

Distr.  
RESTRINGIDA

LC/DEM/R.211  
Serie A N° 291  
26 de mayo de 1994

ORIGINAL: ESPAÑOL

---

**CELADE - Centro Latinoamericano de Demografía**

**EL USO DE VARIABLES SINTOMATICAS EN LA ACTUALIZACION DE  
POBLACION DE AREAS MENORES  
(Versión preliminar)**

Este documento ha sido preparado por Guiomar Bay, Consultora del CELADE. Esta investigación ha sido posible gracias al aporte del Programa de Cooperación e Intercambio CELADE/CANADA apoyado por la Agencia Canadiense para el Desarrollo Internacional (ACDI/CIDA) y del Proyecto Regional del Fondo de Población de las Naciones Unidas (UNFPA). Las opiniones emitidas en este documento son sólo responsabilidad de la autora.

Este documento no fue sometido a la revisión editorial.

## INDICE

1. INTRODUCCION. . . . .	3
2. METODOS Y MODELOS UTILIZADOS. . . . .	5
2.1. Información básica. . . . .	5
2.2. Métodos y modelos. . . . .	6
2.2.1. Distribución por prorratio. . . . .	6
2.2.1. Distribución proporcional. . . . .	8
2.2.3. Correlación de razón. . . . .	9
2.2.4. Correlación de diferencia. . . . .	11
3. APLICACIONES. . . . .	13
3.1. El caso de Chile. . . . .	13
3.1.1. Información básica. . . . .	13
3.1.2. Aplicación de los métodos de prorratio y de distribución proporcional. . . . .	15
3.1.3. Aplicación de los métodos de correlación de razón y de diferencias. . . . .	18
3.1.4. El uso de la estratificación por tamaño de las áreas en los modelos de correlación . . . . .	20
3.1.5. El uso de variables Dummy, incorporando la condición urbano/rural, en los modelos de correlación . . . . .	21
3.1.6. El uso de la estratificación por tamaño y condición urbano/rural en los modelos de correlación. . . . .	23
3.1.7. Conclusiones generales. . . . .	24
3.2. El caso de Costa Rica. . . . .	25
3.2.1. Información básica. . . . .	25
3.2.2. Aplicación de los métodos de prorratio y de distribución proporcional. . . . .	26
3.2.3. Aplicación de los métodos de correlación de razón y de correlación de diferencias. . . . .	28
4. CONCLUSIONES Y RECOMENDACIONES. . . . .	30
5. POSIBLES LINEAS DE INVESTIGACION. . . . .	32
6. BIBLIOGRAFIA. . . . .	34

## 1. INTRODUCCION.

Las estrategias de desarrollo vigentes en los países de América Latina cada vez más demandan 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.

Por lo anterior, el Centro Latinoamericano de Demografía, CELADE, tiene dentro de sus preocupaciones el desarrollar métodos para actualizar la población correspondiente a áreas pequeñas. Actualmente las estimaciones de población para dichas áreas son elaboradas básicamente por medio de extrapolaciones matemáticas de niveles y tendencias, observados en un pasado cercano. No obstante, estos procedimientos en la medida que se alejan del período base, suelen estar lejos de la realidad. Ello como efecto de los cambios de tendencia y nivel de las variables demográficas, sobre todo de la migración en períodos cortos. Por tanto, el uso de metodologías para determinar el número de habitantes basadas en tendencias pasadas son limitadas, a menos que se puedan evaluar y actualizar periódicamente.

Al respecto, como una alternativa útil se propone, **para determinar el número de habitantes en áreas pequeñas y evaluar sus resultados, la aplicación de métodos basados en indicadores indirectos del tamaño de la población, a través de modelos de regresión, cuya variable dependiente es la población a estimar y las variables independientes son, por ejemplo, los nacimientos, las defunciones, la matrícula escolar, etc.** <sup>1/</sup>

---

<sup>1/</sup> Dentro de la bibliografía revisada cabe destacar: Crosetti y Schmitt (1954); Grier y Schmitt (1966); Verma, Basavarajappa y Bender (1982); Teixeira Jardim (1992), y; Long (1993).

Este tipo de metodologías han sido utilizadas en los Estados Unidos, Canadá y, en nuestra región, en el caso de un Estado de Brasil. De los métodos revisados se estudiarán aquellos que pueden aplicarse o adaptarse a la realidad Latinoamericana, considerando especialmente las limitaciones de información.

Otro aspecto pertinente a considerar para la utilización de estos métodos es su fácil aplicación, permitiendo, por un lado, la actualización de información de población en áreas pequeñas y, por otro lado, 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 fueran consideradas en la elaboración de los modelos.

Producto de las investigaciones bibliográficas, y de una evaluación de la información disponible en Chile, este trabajo se circunscribe a la descripción de cuatro métodos -unos sencillos y otros más sofisticados- que podrán aplicarse según la calidad y cantidad de información disponible. Tales métodos requieren registros periódicos de variables como: nacimientos, defunciones, registro de automóviles, impuestos, construcción de viviendas, matrícula escolar, etc.. En la mayoría de los países de América Latina no hay registros de dichas variables, y si los hay suelen tener problemas de omisión diferencial por área, interrupción de registro, alteraciones frecuentes en la división político-administrativa, cambios de sistemas económicos, programas para suprimir déficits (vivienda, educación, salud, servicios básicos, etc.), o simplemente no son tabulados. Sin embargo, con pequeñas alteraciones en los procesos de estos registros, se puede obtener la información necesaria para la aplicación de tales métodos.

Por lo anterior, la aplicación se hará, en un principio, para estimar la población de las comunas chilenas en 1992 y en cantones de Costa Rica en 1984 (años del último censo) y, en el caso de

Chile, su posterior evaluación y comparación con las proyecciones anteriormente realizadas.

## **2. METODOS Y MODELOS UTILIZADOS.**

### **2.1. Información básica.**

El primer paso fue investigar la mayor cantidad de información disponible, tratando de obtener el máximo posible de las referencias acerca de las variables sintomáticas.<sup>2/</sup>

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

Al respecto, cabe destacar la enorme importancia que tiene el "contacto" con la información, su manipulación previa, el análisis crítico de los sistemas de recolección, o sea, lograr una profunda familiarización con la información a utilizar. En efecto, ello permite detectar, por ejemplo, posibles errores de estimación atribuidos a problemas de la calidad de información diferencial por área, o a la compatibilización entre períodos, entre otros. Es más, sin duda alguna, buena parte del tiempo y esfuerzo empleado en la investigación debe orientarse a la evaluación exhaustiva de la información, pues de ella dependerá la adecuada elección del modelo.

---

<sup>2/</sup> Variables sintomáticas están relacionadas con los cambios en el tamaño de la población, tales como: nacimientos, defunciones, permisos para construir o viviendas contruidas, impuesto recolectado (IVA), 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 Seguridad Social, consultas médicas, superficie sembrada o plantada, consumo de energía eléctrica, etc. Los nacimientos y las defunciones son considerados acá como "variables sintomáticas", o sea, como indicadores del tamaño de la población, y no como componentes de la dinámica poblacional.

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

## **2.2. Métodos y modelos.**

De los métodos y modelos estudiados, con la bibliografía disponible y tomando en cuenta la información recolectada y evaluada, se seleccionaron los que a continuación se describen.

Las respectivas propuestas metodológicas se referirán a tres momentos específicos. Ellos son:

- 1) momento 0: corresponde al año del primer censo;
- 2) momento t: corresponde al año del segundo censo; y,
- 3) momento (t+n): corresponde al año de la estimación.

### **2.2.1. Distribución por prorratio.**

Este 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 que se pretende estimar la población del área, y la estimación de la población total de las áreas para dicho año.

Así se tiene que 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;

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

Se pueden destacar como ventajas de este método:

a) necesita información para un único momento (el de la estimación) por lo que puede disponerse de un mayor número de variables sintomáticas a ser analizadas (no se necesita una serie histórica de la variable elegida);

b) no es necesario compatibilizar las áreas geográficas en el tiempo; y

c) no hay cambios de definición o forma de recolección de las variables en el tiempo.

Como desventajas se tiene que:

a) 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;

b) las estimaciones están afectadas por la calidad diferencial -por área- de la variable sintomática.

A pesar de sus desventajas, si hay disponibilidad de un conjunto de variables sintomáticas para el período a estimar, se podría entonces obtener varias estimaciones para la población, cuya media

aritmética -de estas estimaciones- se aproxime más a la realidad.<sup>3/</sup>

### 2.2.1. Distribución proporcional.

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

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

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

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

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

donde:

$P_{(i,t+n)}$ ;  $P_{(.,t+n)}$ ;  $S_{(i,t+n)}$  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 a la formula para que la suma de las  $P_{(i,t+n)}$  sea igual a  $P_{(.,t+n)}$ .

---

<sup>3/</sup> Ver Lahu y Namboodiri (1971), que sugieren la media de una serie de regresiones simples como alternativa a la regresión múltiple. En el caso de Chile y Costa Rica se aplicó la media a las distintas estimaciones obtenidas con cada método, y esto condujo a mejores estimaciones.

Ventajas del método:

- a) es sencillo, dado que se basa en una única variable sintomática;
- b) la incorporación de los cambios de la variable sintomática permite llevar a un mayor número de estimaciones; o sea, se podrían usar 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:

- a) necesita mayor disponibilidad de información que el anterior;
- b) se necesita compatibilización geográfica en los dos momentos, tanto a nivel de población como de la variable sintomática;
- c) las estimaciones están afectadas por cambios en la calidad de la información de la variable sintomática, si esta no es constante en el tiempo.

También en este modelo se puede calcular un conjunto de estimaciones y calcular la media de éstas, como una estimación probablemente más adecuada.

### 2.2.3. Correlación de razón. (Crosetti y Schmitt, 1954).

Este método 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:

$$Y_{(i,t)} = a_{(0)} + a_{(1)} * X1_{(i,t)} + \dots + a_{(n)} * Xn_{(i,t)} + e_i \quad (4)$$

donde:

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

$$X^j_{(i,t)} = \frac{\frac{S^j_{(i,t)}}{S^j_{(.,t)}}}{\frac{S^j_{(i,0)}}{S^j_{(.,0)}}} \quad (6)$$

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

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

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

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

$$P_{(i,t+n)} = \hat{Y}_{(i,t+n)} * \frac{P_{(i,t)}}{P_{(.,t)}} * P_{(.,t+n)} \quad (7)$$

donde:

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

Ventajas del método:

a) la principal ventaja es que se basa en una o más variables cuya evolución explica la mayor parte de la variación de la población;

- b) las estimaciones no están afectadas por errores de calidad en la información, siempre y cuando éstas se mantengan en el tiempo o cambien en forma similar en todas las áreas consideradas;
- c) 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:

- a) necesidad de disponibilidad de un mayor número de información;
- b) necesidad de compatibilizar la información en los dos momentos utilizados para estimar el modelo;
- c) necesidad de compatibilizar la información en los dos momentos necesarios para hacer la estimación.

#### 2.2.4. Correlación de diferencia. (O'Hare, 1976).

Este método es similar al anterior, se basa en el mismo supuesto, y la diferencia consiste en la forma en que se calculan las variaciones; las razones son sustituidas por diferencias, así se tiene:

$$W_{(i,t)} = a_{(0)} + a_{(1)} * Z1_{(i,t)} + \dots + a_{(n)} * Zn_{(i,t)} + \epsilon_{(i)} \quad (8)$$

donde:

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

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

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

$Z_{j(i,t)}$  = 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:

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

donde:

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

Este método presenta las mismas ventajas y desventajas del método anterior. Sin embargo, O'Hare (1976) sugiere calcular las variaciones de proporción por medio de diferencias, pues así se tendría una mayor intercorrelación entre las variables (W y  $Z_j$ ) 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 importante para aquellos casos en que se considera que un modelo único no representa adecuadamente a todas las divisiones geográficas, para las cuales se desea estimar la población. Se podrían generar modelos de regresión para grupos homogéneos de áreas, agrupadas, por ejemplo, según grandes regiones, según predominancia urbana o rural, según la actividad económica principal, etc.

### **3. APLICACIONES.**

#### **3.1. El caso de Chile.**

##### **3.1.1. Información básica.**

Las aplicaciones con datos de Chile se llevaron a cabo para la estimación de 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 ennumerar 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.

1. La división político administrativa de Chile sufrió cambios a todos los niveles en los años 80. A pesar de que las comunas existentes en 1970 siguen existiendo en 1982, prácticamente todas cambiaron sus límites geográficos.

2. La información sobre registros de nacimientos y defunciones eran codificadas y publicadas hasta 1985 a nivel de circunscripciones de registro civil que no corresponden a comunas; sin embargo, para el año de 1985 fue hecha una equivalencia entre las circunscripciones de registro civil y las comunas. Esta equivalencia fue utilizada para obtener los nacimientos y defunciones para los años de 1970 y 1982. Cabe señalar que las circunscripciones de registro civil también sufrieron cambios en sus límites geográficos; tales cambios no fueron evaluados y afectarán la calidad de las estimaciones.

3. La información referente a los permisos para construir o área construida para habitación, actualmente, solo están disponibles para algunas comunas, aunque es factible 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. Ello motivó no considerar esta variable en el presente estudio, aunque debería analizarse su posibilidad en el futuro.

4. La información sobre los registros de automóviles están disponibles por comuna, sin embargo se presentan problemas en compatibilizar la información de comunas existentes en los años base de los modelos de correlación (razón y diferencia). Adicionalmente está el problema de la calidad de la información en las comunas de los grandes centros urbanos, en los cuales se permite el registro del automóvil en una comuna distinta a la de residencia del propietario.

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

6. 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 en períodos previos a elecciones y, adicionalmente, durante un período de aproximadamente 15 años no hubieron elecciones. Sin embargo, podrá ser utilizado en trabajos futuros.

7. El consumo de energía eléctrica dejó de ser publicado en 1989; hasta esa fecha era responsabilidad de la compañía de energía eléctrica la recolección y posterior publicación. Actualmente el Instituto Nacional de Estadística (INE) está diseñando una encuesta para obtener tal información, pudiendo, entonces, ser utilizada a futuro.

8. Las demás variables citadas (afiliados a Asociaciones de Fondos de Pensiones, recaudación de impuestos, etc.) no están disponibles,

o si están no se encuentran a nivel de comunas, aunque sí de provincias o de alguna división distinta a la división político administrativa, o sea, según la forma que fue dividido el país para fines de recolección de la información.

9. La población del censo de 1970 fue adaptada por el INE a la división político administrativa de 1982, trabajo considerado de buena calidad.

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

- 1) La población total por comunas readecuada para 1970, la población censada en 1982 y en 1992;
- 2) Como variables sintomáticas: nacimientos y defunciones para los años de 1970, 1982 y 1992; la matrícula en educación básica para los años de 1982 y 1992; los registros de automóviles para los años de 1982 y 1992, y; los eventos vitales, que corresponde a la suma de nacimientos y defunciones, para los años de 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<sup>4/</sup>

### **3.1.2. 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 aplicaron con cuatro variables: nacimientos (modelo A);

---

<sup>4/</sup> Para mayores antecedentes metodológicos consultar: INE-CELADE, Chile: Proyecciones de población por sexo y edad, Comunas 1980-1995. Fascículo F/CHI 5.

defunciones (modelo B); matrícula en educación básica (modelo C), y; eventos vitales (modelo D).

Además de estos cuatro modelos se calcularon estimaciones promedios: 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 indicadores de la calidad de los resultados de estos modelos, se resume en los cuadros 1 y 2.

CUADRO 1

CHILE: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR COMUNA SEGUN LOS MODELOS DE PRORRATEO - 1992.

MODELO	ERROR ABSOLUTO $\frac{a}{}$ (%)		PORCENTAJE DE COMUNAS CON ERROR ABSOLUTO SUPERIOR AL 10%	80 % DE LAS COMUNAS CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESVIO STANDAR		
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

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

$\frac{a}{}$  (error absoluto) =  $|(población estimada - población censada) \div población censada| * 100$ .

En el cuadro 1 se puede observar que:

- 1) de las variables sintomáticas que se disponían, los eventos vitales proporcionan las mejores estimaciones de población;
- 2) las peores estimaciones son las calculadas con base a las defunciones;
- 3) al considerar los modelos promedio, las mejores estimaciones resultan de los modelos G y H, y;

4) cabe destacar que, pesar de las limitaciones del método y también de las variables sintomáticas disponibles, se lograron estimaciones más precisas que las obtenidas con las proyecciones aludidas.

CUADRO 2

CHILE:ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR COMUNA SEGUN LOS MODELOS DE DISTRIBUCION PROPORCIONAL - 1992.

MODELO	ERROR ABSOLUTO $\alpha$ / (%)		PORCENTAJE DE COMUNAS CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LAS COMUNAS CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESVIO STANDAR		
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

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

$\alpha$ / (error absoluto) =  $|(población estimada - población censada) + población censada| * 100$ .

Observando el cuadro 2 se destaca:

- 1) considerando las estimaciones simples con una variable, la matrícula escolar y los eventos vitales proporcionan estimaciones más plausibles;
- 2) las estimaciones basadas en defunciones resultan mejores que las obtenidas con defunciones en el método de prorrateo;
- 3) si se consideran ahora los modelos promedio (E, F, G y H), las mejores estimaciones las proporcionan los modelos G y H. Los cuales son mucho más adecuadas que los anteriores y que las proyecciones.

Además de esto, se destaca también que, si se dispone de un conjunto de variables sintomáticas cuyas evolución se relacionada

con los cambios en el tamaño de la población, se puede contar con métodos sencillos para la actualización de población.

### **3.1.3. Aplicación de los métodos de correlación de razón y de diferencias.**

En estos métodos se necesita mayor cantidad de información:

- a) la población y variables sintomáticas por área compatible en dos momentos, para estimar el modelo de regresión;
- b) la población y variables sintomáticas por área en un momento y las variables sintomáticas por área en el momento de la estimación, y;
- c) una estimación para el total de las áreas.

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

La aplicación utilizó solamente la información de nacimientos, defunciones y, eventos vitales como variables sintomáticas.

Primeramente se estimaron los modelos de regresión (4) y (8)<sup>5/</sup> considerando el total de comunas, y los nacimientos y las defunciones como variables sintomáticas (modelo A). Posteriormente se estimaron dichas regresiones 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 condidera el total de comunas y como variable sintomática los "eventos vitales" (suma de nacimientos y de defunciones).

---

<sup>5/</sup>Referentes al capítulo 2, secciones 2.2.3 y 2.2.4.

CUADRO 3

CHILE: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR COMUNA SEGUN LOS MODELOS DE CORRELACION - 1992.

METODO/ MODELO	ERROR EBSOLUTO $\bar{a}$ / (%)		PORCENTAJE DE COMUNAS CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LAS COMUNAS CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESVIO STANDAR		
<b>CORRELACION DE RAZON</b>				
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
<b>CORRELACION DE DIFERENCIAS</b>				
A. Nacimientos y defunciones	11.8	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

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

 $\bar{a}$  (error absoluto) =  $|(\text{población estimada} - \text{población censada}) + \text{población censada}| * 100$ .

El cuadro 3 resume los resultados. Al observarlo se destaca:

- 1) las estimaciones encontradas en los diversos modelos son bastante similares al de las proyecciones;
- 2) el método de correlación de diferencias proporciona estimaciones ligeramente mejores que el de correlación de razón;
- 3) el hecho de que el modelo B produzca las peores estimaciones, puede ser consecuencia de la calidad de la información básica que lleva a pensar que la compatibilización hecha para las defunciones en el año 1970 presenta mayores problemas;
- 4) el uso de modelos promedio (modelo D) no condujo a cambios importantes en las estimaciones, y, finalmente;
- 5) 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 realizadas resultaron similares a las proyecciones.

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

#### **3.1.4. El uso de la estratificación por tamaño de las áreas en los modelos de correlación<sup>6/</sup>.**

Para probar los modelos de correlación a conjuntos más homogéneos, se estratificaron las comunas según la cantidad de habitantes en 1982. En efecto se construyeron los modelos y posteriormente la estimación de la población en tres grupos de comunas:

- a) estrato 1 - comunas con población inferior a 10000;
- b) estrato 2 - comunas con población entre 10000 y 50000; y,
- c) estrato 3 - comunas con población superior a 50000 habitantes.

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

---

<sup>6/</sup> Ver Ericksen (1973); Martín y Serow (1978), y; Teixeira Jardim (1992).

CUADRO 4

CHILE: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR COMUNA SEGUN LOS MODELOS DE CORRELACION, ESTRATICADO POR TAMAÑO DE LAS COMUNAS - 1992.

METODO/ MODELO	ERROR ABSOLUTO $\bar{a}$ / (%)		PORCENTAJE DE COMUNAS CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LAS COMUNAS CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESUDIO STANDAR		
<b>CORRELACION DE RAZON</b>				
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
<b>CORRELACION DE DIFERENCIAS</b>				
A. Nacimientos, defunciones y tamaño	14.0	22.6	46.4	18.3
B. Nacimientos y tamaño	13.3	21.2	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

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

$\bar{a}$ / (error absoluto) = [(población estimada - población censada) ÷ población censada] \* 100.

En este cuadro se destaca:

- 1) las mejores estimaciones se obtienen através de la aplicación del método de correlación de razón;
- 2) los modelos B y E presentan las mejores estimaciones; y,
- 3) la estratificación hecha, considerando solamente el tamaño de las comunas, produce los mejores resultados hasta ahora obtenidos.

**3.1.5. El uso de variables Dummy, incorporando la condición urbano/rural, en los modelos de correlación.**

Con el objetivo de mejorar las estimaciones se introdujo el concepto de localidad urbana, incorporando una variable a los modelos de correlación, a la cual se le atribuye valor 1 (uno) si en la comuna predomina la población urbana y 0 (cero) si predomina

la población rural. Los resultados se presentan en el siguiente cuadro.

CUADRO 5

CHILE: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR COMUNA SEGUN LOS MODELOS DE CORRELACION, ESTRATIFICADO POR CONDICION PREDOMINANTE URBANO/RURAL DE LAS COMUNAS - 1992.

METODO/ MODELO	ERROR ABSOOLUTO $\bar{a}$ / (%)		PORCENTAJE DE COMUNAS CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LAS COMUNAS CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESUDIO STANDAR		
<b>CORRELACION DE RAZON</b>				
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
<b>CORRELACION DE DIFERENCIAS</b>				
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	31.3	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

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

$\bar{a}$ / (error absoluto) =  $|(población estimada - población censada)| + población censada \cdot 100$ .

A partir d el cuadro 5 se puede concluir que:

- 1) las mejores estimaciones se obtienen con la aplicación de método de correlación de razón;
- 2) el método de correlación de diferencias tiende a empeorar sus estimaciones;
- 3) los modelos A, B, D y E, muestran las mejores estimaciones.

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

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

CUADRO 6

CHILE: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR COMUNA SEGUN LOS MODELOS DE CORRELACION, ESTRATIFICADO POR TAMAÑO Y CONDICION PREDOMINANTE URBANO/RURAL - 1992.

METODO/ MODELO	ERROR ABSOLUTO $\bar{a}$ / (%)		PORCENTAJE DE COMUNAS CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LAS COMUNAS CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESVIO STANDAR		
<b>CORRELACION DE RAZON</b>				
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
<b>CORRELACION DE DIFERENCIAS</b>				
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

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

$\bar{a}$ / (error absoluto) = [(población estimada - población censada) ÷ población censada] \* 100.

Observando el cuadro 6, se puede decir que:

1) el método de correlación de razón, produce las estimaciones más plausibles;

2) el modelo E, utilizando el método de correlación de razón, presentó las mejores estimaciones de los modelos considerados.

### **3.1.7. Conclusiones generales.**

Se puede concluir, con base al caso chileno, que:

- 1) los modelos que utilizan variables sintomáticas permiten construir un sinnúmero de modelos;
- 2) 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;
- 3) los modelos de correlación de diferencias no condujeron a buenos resultados;
- 4) en los modelos de correlación las estimaciones promedios, basadas en regresiones simples, no introdujeron cambios importantes en la precisión de las estimaciones;
- 5) en los modelos de prorrateo y distribución proporcional resultan mejores las estimaciones promedio; y,
- 6) el modelo E, estimado por el método de correlación de razón, incorporando la estratificación según tamaño, y también la condición de localidad urbana, presentó las mejores estimaciones.

### **3.2. El caso de Costa Rica.**

#### **3.2.1. Información básica.**

Con el objetivo de evaluar y comparar, los métodos antes descritos, siempre en el contexto Latinoamericano, se hizo una aplicación para Costa Rica, esto permite analizar un país de menor tamaño con relación a Chile, y que muestra distintas características en cuanto a problemas de información.

En este caso la aplicación se llevó a cabo para la estimación de la población de los cantones existentes en 1984, comparándolas 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 disponibles en CELADE para dicho país. Además, se pretendió aplicar modelos semejantes al caso anterior con el fin de averiguar si se podrán obtener mejores estimaciones cuando hay mayor compatibilidad de las áreas en el tiempo.

Por lo anterior, considerando las limitaciones aludidas, se utilizó la siguiente información:

- 1) población censada en los tres últimos censos (1963, 1973 y 1984);
- 2) nacimientos y defunciones registrados y ocurridos hasta 10 años antes; y,
- 3) población que asistía a establecimientos regulares de enseñanza al momento del censo, tomada como una aproximación a la matrícula escolar.

Respecto de la información utilizada cabe destacar los siguientes problemas:

- 1) los nacimientos y defunciones, presentan problemas de fluctuaciones debidas a la forma de tabulación, referente a 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.

2) la información sobre asistencia escolar, la cual reemplaza la matrícula escolar, produce un sesgo entre área o localidad donde se ubica el establecimiento escolar y en el que reside el estudiante: la primera, asociada a la residencia del estudiante y, la segunda, al establecimiento de enseñanza. Se espera que la correlación entre matrícula escolar y población sea menor que la estimada entre asistencia escolar y población.

Finalmente, en el caso de Costa Rica, no se compararon las estimaciones para el año de 1984 generadas a través de los distintos métodos con las proyecciones de población a nivel de cantones, por cuanto éstas no estaban disponibles.

### **3.2.2. 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 (nacimientos, defunciones, eventos vitales y asistencia escolar) y también se analizaron las estimaciones promedios, 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 fue observado en el caso chileno).

Al considerar los modelos promedios, las estimaciones más adecuadas se producen con los modelos E y H para el método de prorrateo y con el modelo H para el método de distribución proporcional.

CUADRO 7

COSTA RICA: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR CANTON SEGUN LOS MODELOS DE PRORRATEO - 1984.

MODELO	ERROR ABSOLUTO a/ (%)		PORCENTAJE DE CANTONES CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LOS CANTONES CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESVIO STANDAR		
A. Nacimientos	12.4	11.0	45.7	20.9
B. Defunciones	13.7	10.0	53.1	23.1
C. Matricula 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

Nota: Se consideró la estimación de 81 cantones. La matricula fue aproximada según la asistencia escolar.  
 a/ (error absoluto) =  $|(población\ estimada - poblacion\ censada) \div poblacion\ censada| * 100$ .

CUADRO 8

COSTA RICA: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR CANTON SEGUN LOS MODELOS DE DISTRIBUCION PROPORCIONAL - 1984.

MODELO	ERROR ABSOLUTO a/ (%)		PORCENTAJE DE CANTONES CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LOS CANTONES CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESVIO STANDAR		
A. Nacimientos	14.0	9.1	59.3	22.8
B. Defunciones	18.2	14.9	59.3	30.7
C. Matricula	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

Nota: Se consideró la estimación de 81 cantones. Matricula aproximada según la asistencia escolar.  
 a/ (error absoluto) =  $|(población\ estimada - poblacion\ censada) \div poblacion\ censada| * 100$ .

### 3.2.3. 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 sufrieron grandes cambios geográficos en el período del estudio, lo que permitió incorporar a los modelos de correlación una variable más, o sea, la matrícula escolar.

El análisis de los resultados de la aplicación de 10 modelos se presentan en el cuadro 9, donde se destaca que:

- 1) los modelos de razón producen mejores estimaciones con respecto a los de diferencia;
- 2) el modelo C proporciona las estimaciones más adecuadas con el método de correlación de razón y, el modelo B con el método de correlación de diferencias; y,
- 3) los peores resultados son los obtenidos por medio del modelo E con el método de correlación de razón y, el 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, utilizada en el caso de Chile, no significaron para Costa Rica, mejoría en las estimaciones, por lo tanto no se presentan los resultados referentes a estos modelos. Sin embargo, es posible que otras estratificaciones, más adecuadas a la situación de Costa Rica, produzcan mejores estimaciones.

CUADRO 9

COSTA RICA: ANALISIS DEL ERROR DE LAS ESTIMACIONES DE POBLACION POR CANTON SEGUN LOS MODELOS DE CORRELACION - 1984.

METODO/ MODELO	ERROR ABSOLUTO $a/$ (%)		PORCENTAJE DE CANTONES CON ERROR ABSOLUTO SUPERIOR AL 10%	80% DE LOS CANTONES CON ERROR ABSOLUTO INFERIOR A:
	MEDIA	DESUDIO STANDAR		
<b>CORRELACION DE RAZON</b>				
A. Nacimientos y defunciones	9.6	8.5	39.5	16.7
B. Nac., def., y Matricula	4.7	3.7	7.4	7.3
C. Eventos vitales y matricula	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. Matricula	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
<b>CORRELACION DE DIDERENCIAS</b>				
A. Nacimientos y defunciones	9.9	9.1	37.0	17.4
B. Nac., def. y matricula	5.9	4.9	16.0	9.5
C. Eventos vitales y matricula	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. Matricula	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

Nota: Se consideró la estimación de 81 cantones. La matricula fue aproximada según la asistencia escolar.  
 $a/$  (error absoluto) =  $|(población estimada - población censada) \div población censada| * 100$ .

En resumen, para el caso de Costa Rica, se obtuvieron estimaciones de la población muy razonables con los métodos de distribución porporcional (modelo H) y de correlación de razón (modelos B,C y J), usando como variables sintomáticas nacimientos, defunciones, eventos vitales y matrícula escolar. Esto lleva a pensar que, si no hubiera en el caso chileno el problema de la compatibilización geográfica de las comunas, se obtendrían mejores estimaciones.

#### 4. CONCLUSIONES Y RECOMENDACIONES.

1. Contar con mejores estimaciones de población a nivel de área menor<sup>2/</sup> (Local, Municipal, etc.) está dada por el hecho de que cada día cobra más relevancia en lo económico, político, cultural, etc., el peso de la comunidad. Sea para la elaboración de planes de gobierno, o bien porque ejercen presión a 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 necesario y urgente la actualización y precisión de la población de áreas menores. Por lo tanto es válido todo el esfuerzo que se hace para obtener mejores estimaciones de población.

2. Si se dispusiera de registros frecuentes del movimiento migratorio entre estas áreas, de los nacimientos y de las defunciones, estaría resuelto 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. Contar con métodos alternativos que indiquen la variación de la población de áreas menores tiene gran importancia en la elaboración de estimaciones y proyecciones de población.

3. Censar la población más frecuentemente, sería otra alternativa, pero es poco factible por los altos costos que implica esta operación. Cabe recordar que a pesar de la recomendación de Naciones Unidas para realizar censos cada 10 años, muchos países de América Latina los llevan a cabo en intervalos superiores.

4. Por lo anterior, se cree que una alternativa eficiente y eficaz sea el mejorar los procedimientos de registros administrativos ya

---

<sup>2/</sup> Ver trabajos Flores (1989) y Rueda (1989).

existentes, para su posterior utilización en actualización de las estimaciones de población.

5. En este trabajo, basado en los dos casos estudiados, se demuestra que, a pesar de los problemas existentes de disponibilidad y calidad de información, se puede llegar a estimaciones bastante razonables de la población de las áreas consideradas.

6. Respecto a los métodos aplicados, puede concluirse que hay que hacer todo lo posible para utilizar el **método de correlación de razón**, pues en ambos casos presentó las mejores estimaciones<sup>8/</sup>.

7. Seleccionar el modelo no es tarea fácil, dependerá de la calidad y disponibilidad de información y también de las características particulares de las áreas a estimar. Se considera bastante útil elaborar un estudio previo, como el acá presentado, para una evaluación a priori de la situación en cada caso concreto, y familiarizarse con la información y los distintos métodos, a fin de escoger el modelo "óptimo" a ser utilizado. Además, resultará muy conveniente realizar evaluaciones a posteriori, lo que permitirá introducir cambios en el modelo utilizado.

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

9. Se hace necesaria la integración de las diversas fuentes de información y el organismo responsable de las estimaciones de población. Muchas veces introducir pequeñas modificaciones en el

---

<sup>8/</sup> En su reciente visita al CELADE, Maria Andrassy-Bitto -responsable de las estimaciones de población de áreas menores en Canadá-, sugirió probar la estabilidad de los modelos de correlación (coeficientes de regresión). En caso que se encuentre que los modelos no sean estables en el tiempo, sobretudo en los modelos de correlación de diferencias, se podrá corregirlos mejorando los modelos y, conseqüentemente las estimaciones.

proceso de recolección de información y/o tabulación de la misma, abre una gama de posibilidades de construcción de modelos, sobre todo de criterios para estratificación de las áreas a estimar. De esta forma se podría disponer de un sistema confiable -y de bajo costo- para la realización de estimaciones de población.

10. Finalmente, se puede concluir que es plausible el uso de variables sintomáticas en la estimación de población de áreas menores en América Latina. Otro ejemplo que apunta en esta dirección es la aplicación hecha por Teixeira Jardim <sup>9/</sup> al Estado de Rio Grande do Sul de Brasil; este trabajo establece un punto de partida en la divulgación y utilización de los métodos acá expuestos para América Latina. Es por ello que, basados en estas experiencias y sabiendo que no se tiene un camino fácil por delante, se cree pertinente incorporar a nuestra región experiencias de Estados Unidos y Canadá para tener en un futuro cercano una alternativa más confiable de actualización de población.

##### 5. POSIBLES LINEAS DE INVESTIGACION.

Como una forma de concluir el trabajo, considerando que los numerales anteriores han sido conclusiones y recomendaciones generales, es pertinente enumerar ciertas líneas más específicas a profundizar en el futuro:

1. Se hace necesario profundizar en el conocimiento y divulgación de la teoría estadística implícita en cada método, a fin de tener un mayor soporte en el momento de decidir qué modelo es el más adecuado.<sup>10/</sup>

---

<sup>9/</sup> ver trabajo de Teixeira Jardim (1992).

<sup>10/</sup> Ver Bravo (1993) y Verma y Basavarajappa (1985).

2. Para el caso chileno, mejorar los modelos buscando nuevos criterios de estratificación en las comunas seleccionadas para la estimación, así como, estimar modelos por regiones.

3. Debería promoverse más aplicaciones para otros países de la región, pero con presencia directa 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.

4. Se podría explorar la aplicación de otras metodologías existentes, entre ellas el método "Componentes II", el cual utiliza la ecuación compensadora por cohortes para determinar la población en cada momento, pero considera variables sintomáticas para estimar las migraciones. <sup>11/</sup>

5. Por último, pero quizás lo más urgente, es llevar a cabo una extensa y minuciosa recopilación de los trabajos realizados sobre el tema, hacer una adecuada selección y compilarlos en una publicación. Ello permitiría abrir nuevos horizontes, ya sea para promover la crítica sobre el perfeccionamiento de los métodos o para ampliar el campo de posibilidades de la aplicación.

---

<sup>11/</sup> Batutis, Jr. (1991) "Estimates Preparation: Methods and Procedures". Proposición del SR. John Long en su reciente visita al CELADE.

## 6. BIBLIOGRAFIA.

- BRAVO, Jorge. 1993. Algunas observaciones acerca de los métodos de estimación de población a través de variables sintomáticas. Celade mimeo.
- BASAVARAJAPPA, K. G.; BENDER, Rosemary y VERMA, Ravi B. P. 1982. New approaches to methods of estimating the population of census divisions. Demography Division, Statistics Canada.
- BASAVARAJAPPA, K. G. y VERMA, Ravi B. P. 1985. Recent developments in the estimation of population by the regression method. Statistics Canada. Documento presentado en "International Symposium on Small Area Statistics, Ottawa, Ontario, Canada, May 22-24, 1985".
- CROSETTI, A. H. y R. C. SCHMITT. 1954. Accuracy of the ratio-correlation method for estimating postcensal population. Land Economics 30:279-281.
- ERICKSEN, Eugene P. 1973. A method for combining sample survey data and syntomatic indicators to obtain population estimates for local areas. Demography 10:137-160.
- 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, Octubre 31 a noviembre 2 de 1988.
- GOLDBERG, David, N. K. NAMBOODIRI y V. R. RAO. 1964. A test of the accuracy of ratio correlation population estimates. Land Economics 40:100-102.
- GRIER, John M. y Robert C. SCHMITT. 1966. A method of estimating the population of minor civil divisions. Rural Sociology 31:355-361.
- KMENTA, Jan. 1978. Elementos de econometria. 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 36:187-194.
- LONG, John F. 1993. Methods for postcensal population estimation: flow methods vs. stock methods. General Conference of the International Union for the Scientific Study of Population, Montreal, Canada, August 24 - September 1, 1993.
- MANDELL, Marylou y Jeffrey TAYMAN. 1982. Mesuaring temporal stability in regression models of population estimation. Demography 19:135-146.
- MARTIN, Julia H. y Willian J. SEROW. 1978. Estimating demographic characteristics using the ratio-correlation method. Demography 15:223-233.
- NAMBOODIRI, M. Krishnan. 1972. On the ratio-correlation and related methods of subnational population estimation. Demography 9:443-453.
- NORUSIS, Marija J. 1988. SPSS/PC+ V2.0 - Base manual. SPSS Inc.
- O'HARE, Willian P. 1976. Report of a multiple regression method for making population estimates. Demography 13:367-379.
- O'HARE, Willian P. 1980. A Note on the use of regression methods in population estimates. Demography 17:341-343.
- PURSELL, Donald E. 1970. Improving population estimates with the use of Dummy variables. Demography 7:87-91.
- 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, Octubre 31 a noviembre 2 de 1988.
- SHRYOCK, Henry S. jr. y Meyer ZITTER. 1964. Accuracy of methods of preparing postcensal population estimates for states and local areas. Demography 1:227-241.
- SWANSON, David A. 1980. Improving accuracy in multiple regression estimates of population using principles from causal modelling. Demography 17:413-427.
- SWANSON, David A. y Lucky M. TEDROW. 1984. Improving the mesurement of temporal change in regression models used for county population estimates. Demography 21:373-381.
- TEIXEIRA JARDIM, Maria de Lourdes. 1992. Utilizacao de variáveis sintomáticas para estimar a distribuicao espacial de populacoes: aplicacao aos municípios do Rio Grande do Sul. Anais del VIII Encontro Nacional de Estudos Populacionais - ABEP 1:39-50.

INE, Instituto Nacional de Estadística. 1993. Resultados Generales Censo 1992.

INE, Instituto Nacional de Estadística. 1970, 1982 y 1992 Nacidos vivos y defunciones generales según lugar de residencia. Demografía, 1970:8-17; 1982:8-29; 1991:36-50.

INE, Instituto Nacional de Estadística. 1982, 1992. Parque de vehiculos en circulación. Anuario de Transporte y Comunicación, 1982:118-134; 1990:4-32.

INE, Instituto Nacional de Estadística. 1970. Censo Nacional de Población y Vivienda - Localidades Pobladas. Censo Nacional de Población y Vivienda I a XII región y Región Metropolitana.

INE, Instituto Nacional de Estadística. 1982. Censo Nacional de Población y Vivienda - Localidades Pobladas. Censo Nacional de Población y Vivienda I a XII región y Región Metropolitana.

INE, Instituto Nacional de Estadística. Listado de codigos de comuna en vigencia a partir de 01/07/85 y su correspondencia con antiguo codigo de circunscripción de registro civil. mimeo.

MINISTERIO DE EDUCACION. 1982 y 1992. Matricula Escolar por comunas. Informe de los establecimientos educacionales, remitidos por las las Secretarías Regionales Ministeriales de Educación.

CELADE y INE. CHILE: Proyecciones y estimaciones de población, por sexo y edad. COMUNAS 1980-1995.

DIRECCION GENERAL DE ESTADISTICA. Ministério de Economía, Industria y Comercio. Anuario Estadístico de Costa Rica 1962, 1963, 1964, 1972, 1973 y 1974.

DIRECCION GENERAL DE ESTADISTICA. Ministério de Economía, Industria y Comercio de Costa Rica. Diez Años de Estadística Vital 1978-1987.