

EL USO DE VARIABLES SINTOMÁTICAS EN LA ESTIMACIÓN DE LA POBLACIÓN DE ÁREAS MENORES

Guiomar Bay

Centro Latinoamericano y Caribeño
de Demografía (CELADE)
División de Población

RESUMEN

En este trabajo se presentan cuatro métodos de estimación de la población de áreas menores, con modelos que tienen por base los registros administrativos existentes, y se analiza además el uso de tales métodos en la región.

Los métodos seleccionados fueron: el de distribución por prorrato, el de distribución proporcional, de correlación de razón, y el de correlación de diferencia. Se hace una breve descripción de cada uno y se comenta su posterior aplicación en Chile y Costa Rica.

En las dos aplicaciones se utilizó información del período anterior al último censo y se obtuvieron estimaciones de las áreas consideradas para el momento censal, con lo que fue posible evaluar los modelos considerados. También se enumeran los obstáculos encontrados en la investigación, principalmente en lo que se refiere a la información básica, y se detalla la forma en que se soslayaron tales obstáculos. Estos problemas son similares a los que podrían presentarse en otros países de la región al aplicar esta metodología.

ESTIMACIONES DE POBLACIÓN)
(METODOLOGÍA)

(MODELOS)
(DATOS CENSALES)

**USE OF SYMPTOMATIC VARIABLES
FOR ESTIMATING POPULATION
IN SMALL AREAS**

ABSTRACT

This study presents four methods for estimating population in small areas, using models based on existing administrative records. It also analyses the use of such methods in the region.

Four methods were selected: pro rata distribution, proportional distribution, ratio correlation, and difference correlation. Each one is briefly described and its subsequent use in Chile and Costa Rica is discussed.

In both countries, data from the period before the latest census were used, and estimates were obtained for the areas covered at the time of the census; this enabled the models in use to be evaluated. The study also lists the obstacles encountered during the project, mostly regarding the basic data, and the ways in which the obstacles were overcome are described. These problems are similar to those which may arise if the same methods are applied in other countries of the region.

(POPULATION ESTIMATES)
(METHODOLOGY)

(MODELS)
(CENSUS DATA)

INTRODUCCIÓN

Las actuales estrategias de desarrollo de los países latinoamericanos demandan cada vez más información desagregada geográficamente. Asimismo, las políticas y programas destinados a atender las necesidades básicas de la población dan mayor responsabilidad a los gobiernos locales, lo que conduce a una mayor demanda de datos a ese nivel.

El CELADE tiene como una de sus preocupaciones el desarrollo de métodos que permitan actualizar las cifras de población de áreas pequeñas. Actualmente, las estimaciones de población para tales áreas se realizan básicamente por medio de extrapolaciones matemáticas de niveles y tendencias observados en un pasado cercano. No obstante, estos procedimientos, en la medida en que se alejan del período base, suelen arrojar resultados que distan de la realidad, a causa de los cambios que experimentan las tendencias y niveles de las variables demográficas, sobre todo por efecto de la migración en períodos cortos. Por tanto, las metodologías para determinar el número de habitantes basadas en tendencias pasadas tienen limitaciones, a menos que se puedan evaluar y actualizar periódicamente.

Como una alternativa útil para determinar el número de habitantes en áreas pequeñas, y para evaluar los resultados, se propone la aplicación de métodos basados en indicadores indirectos del tamaño de la población, mediante la utilización de modelos de regresión, cuya variable dependiente es la población que se desea estimar y cuyas variables independientes son, por ejemplo, los nacimientos, las defunciones, o la matrícula escolar (véanse en particular Crosetti y Schmitt (1954); Grier y Schmitt (1966); Basavarajappa y Verma (1982); Teixeira Jardim (1992), y Long (1993)).

Metodologías de este tipo han sido utilizadas en Brasil, en Canadá y en los Estados Unidos. En el presente trabajo se estudiarán los métodos que pueden aplicarse o adaptarse a la realidad latinoamericana, considerando especialmente las limitaciones de información.

Otro aspecto que debe tomarse en cuenta para la utilización de estos métodos es su fácil aplicación, ya que permiten, por una parte, actualizar la información sobre la población de áreas pequeñas y, por

otra, detectar cambios en la tendencia del movimiento y crecimiento intercensal de la población. Asimismo, cabe resaltar que también podrían obtenerse estimaciones de población para nuevas divisiones administrativas, aunque éstas no se consideraran en la elaboración de los modelos.

Producto de investigaciones bibliográficas y de una evaluación de la información disponible en Chile, este trabajo se circunscribe a describir cuatro métodos –unos sencillos y otros más sofisticados– que pueden aplicarse, dependiendo de la calidad y cantidad de la información disponible. Estos métodos requieren registros periódicos de variables tales como nacimientos, defunciones, inscripción de automóviles, impuestos, construcción de viviendas, o matrícula escolar. En la mayoría de los países de América Latina y el Caribe no hay registros de estas variables, y cuando existen suelen presentar problemas a causa de omisión diferencial por área, interrupción del registro, alteraciones frecuentes en la división político-administrativa, cambio de sistema económico, aplicación de programas para suprimir déficit (vivienda, educación, salud, servicios básicos), o simplemente porque no son tabulados. Sin embargo, con pequeñas modificaciones en el procesamiento de estos registros, se puede obtener la información necesaria para aplicar los métodos en cuestión.

La aplicación se hará, en principio, para estimar la población de las comunas existentes en Chile en 1992 y de algunos cantones de Costa Rica en 1984 (años del último censo); en el caso de Chile, los resultados serán posteriormente evaluados y comparados con proyecciones anteriores.

I. MÉTODOS UTILIZADOS

1. Información básica

El primer paso consistió en obtener la mayor cantidad de información, tratando de lograr el máximo posible de referencias acerca de las variables sintomáticas.¹

¹ Las variables sintomáticas están relacionadas con los cambios en el tamaño de la población. Entre ellas figuran elementos tales como: nacimientos, defunciones, permisos para construir o viviendas construidas, impuesto sobre el valor agregado (IVA) recaudado, registro de automóviles o licencias de conducir, registro electoral, matrícula escolar (por edad y nivel), valor de los depósitos bancarios, tasa de ocupación, mano de obra agrícola, afiliados a sistemas de seguridad social, consultas médicas, superficie sembrada o plantada, consumo de energía eléctrica. Los nacimientos y las defunciones son considerados aquí como variables sintomáticas, o sea, como indicadores del tamaño de la población y no como componentes de la dinámica poblacional.

A continuación fue necesario averiguar la disponibilidad periódica de cada una de esas variables sintomáticas, su calidad diferencial por área, las posibles interrupciones o la creación de nuevos registros, los cambios estacionales importantes, los incentivos especiales a ciertos registros, las modificaciones en las leyes o en la definición del hecho registrado, y otros aspectos semejantes.

Cabe destacar la enorme importancia que tiene el contacto directo con la información, su manipulación previa, el análisis crítico de los sistemas de recolección, o sea, el lograr una profunda familiarización con la información que se utilizará. Ello permitirá detectar, por ejemplo, posibles errores de estimación, atribuibles, entre otros, a problemas de la calidad de información diferencial por área o a la compatibilización entre períodos. Sin duda alguna, buena parte del tiempo y el esfuerzo empleados en la investigación debe dedicarse a la evaluación exhaustiva de la información, pues de ello dependerá la adecuada elección del modelo.

Luego de los pasos anteriormente descritos, se completó lo que se podría llamar el inventario de la información disponible, con lo cual se obtuvo una lista de las variables disponibles por área y períodos y también la población de las mismas áreas en los últimos dos censos.

2. Propuestas metodológicas

Las propuestas metodológicas se referirán a tres momentos específicos. Éstos son: i) momento 0, que corresponde al año del primer censo; ii) momento t , que corresponde al año del segundo censo, y iii) momento $(t+n)$, que corresponde al año de la estimación.

i) *Distribución por prorratio*

Éste es el más sencillo de los métodos considerados en este trabajo. Se basa en el supuesto de que la distribución de la población por área es idéntica a la distribución de la variable sintomática. Para su aplicación se necesita solamente la distribución proporcional de la variable sintomática para el año en que se pretende estimar la población del área y, además, la estimación de la población total de las áreas para ese año. La población estimada del área i en el momento $t+n$ está dada por

$$P_{(i,t+n)} = \frac{S_{(i,t+n)}}{S_{(.,t+n)}} * P_{(.,t+n)} \quad [1]$$

donde

$P_{(i,t+n)}$ = población del área i en el momento $t+n$;

$P_{(.,t+n)}$ = estimación de la población total (suma de las i áreas) en el momento $t+n$;

$S_{(i,t+n)}$ = valor de la variable sintomática del área i en el momento $t+n$, y

$S_{(.,t+n)}$ = valor de la variable sintomática para el total de las áreas en el momento $t+n$.

Las principales ventajas de este método son: i) necesita información para un único momento (el de la estimación), por lo que puede disponerse de un mayor número de variables sintomáticas para el análisis (no se necesita una serie histórica de la variable elegida); ii) no es necesario compatibilizar las áreas geográficas en el tiempo, y iii) no hay cambios de definición o de forma de recolección de las variables en el tiempo.

Las desventajas son: i) el supuesto en que se basa el método no es muy bueno, ya que es difícil encontrar en la práctica una variable sintomática cuya distribución sea igual o muy similar a la de la población, y ii) las estimaciones están afectadas por la calidad diferencial —por área— de la variable sintomática.

A pesar de sus desventajas, si se dispone de un conjunto de variables sintomáticas para el período de estimación, se pueden obtener aquellas estimaciones para la población cuya media aritmética se aproxime más a la realidad.²

ii) *Distribución proporcional*

Este método supone que la población varía en igual proporción que la variable sintomática. Difiere del anterior en cuanto relaciona los cambios ocurridos en la variable sintomática y en la población entre dos momentos (en general, entre el último censo y el año deseado).

Para aplicarlo se necesita la información referente a la variable sintomática, por área, en dos momentos (censo y año deseado); la población por área en el momento inicial (censo), y una estimación de la población del total de las áreas para el año en cuestión.

² Lalu y Namboodiri (1971) proponen la media de una serie de regresiones simples como alternativa a la regresión múltiple. En los casos de Chile y Costa Rica, se aplicó la media a las distintas estimaciones obtenidas con cada método, lo que condujo a mejores estimaciones.

La población estimada del área i en el momento $t+n$ está dada por:

$$P_{(i,t+n)} = P_{(i,t)} * \frac{S_{(i,t+n)}}{S_{(i,t)}} * F_a \quad [2]$$

$$F_a = \frac{P_{(i,t+n)}}{\sum \left[P_{(i,t)} * \frac{S_{(i,t+n)}}{S_{(i,t)}} \right]} \quad [3]$$

donde

$P_{(i,t+n)}$; $P_{(.,t+n)}$; $S_{(i,t+n)}$ ya fueron definidos anteriormente;

$P_{(i,t)}$ = población del área i en el momento t ;

$S_{(i,t)}$ = variable sintomática en el momento t , y

F_a = factor de ajuste de la fórmula para que la suma de las $P_{(i,t+n)}$ sea igual a $P_{(.,t+n)}$.

Ventajas del método: i) es sencillo, dado que se basa en una sola variable sintomática; ii) la incorporación de los cambios de la variable sintomática permite llegar a un mayor número de estimaciones; o sea, se podrían utilizar variables cuya distribución no se asemeje a la de la población, siempre que su variación sea un buen indicador de los cambios de tamaño de la población.

Desventajas del método: i) necesita mayor cantidad de información que el anterior; ii) se necesita una compatibilización geográfica en los dos momentos, tanto a nivel de población como de la variable sintomática, y iii) las estimaciones están afectadas por cambios en la calidad de la información de la variable sintomática, si ésta no es constante en el tiempo.

Con este método se puede calcular tanto un conjunto de estimaciones como su media, lo cual redundará probablemente en una estimación más adecuada.

iii) *Correlación de razón*

Este método –desarrollado por Crosetti y Schmitt (1954)– se basa en el supuesto de que la evolución de la población está correlacionada con la variación de un conjunto de variables sintomáticas; la correlación se estima por medio de un modelo de regresión.

El modelo puede ser descrito de la siguiente forma:

$$Y_{(i,t)} = a_{(0)} + a_{(1)} * Xl_{(i,t)} + \dots + a_{(n)} * Xn_{(i,t)} + e_i \quad [4]$$

donde

$$Y_{(i,t)} = \frac{[P_{(i,t)}] / [P_{(.,t)}]}{[P_{(i,0)}] / [P_{(.,0)}]} \quad [5]$$

$$Xj_{(i,t)} = \frac{[Sj_{(i,t)}] / [Sj_{(.,t)}]}{[Sj_{(i,0)}] / [Sj_{(.,0)}]} \quad [6]$$

$Y_{(i,t)}$ = razón entre la proporción de población del área i en el momento 0 y el momento t ;

$Xj_{(i,t)}$, ($j = 1, \dots, n$) = razón entre la proporción de la variable j del área i en el momento 0 y el momento t , y

$e_{(i)}$ = error de la estimación según el modelo.

La población del área i en el momento $t+n$ se encuentra mediante la estimación del modelo de la ecuación [4], basado en el período $0, t$, y la posterior estimación de $Y_{(i,t+n)}$ con base en el período $t, t+n$, o sea:

$$P_{(i,t+n)} = Y_{(i,t+n)} * \frac{P_{(i,t)}}{P_{(.,t)}} * P_{(.,t+n)} \quad [7]$$

donde

$Y_{(i,t+n)}$ = razón de proporción de la población del área i estimada, para el período $t, t+n$, por el modelo de regresión.

Ventajas del método: i) la principal es que se basa en una o más variables cuya evolución explica la mayor parte de la variación de la población; ii) las estimaciones no debieran verse afectadas por errores en la calidad de la información, siempre y cuando los datos se mantengan en el tiempo o cambien en forma similar en todas las áreas consideradas, y iii) la estimación del modelo es independiente de las estimaciones realizadas, o sea, las áreas utilizadas para estimar el modelo no necesariamente deben ser las mismas áreas que se pretende estimar.

Desventajas: i) necesidad de disponer de mayor cantidad de información; ii) necesidad de compatibilizar la información en los dos momentos elegidos para estimar el modelo, y iii) necesidad de compatibilizar la información en los dos momentos necesarios para hacer la estimación.

iv) *Correlación de diferencia*

Este método –desarrollado por O’Hare (1976)– es similar al anterior y se basa en el mismo supuesto. La diferencia consiste en la forma de calcular las variaciones, pues las razones son sustituidas por diferencias. Así:

$$W_{(i,t)} = a_{(0)} + a_{(1)} * Zl_{(i,t)} + \dots + a_{(n)} * Zn_{(i,t)} + e_{(i)} \quad [8]$$

donde

$$W_{(i,t)} = \frac{P_{(i,t)}}{P_{(.,t)}} - \frac{P_{(i,0)}}{P_{(.,0)}} \quad [9]$$

$$Zj_{(i,t)} = \frac{Sj_{(i,t)}}{Sj_{(.,t)}} - \frac{Sj_{(i,0)}}{Sj_{(.,0)}} \quad [10]$$

$W_{(i,t)}$ = diferencia entre la proporción de la población del área i entre el momento 0 y el momento t ;

$Zj_{(i,t)}$ = diferencia entre la proporción de la variable j del área i entre el momento 0 y el momento t , y

$e_{(i)}$ = error de la estimación según el modelo.

La población del área i en el momento $t+n$ estará dada por:

$$P_{(i,t+n)} = \left[W_{(i,t+n)} + \frac{P_{(i,t)}}{P_{(.,t)}} \right] * P_{(.,t+n)} \quad [11]$$

donde

$W_{(i,t+n)}$ = diferencia de proporciones de la población del área i estimada, para el período $t, t+n$, por el modelo de regresión.

Este método presenta las mismas ventajas y desventajas que el método anterior. Sin embargo, O’Hare (1976) propone calcular las variaciones de proporción por medio de diferencias, pues así se puede lograr una mayor intercorrelación entre las variables (W y Zj) y, además, las estimaciones estarían menos afectadas por cambios temporales de las variables.

Otra característica importante de los métodos de correlación es que se pueden construir modelos distintos para subgrupos de áreas. Esto es especialmente útil en aquellos casos en que se considera que un modelo único no representa adecuadamente todas las divisiones geográficas cuya población se desea estimar. De este modo, se podrían construir modelos de regresión para grupos homogéneos de áreas, agrupadas, por ejemplo, según grandes regiones, predominancia urbana o rural, actividad económica principal o conforme a otros parámetros.

II. APLICACIONES

1. El caso de Chile

i) *Información básica*

Los métodos mencionados se aplicaron en este caso para estimar la población de las comunas existentes en 1992; posteriormente se verificó la precisión de las estimaciones obtenidas para 1992 comparándolas con los resultados del censo de ese año.

Conviene enumerar algunos de los obstáculos encontrados en la investigación, y la forma en que se soslayaron para poder aplicar los respectivos métodos o modelos. Como podrá verse, es posible que en otros países se presenten problemas similares.

Primero, la división politicoadministrativa de Chile experimentó cambios en todos los niveles en los años ochenta. A pesar de que las comunas existentes en 1970 seguían existiendo en 1982, prácticamente todas cambiaron sus límites geográficos.

Segundo, la información sobre registro de nacimientos y defunciones era codificada y publicada hasta 1985 a nivel de circunscripciones de registro civil que no correspondían a comunas; sin embargo, en la presente investigación se hizo una equivalencia entre las circunscripciones de registro civil y las comunas para el año 1985. Esta equivalencia fue utilizada para obtener los nacimientos y defunciones correspondientes a 1970 y 1982. Cabe señalar que los límites geográficos de las circunscripciones de registro civil también experimentaron modificaciones; tales cambios no fueron evaluados y ello afectará la calidad de las estimaciones.

Tercero, la información referente a los permisos de construcción o superficie construida para vivienda actualmente sólo está disponible para algunas comunas, aunque es posible obtenerla para las demás. No obstante, para el año 1970 se presenta el problema de la compatibilización de las comunas con las existentes en 1982. Debido a ello, se optó por no considerar esa variable en el presente estudio, aunque debería analizarse la posibilidad de hacerlo en el futuro.

Cuarto, aunque la información sobre el registro de automóviles está disponible por comuna, se presentan problemas para compatibilizar la información de las comunas existentes en los años base de los modelos de correlación (razón y diferencia). Está, además, el problema de la calidad de la información en las comunas de los grandes centros urbanos, en los cuales se permite inscribir el automóvil en una comuna distinta a la de residencia del propietario.

Quinto, la información sobre matrícula escolar está disponible por comunas, pero presenta también el problema de compatibilización en el tiempo. En este trabajo se consideró la matrícula en educación básica (de primero a octavo año).

Sexto, el registro electoral no ha sido actualizado de forma periódica y, además, la inscripción en él no es obligatoria. Generalmente se inscribe un gran número de personas poco antes de las elecciones; además, durante aproximadamente 15 años no hubo elecciones. Sin embargo, los datos del registro electoral podrán ser utilizados en trabajos futuros.

Séptimo, el consumo de energía eléctrica dejó de ser publicado en 1989; hasta esa fecha la compañía de energía eléctrica se encargaba de llevar el registro y publicarlo. Actualmente, el Instituto Nacional de Estadísticas (INE) está diseñando una encuesta para obtener tal información, la cual, por lo tanto, podrá ser utilizada en el futuro.

Octavo, las demás variables citadas (afiliadas a las administradoras de fondos de pensiones (AFP), recaudación de impuestos, y otras) no están disponibles, o, si lo están, no se encuentran a nivel de comunas, aunque sí de provincias o de alguna división distinta a la división politicoadministrativa, o sea, según la forma en que fue dividido el país para fines de recolección de la información.

Noveno, la población del censo de 1970 fue adaptada por el INE a la división politicoadministrativa de 1982, trabajo considerado de buena calidad.

Debido a los problemas o restricciones mencionados, finalmente se trabajó con la siguiente información:

En primer lugar, la población total por comunas readecuada para 1970, así como la población censada en 1982 y en 1992; segundo, como variables sintomáticas, los nacimientos y defunciones para los años 1970, 1982 y 1992; la matrícula en educación básica para los años 1982 y 1992; la inscripción de automóviles para los años 1982 y 1992, y los acontecimientos vitales, que corresponden a la suma de nacimientos y defunciones, para los años 1970, 1982 y 1992.

Además de esta información, y con el propósito de hacer comparaciones, se utilizaron las proyecciones de población de comunas realizadas por el INE/CELADE, que fueran elaboradas por el método de relación de cohortes.³

³Para mayores antecedentes metodológicos, véase INE/CELADE (1989).

Cuadro 1
**CHILE: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE
 POBLACIÓN POR COMUNA SEGÚN LOS MODELOS
 DE PRORRATEO, 1992**

Modelo	Error absoluto ^a (porcentajes)		Porcentaje de comunas con error absoluto superior al 10%	80% de las comunas con error absoluto inferior a
	Media	Desvia- ción estándar		
A. Nacimientos	12.0	12.1	45.1	19.2
B. Defunciones	22.4	19.1	71.2	35.3
C. Matrícula en educación básica	12.7	10.6	51.9	19.9
D. Eventos vitales	9.8	11.4	35.6	13.9
E. Promedio modelos A y B	11.8	12.7	43.4	18.4
F. Promedio modelos A, B y C	10.7	9.4	40.0	15.7
G. Promedio modelos A y C	9.0	7.8	34.6	13.2
H. Promedio modelos C y D	9.1	7.9	35.3	12.8
Proyecciones	11.8	13.2	40.3	17.8

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: Considera la estimación para un total de 295 comunas.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

ii) *Aplicación de los métodos de prorrateo y de distribución proporcional*

Como se dijo anteriormente, para la aplicación de estos métodos se necesita información referente a una variable sintomática. Siendo así, se utilizaron cuatro variables: nacimientos (modelo A), defunciones (modelo B), matrícula en educación básica (modelo C) y eventos vitales (modelo D).

Además de estos cuatro modelos, se calcularon estimaciones promedio: modelo E (nacimientos y defunciones), modelo F (nacimientos, defunciones y matrícula escolar), modelo G (nacimientos y matrícula), y modelo H (eventos vitales y matrícula escolar).

El análisis de los indicadores de la calidad de los resultados de estos modelos se resume en los cuadros 1 y 2.

Del cuadro 1 puede desprenderse que i) de las variables sintomáticas, los eventos vitales proporcionan las mejores estimaciones de población; ii) las peores estimaciones son las calculadas sobre la base de las defunciones; iii) al considerar los modelos promedio, las mejores estimaciones resultan de los modelos G y H, y iv) a pesar de las limitaciones impuestas por el método y por las variables sintomáticas disponibles, se lograron estimaciones más precisas que las obtenidas con las proyecciones aludidas.

Cuadro 2
**CHILE: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE
 POBLACIÓN POR COMUNA SEGÚN LOS MODELOS
 DE DISTRIBUCIÓN PROPORCIONAL, 1992**

Modelo	Error absoluto ^a (porcentajes)		Porcentaje de comunas con error absoluto superior al 10%	80% de las comunas con error absoluto inferior a
	Media	Desvia- ción estándar		
A. Nacimientos	15.5	22.5	49.5	21.6
B. Defunciones	20.0	27.8	56.6	26.9
C. Matrícula en educación básica	13.1	13.3	48.1	17.9
D. Eventos vitales	13.7	20.8	42.7	19.7
E. Promedio modelos A y B	14.5	21.0	43.1	20.6
F. Promedio modelos A, B y C	10.9	14.7	34.2	15.7
G. Promedio modelos A y C	9.8	13.3	32.5	14.0
H. Promedio modelos C y D	9.4	12.7	28.5	13.6
Proyecciones	11.8	13.2	40.3	17.8

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: Considera la estimación para un total de 295 comunas.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

De lo que se muestra en el cuadro 2 cabe destacar que i) dentro de los modelos simples, esto es, con una variable, aquellos basados en la matrícula escolar y los eventos vitales son los que proporcionan mejores estimaciones; ii) las estimaciones basadas en defunciones resultan mejores que las obtenidas con defunciones en el método de prorrateo, y iii) si se consideran los modelos promedio (E, F, G y H), las mejores estimaciones las proporcionan los modelos G y H, que son mucho más adecuadas que las anteriores y que las proyecciones.

Se destaca también que si se dispone de un conjunto de variables sintomáticas cuya evolución pueda ser relacionada con los cambios en el tamaño de la población, se puede contar con métodos sencillos para la actualización de la población.

iii) *Aplicación del método de correlación de razón y del método de diferencias*

En estos métodos se necesita mayor cantidad de información: i) es preciso conocer la población y las variables sintomáticas por área compatible en dos momentos, para estimar el modelo de regresión; ii) la

población y las variables sintomáticas por área en un momento dado y las variables sintomáticas por área en el momento de la estimación; iii) se requiere una estimación para el total de las áreas.

La mayor cantidad de información es la gran limitación que encuentra la aplicación de estos métodos. En este trabajo, en particular, destinado a evaluar el resultado de los métodos, se necesitó, además de la información mencionada anteriormente, la población censada por áreas en el momento estimado.

En la aplicación se utilizó solamente la información relativa a nacimientos, defunciones y eventos vitales como variables sintomáticas.

Primeramente se estimaron los modelos de regresión (4) y (8) considerando el total de comunas, y los nacimientos y las defunciones como variables sintomáticas (modelo A). Posteriormente esas regresiones se estimaron considerando el total de comunas y las variables sintomáticas separadamente (modelos B y C) y el promedio de éstas (modelo D), incorporando así la alternativa a la regresión múltiple sugerida por Lalu y Namboodiri (1971). Finalmente se estimó el modelo E, que considera el total de las comunas y, como variable sintomática, los eventos vitales (suma de nacimientos y de defunciones).

En el cuadro 3 se resumen los resultados, en los que se destaca que: i) las estimaciones encontradas en los diversos modelos son bastante similares a las proyecciones; ii) el método de correlación de diferencias proporciona estimaciones ligeramente mejores que el método de correlación de razón; iii) el hecho de que el modelo B produzca las peores estimaciones puede ser consecuencia de la calidad de la información básica, lo cual lleva a pensar que la compatibilización hecha para las defunciones en el año 1970 presenta problemas mayores; iv) el uso de modelos promedio (modelo D) no condujo a cambios importantes en las estimaciones, y v) a pesar de los problemas de información (sobre todo la compatibilización de los nacimientos y defunciones en los años considerados para estimar el modelo, 1970 y 1982, y en los años considerados en la estimación, 1982 y 1992), las estimaciones resultaron similares a las proyecciones.

Considerando las limitaciones de disponibilidad de información y los problemas existentes en ésta, se puede concluir que los métodos de correlación permiten elaborar una gran cantidad de modelos, con resultados razonablemente buenos.

Cuadro 3
**CHILE: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE
 POBLACIÓN POR COMUNA SEGÚN LOS MODELOS
 DE CORRELACIÓN, 1992**

Método/modelo	Error absoluto ^a (porcentajes)		Porcentaje de comunas con error absoluto superior al 10%	80% de las comunas con error absoluto inferior a
	Media	Desvia- ción estándar		
Correlación de razón				
A. Nacimientos y defunciones	12.0	11.0	49.2	19.3
B. Nacimientos	11.7	10.8	49.5	18.4
C. Defunciones	13.1	12.2	52.2	20.4
D. Promedio modelos A y B	12.3	11.4	49.8	19.8
E. Eventos vitales	11.4	10.4	47.8	18.4
Correlación de diferencias				
A. Nacimientos y defunciones	18.8	12.8	41.0	17.7
B. Nacimientos	11.8	13.4	44.1	19.1
C. Defunciones	15.3	16.2	52.9	22.9
D. Promedio modelos A y B	12.1	13.2	43.1	17.9
E. Eventos vitales	11.4	12.7	41.4	17.8
Proyecciones	11.8	13.2	40.3	17.8

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: Considera la estimación para un total de 295 comunas.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

iv) *El uso de la estratificación por tamaño de las áreas en los modelos de correlación* ⁴

Para probar los modelos de correlación en conjuntos más homogéneos, se estratificaron las comunas según la cantidad de habitantes que tenían en 1982. Así, éstas quedaron divididas en tres grupos: estrato 1, comunas con una población inferior a 10 000 habitantes; estrato 2, comunas con una población de entre 10 000 y 50 000 habitantes, y estrato 3, comunas con una población superior a 50 000 habitantes. Posteriormente se construyeron los modelos y se hicieron las estimaciones para cada estrato.

⁴ Véanse Ericksen (1973); Martin y Serow (1978), y Teixeira Jardim (1992).

Cuadro 4
**CHILE: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE
 POBLACIÓN POR COMUNA SEGÚN LOS MODELOS
 DE CORRELACIÓN, ESTRATIFICADO POR
 TAMAÑO DE LAS COMUNAS, 1992**

Método/modelo	Error absoluto ^a (porcentajes)		Porcentaje de comunas con error absoluto superior al 10%	80% de las comunas con error absoluto inferior a
	Media	Desvia- ción estándar		
Correlación de razón				
A. Nacimientos, defunciones y tamaño	9.3	9.7	32.2	14.3
B. Nacimientos y tamaño	8.9	9.1	29.5	13.4
C. Defunciones y tamaño	10.1	10.4	35.5	14.9
D. Promedio modelos A y B	9.2	9.6	31.9	14.2
E. Eventos vitales y tamaño	8.7	8.9	29.8	13.0
Correlación de diferencias				
A. Nacimientos, defunciones y tamaño	14.0	22.6	46.4	18.3
B. Nacimientos y tamaño	13.3	21.1	45.4	18.4
C. Defunciones y tamaño	24.0	46.8	59.3	31.3
D. Promedio modelos A y B	17.5	33.3	49.8	19.6
E. Eventos vitales y tamaño	12.8	19.9	41.4	16.6
Proyecciones	11.8	13.2	40.3	17.8

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: considera la estimación para un total de 295 comunas.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

De esta forma, se utilizaron los mismos modelos que antes (A, B, C, D y E), pero incorporando esta vez la variable "tamaño" (que adopta los valores 1, 2 ó 3 según el estrato a que pertenezca la comuna). Un resumen del análisis de los resultados puede observarse en el cuadro 4.

En el cuadro 4 se destaca que i) las mejores estimaciones se obtienen por medio del método de correlación de razón; ii) los modelos B y E entregan las mejores estimaciones; y iii) la estratificación en que se toma en cuenta solamente el tamaño de las comunas produce los mejores resultados obtenidos hasta ahora.

v) *El uso de variables mudas (dummy), para incorporar la condición urbano/rural en los modelos de correlación*

Con el objeto de mejorar las estimaciones se introdujo el concepto de localidad urbana, incorporando para ello a los modelos de correlación

Cuadro 5
**CHILE: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES
 DE POBLACIÓN POR COMUNA, DE ACUERDO A LOS
 MODELOS DE CORRELACIÓN, ESTRATIFICADO
 SEGÚN CONDICIÓN URBANO/RURAL
 PREDOMINANTE EN LAS COMUNAS, 1992**

Método/modelo	Error absoluto ^a (porcentajes)		Porcentaje de comunas con error absoluto superior al 10%	80% de las comunas con error absoluto inferior a
	Media	Desvia- ción estándar		
Correlación de razón				
A. Nacimientos, defunciones y condición urbano/rural	9.2	9.2	37.3	13.6
B. Nacimientos y condición urbano/rural	9.0	9.1	35.9	13.6
C. Defunciones y condición urbano/rural	9.9	9.9	39.7	14.8
D. Promedio modelos A y B	9.3	9.3	36.9	13.7
E. Eventos vitales y condición urbano/rural	8.9	8.8	35.3	13.8
Correlación de diferencias				
A. Nacimientos, defunciones y condición urbano/rural	13.9	20.8	44.7	19.3
B. Nacimientos y condición urbano/rural	13.4	19.7	45.1	17.9
C. Defunciones y condición urbano/rural	19.3	31.1	55.3	27.1
D. Promedio modelos A y B	15.0	24.8	46.4	20.8
E. Eventos vitales y condición urbano/rural	12.9	18.6	45.8	16.7
Proyecciones	11.8	13.2	40.3	17.8

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: Considera la estimación para un total de 295 comunas.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

una variable, a la cual se le atribuye valor 1 si en la comuna predomina la población urbana, y 0 si predomina la población rural. Los resultados se presentan en el cuadro 5.

De la información del cuadro 5 se puede concluir que i) las mejores estimaciones se obtienen con el método de correlación de razón; ii) las estimaciones obtenidas con el método de correlación de diferencias tienden a empeorar, y iii) los modelos A, B, D y E arrojan las mejores estimaciones.

Cuadro 6
**CHILE: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES DE
 POBLACIÓN POR COMUNA SEGÚN LOS MODELOS DE
 CORRELACIÓN, ESTRATIFICADO POR TAMAÑO
 Y CONDICIÓN URBANO/RURAL
 PREDOMINANTE, 1992**

Método/modelo	Error absoluto ^a (porcentajes)		Porcentaje de comunas con error absoluto superior al 10%	80% de las comunas con error absoluto inferior a
	Media	Desvia- ción estándar		
Correlación de razón				
A. Nacimientos, defunciones, tamaño y condición urbano/rural	9.0	9.3	30.2	13.6
B. Nacimientos, tamaño y condición urbano/rural	8.4	8.9	27.1	13.7
C. Defunciones, tamaño y condición urbano/rural	9.7	10.0	32.9	15.5
D. Promedio modelos A y B	8.9	9.3	30.8	13.3
E. Eventos vitales, tamaño y condición urbano/rural	8.1	8.5	26.8	12.5
Correlación de diferencias				
A. Nacimientos, defunciones, tamaño y condición urbano/rural	15.2	30.7	44.7	19.2
B. Nacimientos, tamaño y condición urbano/rural	14.7	29.5	44.1	17.2
C. Defunciones, tamaño y condición urbano/rural	24.8	55.7	53.9	31.1
D. Promedio modelos A y B	18.3	42.2	46.1	21.2
E. Eventos vitales, tamaño y condición urbano/rural	14.1	27.6	42.0	17.2
Proyecciones	11.8	13.2	40.3	17.8

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: considera la estimación para un total de 295 comunas.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

vi) *El uso de la estratificación por tamaño y por condición urbano/rural en los modelos de correlación*

Como una última alternativa, se probó la incorporación de la variable localidad urbana a los modelos estratificados anteriormente descritos. Los resultados se presentan en el cuadro 6.

Observando el cuadro 6, se puede decir que el método de correlación de razón produce las mejores estimaciones y, que el modelo E,

utilizando el método de correlación de razón, fue el que presentó las mejores estimaciones entre todos los modelos considerados.

vii) *Conclusiones generales*

Sobre la base del caso de Chile, se puede concluir que: i) la utilización de variables sintomáticas permite construir un sinnúmero de modelos; ii) los modelos de correlación de razón son probablemente los más adecuados y mejoran sus estimaciones a medida que se incorporan modificaciones con vistas a homogeneizar las áreas en estudio; iii) los modelos de correlación de diferencias no condujeron a buenos resultados; iv) en los modelos de correlación las estimaciones promedio, basadas en regresiones simples, no introdujeron cambios importantes en cuanto a la precisión de las estimaciones; v) en los modelos de prorrateo y distribución proporcional resultan mejores las estimaciones promedio, y vi) el modelo E, estimado por el método de correlación de razón, es el que arrojó las mejores estimaciones cuando se le incorporaron la estratificación según tamaño y la condición de localidad urbana.

2. El caso de Costa Rica

i) *Información básica*

Con el objetivo de evaluarlos y compararlos, se aplicaron los métodos antes descritos a determinados aspectos de Costa Rica, lo cual permite a la vez analizar un país de menor tamaño que Chile y que muestra distintas características en cuanto a problemas de información.

En este caso la aplicación se llevó a cabo para estimar la población de los cantones existentes en 1984, comparándola con la población censada en ese año, que corresponde al último censo. En esta aplicación no se tuvo contactos directos con las fuentes de información, lo cual limitó el estudio a las publicaciones relativas a Costa Rica disponibles en el CELADE. Se aplicaron aquí modelos semejantes al caso anterior, con el fin de averiguar adicionalmente si se podían obtener mejores estimaciones cuando hay mayor compatibilidad temporal entre las áreas.

Dadas las limitaciones aludidas, se utilizó la siguiente información:

i) la población censada en los tres últimos censos (1963, 1973 y 1984); ii) los nacimientos y defunciones registrados y ocurridos hasta 10 años antes, y iii) la población que asistía a establecimientos regulares de enseñanza en el momento de los distintos censos, tomada como una aproximación a la matrícula escolar.

Cuadro 7
**COSTA RICA: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES
 DE POBLACIÓN POR CANTÓN SEGÚN LOS
 MODELOS DE PRORRATEO, 1984**

Método/modelo	Error absoluto ^a (porcentajes)		Porcentaje de cantones con error absoluto superior al 10%	80% de los cantones con error absoluto inferior a
	Media	Desvia- ción estándar		
A. Nacimientos	12.4	11.0	45.7	20.9
B. Defunciones	13.7	10.0	53.1	23.1
C. Matrícula escolar	12.7	7.9	56.8	20.0
D. Eventos vitales	10.8	9.9	35.8	17.7
E. Promedio modelos A y B	8.1	8.1	27.2	14.3
F. Promedio modelos A, B y C	7.9	6.5	29.6	13.4
G. Promedio A y C	8.4	6.4	34.6	13.3
H. Promedio C y D	8.1	6.2	29.6	13.4

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: Se consideró la estimación de 81 cantones. La matrícula fue aproximada según la asistencia escolar.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

Con respecto a la información utilizada, cabe destacar los siguientes problemas: i) los nacimientos y defunciones presentan problemas de fluctuación debidos a la forma de tabulación, que se refiere a los hechos registrados y no a los ocurridos. Para minimizar el problema se calculó un promedio de tres años como una estimación de los hechos ocurridos en el año central; ii) la información sobre asistencia escolar, que reemplaza la relativa a matrícula escolar, produce un sesgo entre el área o localidad donde se ubica el establecimiento escolar y el área en que reside el estudiante: primero, en su asociación con la residencia del estudiante y, segundo, con el establecimiento de enseñanza. Se espera que la correlación entre matrícula escolar y población sea menor que la estimada entre asistencia escolar y población.

En este caso, las estimaciones correspondientes a 1984 generadas por medio de los distintos métodos no se compararon con las proyecciones de población a nivel de cantones, por cuanto éstas no estaban disponibles.

ii) *Aplicación de los métodos de prorrateo y de distribución proporcional*

Siguiendo la misma línea de investigación utilizada en el caso anterior, se hicieron estimaciones con cada una de las variables sintomáticas

Cuadro 8
**COSTA RICA: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES
 DE POBLACIÓN POR CANTÓN SEGÚN LOS MODELOS DE
 DISTRIBUCIÓN PROPORCIONAL, 1984**

Método/modelo	Error absoluto ^a (porcentajes)		Porcentaje de cantones con error absoluto superior al 10%	80% de los cantones con error absoluto inferior a
	Media	Desvia- ción estándar		
A. Nacimientos	14.0	9.1	59.3	22.8
B. Defunciones	18.2	14.9	59.3	30.7
C. Matrícula	9.9	8.2	38.3	16.1
D. Eventos vitales	11.8	9.9	45.7	18.9
E. Promedio modelos A y B	14.4	10.8	51.9	22.7
F. Promedio modelos A, B y C	8.9	6.4	40.7	14.4
G. Promedio modelos A y C	7.2	4.5	28.4	11.2
H. Promedio modelos C y D	5.3	4.3	13.6	8.4

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: Se consideró la estimación de 81 cantones. La matrícula fue aproximada según la asistencia escolar.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

(nacimientos, defunciones, eventos vitales y asistencia escolar), y se analizaron también las estimaciones promedio, las cuales se presentan en los cuadros 7 y 8. Considerando los modelos A, B, C y D, puede concluirse que de las variables estudiadas los eventos vitales y la matrícula escolar son las que proporcionan mejores estimaciones y las defunciones las peores (esto también se observó en el caso chileno). Al considerar los modelos promedio, las estimaciones más adecuadas se producen con los modelos E y H cuando se aplica el método de prorrateo y con el modelo H cuando se utiliza el método de distribución proporcional.

iii) *Aplicación de los métodos de correlación de razón y de correlación de diferencias*

El caso de Costa Rica presenta la ventaja de que los cantones no experimentaron grandes cambios geográficos en el período en estudio, lo que permitió incorporar a los modelos de correlación una variable más, o sea, la matrícula escolar.

Los resultados de la aplicación de 10 modelos se presentan en el cuadro 9, donde se destaca que: i) los modelos de razón producen mejores estimaciones que los de diferencia; ii) el modelo C proporciona

Cuadro 9
**COSTA RICA: ANÁLISIS DEL ERROR DE LAS ESTIMACIONES
 DE POBLACIÓN POR CANTÓN SEGÚN LOS
 MODELOS DE CORRELACIÓN, 1984**

Método/modelo	Error absoluto ^a (porcentajes)		Porcentaje de cantones con error absoluto superior al 10%	90% de los cantones con error absoluto inferior a
	Media	Desvia- ción estándar		
Correlación de razón				
A. Nacimientos y defunciones	9.6	8.5	39.5	16.7
B. Nacimientos, defunciones y matrícula	4.7	3.7	7.4	7.3
C. Eventos vitales y matrícula	4.2	3.5	6.2	6.3
D. Nacimientos	9.5	8.4	37.0	17.2
E. Defunciones	12.6	10.2	49.4	18.9
F. Matrícula	7.8	6.5	24.7	12.1
G. Eventos vitales	9.2	8.6	33.3	17.1
H. Promedio modelos D y E	10.7	8.7	50.6	15.5
I. Promedio modelos D, E y F	6.7	5.9	18.5	10.0
J. Promedio modelos F y G	4.6	4.3	7.4	7.3
Correlación de diferencias				
A. Nacimientos y defunciones	9.9	9.1	37.0	17.4
B. Nacimientos, defunciones y matrícula	5.9	4.9	16.0	9.5
C. Eventos vitales y matrícula	18.1	21.4	51.9	29.7
D. Nacimientos	9.8	9.1	35.8	17.8
E. Defunciones	13.6	11.1	53.1	20.0
F. Matrícula	8.1	8.2	25.9	10.9
G. Eventos vitales	10.2	8.9	39.5	18.1
H. Promedio modelos D y E	10.3	8.6	44.4	16.5
I. Promedio modelos D, E y F	8.2	7.0	29.6	11.7
J. Promedio modelos F y G	6.0	5.5	16.0	8.9

Fuente: Elaborado por la autora sobre la base de cifras oficiales.

Nota: Se consideró la estimación de 81 cantones. La matrícula fue aproximada según la asistencia escolar.

^a Error absoluto = [(población estimada - población censada) / población censada] * 100.

las estimaciones más adecuadas con el método de correlación de razón, y el modelo B con el método de correlación de diferencias, y iii) los peores resultados son los obtenidos por medio del modelo E con el método de correlación de razón, y por medio del modelo C con el método de correlación de diferencias.

Asimismo, cabe destacar que si se incorpora al modelo A un mejor indicador de la evolución de la población, por ejemplo, la matrícula escolar, las estimaciones mejoran considerablemente.

La estratificación según el tamaño de los cantones y la condición de localidad urbana, utilizadas en el caso de Chile, no significaron en el caso de Costa Rica mejoría de las estimaciones, por lo cual no se presentan aquí los resultados respectivos. Sin embargo, es posible que otras estratificaciones, más adecuadas a la situación de Costa Rica, produzcan mejores estimaciones.

En el caso de Costa Rica se lograron estimaciones de población muy razonables con los métodos de distribución proporcional (modelo H) y de correlación de razón (modelos B, C y J), usando como variables sintomáticas los nacimientos, las defunciones, eventos vitales y la matrícula escolar. En consecuencia, si en Chile no existiera el problema de la compatibilización geográfica de las comunas, se obtendrían mejores estimaciones.

III. CONCLUSIONES Y RECOMENDACIONES

A. La necesidad de contar con mejores estimaciones de población a nivel de área menor⁵ (local, municipal, u otras) está dada por el hecho de que cada día cobran más peso en lo económico, político, cultural, las comunidades pequeñas. Sea para la elaboración de planes de gobierno, o bien porque ejercen presión sobre las autoridades centrales en beneficio del desarrollo del área, o por las prioridades de descentralización y desarrollo local, lo cierto es que cada vez será más urgente actualizar y precisar los datos de población de las áreas menores, por lo cual es válido todo esfuerzo que se haga por obtener mejores estimaciones de población.

B. Si se dispusiera de registros frecuentes del movimiento migratorio entre estas áreas, de los nacimientos y de las defunciones, estaría resuelto

⁵ Véanse Flores (1989) y Rueda (1989).

el problema de la actualización de las estimaciones de población, al estar determinados los tres componentes de la variación del tamaño de la población. Los nacimientos y las defunciones se registran frecuentemente, pero no ocurre lo mismo con la migración. Por ello, el contar con métodos alternativos que indiquen la variación de la población de las áreas menores tiene gran importancia en la elaboración de estimaciones y proyecciones de población.

C. Censar la población en forma más frecuente sería otra alternativa, pero es poco factible a causa de los altos costos que implica esta operación. Cabe recordar que a pesar de la recomendación de las Naciones Unidas de realizar censos cada 10 años, muchos países de América Latina y el Caribe los llevan a cabo en intervalos más amplios.

D. Por lo dicho, una alternativa eficiente y eficaz es mejorar los procedimientos de registro administrativo ya existentes, para su posterior uso en la actualización de las estimaciones de población.

E. En este trabajo se demuestra que a pesar de los problemas de disponibilidad y de calidad de la información, se puede llegar a estimaciones bastante razonables de la población de las áreas pequeñas.

F. Con respecto a los métodos aplicados, puede concluirse que hay que hacer todo lo posible por utilizar el método de correlación de razón, pues en ambos casos presentó las mejores estimaciones.⁶

G. Seleccionar el modelo no es tarea fácil, pues ello depende de la calidad y disponibilidad de información y también de las características particulares de las áreas a estimar. Para escoger el modelo "óptimo" parece bastante útil elaborar un estudio previo, como el aquí presentado, para una evaluación a priori de la situación en cada caso concreto, así como familiarizarse con la información y los distintos métodos. Además, es muy conveniente hacer evaluaciones a posteriori, lo que permitirá introducir cambios en el modelo utilizado.

⁶En su reciente visita al CELADE, Maria Andrassy-Bitto —encargada de las estimaciones de población de áreas menores en Canadá— sugirió probar la estabilidad de los modelos de correlación (coeficientes de regresión). En caso de que se encuentre que los modelos, sobre todo los modelos de correlación de diferencias, no son estables en el tiempo, será posible corregirlos y perfeccionarlos, con lo cual mejorarán también las estimaciones.

H. Otro aspecto importante es el control permanente de la información básica, o sea, de los indicadores utilizados, lo que podría llevar incluso a cambiar de modelo de un período intercensal a otro.

I. Se hace necesario integrar las diversas fuentes de información y el organismo encargado de las estimaciones de población. Muchas veces el introducir pequeñas modificaciones en el proceso de recolección o de tabulación de la información abre una amplia gama de posibilidades en lo que a construcción de modelos se refiere, y sobre todo en lo tocante a los criterios para estratificar las áreas a estimar. De esta forma se podría disponer de un sistema confiable –y de bajo costo– para realizar las estimaciones de población.

J. Finalmente, parece recomendable utilizar variables sintomáticas en la estimación de la población de áreas menores en América Latina y el Caribe. Otro ejemplo que apunta en esta dirección es la aplicación hecha por Teixeira Jardim (1992) en lo concerniente al Estado de Rio Grande do Sul de Brasil, trabajo que establece un punto de partida dentro de América Latina y el Caribe en cuanto a la divulgación y utilización de los métodos aquí expuestos. Es por ello que, sobre la base de estas experiencias y sabiendo que no se tiene un camino fácil por delante, parece pertinente incorporar a la región las experiencias de Canadá y de los Estados Unidos, a fin de tener en un futuro cercano una alternativa más confiable para la actualización de los datos de población.

K. Además de las conclusiones y recomendaciones generales recién expuestas, es pertinente enumerar ciertas líneas de investigación más específicas que convendría profundizar en el futuro: i) se hace necesario profundizar en el conocimiento y divulgación de la teoría estadística implícita en cada método, a fin de tener una base sólida para decidir cuál es el modelo más adecuado (véanse al respecto Bravo (1993), y Basavarajappa y Verma (1985)); ii) en el caso de Chile, es preciso mejorar los modelos, para lo cual es menester encontrar nuevos criterios de estratificación de las comunas seleccionadas para la estimación así como definir modelos por regiones; iii) debería promoverse la realización de aplicaciones semejantes en otros países de la región, pero con presencia directa del investigador, a fin de mejorar la información. Esto permitiría enriquecer el campo de evaluación de los distintos métodos y modelos utilizados, así como buscar nuevos modelos para mejorar los resultados; iv) se podría explorar la aplicación de otras metodologías

existentes, entre ellas el método llamado Componentes II (Batutis, 1991), que utiliza la ecuación compensadora por cohortes para determinar la población en cada momento, pero considera variables sintomáticas para estimar las migraciones, y v) por último, quizás lo más urgente es llevar a cabo una extensa y minuciosa recopilación de los trabajos realizados sobre el tema, hacer una selección adecuada y compilarlos en una publicación. Ello permitiría abrir nuevos horizontes, ya sea para promover la discusión sobre el modo de perfeccionar los métodos, ya sea para ampliar el campo de sus posibles aplicaciones.

BIBLIOGRAFÍA

- Bravo, Jorge (1993), "Algunas observaciones acerca de los métodos de estimación de población a través de variables sintomáticas", Santiago de Chile, Centro Latinoamericano de Demografía (CELADE), inédito.
- Basavarajappa, K. G. y Ravi B. P. Verma (1985), "Recent developments in the estimation of population by the regression method", documento presentado en el Simposio Internacional sobre Estadísticas en Áreas Pequeñas (Ottawa, Ontario, 22 al 24 de mayo de 1985), Statistics Canada.
- Basavarajappa, K. G., Rosemary Bender y Ravi B. P. Verma (1982), "New approaches to methods of estimating the population of census divisions", Ottawa, Demography Division, Statistics Canada.
- Batutis, Jr. (1991), "Estimates preparation: methods and procedures", en comunicación personal de John Long en su visita al Centro Latinoamericano de Demografía (CELADE).
- Crosetti, A. H. y R. C. Schmitt (1954), "Accuracy of the ratio-correlation method for estimating postcensal population", *Land Economics*, N° 30.
- Ericksen, Eugene P. (1973), "A method for combining sample survey data and syntomatic indicators to obtain population estimates for local areas", *Demography*, N° 10.
- Florez, Carmen Elisa (1989), "Las proyecciones de población y la estructura económica y social", documento presentado en el Seminario Internacional sobre Proyecciones Subnacionales de Población (Girardot, 31 de octubre al 2 de noviembre de 1988), Santafé de Bogotá, Centro de Estudios sobre Desarrollo Económico (CEDE), Universidad de los Andes.
- Goldberg, David, N. K. Namboodiri y V. R. Rao (1964), "A test of the accuracy of ratio correlation population estimates", *Land Economics*, N° 40.
- Grier, John M. y Robert C. Schmitt (1966), "A method of estimating the population of minor civil divisions", *Rural Sociology*, N° 31.
- INE (Instituto Nacional de Estadísticas) (1992a), *Censo de población y vivienda: resultados generales Chile 1992*, Santiago de Chile.
- (1992b), "Nacidos vivos y defunciones generales según lugar de residencia", *Demografía*, Santiago de Chile.
- (1992c), *Parque de vehículos en circulación. Anuario de transporte y comunicación*, Santiago de Chile.

- (1982a), “Nacidos vivos y defunciones generales según lugar de residencia”, *Demografía*, Santiago de Chile.
- (1982b), *Parque de vehículos en circulación. Anuario de transporte y comunicación*, Santiago de Chile.
- (1982c), *Censo nacional de población y vivienda: localidades pobladas. Censo nacional de población y vivienda: I a XII región y Región Metropolitana*, Santiago de Chile.
- (1970a), “Nacidos vivos y defunciones generales según lugar de residencia”, *Demografía*, Santiago de Chile.
- (1970b), *Censo nacional de población y vivienda: localidades pobladas. Censo nacional de población y vivienda: I a XII región y Región Metropolitana*, Santiago de Chile.
- (s/f), *Listado de códigos de comuna en vigencia a partir de 01/07/85 y su correspondencia con antiguo código de circunscripción de registro civil*, Santiago de Chile, inédito.
- INE/CELADE (Instituto Nacional de Estadísticas/Centro Latinoamericano de Demografía) (1989), *Chile: Proyecciones y estimaciones de población, por sexo y edad. Comunas 1980-1995*, Santiago de Chile.
- Kmenta, Jan (1978), *Elementos de econometría*, traducción de Carlos Roberto Vieira Araújo, Editora Atlas S.A.
- Lalu, N.M. y N. Krishnan Namboodiri (1971), “The average of several simple regression estimates as an alternative to the multiple regression estimate in postcensal and intercensal population estimation: a case study”, *Rural Sociology*, vol. 36.
- Long, John F. (1993), “Methods for postcensal population estimation: flow methods vs. stock methods”, *International Union for the Scientific Study of Population. International Population Conference, 1993*, Liège, Unión Internacional para el Estudio Científico de la Población (UIECP).
- Mandell, Marylou y Jeffrey Tayman (1982), “Measuring temporal stability in regression models of population estimation”, *Demography*, vol. 19.
- Martin, Julia H. y William J. Serow (1978), “Estimating demographic characteristics using the ratio-correlation method”, *Demography*, vol. 15.
- Ministerio de Economía, Industria y Comercio (1992), *Diez años de estadística vital 1978-1987: población, nacimientos, defunciones, contrayentes*, San José de Costa Rica, Dirección General de Estadística y Censos (DGEC).
- (1975), *Anuario estadístico de Costa Rica, 1974*, San José de Costa Rica, Dirección General de Estadística y Censos (DGEC).
- (1974), *Anuario estadístico de Costa Rica, 1973*, San José de Costa Rica, Dirección General de Estadística y Censos (DGEC).
- (1973), *Anuario estadístico de Costa Rica, 1972*, San José de Costa Rica, Dirección General de Estadística y Censos (DGEC).
- (1965), *Anuario estadístico de Costa Rica, 1964*, San José de Costa Rica Dirección General de Estadística y Censos (DGEC).
- (1964), *Anuario estadístico de Costa Rica, 1963*, San José de Costa Rica, Dirección General de Estadística y Censos (DGEC).
- (1963), *Anuario estadístico de Costa Rica, 1962*, San José de Costa Rica, Dirección General de Estadística y Censos (DGEC).
- Ministerio de Educación Pública (1992), *Matrícula escolar por comunas*, Informe de los establecimientos educacionales, remitidos por las secretarías regionales ministeriales de educación, Santiago de Chile.

- (1982), *Matrícula escolar por comunas*, Informe de los establecimientos educacionales, remitidos por las secretarías regionales ministeriales de educación, Santiago de Chile.
- Namboodiri, N. Krishnan (1972), "On the ratio-correlation and related methods of subnational population estimation", *Demography*, N° 9.
- Norusis, Marija J. (1988), *SPSS/PC+ V2.0: base manual*, SPSS Inc.
- O'Hare, William P. (1980), "A note on the use of regression methods in population estimates", *Demography*, N° 17.
- (1976), "Report of a multiple regression method for making population estimates", *Demography*, N° 13.
- Pursell, Donald E. (1970), "Improving population estimates with the use of dummy variables", *Demography*, N° 7.
- Rueda, José Olinto (1989), "Las proyecciones subnacionales de población y la planificación del desarrollo", documento presentado en el Seminario Internacional sobre Proyecciones Subnacionales de Población (Girardot, 31 de octubre al 2 de noviembre de 1988), Serie OI, N° 42, Santafé de Bogotá, Departamento Administrativo Nacional de Estadística (DANE).
- Shryock, Henry S. jr. y Meyer Zitter (1964), "Accuracy of methods of preparing postcensal population estimates for states and local areas", *Demography*, N° 1.
- Swanson, David A. (1980), "Improving accuracy in multiple regression estimates of population using principles from causal modelling", *Demography*, N° 17.
- Swanson, David A. y Lucky M. Tedrow (1984), "Improving the measurement of temporal change in regression models used for country population estimates", *Demography*, N° 21.
- Teixeira Jardim, Maria de Lourdes (1992), "Utilização de variáveis sintomáticas para estimar a distribuição espacial de populações: aplicação aos municípios do Rio Grande do Sul", *Anais del VIII Encontro Nacional de Estudos Populacionais*, N° 1, Asociación Brasileña de Estudios Poblacionales (ABEP).