

Estimación de población municipal

1985-2005

REPÚBLICA DE COLOMBIA



DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA

Héctor Maldonado Gómez

Director

Carlos Eduardo Sepúlveda Rico

Subdirector

Alfredo Vargas Abad

Secretario General

Eduardo Efraín Freire Delgado

Director de Metodología y Producción Estadística

Javier Alberto Gutiérrez López

Director de Regulación, Planeación, Estandarización
y Normalización

Bernardo Guerrero Lozano

Director de Censos y Demografía

Ana Victoria Vega Acevedo

Directora de Síntesis y Cuentas Nacionales

Nelcy Araque García

Directora de Geoestadística

Carolina Gutiérrez Hernández

Directora de Difusión, Mercadeo y Cultura Estadística

Coordinadora de Demografía: Myriam Cifuentes Noyes

Asesores: Edgar Sardi Perea, Rafael Cubillos López

Equipo de Trabajo: Erika María Vargas, Jorge Cabezas Zabala, Diana Nayibe Rucinke Gonzalez, Edwin Yovani Castelblanco Cristancho, Diana Marcela Reyes Toledo, Jacqueline Martínez García, Charles Erasmo Daza Malagon, Mariana Francisca Ospina Bohorquez, Luis Alejandro Montenegro Ramirez, Leonel Gutiérrez Ceballos, Diana Rubriche Cárdenas.

Diseño y diagramación: Daniel Andres Velásquez

Departamento Administrativo Nacional de Estadística (DANE)

Carrera 59 No 26-70 Int. 1 CAN

E-mail: dane@dane.gov.co

Bogotá, D.C

Impresión: Imprenta Nacional

CONTENIDO

INTRODUCCIÓN	pág 7
1. ANTECEDENTES	9
2. DISEÑO	10
2.1. Marco de referencia	10
2.1.1. Objetivos	10
2.1.2. Marco teórico	10
2.1.3. Referentes internacionales	11
3. PRODUCCIÓN ESTADÍSTICA	13
3.1. Actividades preparatorias	13
3.1.1. Definición y revisión de los techos iniciales nacional y departamental por área, para el proceso de estandarización	13
3.1.2. Evaluación y ajuste del nivel de población municipal por área	13
3.1.3. Revisión de las estructuras municipales por sexo y grupos de edad	16
3.2. Métodos para la estimación de población municipal	18
3.2.1. Estimación de niveles	18
3.2.2. Estimación de estructuras	19
3.2.3. Estimación de la población municipal, 1985-2005	19
3.2.4. Desarrollos tecnológicos para la interpolación de población	19
3.3. Criterios de control de la calidad	20
4. DIFUSIÓN	22
4.1. Productos e instrumentos de difusión	22
GLOSARIO	23
BIBLIOGRAFÍA	24
ANEXOS	25
Anexo A. El modelo de interpolación con Splines	25

INTRODUCCIÓN

Los resultados del Censo General 2005 evidenciaron una dinámica demográfica diferente a la considerada en las últimas proyecciones basadas en el Censo 1993; esto se refleja en la distribución espacial de la población, en su estructura etaria, en su distribución por sexo y en los principales indicadores determinantes de la dinámica poblacional.

La población proyectada para el año 2005, obtenida con base en la conciliación del Censo 1993, superó en un poco más de tres millones de personas la población resultante del Censo 2005, después del proceso de conciliación. Dicha diferencia se explica por dos grandes factores: una mayor población base para las proyecciones, derivada a su vez de un sobreajuste en la conciliación del Censo 1993 y, en segundo término, la hipótesis de un saldo neto migratorio internacional igual a cero, esto es, que en un periodo determinado el número de personas que llega al país es igual al que sale con destino a otro país para vivir permanentemente o durante un periodo superior a seis meses.

La conciliación censal que se efectuó con posterioridad al Censo 2005, y que se extendió a los censos de 1985 y 1993, no sólo permitió identificar estas diferencias respecto a las proyecciones basadas en el Censo de 1993, sino que derivó posteriormente en el ejercicio de ajustar las cifras a nivel municipal para cada uno de los censos del período de conciliación; ejercicio que incluyó la revisión a nivel de área y por estructura poblacional por sexo y grupos de edad. En su momento, las cifras censales municipales del 2005 se ajustaron a nivel total y por área, dejando para una revisión posterior la validación de sus estructuras poblacionales.

Con estos cambios en las cifras poblacionales de referencia, y ante la necesidad de brindar al usuario información actualizada, el DANE se comprometió con la tarea de revisar y actualizar las estimaciones de población municipales para el periodo 1985-2005. Este proyecto cubría no sólo la estimación del nivel total de población municipal, sino que consideró hacerlo por área (cabecera-resto o urbano-rural) y según estructuras por sexo y edad, de tal manera que el producto final cubriera demandas de información más específicas en cuanto a salud pública, sistema educativo, empleo, seguridad social, y seguimiento de los indicadores de los Objetivos de Desarrollo del Milenio, entre otras.

De otro lado, es importante resaltar que es la primera vez que el DANE lleva a cabo un ejercicio de tal magnitud, es decir, generar estimativos de población a nivel municipal, con desagregación por sexo y grupos de edad. Incluso en la práctica internacional tal detalle en las estimaciones no es común, por lo que la metodología y sus estimaciones a nivel municipal se convierten en un nuevo referente. En muchas latitudes las estimaciones intercensales de población a nivel subnacional, realizadas por la entidad estadística central, tienen un desglose sólo hasta cierto nivel regional.

Es preciso hacer una nota acerca de las limitaciones y advertencias inherentes a un ejercicio de esta naturaleza. Las cifras resultantes son un estimativo de un proceso más o menos armonioso de la población y no pueden ser tomadas como la verdadera población en cada momento del tiempo. Sin embargo, sí se realizó un gran esfuerzo para que las cifras reflejaran la creación de nuevos municipios durante el período de estimación, y el correspondiente ajuste en sus entes segregantes. Una anotación adicional es que bajo la premisa de que los procesos demográficos son armoniosos para las distintas instancias del ejercicio (según área y estructuras), los ajustes y estimaciones realizadas implícitamente tenían que ver con eventos que alejaban la población de tal regularidad, como lo son los errores de subnumeración o por declaración de la edad, e incluso traslajos en la distribución territorial de la población censal. De otra parte, se debe tener en cuenta que la existencia de municipios con un reducido volumen de población (40% son municipios con menos de 10.000 habitantes) puede dar lugar a estructuras poblacionales con ciertos sesgos, por lo que se trató de homogenizar su trayectoria de largo plazo; sin embargo, para casos específicos de surgimiento de población en lugares especiales de alojamiento —LEA¹— se hizo cierta transición entre un censo y otro.

¹ Lugares que se caracterizan por agrupar o resguardar a un grupo de población específico que viven colectivamente, generalmente no parientes, y participan de una vida en común por razones de estudio, trabajo, culto religioso, disciplina militar, labores administrativas, procesos de rehabilitación carcelaria, recreación, carencia de un hogar, entre otras.

El presente documento busca dar a conocer los principales aspectos conceptuales, metodológicos y técnicos aplicados en la estimación de la población municipal para el periodo 1985-2005. En el capítulo 1 se presentan los antecedentes de la investigación; en el capítulo 2 se exponen los aspectos relevantes del diseño de este proyecto, en particular los objetivos y marco teórico de la investigación; el capítulo 3 presenta los elementos y características técnicas del proceso de producción estadística, como son la evaluación y el ajuste de las cifras censales consideradas, 1985, 1993 y 2005; y la estimación interanual de niveles y estructuras. En el capítulo 4 se mencionan las series disponibles de población producto del proyecto. Se espera que este documento contribuya a la comprensión del proceso seguido, y que oriente el uso adecuado de los datos.

1. ANTECEDENTES

A nivel de proyecciones subnacionales, el DANE realizó en su momento proyecciones municipales de población 1995-2005, discriminadas por área, basadas en los resultados del Censo 1993. Estas proyecciones estuvieron vigentes y fueron el referente nacional para múltiples investigaciones hasta la realización de nuevos estimativos en el 2007 basados en el Censo 2005. A este nivel tales proyecciones no presentaron una apertura por sexo y grupos de edad. Adicionalmente, a nivel municipal no se realizaron estimaciones intercensales respecto al Censo 1985, pero sí se hicieron para el agregado nacional.

Con posterioridad a la publicación de los datos censales ajustados del 2005, el DANE empezó a trabajar en la producción de proyecciones de áreas mayores (departamentos) para el período 2005 al 2020, así como también para los municipios, llegándose a discriminar por sexo, grupos de edad y edades simples 0-26 años.

Estas cifras han cumplido los requerimientos tanto internos como externos, pero estas necesidades de información se han hecho extensivas a períodos anteriores para diferentes propósitos como la revisión de las metas del milenio y otras necesidades a nivel de gobierno. Por esta razón, el DANE desarrolló el proyecto de estimación de la población para los períodos intercensales 1985-1993 y 1993-2005, cuya metodología se presenta en este documento.

La metodología seguida para estas estimaciones municipales intercensales descansa fundamentalmente en los métodos matemáticos de interpolación no lineal, pero con los referentes demográficos que impuso la conciliación departamental realizada con posterioridad al Censo 2005. En la revisión de la literatura de otros países no se encontraron antecedentes de estimativos intercensales a nivel municipal, por lo menos realizados por las oficinas centrales de estadística, por lo que el trabajo presentado se convierte en un nuevo referente en la estimación de áreas menores de menor nivel, como son los municipios.

2. DISEÑO

2.1. MARCO DE REFERENCIA

2.1.1. Objetivos

Objetivo general. Producir las estimaciones retrospectivas de población a nivel municipal por sexo y grupos quinquenales de edad, según área de residencia, tomando como base los datos ajustados de los censos 1985, 1993 y 2005.

Objetivos específicos.

- Evaluar y analizar las fuentes de información básica para la elaboración de las estimaciones retrospectivas a nivel municipal (censos 1985, 1993 y 2005), teniendo en cuenta la prevalencia de población indígena en ciertas áreas así como la existencia de grupos particulares, como los LEA.
- Elaborar las estimaciones de población municipal mediante la aplicación de modelos matemáticos que permitan establecer los volúmenes y estructuras de la población por sexo, edad y área, de cada uno de los municipios que conforman cada departamento, acorde con su etapa de transición demográfica² obtenida en el proceso de conciliación censal.
- Actualizar la información sobre el volumen de población existente a nivel municipal, para el periodo 1985-2005, por grupos de edad y sexo según área³.
- Proveer información indispensable para la planeación y gestión local.

Los municipios considerados son los 1.119 existentes en el momento del Censo 2005, incluyendo los denominados corregimientos departamentales⁴. Para estos efectos se han utilizado tanto los códigos geográficos como los nombres de la División Política Administrativa Colombiana vigentes en el momento del Censo 2005.

2.1.2. Marco teórico. Al igual que en las proyecciones de población, para el ejercicio de estimación de población municipal es necesario contar con una población base, ya que este ejercicio descansa en los métodos de interpolación. Realmente, se debe contar con al menos dos puntos censales, los que han de ser ajustados en sus niveles y estructuras poblacionales con el fin hacerlos consistentes con los valores de la conciliación demográfica realizada a nivel departamental.

En cuanto al ajuste de los niveles poblacionales, este debe consultar su distribución territorial, por cuanto la dinámica se diferencia claramente entre la población residente en las áreas urbanas y rurales. Por esta razón, el ajuste de los niveles de población involucra la revisión de la evolución del porcentaje de urbanismo, que corresponde a la relación de la cantidad de población que habita en la cabecera municipal con respecto a la población total. Sin embargo, esta relación resulta afectada cuando, por efectos de la segregación de un nuevo municipio, se reduce la zona rural del municipio que cede terreno. Así, la dinámica de este indicador está determinada, en general, por la dinámica relativa del área urbana frente a la rural; en particular, para una no despreciable cantidad de municipios por sus efectos segregacionistas⁵. Por lo anterior, es necesario estandarizar los niveles de las series total, cabecera y resto de cada municipio; así se permite garantizar las tendencias de la serie de largo plazo y, por tanto, su empalme con la serie proyectada 2005-2020.

² Ver glosario.

³ Esto incluye tanto a los nacionales como a los extranjeros residentes en el momento del censo, pues los censos se han realizado bajo el concepto de Jure: se excluye a los nacionales que residen en el exterior.

⁴ Es una forma de división política de los departamentos que incluye un núcleo de población y sus alrededores pero no son elevados a la categoría de municipio. Éstos se diferencian de corregimientos corrientes en que no son jurisdicción de ningún municipio, sino que son gobernados por el departamento. El país cuenta con 20 corregimientos departamentales localizados en los departamentos de Amazonas, Vaupés y Guainía.

⁵ En el periodo 1985-2007 se crearon noventa municipios, lo que a su vez afectó el territorio de un número mayor de municipios.

En cuanto a los ajustes de las estructuras poblacionales, dada la precariedad de la información demográfica disponible a nivel municipal para la estimación de las componentes, se hace necesaria la utilización de otras metodologías para ajustar sus estructuras. Considerando que los censos de población, además de los problemas de cobertura, presentan problemas de contenido, como la mala declaración de la edad o subregistro de ciertos grupos de edad, es necesario realizar la evaluación de las estructuras etarias de manera que reflejen la dinámica y estructura demográfica de su respectivo departamento.

La estimación de la población municipal mediante métodos de interpolación parte del supuesto de que, salvo eventos catastróficos, las poblaciones tienen un comportamiento continuo y armonioso, reflejando paulatinamente los cambios demográficos que se presentan en su interior; estas características del crecimiento poblacional son válidas tanto para las épocas de explosión demográfica como para las de transición demográfica. Bajo esta premisa, las poblaciones estimadas entre dos puntos en el tiempo guardarán esta relación de continuidad entre dichos puntos. Más aún, llevando esta premisa a nivel de los grupos de edad, esta continuidad representa la evolución de la estructura de población en alguna fase de su desarrollo demográfico (por ejemplo, transición o explosión).

Uno de los métodos que permite guardar esta continuidad del proceso demográfico, particularmente cuando se tienen más de dos puntos a interpolar, como en el presente estudio con tres puntos censales ajustados, es el método de “Splines”. En términos simples, un “Spline” es una curva que interpola los puntos observados de manera suavizada mediante la aplicación de un polinomio. En general, estas curvas se definen con un polinomio de bajo grado, que evita oscilaciones fuertes e introduce un suavizamiento a la serie estimada.

Una interpolación como la enunciada permitiría seguir los supuestos del crecimiento poblacional armonioso y continuo. Más específicamente, la suavización de las series, implícitamente, deriva en tasas de crecimiento más armoniosas que las que se obtienen en procesos de interpolación lineal. Una ventaja adicional del método de “Splines” es que el polinomio estimado pasa por los puntos censales ajustados de entrada. En el Anexo A se presenta una exposición matemática de este método.

2.1.3. Referentes internacionales. Ciertamente muchos países realizan estimaciones de población a nivel subnacional⁶ que, en general, hacen referencia a la proyección de estos agregados más que a estimativos hacia atrás, o bien se presentan estimaciones retrospectivas para agregados nacionales o regionales. Sin embargo, no se tiene conocimiento de estimaciones al nivel de desagregación que se realiza en el presente estudio, es decir, estimaciones a nivel municipal discriminadas por grupos quinquenales de edad y sexo, con un cálculo de estimación adicional por área⁷.

En todo caso, ya se trate de proyecciones o estimaciones retrospectivas agregadas por regiones, en general lo que se observa es que dichos cálculos se apoyan en el método de componentes⁸, lo que permite estimar las estructuras poblacionales y garantizar su coherencia en la evolución temporal. Tal vez uno de los aspectos que representa mayor incertidumbre es el componente migratorio. En algunos casos (por ejemplo de algunas comunidades españolas y en Australia) se utilizan modelos regionales donde se separa la migración internacional de la interna; para este último componente se estiman modelos “intra-regionales” que permiten estimar la migración entre ellas, y dentro de éstas, esta migración se distribuye en subregiones utilizando modelos tipo “migration pool”⁹. Cabe anotar que este nivel de refinamiento en los cálculos no llega a nivel municipal, es decir, no se realizan por componentes ni por ningún otro método conocido.

En España, los estimativos de población retrospectivos se hacen a nivel de comunidades autónomas (18) y provincias (52)¹⁰, elaborados desde los censos de 1970 con desagregaciones por sexo y grupos quinquenales de edad, pero que no incluyen estimaciones a nivel municipal. En resumen, por lo menos desde la producción de la entidad estadística nacional (INE), a nivel municipal solo existen las estadísticas censales y de los padrones anuales, pero no se conocen ejercicios de estimaciones ni proyecciones a nivel municipal. Sin embargo,

⁶ Como es el caso, por ejemplo, de Estados Unidos y países suramericanos, centroamericanos y europeos.

⁷ Para el caso colombiano, se tiene que el 40% son entes territoriales con menos de 10 mil personas, el 90% comprende poblaciones hasta 50 mil habitantes.

⁸ Para lo que se deben introducir supuestos acerca de la evolución de la fecundidad, mortalidad y migración para cada región a estimar.

⁹ El término “migration pool” consiste en asumir unas tasas de distribución dentro de una región, las cuales se aplican al estimativo regional de migración. Ver New South Wales state and regional population projections 2001-2051, Australian Population Association, Sep 2004, o los estimativos para la comunidad Cataluña Obtenció de projeccions de població municipals a partir de les projeccions de població comarcals (base 2008), Institut d'Estadística de Catalunya, març 2010.

¹⁰ Estos cálculos se conocen como estimaciones intercensales y se presentan para cada par de censos.

teniendo en cuenta la independencia de las comunidades, cada una puede elaborar un mayor detalle de sus proyecciones, como es el caso de Cataluña, que realiza proyecciones para su capital (Barcelona) y para sus cuarenta y una comarcas, a nivel total.

Algo similar sucede en Estados Unidos, pero con menos detalle que el caso español. La oficina de censos cuenta con estimativos y proyecciones a nivel de Estados desde 1995 hasta 2025 en la penúltima versión de estimativos, pero de 2000 hasta 2030 en su última entrega. A nivel más desagregado cuenta solo con estimaciones de la población corriente (año 2009) por condado, pero solo a nivel agregado (no se publica un mayor detalle); no hay estimativos por ciudad (salvo los casos donde se asimilan las ciudades a condados). La única información disponible a nivel municipal es la derivada del censo del año 2000.

Así que, para subniveles de áreas menores más allá de lo regional, no se tiene conocimiento de trabajos efectuados, para el cálculo de la población por grupos quinquenales de edad, sexo y área, de manera simultánea, como se presenta en este documento metodológico para el caso colombiano.

3. PRODUCCIÓN ESTADÍSTICA

3.1. ACTIVIDADES PREPARATORIAS

A continuación se resumen las actividades relacionadas con la preparación de los insumos necesarios para la estimación retrospectiva de la población municipal por área. Entre las actividades se encuentran:

- Determinación de los techos departamentales por área, para la definición de las series estandarizadas municipales tanto en su nivel como en su estructura.
- Evaluación y ajuste del volumen y estructura de población por municipio y área.
- Calibración de las áreas menores con los correspondientes techos departamentales.

3.1.1. Definición y revisión de los techos iniciales nacional y departamental por área, para el proceso de estandarización. El proceso de conciliación demográfica entregó estimativos totales por sexo y grupos de edad para los departamentos y el agregado nacional del período 1985-2005, estimando igualmente la población ajustada para los años censales 1985 y 1993.

Posteriormente se llevó a cabo la estimación de los techos por área para el nivel nacional y departamental: para la cabecera a nivel nacional, por el método de componentes; y el área resto se obtiene por diferencia entre las áreas total y cabecera.

Para los departamentos, dadas las deficiencias de información insumo por área, se ajustaron los pivotes censales 1985 y 1993, mediante un proceso iterativo en el cual se tiene en cuenta la evaluación, el ajuste y las tendencias por área de los municipios de la serie estandarizada, garantizando estimaciones consistentes a largo plazo (1951-2005). Adicionalmente, se evaluó la trayectoria del porcentaje de urbanismo, así como su estructura por sexo y grupos de edad buscando la coherencia con la respectiva etapa de transición demográfica. Con lo anterior, se obtienen los techos departamentales por área, para los tres puntos censales, los cuales sirven de referencia para el ajuste a nivel municipal.

3.1.2. Evaluación y ajuste del nivel de población municipal por área. A partir de los resultados de la conciliación de la población total por departamentos y la estimación de su distribución urbano-rural, se evalúan y ajustan los valores censales municipales para los años 1985 y 1993¹¹, con la finalidad de ajustar las cifras municipales a los techos departamentales y generar niveles de población coherentes en su dinámica de largo plazo. En este proceso se implementaron dos fases: la primera consistió en realizar la evaluación de los tres pivotes censales por área; la segunda consistió en el ajuste y calibración de los municipios con el diagnóstico realizado en la fase anterior.

Fase 1: Evaluación de niveles. Consiste en evaluar la trayectoria de población de cada municipio en un horizonte de largo plazo, a partir de las cifras de los últimos seis censos realizados en el país¹². Teniendo en cuenta que las series están afectadas por la creación de nuevos municipios, se estima la población estandarizada municipal desde el año 1951. Dicha estandarización consiste en estimar hacia atrás el nivel de la población para los nuevos municipios, ajustando por lo tanto la población de los municipios segregantes. De esta manera se cuenta con un antecedente histórico para los nuevos municipios y una evolución más estable de la población de los municipios segregantes (ver 3.2.3).

Un segundo proceso es realizar una interpolación lineal para obtener el punto en el año 1985 y 1993, con el objetivo de tener un referente de una tendencia teórica esperada. Igualmente, se incluyen como parte del análisis el comportamiento de indicadores tales como el porcentaje de urbanismo y la tasa de crecimiento intercensal

¹¹ Este proceso de ajuste de los niveles cabecera-resto para el Censo 2005 se surtió durante el año 2006, luego de finalizado el operativo censal. Sin embargo, para esta ocasión sí se revisaron las estructuras poblacionales por área.

¹² Censos de 1951, 1964, 1973, 1985, 1993, 2005, los cuales fueron previamente homologados a la división territorial existente en el 2005 (1.119 municipios).

de acuerdo con la información disponible. Teniendo en cuenta lo anterior, se aplican los siguientes criterios:

- Si el punto censal es menor que el valor esperado, se evalúa si corresponde a un municipio segregante, caso en el cual se adopta el valor censal.
- Respecto de la tasa de urbanismo¹³, se estableció que en general las áreas más urbanizadas registran una curva creciente o estable en el tiempo, mientras que en los municipios predominantemente rurales, el crecimiento de la tasa se da de forma más paulatina. Ante una alteración inexplicable en este patrón de comportamiento, el nivel del ente territorial debe ser ajustado.
- Si la tasa promedio de crecimiento¹⁴ es oscilatoria o presenta un punto de inflexión, el municipio debe ser ajustado.

Para interpolar los pivotes 1985 y 1993 derivados como puntos intermedios en la trayectoria de largo plazo 1973-2005, se utilizan los siguientes algoritmos:

Valor esperado año 1985, $PI_{1985} = (P_{1993} - P_{1973} / T_2)(T_1) + P_{1973}$; y

Valor esperado año 1993, $PI_{1993} = (P_{2005} - P_{1985} / T_2)(T_1) + P_{1985}$

Donde:

PI_{1985} , población total interpolada para el año 1985

PI_{1993} , población total interpolada para el año 1993

P_{2005} , población total estandarizada en el año 2005.

P_{1993} , población total estandarizada en el año 1993.

P_{1985} , población total estandarizada en el año 1985.

P_{1973} , población total estandarizada en el año 1973.

T^1 , número de años comprendidos entre el año de interés (1985, 1993) y el pivote inferior (1973, 1985).

T^2 , Número de años comprendidos entre los dos pivotes.

Con el fin de identificar la desviación entre el valor censado y el interpolado o referente, se estima la diferencia porcentual entre estos dos valores, obteniendo así los municipios que deben ser ajustados. Este indicador se define como:

$$\text{Dif\%} = (P - PI) / PI * 100$$

Donde:

P es la población inicial, la cual será objeto de ajuste.

PI es la población interpolada, y que sirve como punto de referencia.

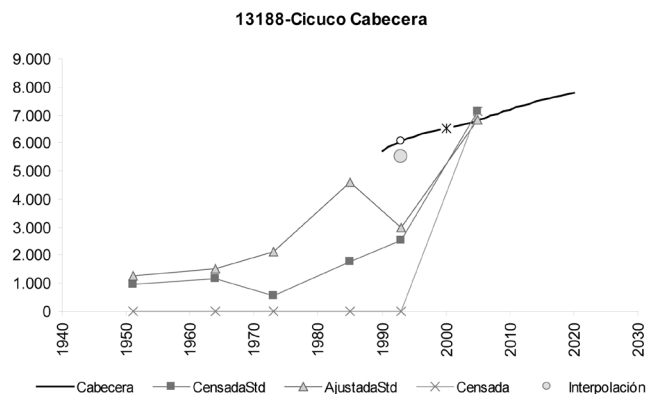
En el Gráfico 1 se muestra el proceso planteado anteriormente. En el panel A se observa una subvaloración en las cifras censales dado que el valor interpolado es superior y es acorde con la tendencia observada en el Censo 2005. Por el contrario, en el panel B se tiene que el valor esperado de 1993 es inferior al pivote ajustado inicial.

¹³ Tasa de urbanismo definida como la proporción de población total que reside en cabecera.

¹⁴ Para este ejercicio se calcularon las tasas promedio de crecimiento intercensales para los años 1951,1964, 1973, 1985 y 1993; y las tasas estimadas para los quinquenios contemplados en el periodo 1990-2020.

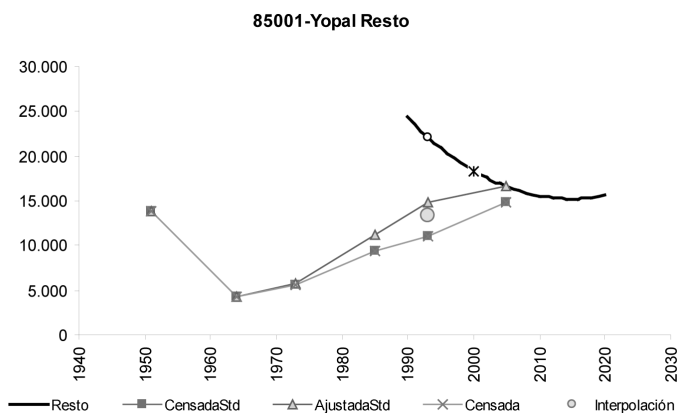
Gráfico 1. Pivotes deficientes en el municipio de Cicuco-cabecera y Yopal-resto 1993

Panel A



Fuente: DANE.

Panel B

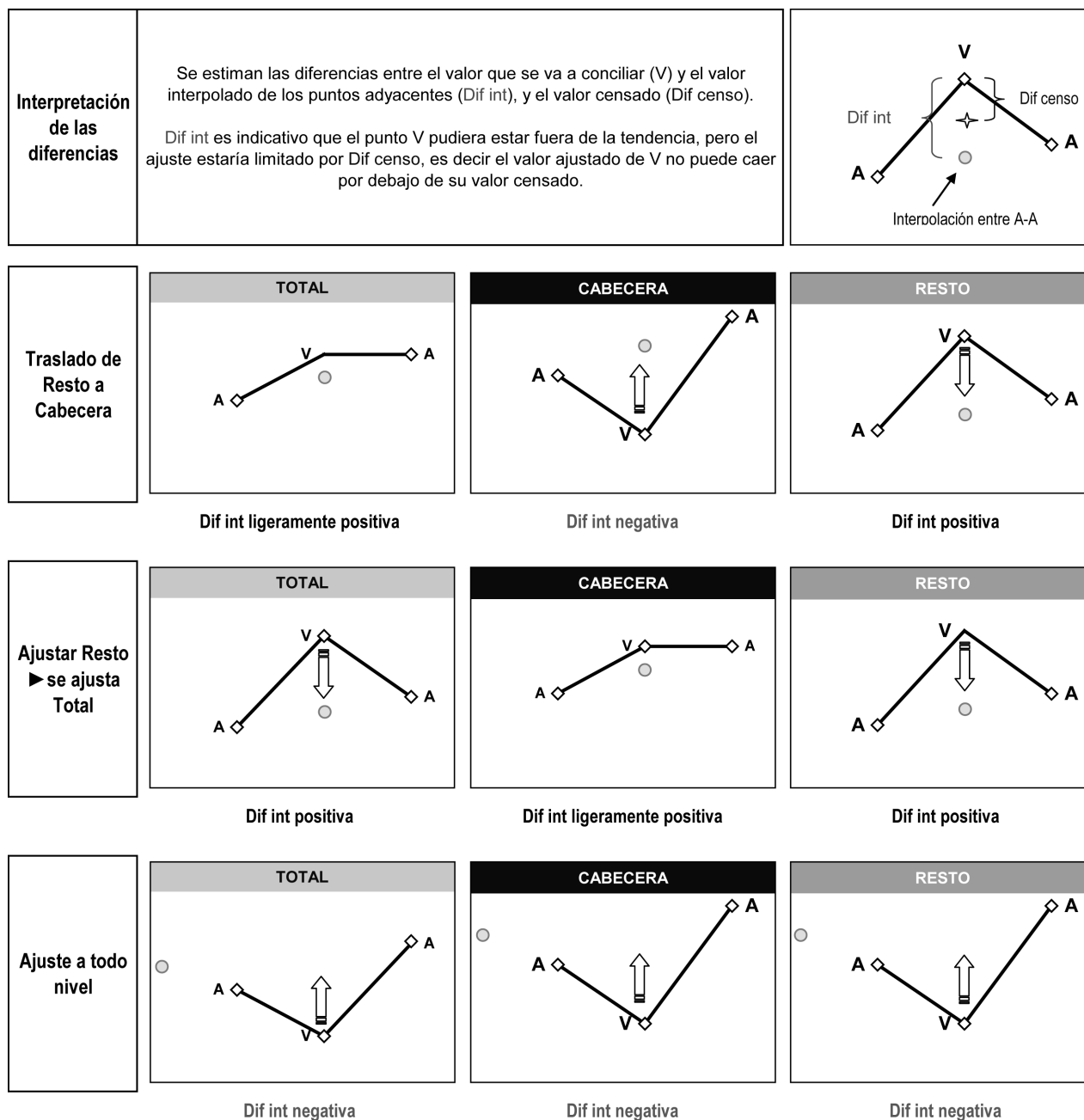


Fuente: DANE.

Fase 2: Ajuste y calibración de los niveles municipales. Los municipios a ser ajustados fueron priorizados en función de la diferencia porcentual entre el valor censado y el interpolado, lo cual define valores críticos diferenciales para cada municipio. Si esta diferencia es superior al valor crítico, entonces se reduce la población ajustada (caso del panel B, ajustada estándar); por el contrario, si la diferencia es inferior al valor crítico se debe incrementar la población ajustada (caso del panel A). Se hace este proceso de manera iterativa hasta que la suma de los niveles municipales por área sea consistente con el respectivo techo departamental.

Se debe señalar que se realizaron diferentes tipos de ajustes. En uno de ellos, se trasladó población de resto a cabecera cuando se identificó que el nivel total del municipio era el adecuado pero que el de las otras dos áreas no; también se presentó el caso donde el nivel de cabecera era consistente con la tendencia poblacional, pero no así para el total y el resto. En este último caso se ajusta la población de resto, y el total se obtiene por suma. Finalmente, se consideró la situación en donde todos los niveles presentan deficiencias, por lo cual se hizo necesario ajustar todas las áreas (ver Gráfico 2).

Gráfico 2. Interpretación visual de las diferencias identificadas entre el dato ajustado censal y el interpolado



Fuente: DANE

Este proceso impactó 200 municipios afectados por segregación durante el período de análisis.

3.1.3. Revisión de las estructuras municipales por sexo y grupos de edad. Una vez estimado el nivel de la población municipal y calibrado a los techos departamentales, se revisan y ajustan las estructuras, de los tres últimos censos, para los 1.119 municipios.

Para cumplir con este objetivo se definen tres fases fundamentales. La primera, la evaluación de las estructuras de población por área; la segunda, el ajuste de las estructuras, y en la última etapa, se calibra para llevar las cifras a los techos departamentales.

Fase 1: Evaluación de estructuras de población municipales. Se definieron los siguientes criterios para llevar a cabo la evaluación de cada estructura, considerando que la evolución tiene inmerso el proceso de transición demográfica, y por ende refleja dicho fenómeno en la composición de la población de los años bajo análisis:

- Identificar los grupos de edad quinquenales que no presentan población en alguna de las áreas analizadas, asociado a omisión censal. Se hace énfasis para el caso típico de sub-enumeración de la población de 0-4 años.
- Evaluar la distribución de los grupos de edad para un mismo año censal, con el fin de identificar casos donde la población adulta o de más de 65 años es porcentualmente mayor a grupos de menor edad.
- Comparar la distribución de los diferentes grupos de edad de las estructuras de población de los años 1985, 1993 y 2005, bajo la óptica de la transición demográfica; especialmente, se busca observar la coherencia en la evolución de los grupos 0-4 y 60 y más años, en los cuales se acentúa la reducción de la fecundidad y el incremento del envejecimiento, respectivamente.
- Observar la coherencia de los índices de masculinidad¹⁵ cabecera-resto entre los tres censos.
- Tener en cuenta algunas consideraciones particulares para zonas con prevalencia de población indígena así como para poblaciones con una significativa participación de los LEA.

Como resultado de la aplicación de los anteriores criterios, se identifican estructuras poblacionales deficientes, como son, por ejemplo, grupos de edad sin población, índice de masculinidad en el grupo de edad de cero a cuatro años superior a 500 o un retroceso de la transición demográfica.

Fase 2: Ajuste de las estructuras municipales por área de residencia. En esta fase se buscó eliminar las inconsistencias de la estructura de población¹⁶. Para el caso específico del año 2005, la estructura total municipal no fue modificada debido al trabajo realizado con anterioridad en el marco de las proyecciones municipales; en la revisión para dicho año, solo se incluyen cabecera y resto. Para los censos 1985 y 1993, se ajustan tanto el total municipal como sus áreas.

En el proceso se encontraron una serie de casos particulares generados por errores de cobertura y contenido de los censos que producen una estructura municipal deficiente, determinada por grupos de edad sin población o una de las áreas sin población. En estos casos, y si su nivel fuera inferior a 3.000 habitantes, se le imputó la estructura del municipio más cercano geográficamente, perteneciente al mismo conglomerado según dinámica demográfica, construido para el censo 2005, o en algunos casos se le asignó la estructura departamental para el área que se estuviera ajustando.

También se recurrió al ajuste manual de las estructuras. En este módulo se ajustaban los pesos relativos de cada grupo de edad entre cabecera y resto, dejando inalterado el índice de masculinidad del total del municipio y los niveles ajustados por área obtenidos en la sección 3.2.1.

En algunos casos se ajustaron las estructuras mediante técnicas de suavización matemática tales como los ajustes de Arriaga y de Strong. Estos dos métodos de ajuste suavizan la transición entre un grupo de edad y otro, garantizando la preservación del total de población del ente territorial en cuestión. La metodología de esas dos técnicas se tomó de las hojas de cálculo del PAS (Bureau de Censos de los Estados Unidos).

Finalmente, se debe tener en cuenta que dentro del proceso de ajuste de las estructuras poblacionales de los municipios, se tuvo en cuenta la existencia de los LEA, afectando el volumen de población para uno o varios rangos de edad, pero que por sus características no afecta la dinámica poblacional propia del ente territorial. Excepcionalmente, para el año 2005 se ajustaron algunas estructuras municipales a nivel del total de la población¹⁷.

De este proceso de ajuste se excluyeron los 34 municipios que conforman lo que se denominó el “Grupo Amazonas”¹⁸.

¹⁵ Ver glosario.

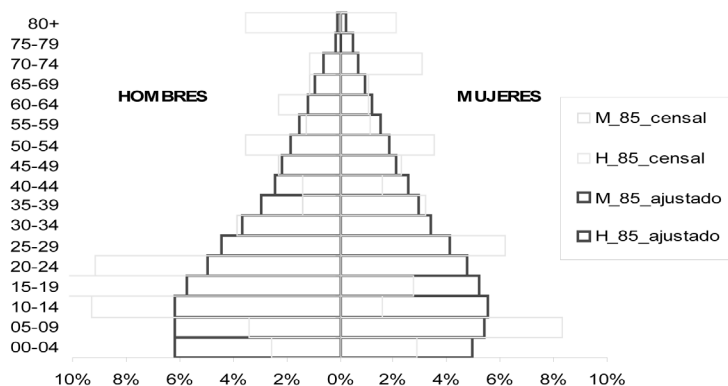
¹⁶ Para información más detallada, se puede consultar el Manual de usuario para la plantilla de ajustes 1985, 1993, 2005.

¹⁷ Aunque a nivel del total del municipio se mencionó que no se harían ajustes, en esta revisión se detectó que era necesario reincorporar sus LEA a 11 poblaciones. Este ajuste implicó afectar la estructura cerca de 90 municipios a fin de preservar los techos municipales.

¹⁸ Conformado por los departamentos de Amazonas, Guainía, Guaviare, Vaupés y Vichada.

En el Gráfico 3, se puede observar la estructura poblacional de la cabecera del municipio de Anzá para el Censo 1985, contrastando la estructura censal inicial y la estructura luego de los ajustes realizados bajo los lineamientos mencionados anteriormente.

Gráfico 3. Estructura poblacional del municipio de Anzá cabecera 1985



Fase 3: Calibración de estructuras municipales. Si bien los ajustes de las estructuras realizados en la Fase 2 mejoran la condición de las pirámides, no se garantizaba que la suma de estos grupos fuera igual al techo municipal ni al departamental. Con este fin se aplicaron dos métodos:

- Prorrato: toma el valor de la participación porcentual de cada uno de los grupos de edad de cada municipio frente al total y, de acuerdo a él, le asigna una porción de las diferencias frente al techo. Este mecanismo busca que a cada grupo de edad se le haga una asignación de acuerdo a su tamaño; también, busca evitar que las estructuras con diferencias muy pequeñas se vean afectadas dentro del prorrato.
- Tabla cuadrada: mecanismo por el cual se hacen aproximaciones sucesivas a los totales marginales de un conjunto de datos, con tres restricciones a cumplir: i) ajuste al techo municipal; ii) ajuste de los techos departamentales para cada grupo de edad; iii) ajuste de cabecera-resto a la estructura total del municipio¹⁹.

3.2. MÉTODOS PARA LA ESTIMACIÓN DE POBLACIÓN MUNICIPAL

Para llevar a cabo la estimación de población municipal para el período 1985-2005, y una vez ajustados los insumos necesarios, se utilizó la “Interpolación por Splines” entre los años censales (ver Anexo A. El modelo de interpolación con splines). Este método se aplicó para: i) estimar los niveles interpolados de cabecera y resto (el total resultaría por suma); ii) generar las cifras interpoladas de grupos de edad y sexo.

3.2.1. Estimación de niveles. El proceso de estimación de niveles para los departamentos y municipios se llevó a cabo por metodologías similares, ya que implicó el uso de la función de interpolación Splines, considerando como puntos base para la interpolación los censos de 1985, 1993 y 2005.

En el caso de los 46 municipios de inclusión forzosa del marco muestral de la Gran encuesta integrada de hogares (GEIH), la estimación de la serie considera como pivotes el año 1985, al igual que los años 2000 y 2001, logrando así una consistencia a largo plazo con los datos estimados en un momento anterior.

Para los restantes municipios, la información considerada como puntos base para la estimación fueron los años censales 1985, 1993 y 2005. Con dicha información se realizó el proceso de estimación de los niveles municipales por área²⁰, sosteniendo como referencia el 30 de junio de los años comprendidos entre 1985 y 2005. Debido a que estas estimaciones no eran consistentes con los techos departamentales, fue necesario realizar un proceso de prorrato, excluyendo los municipios de inclusión forzosa de la GEIH.

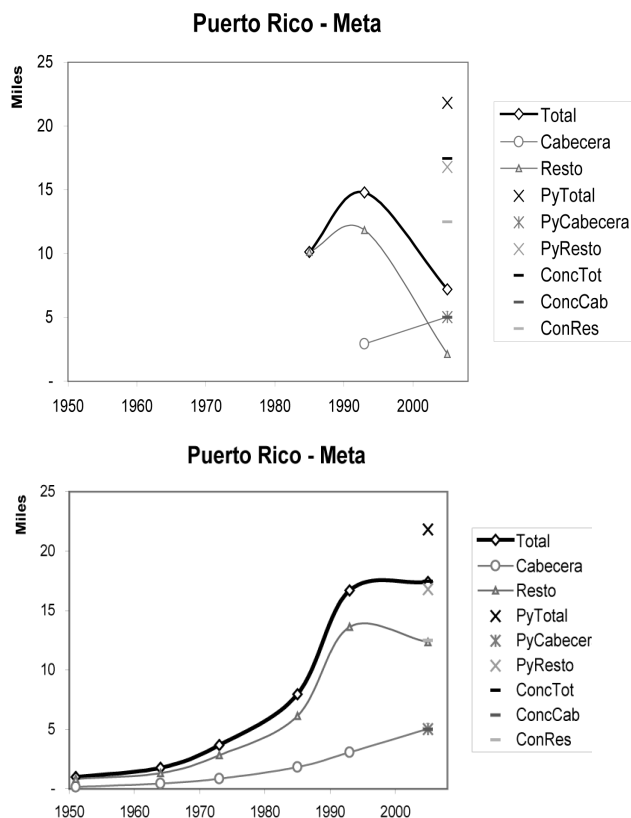
¹⁹ Esta última restricción se aplicó exclusivamente para el año 2005, a fin de respetar los valores censales ajustados. De otro lado, para los años 1985 y 1993, se aplicaron las dos primeras restricciones.

3.2.2. Estimación de estructuras. Para la estimación de las estructuras poblacionales municipales de los años intercensales, se utilizó el método de interpolación por Splines, considerando cada grupo de edad de manera independiente. Para dicho ejercicio se contó con la información de los pivotes censales por área, 1985, 1993 y 2005. Luego, en una etapa posterior, debido a la existencia de diferencias remanentes de los grupos de edad frente a los techos departamentales, se programó el prorrateo de tales diferencias haciendo uso de la Tabla Cuadrada (ver: Revisión de las estructuras municipales por sexo y grupos de edad).

3.2.3. Estimación de la población municipal, 1985-2005. Como ya se mencionó, debido a la creación de nuevos municipios en el periodo de estudio, el DANE se vio en la necesidad de adoptar dos series de población para facilitar el ejercicio de estimación: i) la estandarizada, construida con el fin de contar con un antecedente histórico para los nuevos municipios y una evolución más estable de la población de los municipios segregantes; ii) la serie final publicable, la cual revierte de alguna manera la operación realizada con la población estandarizada una vez concluido el proceso de retroproyección. En este caso, se genera la población desde el año de creación del municipio, devolviendo la de años anteriores a sus municipios segregantes, afectando en su conjunto la distribución de población cabecera-resto. Esta operación permite hacerse una idea de la magnitud de los municipios desde el momento de su creación, pero no permite hacer un análisis para proyección (lo que sí se posibilita con la serie estandarizada).

En el Gráfico 4 se ilustra el resultado de la estandarización para un municipio. El recuadro de la izquierda muestra la serie desde el momento de creación de Puerto Rico-Meta, en el año 1984, serie que presenta una concavidad pronunciada hacia el año 1993. Al momento de llevarse a cabo el proceso de estimación, esto genera resultados con una tendencia decreciente a partir de dicho año. El recuadro de la derecha muestra la población estandarizada, que en este caso particular considera una población hipotética para los períodos que antecedieron a su creación como municipio, dando como resultado una serie con una tendencia más suavizada a partir del año 1993.

Gráfico 4. Población estandarizada, Puerto Rico-Meta, 1950-2005



Fuente: DANE.

A nivel de grupos de edad y sexo, el estimativo inicial se realizó con las cifras de la población estandarizada. Una vez validada toda la información municipal para todos los departamentos, se generó la población en su versión final. A este nivel no se incluyeron en la revisión municipal los cinco departamentos del Grupo Amazonas, para los cuales sólo se cuenta con una estimación agregada, tanto en niveles como en grupos de edad y sexo.

3.2.4. Desarrollos tecnológicos para la interpolación de población. Para llevar a cabo la interpolación de población (por grupos de edad, sexo y área), y producir las estimaciones de población a 30 de junio, se hicieron desarrollos basados en el programa R, articulados con Excel, para hallar los niveles de población anual intermedia entre los períodos censales, utilizando la rutina de interpolación Spline (específicamente el método Spline Natural).

Por la necesidad de realizar un proceso de corridas manuales en R²¹, lo cual estaba sujeto a un amplio margen de error, se programaron corridas en lotes, conocidas también como corridas en Batch²², que consistió en elaborar una macro²³ en Excel para permitir la construcción de estos códigos y agregarlos en un archivo de lotes único para ser ejecutados con R, permitiendo no sólo agilizar la actividad sino brindar una disminución en el margen de error de todo el proceso.

Finalmente, es importante mencionar que no se requiere prorratar diferencias, debido a que i) el método Spline garantiza que el valor interpolado pasa exactamente por el valor anual de pivote, ii) los agregados del total de población se obtienen por sumatoria, es decir, el total de hombres y mujeres se obtiene como la suma de estos dos.

3.3 CRITERIOS DE CONTROL DE LA CALIDAD

Para el caso de las estimaciones municipales, por edad y sexo, el objetivo es garantizar la consistencia del total municipal por sexo y edad, permitiendo identificar irregularidades en la evolución de las estructuras poblacionales para el período estimado, teniendo en cuenta los valores del techo obtenidos para cada departamento, por área y grupos de edad. Los criterios tenidos en cuenta se listan en la Tabla 1.

²¹ Lenguaje para el cálculo estadístico y la generación de gráficos, conocido como "GNU S"; es de libre uso.

²² Ver glosario.

²³ Ver glosario.

Tabla 1. Criterios para la revisión de las estructuras de población

Característica por evaluar	Comportamiento esperado	Elementos para tener en cuenta en el ajuste gráfico de estructuras
Perfil de la estructura poblacional	Estructura continua, triangular, rectangular o con reducción en la base pero sin cambios bruscos o escalonados entre grupos de edad.	
Comportamiento de la base	Evolución continua de la participación de la base (menores de 15 años), cumpliendo con el supuesto de que la fecundidad no aumenta en la transición.	Que la base mantenga la tendencia observada entre las estructuras censadas 1993 y 2005.
Grandes grupos de edad	La participación de los menores de 15 años corresponde a la tercera parte de la población y que este indicador vaya en descenso.	Que la participación de los grupos no experimente cambios porcentuales altos. La participación de menores de 15 años debe ser revisada si no se encuentra en el intervalo 20% -40%.
	Respecto a la población mayor de 65 años se espera que ésta no esté por debajo del 3%, salvo casos especiales.	Chequear que el peso porcentual de mayores de 65 años no sea inferior al 3% o superior al 10%.
Razón niños-mujer	La participación del grupo de 0-4 años equivale a una tercera parte de las mujeres en edad reproductiva (15-49 años).	
Razón de dependencia por edad	La cifra no supera las dos terceras partes de la población total.	El indicador se debe revisar si está por encima del 0,66.
Índice de masculinidad (IM) por grupos de edad	El IM está por debajo de 100 a partir de los 60 años de edad por tardar; y que no presente grandes variaciones. Cuando la variabilidad sea alta se revisa la diferencia en valores absolutos entre hombres y mujeres y si ésta no es superior a 100 personas se puede validar de no ser posible su suavizamiento.	No alterar significativamente el IM entre la estructura censada y la ajustada.
Distribución cabecera-resto		Que cada estructura conserve su identidad demográfica.
Omisión censal		Una alta omisión censal a nivel municipal se puede reflejar en omisiones por grupos de edad.

Fuente: DANE.

La revisión de los resultados se efectuó en la totalidad de los 1.085 municipios y 34 corregimientos departamentales existentes al año 2005, lo cual permitió identificar aquellos que presentaban inconsistencias según los parámetros definidos en la tabla anterior.

4. DIFUSIÓN

4.1. PRODUCTOS E INSTRUMENTOS DE DIFUSIÓN

La información de las estimaciones y proyecciones de población se encuentra disponible para facilidad del usuario en la página web de la entidad —www.dane.gov.co—, en formato Excel (xls) y corresponde a:

- Estimaciones de población municipales por sexo y grupos de edad, según área, 1985-2005, en el link: http://www.dane.gov.co/files/investigaciones/poblacion/proyepobla06_20/s_19852004_Ajustadosgruposedad.xls.
- Estimaciones y proyecciones de población nacional, departamentales y municipales, por sexo y grupos de edad, 1985-2020, en el link: http://www.dane.gov.co/files/investigaciones/poblacion/proyepobla06_20/p_20052020_Ajustadosgruposedad.xls

GLOSARIO

Censo de derecho o “de jure”. En esta clase de censo la persona es enumerada en el lugar donde reside habitualmente.

Corridos en Batch. Es un archivo de procesamiento por lotes. Se trata de archivos de texto sin formato, guardados con la extensión BAT que contienen un conjunto de comandos. Cuando se ejecuta este archivo, los comandos contenidos son ejecutados en grupo, de forma secuencial, permitiendo automatizar diversas tareas.

Estimación de población. Número de personas que se calcula que tiene o tuvo una población en un momento específico del tiempo, ya sea globalmente o de una categoría más reducida. Tal volumen no es el producto de una medición directa, pero para obtenerla, se tuvo en cuenta alguna información sobre la población (Rincón, 2000: 8).

Índice de masculinidad. Es el número de hombres por cada 100 mujeres.

Macro. Es una serie de instrucciones que se almacenan para que se puedan ejecutar de forma secuencial mediante una sola llamada u orden de ejecución. Esto permite la automatización de tareas repetitivas

Población estandarizada. Se refiere a la estimación de población entre 1985 y el año de creación para aquellos entes territoriales que fueron creados después del año 1986, asumiendo que dicha población habitó en ese territorio desde tiempo atrás.

Proyección de población. Es el resultado de un conjunto de estimaciones demográficas, matemáticas o de otro tipo, por medio de las cuales se busca establecer las tendencias de las variables determinantes de la dinámica poblacional y, con ello, la derivación de los volúmenes de población y de sus principales características (Rincón, 2000: 6).

Razón niños-mujer. Representa el número de niños menores de 5 años por cada mujer en edad fértil y es obtenido como el resultado de la población de 0 a 4 años de edad entre el número total de mujeres de 15 a 49 años.

Razón de dependencia por edad. Es la razón de personas en edades en las que “dependen” (generalmente personas menores de 15 y mayores de 64 años) de personas “económicamente productivas” (entre 15 y 64 años de edad) en una población.

Transición demográfica. En general, se refiere a los cambios de natalidad y mortalidad que se presentan en una sociedad como resultado del proceso de industrialización o modernización.

BIBLIOGRAFÍA

Australian Population Association (APA). (2004). New South Wales state and regional population projections 2001-2051. 12th Biennial Conference. Canberra: APA.

Congreso de Colombia. Ley 139 de 1994.

Departamento Administrativo Nacional de Estadística (DANE). (2007). Cambios demográficos: análisis de contexto de los cambios demográficos. Bogotá: DANE.

—. XIII Censo nacional de población. Julio 15 de 1964. Resumen general. Bogotá: DANE.

—. XIV Censo nacional de población y vivienda. Octubre 24 de 1973. Resumen nacional. Bogotá: DANE.

—. Censo General 2005. Conciliación demográfica.

—. (2008). Metodología Estimaciones y Proyecciones de población, por área, sexo y edad para los dominios de la Gran Encuesta Integrada de Hogares (2000-2009). Bogotá: DANE.

García, J. (2000). Tutorial de análisis numérico. Interpolación: Splines cúbicos. España: Universidad de Las Palmas de Gran Canaria, Departamento de Informática y Sistemas.

Institut d'Estadística de Catalunya (Indecat). (2010). Obtenció de projeccions de població municipals a partir de les projeccions de població comarcals (base 2008). Barcelona: Indecat.

Rincón, M. (agosto de 2000). Teoría y métodos para la preparación de estimaciones y proyecciones de población. Bogotá.

Siegel, J. & Swanson, D. (Eds.). (2004) The methods and materials of demography. USA: Elsevier Academic Press.

ANEXOS

Anexo A. El modelo de interpolación con splines

Para llevar a cabo la estimación de los niveles de población para el periodo 1985-2005, teniendo la información de tres puntos censales, 1985, 1993 y 2005 en las áreas total, cabecera-resto, se utilizaron polinomios algebraicos, es decir el conjunto de funciones de la forma $P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 x^0$, donde n es un entero no negativo y $a_n, a_{n-1}, \dots, a_1, a_0$ son constantes reales.

La interpolación polinómica se basa en la sustitución de una función o de un conjunto de valores por un polinomio que toma dichos valores. Cuando el número de puntos aumenta, también aumenta el grado del polinomio, que se hace más oscilante (lo cual se traduce en un aumento de los errores).

Un enfoque alternativo para conseguir mejores aproximaciones polinomiales es el uso de polinomios de grado menor en subintervalos, es decir la interpolación segmentaria o interpolación por Splines. La idea central es que en vez de usar un solo polinomio para interpolar los datos, se utilicen segmentos de polinomios, cada uno definido en un intervalo, que se unen entre sí bajo ciertas condiciones de continuidad para realizar la interpolación. La aproximación polinómica por segmentos más común utiliza polinomios de grado tres entre cada par de puntos consecutivos y recibe el nombre de interpolación de trazadores cúbicos o Splines cúbicos.

El uso de polinomios cúbicos se debe a que son los de mínimo orden que pueden garantizar la generación de curvas sintéticas²⁴ con órdenes de continuidad de hasta segundo orden; además son los de menor grado que permiten inflexión dentro de un segmento o trozo de curva compuesta, característica importante en las aplicaciones y que resulta una técnica útil cuando se quiere representar un comportamiento no lineal de datos. Las curvas sintéticas más utilizadas son las constituidas por tramos de polinomios del mismo orden encadenados por condiciones de borde que garantizan una adecuada continuidad y suavidad.

Una de las principales ventajas en el uso de este tipo de curvas es la posibilidad de modificar la forma de las mismas haciendo cambios en los datos de entrada, ya que los puntos-dato influyen decisivamente en la forma y orientación de las curvas sintéticas. Por esta razón los puntos-dato, en el caso de las curvas sintéticas, son llamados también puntos de control.

A continuación se expone a mayor profundidad, la argumentación matemática del modelo.

Dados $n + 1$ puntos $a = x_0 < x_1 < \dots < x_n = b$, y dada una función f definida en $[a, b]$, una función Spline cúbica S que interpola la función f , es una función que cumple con las siguientes condiciones:

$s(x)$ es un polinomio cúbico denotado por $q_i(x)$ en el intervalo $[x_i, x_{i+1}]$, así:

$$s(x) = q_i(x) \text{ en } [x_i, x_{i+1}], \quad i = 0, 1, \dots, n-1.$$

Para que $s(x)$ interpole en los puntos P_0, P_1, \dots, P_n , los $q_i(x)$ han de verificar:

$$\begin{cases} q_i(x_i) = y_i \\ q_i(x_{i+1}) = y_{i+1} \end{cases} \quad i = 0, 1, \dots, n-1. \quad (1)$$

En los puntos en común de cada sub-intervalo se cumple $q_{i-1}(x_i) = q_i(x_i)$ para $i = 1, \dots, n-1$, asegurando la continuidad.

²⁴ Son aquellas obtenidas por combinación de diferentes funciones (generalmente polinomios de grado bajo) y que, basadas en los puntos-dato (puntos digitalizados empleados matemáticamente para la interpolación o aproximación) iniciales, permiten interpolar o aproximar dichos valores manteniendo un adecuado orden de continuidad de acuerdo a los requerimientos o a la naturaleza de los datos.

Los polinomios $q_i(x)$ tienen la misma pendiente y la misma concavidad en los nodos que los unen, es decir

$$\begin{cases} q'_{i-1}(x_i) = q'_i(x_i) \\ q''_{i-1}(x_{i+1}) = q''_i(x_{i+1}) \end{cases} \quad i = 1, \dots, n-1. \quad (2)$$

Si $s(x)$ es cúbica a trozos en el intervalo $[x_0, x_n]$, su derivada segunda $s''(x)$ es lineal en el mismo intervalo e interpola en los puntos $(x_i, s''(x_i))$ y $(x_{i+1}, s''(x_{i+1}))$ en $[x_i, x_{i+1}]$. Por tanto, $q_i(x)$ es un polinomio de grado uno que interpola en los puntos $(x_i, s''(x_i))$ y $(x_{i+1}, s''(x_{i+1}))$, y está dado por:

$$q''_i(x) = \frac{s''(x_i)(x_{i+1} - x) + s''(x_{i+1})(x - x_i)}{x_{i+1} - x_i}, \quad i = 1, \dots, n-1.$$

Denotando con $\alpha_i = s''(x_i)$, $\alpha_{i+1} = s''(x_{i+1})$, $i = 0, \dots, n$,

y $i = 1, \dots, n-1$, $i = 1, \dots, n-1$,

se tiene

$$q''_i(x) = \frac{\alpha_i(x_{i+1} - x)}{h_i} + \frac{\alpha_{i+1}(x - x_i)}{h_i}, \quad i = 0, 1, \dots, n-1, \quad (3)$$

donde h_i y α_i son constantes. Integrando dos veces:

$$q_i(x) = \frac{\alpha_i(x_{i+1} - x)^3}{6h_i} + \frac{\alpha_{i+1}(x - x_i)^3}{6h_i} + C_i x + D_i, \quad i = 0, 1, \dots, n-1, \quad (4)$$

Donde el término lineal se puede escribir como:

$$C_i x + D_i = A_i(x - x_i) + B_i(x_{i+1} - x),$$

Donde A_i y B_i son constantes arbitrarias, luego se tiene que

$$q_i(x) = \frac{\alpha_i(x_{i+1} - x)^3}{6h_i} + \frac{\alpha_{i+1}(x - x_i)^3}{6h_i} + A_i(x - x_i) + B_i(x_{i+1} - x), \quad i = 0, 1, \dots, n-1, \quad (5)$$

Aplicando a (5) las condiciones (1), se obtiene:

$$y_i = \frac{\alpha_i(x_{i+1} - x_i)^3}{6h_i} + B_i(x_{i+1} - x_i) = \frac{\alpha_i h_i^2}{6} + B_i h_i \quad (6)$$

$$y_{i+1} = \frac{\alpha_{i+1}(x_{i+1} - x_i)^3}{6h_i} + A_i(x_{i+1} - x_i) = \frac{\alpha_{i+1} h_i^2}{6} + A_i h_i \quad (7)$$

Despejando de (6) y (7) A_i y B_i , se tiene:

$$A_i = \frac{y_{i+1}}{h_i} - \frac{\alpha_{i+1} h_i}{6}$$

$$B_i = \frac{y_i}{h_i} - \frac{\alpha_i h_i}{6}$$

Al reemplazar A_i y B_i en (5) tenemos la ecuación del Spline $q_i(x)$:

$$q_i(x) = \frac{\alpha_i(x_{i+1} - x)^3}{6h_i} + \frac{\alpha_{i+1}(x - x_i)^3}{6h_i} + \frac{y_{i+1}(x - x_i)}{h_i} - \frac{\alpha_{i+1}h_i(x - x_i)}{6} \\ + \frac{y_i(x_{i+1} - x)}{h_i} - \frac{\alpha_i h_i(x_{i+1} - x)}{6}, \quad i = 0, 1, \dots, n-1. \quad (8)$$

Derivando esta última se genera:

$$q'_i(x) = \frac{-\alpha_i(x_{i+1} - x)^2 + \alpha_{i+1}(x - x_i)^2}{2h_i} + \frac{y_{i+1} - y_i}{h_i} + \frac{\alpha_i - \alpha_{i+1}}{6} h_i \quad (9)$$

Aplicando a (9) las condiciones (2), la ecuación queda definida de la siguiente manera:

$$q'_i(x_i) = \frac{y_{i+1} - y_i}{h_i} - \frac{\alpha_i h_i}{3} - \frac{\alpha_{i+1}}{6} h_i \quad (10)$$

$$q'_i(x_{i+1}) = \frac{y_{i+1} - y_i}{h_i} + \frac{\alpha_{i+1}}{3} h_i + \frac{\alpha_i}{6} h_i \quad (11)$$

Utilizando la condición de continuidad $q'_{i-1}(x_i) = q'_i(x_i)$, se reemplaza i por $i-1$ en (11) para obtener $q'_{i-1}(x_i)$ e igualando a (10) se obtiene:

$$h_{i-1}\alpha_{i-1} + 2(h_{i-1} + h_i)\alpha_i + h_i\alpha_{i+1} = 6\left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right), \quad i = 1, \dots, n-1.$$

La ecuación anterior, genera un sistema de $n-1$ ecuaciones lineales con $n+1$ incógnitas $\alpha_0, \alpha_1, \dots, \alpha_n$, lo cual produce un sistema subdeterminado que tiene infinitas soluciones.

Existen varias estrategias para eliminar α_0 de la primera ecuación y α_n de la $(n-1)$ -ésima ecuación, produciendo un sistema tridiagonal de orden $(n-1)$ en las variables $\alpha_1, \alpha_2, \dots, \alpha_{n-1}$, que se puede resolver mediante eliminación gaussiana sin pivoteo.

Las siguientes alternativas permiten resolver el anterior sistema de ecuaciones

$\alpha_0 = s''(x_0)$ y $\alpha_n = s''(x_n)$. Si se supone que $\alpha_0 = 0$ y $\alpha_n = 0$ se denomina Spline natural.

O bien

$s''(x_0) = f'(x_0)$ y $s''(x_n) = f'(x_n)$, se denomina Spline cúbico.

