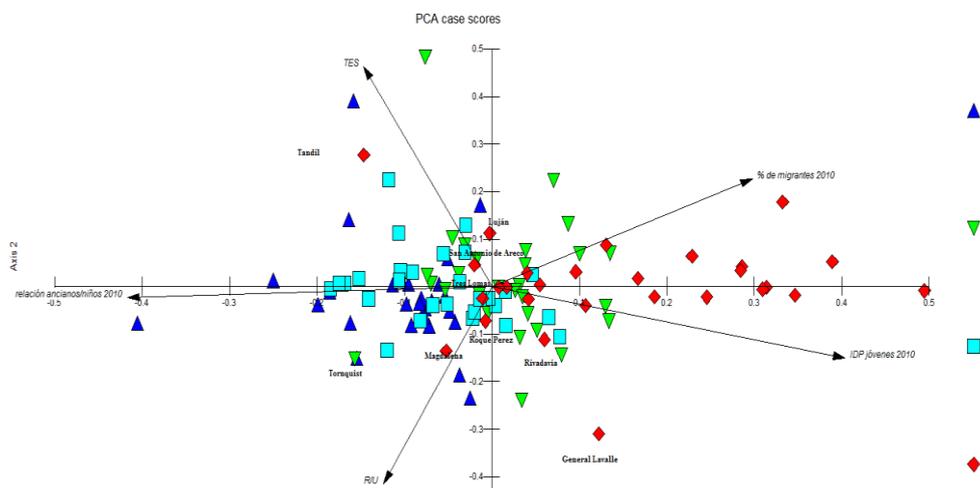
Cuadro 4. PCA variable loadings

	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5
A-% de migrantes 2010	0,428	0,364	0,505	0,655	0,009
B-IDP jóvenes 2010	0,611	-0,213	-0,084	-0,206	-0,729
C-relación ancianos/niños 2010	-0,544	0,364	-0,234	0,343	-0,632
D-R/U	0,035	0,737	0,252	-0,625	-0,038
E-TES	-0,383	-0,384	0,787	-0,141	-0,259

En el Gráfico 2 es posible detectar casos atípicos. Hay partidos clasificados como de alto crecimiento (color rojo en el gráfico) que se localizan en los ejes que identifican a los partidos de crecimiento potencial bajo, es decir, envejecidos, rurales, etc.

Tal es el caso de partidos como Magdalena, Roque Pérez, Tres Lomas, San Antonio de Areco, y Tornquist, ubicados en la zona central del gráfico pero del lado izquierdo, sector de los partidos con bajo potencial de crecimiento.

Gráfico 3. Dispersión de los partidos en el plano factorial de los dos primeros CP del análisis con primer set de variables seleccionadas. Identificación de casos atípicos. Período 2010-2015. Partidos del Interior



Fuente: Elaboración propia en base a INDEC (2005, 2013, 2015)

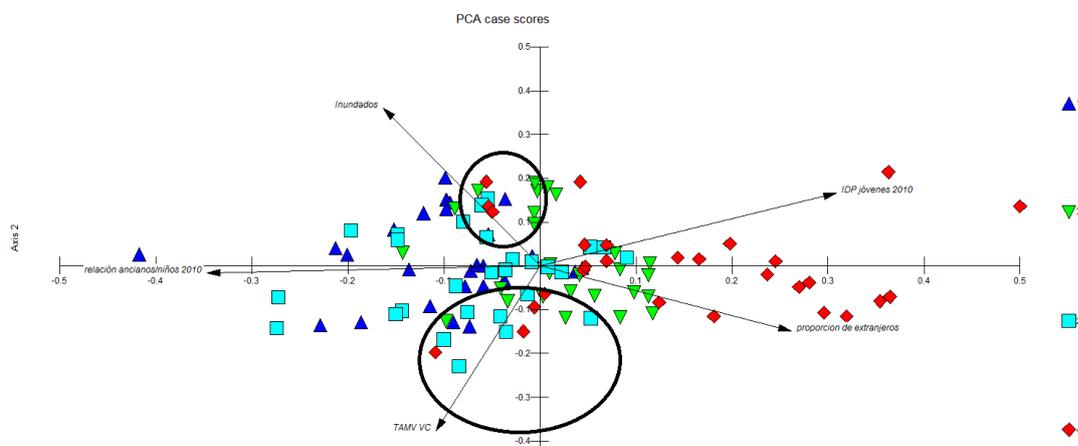
Es evidente que las variables incluidas en el análisis no alcanzan a explicar el crecimiento experimentado en el período en algunos partidos.

A continuación se incluyen en ACP otras variables considerando por ejemplo los partidos en los que se instalaron unidades penitenciarias en el período intercensal, que por tanto aumentaron en un momento específico el stock poblacional; los partidos que estuvieron afectados por las inundaciones en el año 2001, que por ende tuvieron un déficit de cobertura censal en ese año; etc.

Para esto se incluyeron en el análisis las variables:

- TAMV VC (tasa anual media de variación de la población en viviendas colectivas).
- Inundados (partidos que sufrieron inundaciones en 2001 o no).

Gráfico 4. Dispersión de los partidos en el plano factorial de los dos primeros CP del análisis con segundo set de variables seleccionadas. Período 2010-2015. Partidos del Interior



Fuente: Elaboración propia en base a INDEC (2005, 2013, 2015)

Como puede verse en el gráfico, la inclusión de las variables “Inundados” y TAMV VC”, modifica el agrupamiento y los partidos afectados por estos fenómenos se corren a los cuadrantes dominados por estas variables. Ahora en el Eje 1, hacia abajo, se encuentran los partidos que abrieron una unidad penitenciaria o incrementaron su población penal en los años anteriores al censo 2010. Magdalena, Saavedra, Benito Juárez, y Tornquist tuvieron crecimientos derivados de esta situación. Hacia arriba los partidos que estuvieron afectados por inundaciones en el año 2001, por lo cual hubo un déficit en la cobertura de ese año. Allí se encuentran los partidos de Tres Lomas, Salliqueló, Roque Pérez, Rivadavia, Hipólito Irigoyen, Pellegrini, entre otros.

Estos resultados son útiles tanto al momento de realizar las proyecciones con IR sin sobreestimaciones o subestimaciones, o bien para corregirlas si no se detectaron estos desvíos oportunamente.

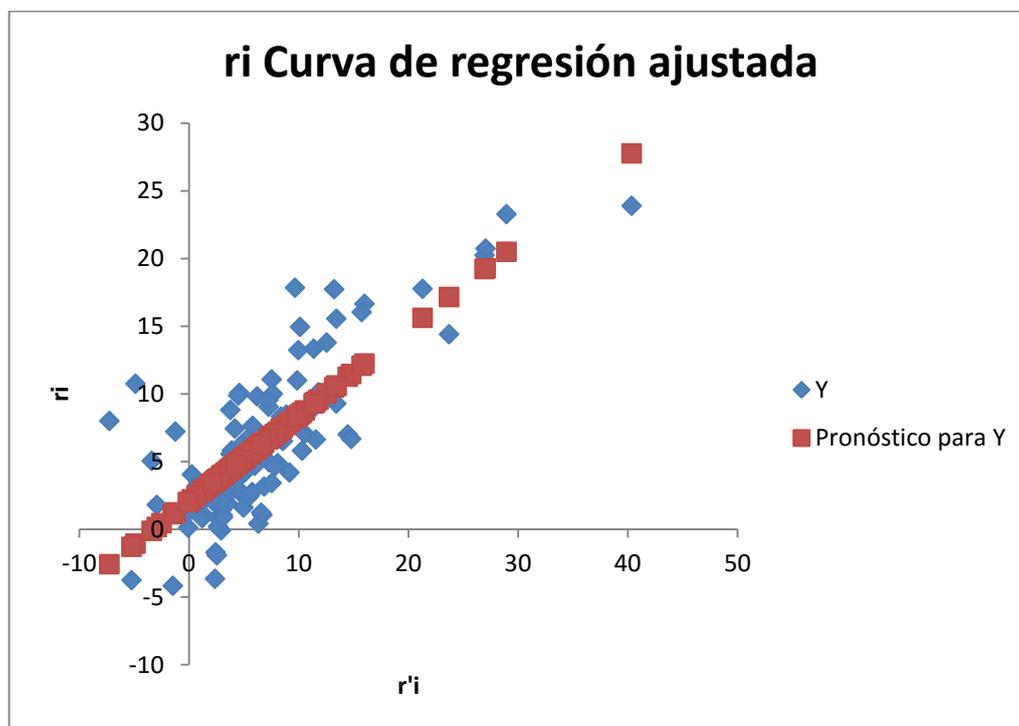
En el caso de la provincia de Buenos Aires estudiado, el análisis permitió encontrar las herramientas adecuadas para realizar correcciones en los siguientes casos:

- 1- Partidos en los que se abrieron unidades penitenciarias entre 2001 y 2010. En estos casos la población penal es un stock fijo, por lo tanto se procede a restarla de la población base para las proyecciones en el año 2010. Una vez proyectada la población se suma a la misma el stock de población penal.
- 2- Partidos afectados por inundaciones en alguna de las mediciones. En estos casos se corrige la población base para las proyecciones del año 2001. Dicha

corrección consiste en aumentar la población base en un valor estimado en función del déficit de cobertura detectado y luego se procede a proyectar con IR.

Se repite el análisis de regresión lineal incorporando la tasa proveniente de IR (r_i) corregida, y se observan mejoras en el ajuste de la recta de regresión y en el coeficiente de determinación R^2 , el cual pasa de 0,45 a 0,61, indicando que las variables incluidas explican parte de la variación en la tasa de crecimiento poblacional.

Gráfico 5 Modelo de regresión simple entre la Tasa de Variación de la Población por Partidos del método de IR (TV_IR) corregida y la Tasa de Variación del método de Componentes Demográficos (TV_CD) Período 2010-2015. Partidos del Interior



Fuente: Elaboración propia en base a INDEC (2005, 2013, 2015) MS (s.f)

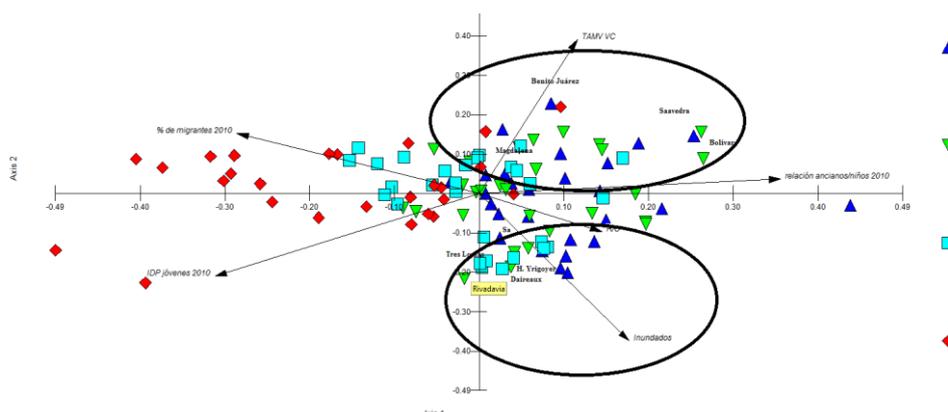
Cuadro 2.

<i>Estadísticas de la regresión</i>	
Coefficiente de determinación R ²	0,6147473
R ² ajustado	0,61118015
Error típico	3,48901075
Observaciones	110

Fuente: Elaboración propia en base a INDEC (2005, 2013)

En un nuevo procesamiento con ACP se observa que con la corrección los partidos afectados por los fenómenos detectados se agrupan ahora conforme a sus características poblacionales estructurales y dinámicas.

Gráfico 6. Dispersión de los partidos en el plano factorial de los dos primeros CP del análisis con segundo set de variables seleccionadas. Identificación de casos atípicos. Período 2010-2015. Partidos del Interior



Fuente: Elaboración propia en base a INDEC (2005, 2013, 2015)

V. CONSIDERACIONES FINALES Y DISCUSIÓN

Es común realizar proyecciones de población para áreas menores usando modelos matemáticos o semidemográficos. En general estas deben ajustarse a un área mayor que se proyecta usando técnicas demográficas, es decir a partir de la proyección de los componentes del crecimiento demográfico. A diferencia de las áreas mayores, en las menores no se disponen de los datos necesarios para proyectar las tasas, por lo que se utilizan las tendencias pasadas para distribuir el crecimiento del área mayor.

Las técnicas semidemográficas de elaboración de proyecciones de población brindan soluciones eficientes en el corto y mediano plazo, en la medida que las causas

del crecimiento de la población de las áreas menores no hayan sido afectadas por factores exógenos a los demográficos.

Existen muchos motivos por los cuales la población en general puede sufrir cambios no relacionados con el comportamiento de los componentes demográficos en el pasado, esto es, el comportamiento de los nacimientos, defunciones y saldos migratorios pasados. Esta metodología surgió a partir de la necesidad de evaluar las proyecciones de población por partido realizadas por otro organismo. En Argentina desde hace 30 años, estas se producen al finalizar cada Censo Nacional de Población. En una provincia como la de Buenos Aires, que cuenta con 134 partidos, es necesario disponer de herramientas objetivas para hacer esta evaluación.

Uno de los casos detectados en este trabajo es por ejemplo la instalación de una unidad penitenciaria que en el período intercensal produjo un aumento en el volumen poblacional en un partido. Si bien este último representa un aumento objetivo, por sus características (capacidad del penal limitada) dicho crecimiento no seguirá a futuro un patrón acorde con las tendencias demográficas. Cuando el área menor tiene un volumen poblacional reducido, su impacto puede ser muy grande y el resultado estimado desproporcionado con respecto a la situación real. Además, dado que este tiene que ajustarse a un área mayor cuyo crecimiento fue obtenido por otra metodología, la forma de mantener tasas mayores es absorbiendo el crecimiento de otras áreas que tuvieron un crecimiento menor, lo que altera su participación relativa en la población total.

El ACP es útil para ensayar la inclusión de otras variables que posibilitan el ajuste del modelo para tener en cuenta al realizar las proyecciones por IR o bien corregirlas posteriormente. Conociendo los partidos en los cuales se dan anomalías entre el crecimiento observado y el potencial es más fácil indagar puntualmente sobre hechos ocurridos durante el período intercensal o en los momentos censales que puedan afectar el crecimiento de la población y/o su medición. Una vez determinados esos factores se pueden proponer correcciones en el momento de efectuar las proyecciones por el método de IR de acuerdo a las interferencias detectadas.

Falta indagar acerca de las posibilidades de la metodología evaluar escenarios futuros a partir de añadir otras variables, por ejemplo variables sintomáticas, que puedan alertar sobre variaciones en períodos intercensales.

Bibliografía

Cabrera, Mariana (2012) *Estimación de Población en áreas menores con métodos que utilizan variables sintomáticas*. Informe final diciembre 2011. Comisión Nacional de Población. República Oriental del Uruguay. Recuperado de

http://www.unfpa.org.uy/userfiles/publications/68_file1.pdf

INDEC (2005) *Censo Nacional de Población, Hogares y Viviendas 2001*. CEPAL / CELADE- Redatam + SP (base de datos)

INDEC (2013) *Censo Nacional de Población, Hogares y Viviendas 2010*. CEPAL / CELADE- Redatam + SP (base de datos)

INDEC (2015) *Estimaciones y proyecciones de población por sexo, departamento y año calendario 2010-2025*. Serie Análisis Demográfico N° 38. ISBN 978-950-896-447-2. Recuperado de

http://www.indec.gov.ar/nivel4_default.asp?id_tema_1=2&id_tema_2=24&id_tema_3=119

Ministerio de Salud de la provincia de Buenos Aires (s.f) Serie de nacimientos y defunciones período 2000- 2010. Recuperado de <http://www.ms.gba.gov.ar/estadisticas/>

Valderrey Sanz, Pablo (2010) *SPSS 17 Extracción de conocimientos a partir del análisis de datos*. Alfa-omega- RA-MA. ISBN 978-607-707-027-6 México.