

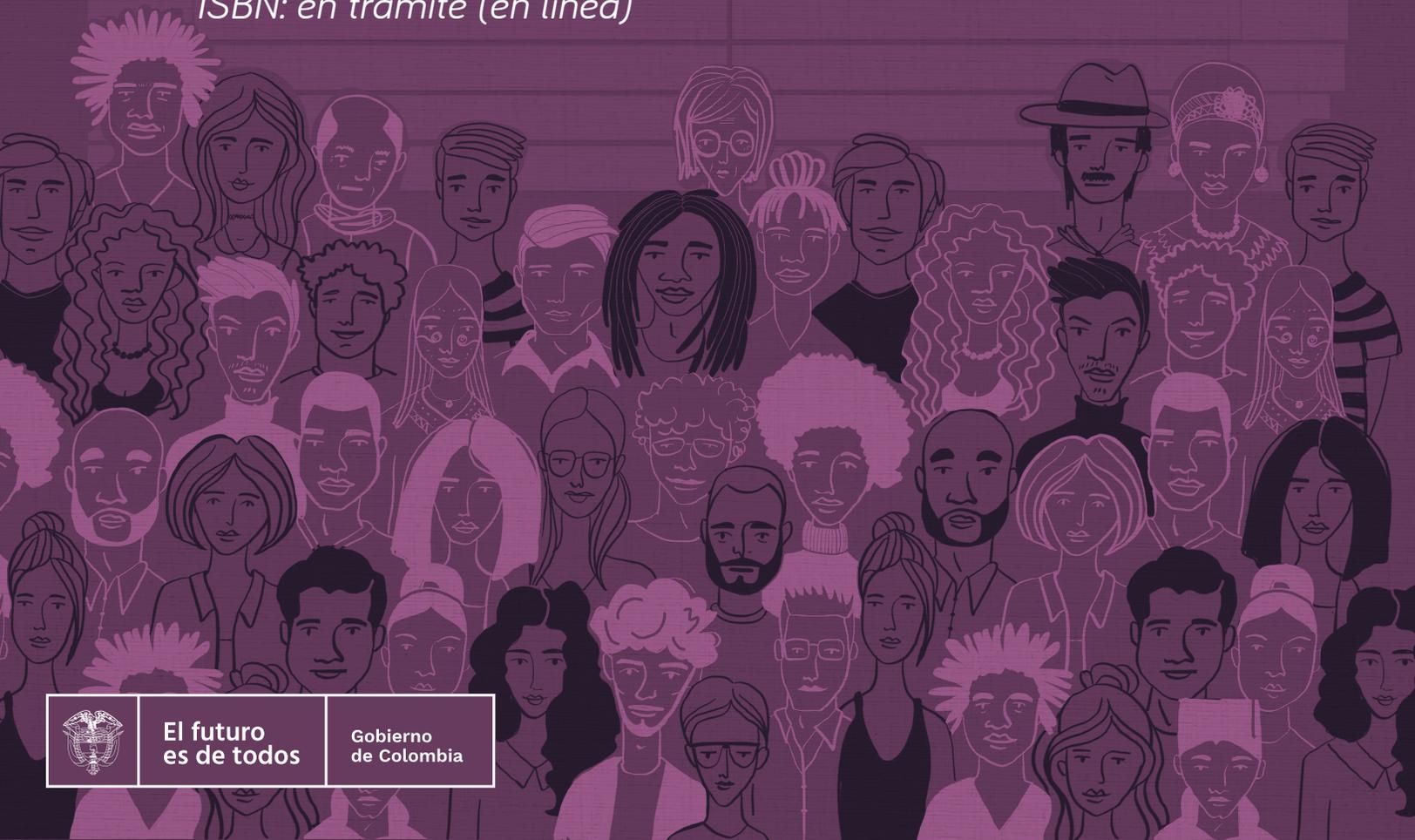
Metodologías

Demográficas Aplicadas

Número 1

Estimación de las curvas de mortalidad a nivel subnacional: *Colombia 2017*

ISBN: en trámite (en línea)



DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA DANE

JUAN DANIEL OVIEDO ARANGO

Director

RICARDO VALENCIA RAMÍREZ

Subdirector

MARÍA FERNANDA DE LA OSSA ARCHILA

Secretaria General

DIRECTORES

ANGELA VEGA LANDAETA

Dirección de Censos y Demografía

MAURICIO ORTIZ GONZÁLEZ

Dirección de Difusión y Cultura Estadística

SANDRA LILIANA MORENO MAYORGA

Dirección de Geoestadística

HORACIO CORAL DÍAZ

Dirección de Metodología y Producción
Estadística

JULIETH ALEJANDRA SOLANO VILLA

Dirección de Regulación, Planeación,
Estandarización y Normalización

JUAN PABLO CARDOSO TORRES

Dirección de Síntesis y Cuentas Nacionales

FABIAN RICARDO MEJIA OSPINA

Dirección de Recolección y Acopio

EQUIPO DE DIRECCIÓN DE CENSOS Y DEMOGRAFÍA

Lina María Sánchez Céspedes

Marcela Pantoja Echeverri

Glenn Harry Amaya Cruz

Juan Sebastián Oviedo Mozo

DIRECCIÓN DE DIFUSIÓN Y CULTURA ESTADÍSTICA EDICIÓN DE CONTENIDO, DISEÑO Y DIAGRAMACIÓN

Brandon Steve Rojas Guerra

Gladys Adriana Quintero Hernández

Gabriela Wiesner Montaña

Maria Jimena Muñoz Saldarriaga

© **DANE, 2022**

Prohibida la reproducción total o parcial sin permiso
o autorización del Departamento Administrativo
Nacional de Estadística, Colombia.

Contenido

1.	Introducción	8
2.	Metodología	11
2.1.	Estimación de las curvas de mortalidad de las cabeceras departamentales	12
2.2.	Estimación de las curvas de mortalidad para los restos departamentales	23
3.	Resultados	25
3.1.	Estimación de las curvas de mortalidad en las zonas urbanas departamentales	26
3.2.	Estimación de las curvas de mortalidad en el área resto de los departamentos	31
4.	Comparación de los resultados con otros métodos	33
5.	Conclusiones	36

Lista de gráficos

Gráfico 1.	Estimaciones del ${}_1q_0$ y ${}_4q_1$ con los Censos de 1993, 2005 y 2018 – Colombia	21
Gráfico 2.	Mortalidad infantil nacional estimada con el Método de Brass-Coale (1968) utilizando los últimos cuatro censos y ajuste al nivel de 2017	22
Gráfico 3.	Probabilidad de morir a la edad x , q_x , para la cabecera de La Guajira estimado con las defunciones sin corregir y corregidas con el factor del CNPV2018	27
Gráfico 4.	Métodos de suavizamiento de la curva de ${}_1q_x$, La Guajira – cabecera	28
Gráfico 5.	Curva q_x seleccionada para la cabecera de La Guajira – cabecera	29
Gráfico 6.	Mortalidad infantil de La Guajira estimada con el Método de Brass-Coale (1968) utilizando los últimos tres censos (1993, 2005 y 2018) y ajuste del método al nivel nacional de 2017	30
Gráfico 7.	Curvas ${}_1q_x$ para cabecera y resto por sexo para La Guajira	32
Gráfico 8.	Comparación de las curvas de mortalidad estimada para La Guajira con la de otros métodos	35

Lista de figuras

Figura 1.	Algoritmo para el cálculo de las tablas de vida a nivel departamental a partir de las defunciones corregidas	13
Figura 2.	Algoritmo del suavizamiento o estimación del patrón de la mortalidad por Heligman-Pollard	15
Figura 3.	Algoritmo del suavizamiento o estimación del patrón de la mortalidad por Alexander-Zagheni-Barbieri	16
Figura 4.	Paso a paso de la estimación de la mortalidad infantil en el nivel departamental – cabecera y resto	17
Figura 5.	Algoritmo para calcular la probabilidad de morir en el primer año ($n=1$) y en los primeros 5 años ($n=5$) de vida	19
Figura 6.	Algoritmo para estimar las curvas de mortalidad de los restos de los departamentos	24

Lista de tablas

Tabla 1. | Criterios para escoger las tablas de vida modelo _____ **20**

Tabla 2. | Clusters de departamentos para hacer los modelos probabilísticos _____ **31**

Resumen

Las proyecciones de población en Colombia se hacen a nivel departamental, por sexo y por tipo de área, cabecera y resto (o urbano y rural). El objetivo de este documento es explicar en detalle la metodología a partir de la que se realizó la estimación de las curvas de mortalidad a nivel subnacional para 2017. Durante este proceso, el equipo del Departamento Administrativo Nacional de Estadística - DANE resolvió diferentes inconvenientes, que seguramente los institutos de estadística de otros países de la región enfrentaran. Por ejemplo, se encontró que los factores de corrección del Censo, estimados con la pregunta sobre la expedición o no del certificado de defunción,

corrigen consistentemente el subregistro de las muertes de los Registros de Estadísticas Vitales para la cabecera, pero no para el resto. Para esta área, el número corregido de muertes en algunos casos no fue suficiente y en otros excesivo. Para superar este obstáculo, se propuso un método que arrojó resultados consistentes: estimar la probabilidad de morir a la edad x con el Censo para determinar la relación $q_x(\text{resto})/q_x(\text{cabecera})$.

Palabras clave:

Curvas de mortalidad a nivel subnacional, proyecciones de población, subregistro de las muertes en los Registros de Estadísticas Vitales.

Abstract

Population projections in Colombia are estimated at departmental level, by sex and by type of area, urban and rural. The objective of the document is to explain in detail the estimation of the mortality curves at subnational level for 2017. During this process, the team of Departamento Administrativo Nacional de Estadística - DANE overcome different obstacles, which surely other statistical offices in the region will face. For instance, we found that the Census correction factors, calculated with the question about whether the death certificate was issuance, consistently correct the underreporting of

deaths that are registered at Vital Statistics for urban areas, but not for rural. For this area, the corrected number of deaths was in some cases insufficient and in others excessive. To solve this problem, a method was proposed that yielded consistent results: estimating the probability of dying at age x with the Census to determine the ratio $q_x(\text{rural})/q_x(\text{urban})$.

Keywords:

Subnational mortality curves, population projections, underreporting of deaths at vital statistics or records

1.

Introducción



Introducción

De acuerdo con el Fondo de Población de las Naciones Unidas - UNFPA más de 150 países programaron la realización del Censo entre 2018 y 2021: 20 países en 2018, 17 en 2019, 48 en 2020 y 71 en 2021. Colombia fue uno de los países que lo realizó en 2018, y varios de los que lo programaron para 2020 tuvieron que postponerlo, debido a la pandemia causada por el COVID-19. Esto significa, que un número importante de países están o estarán, en el futuro cercano, en el proceso de la estimación de las proyecciones de población. En el caso de Colombia, las proyecciones se hacen a nivel departamental, según área (cabecera y resto) y sexo. Es decir, cada uno de los componentes demográficos, fecundidad, mortalidad y migración, se estiman a este nivel de desagregación, lo cual dificulta su cálculo. El objetivo del documento es compartir la experiencia del equipo de investigación de la Dirección de Censos y Demografía con relación a la estimación de las curvas de mortalidad para el nivel de desagregación mencionado. Dentro de los desafíos de información y metodológicos a los que se pueden enfrentar se encuentran:

- a. Corregir el subregistro de las Estadísticas Vitales
- b. Estimar las curvas de mortalidad a nivel subnacional, no solo a nivel departamental sino por cabecera y resto de cada departamento (cabecera es equivalente al área urbana y resto se aproxima conceptualmente al área rural)¹.
- c. Estimar las curvas de mortalidad por edades simples, y no solo por grupos quinquenales, como tradicionalmente se ha hecho en Colombia.

Para presentar como se solventó cada uno de estos desafíos, este documento expone la metodología utilizada en dos partes: la primera describe la estimación de las curvas de mortalidad para la cabecera, y la segunda para el resto.

La metodología para la estimación de las curvas de cabecera consistió en:

1. Estimación de la tasa central de mortalidad a la edad x , ${}_1m_x$, y de la probabilidad de morir a la edad x , ${}_1q_x$, por sexo, edad simple y departamento a partir de las Estadísticas Vitales considerando el factor de corrección del Censo Nacional de Población y Vivienda 2018 - CNPV-2018. Este factor se calculó con la pregunta que informa si una muerte cuenta o no con certificado de defunción.
2. Suavizamiento de los valores estimados en el paso anterior. Se utilizaron los siguientes métodos de suavizamiento: a) Heligman-Pollard basado en el artículo original de 1980. Este suavizamiento se hizo considerando y sin considerar los valores atípicos. b) Se llevó a cabo el procedimiento propuesto por Alexander et al. (2017) "A Flexible Bayesian Model for Estimating Subnational Mortality". En conclusión, se compararon tres curvas: dos de Heligman-Pollard y una del Bayesiano. El criterio de elección de la curva final fue la minimización de la sumatoria de los errores al cuadrado.
3. Estimación de la mortalidad infantil por dos métodos: a) método directo descrito en el numeral 1. y b) el método indirecto de Brass y Coale (1969). Este último método se

¹ La cabecera departamental hace referencia a un área geográfica definida por un perímetro urbano, mientras que el área resto se refiere a zonas geográficas donde las viviendas se encuentran dispersas en el terreno. Para más información, consultar: https://www.dane.gov.co/files/inf_geo/4Ge_ConceptosBasicos.pdf

implementó a partir de los Censos de 1993, 2005 y 2018, obteniéndose la curva de decrecimiento de la ${}_1q_0$ desde 1985. La familia de tabla de vida para cada departamento se determinó considerando, principalmente, el consumo de leche materna durante el primer año como aconseja el Capítulo 31 de Tools for Demographic Estimation (TDE). Esta información fue tomada de la Encuesta Nacional de Situación Nutricional 2015 -ENSIN-2015. La selección de la probabilidad de morir en el primer año de vida, ${}_1q_0$ final, se basó en los rankings históricos de los departamentos.

Uno de los inconvenientes que enfrentó el equipo del DANE fue que el subregistro en el área resto no fue captado correctamente por la pregunta del certificado de defunción del CNPV-2018. Debido a esto, se llevó a cabo un procedimiento para estimar las funciones de mortalidad en los restos departamentales:

1. Se estimaron modelos logísticos de probabilidad de morir en 2017 uniendo las bases de personas vivas y fallecidas del CNPV-2018. Así, se obtuvo el factor $q_x(\text{resto})/q_x(\text{cabecera})$ por sexo y departamento.

2. Este factor se utilizó para calcular las ${}_1q_x$ del resto a partir de las ${}_1q_x$ de la cabecera.

Finalmente, las curvas calculadas se compararon con las estimadas por dos métodos demográficos: el método de la ecuación del equilibrio del crecimiento de Brass (The Brass Growth Balance Method) y el método del equilibrio del crecimiento generalizado (The Generalized Growth Balance Method). Las curvas estimadas a partir del segundo método, el cual tiene en cuenta la migración, fueron similares a las calculadas por el método implementado por el DANE.

El documento está organizado en cinco secciones, incluyendo esta introducción. La segunda sección describe el procedimiento de la estimación de las curvas de mortalidad tanto para las cabeceras y los restos departamentales. Por su parte, la sección 3 muestra los resultados de cada uno de los pasos explicados en la sección 2. La sección 4 presenta la comparación de los resultados con los dos métodos demográficos mencionados en el párrafo anterior. Finalmente, la sección 5 presenta las conclusiones.

2.

Metodología



Metodología

2.1 | Estimación de las curvas de mortalidad de las cabeceras departamentales

Esta sección del documento expone la metodología utilizada para estimar las curvas de mortalidad de las cabeceras departamentales, lo cual abarca la descripción de los métodos demográficos y estadísticos utilizados. La sección se divide en tres partes: la primera explica cómo se corrigió el nivel de las defunciones. La segunda presenta los métodos de suavizamiento utilizados para establecer el patrón. La tercera explica el cálculo de la mortalidad infantil.

a. Estimación de las tablas de vida a nivel departamental a partir de las defunciones corregidas

La Figura 1 muestra el algoritmo que se utilizó para estimar las tablas de vida a nivel departamental. El primer paso fue repartir proporcionalmente las defunciones que tenían ausencia de información en las Estadísticas Vitales para cada uno de estos tres años: 2016, 2017 y 2018. Por ejemplo, si 100 observaciones no tenían la información de edad para un departamento, área y sexo, estas se repartían proporcionalmente considerando la distribución de esta variable dentro de esta desagregación. Posteriormente, en el paso 2, se hizo el promedio de las defunciones de los tres años, con el fin de obtener el número de defunciones para 2017.

En el paso 3 se estimaron los factores de corrección del subregistro de las defunciones en las Estadísticas Vitales. Estos factores de corrección

se calcularon con base en la pregunta 26 del formulario del CNPV-2018: *¿Cuántas personas que eran miembros de este hogar fallecieron en el 2017?*. Para cada una de las personas fallecidas se preguntó por el sexo, la edad, y si se había expedido el certificado de defunción. Se estimaron dos factores de corrección, el primero fue el inverso de la proporción de las defunciones con certificado y el segundo fue la fórmula utilizada en 2005: $1 + (\text{difuntos sin registro} / (\text{difuntos sin registro} + \text{difuntos con registro}))$. Esta fórmula da valores cercanos al inverso de la proporción de las defunciones con certificado cuando la proporción de defunciones sin certificado es pequeña, menor al 20%, como usualmente sucede en las cabeceras municipales.

Al tener el número de defunciones para 2017 y el factor de corrección del CNPV-2018, se calculó el número de defunciones corregidas por edad, sexo, área y departamento al multiplicar ambas estimaciones (paso 4): *Número de defunciones corregidas = Número de defunciones de las Estadísticas Vitales * Factor de corrección*

A partir de las defunciones corregidas se estimaron las m_x , en el paso 5. Las ${}_1m_x$ se suavizaron con el método de tasas acumuladas de Arriaga (2011), paso 6. Por último, se calcularon los factores de separación y el resto de las funciones de la tabla de vida. Los productos finales de este procedimiento fueron: 1) Curvas ${}_1q_x$ y ${}_1m_x$ por departamento, área y sexo y 2) Curvas ${}_5q_x$ y ${}_5m_x$ para la misma desagregación.

Figura 1. | Algoritmo para el cálculo de las tablas de vida a nivel departamental a partir de las defunciones corregidas



Fuente: Dirección de Censos y Demografía, DANE

b. Suavizamiento y estimación del patrón

Después de estimar la tabla de vida con las defunciones corregidas, se suavizaron las curvas y se estimó el patrón de la mortalidad. Para esto se usaron dos métodos: “The age pattern of mortality” propuesto por L. Heligman y H.J. Pollard en 1980 y el método propuesto por Alexander et al. (2017) “A Flexible Bayesian Model for Estimating Subnational Mortality”. El paso a paso del primer método se ilustra en la Figura 2, mientras que el del segundo en la Figura 3.

La Figura 2 muestra que el primer paso para llevar a cabo el Heligman-Pollard fue identificar los valores atípicos en las ${}_1q_x$ por edad. Para este fin se estimó la diferencia $\ln({}_1q_x) - \ln({}_1q_0) = d_x$ para todas las edades por sexo y departamento. Posteriormente, para cada edad y sexo, se estimó los cuantiles de la variable d_x a nivel departamental. El criterio para identificar un valor atípico para una edad x por sexo fue que esta estuviera por fuera del intervalo:

$$[Q_1(d_x) - 1.5(Q_3(d_x) - Q_1(d_x)), \\ Q_3(d_x) + 1.5(Q_3(d_x) - Q_1(d_x))]$$

Donde, $Q_1(d_x)$ es el primer cuantil de d_x y $Q_3(d_x)$ es el tercer cuantil. Por ejemplo, se tomaron los valores departamentales de la diferencia d_x para la edad x de las mujeres que residían en cabecera; a estos valores se les calculo los valores del primer y tercer cuartil. Los valores departamentales que se identificaron como atípicos para la edad x y este grupo de población, mujeres de cabecera, son los que estuvieron por fuera del intervalo estimado con los cuantiles de este grupo.

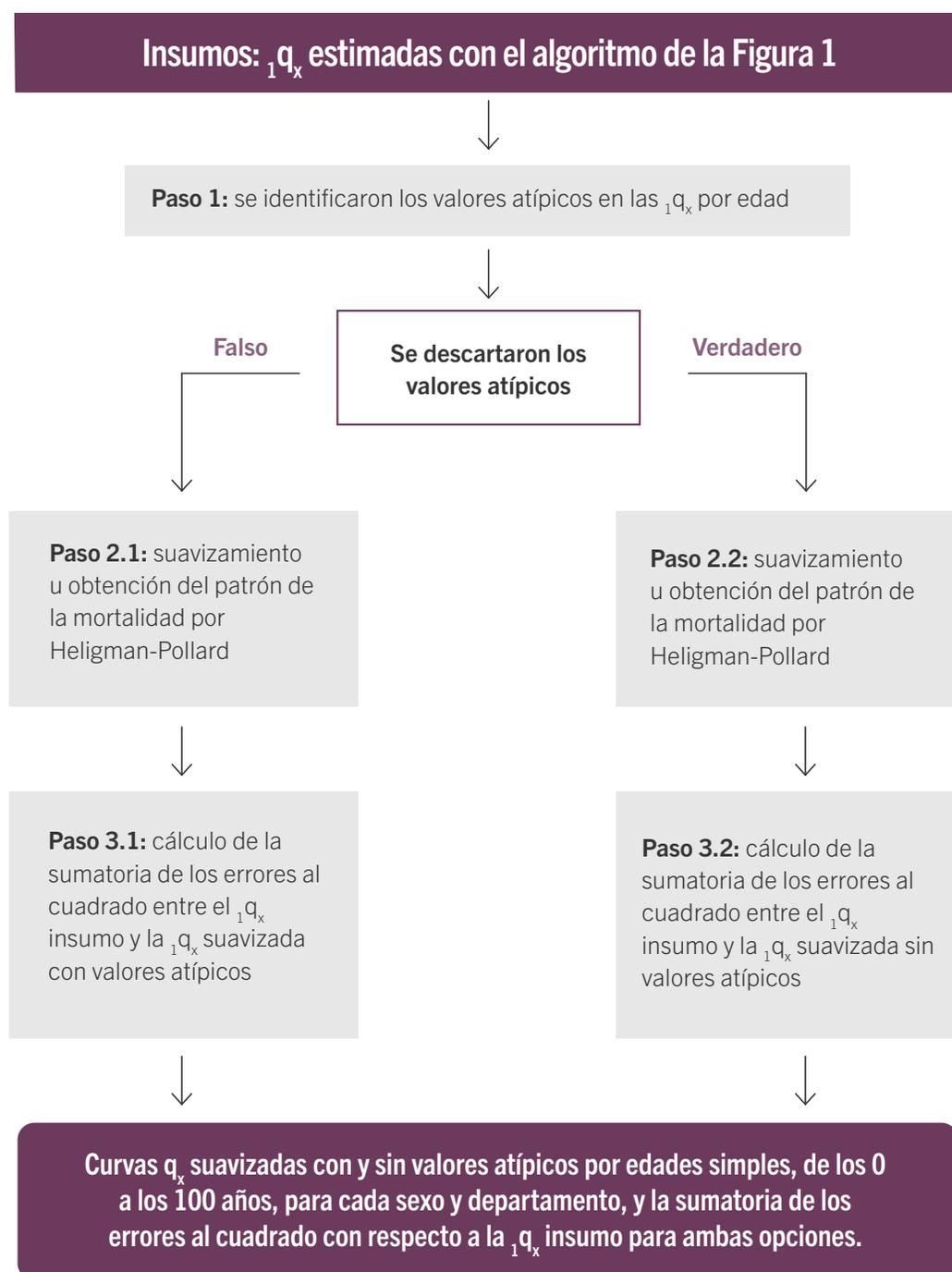
El segundo paso se divide en dos, como muestra la Figura 2. El paso 2.1 es obtener el patrón de la mortalidad utilizando método de Heligman-Pollard considerando los valores atípicos de ${}_1q_x$ y el 2.2

es realizar esta estimación sin considerar dichos valores. De igual forma, el paso 3 se divide en dos, en el 3.1 se calcula la sumatoria de los errores al cuadrado entre el ${}_1q_x$ insumo (resultado del algoritmo de la Figura 1) y el ${}_1q_x$ suavizado con valores atípicos y en el 3.2 se halla el mismo indicador, pero con el ${}_1q_x$ suavizado sin valores atípicos.

El resultado final del algoritmo de la Figura 2 es: curvas ${}_1q_x$ suavizadas con y sin valores atípicos por edades simples, de los 0 a los 100 años, para cada sexo y departamento, y sumatoria de los errores al cuadrado con respecto a la ${}_1q_x$ insumo, que se obtuvo con el algoritmo de la Figura 1, para ambas opciones (con y sin valores atípicos).

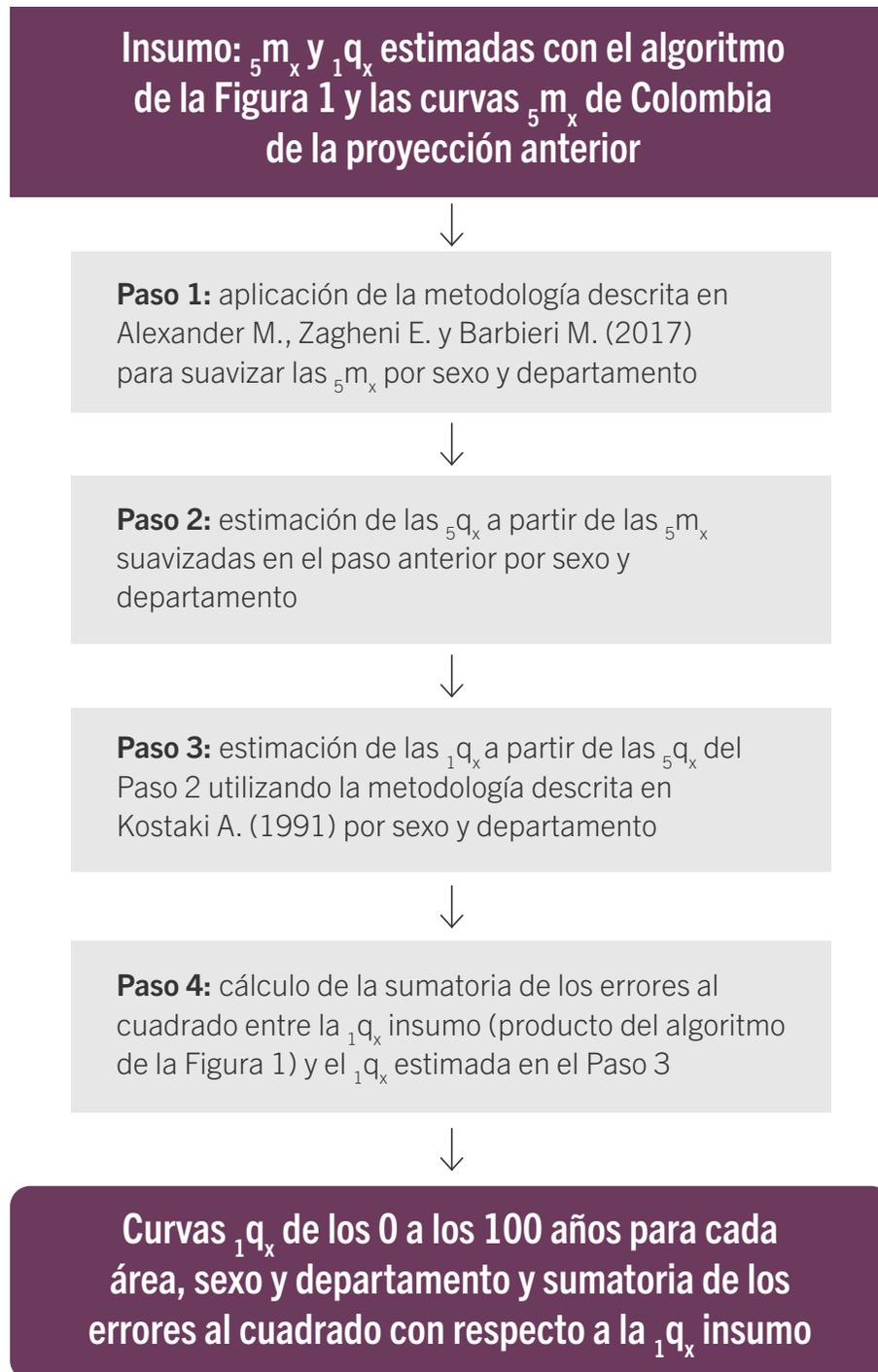
La Figura 3 presenta cómo se implementó el método de Alexander et al. (2017). Para su implementación se utilizaron dos productos del algoritmo de la Figura 1, ${}_5m_x$ y ${}_1q_x$, y las curvas de mortalidad estimadas para Colombia en las proyecciones de 2005 (este es uno de los insumos del método Bayesiano de Alexander et al.). El primer paso fue aplicar el método por departamento y sexo utilizando las respectivas ${}_5m_x$, de esta manera se logró suavizar el patrón de las ${}_5m_x$. El paso siguiente fue estimar las ${}_5q_x$ a partir de las ${}_5m_x$ suavizadas en el paso anterior, esto también se hizo para cada una de las desagregaciones. El paso 3 consistió en estimar las ${}_1q_x$ con base en las ${}_5q_x$ del paso 2; para esto se usó la metodología propuesta por Kostaki (1991). Por último, se calculó la sumatoria de los errores al cuadrado de las ${}_1q_x$ estimadas en el paso 3 con relación a las ${}_1q_x$ resultado del algoritmo de la Figura 1. Por tanto, los productos finales del algoritmo de la Figura 3 son: curvas ${}_1q_x$ de los 0 a los 100 años para cada sexo y departamento, y la sumatoria de los errores al cuadrado con respecto a la ${}_1q_x$ estimadas en el algoritmo de la Figura 1.

Figura 2. | Algoritmo del suavizamiento o estimación del patrón de la mortalidad por Heligman-Pollard



Fuente: Dirección de Censos y Demografía, DANE

Figura 3. | Algoritmo del suavizamiento o estimación del patrón de la mortalidad por Alexander-Zagheni-Barbieri



Fuente: Dirección de Censos y Demografía, DANE

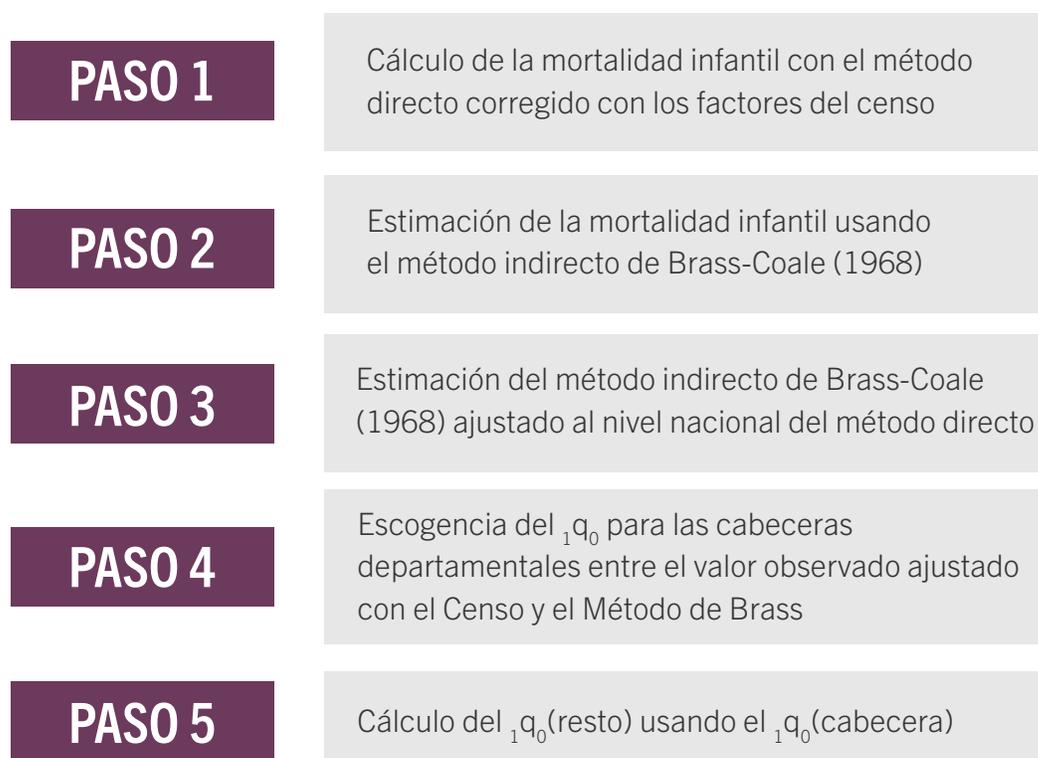
Al finalizar los algoritmos de las Figuras 2 y 3, se obtienen tres curvas ${}_1q_x$ para cada desagregación con la respectiva sumatoria de los errores al cuadrado: dos de Heligman-Pollard, con y sin valores atípicos, y una del Bayesiano de Alexander et al. El criterio de elección de la curva final por sexo para las cabeceras departamentales fue la minimización de la sumatoria de los errores al cuadrado.

c. Estimación de la mortalidad infantil en 2017

La estimación de la mortalidad infantil se hizo siguiendo los pasos ilustrados en la Figura 4. Los pasos 1, 2 y 3 se realizaron tanto para la cabecera

y el resto en el nivel departamental. Así como en el cálculo de la mortalidad de mayores de 0, los resultados del resto no fueron consistentes, presentando niveles de mortalidad menores a los de las cabeceras o valores extremadamente altos. Debido a esto, solo los resultados de cabecera de los pasos 1, 2 y 3 se tuvieron en cuenta. Por esto, en el paso 4 se escogió el valor más adecuado para la cabecera de los departamentos entre los obtenidos por los métodos directo corregido e indirecto ajustado, pasos 1 y 3, respectivamente. Posteriormente, en el paso 5, se calcularon los valores del resto a partir de los de cabecera usando la razón $q_x(\text{resto})/q_x(\text{cabecera})$, estimada a partir de modelos de probabilidad logísticos.

Figura 4. | Paso a paso de la estimación de la mortalidad infantil en el nivel departamental – cabecera y resto



Fuente: Dirección de Censos y Demografía, DANE

El procedimiento que se llevó a cabo, en el paso 2 de la Figura 4, para estimar la mortalidad infantil a partir del método indirecto de Brass-Coale (1968) fue el descrito por Kennet Hill (2013) en el libro TDE. Este procedimiento se utilizó para los cuatro últimos censos: 1985, 1993, 2005 y 2018. No obstante, solo los resultados de los últimos tres se consideraron, debido a que en estos preguntaron por separado acerca de la sobrevivencia de los hijos y de las hijas. Las preguntas que se utilizaron fueron:

1. ¿Ha tenido algún hijo o hija que haya nacido vivo?

Sí ¿Cuántos hombres?

¿Cuántas mujeres?

No

2. ¿De los hijos e hijas, que nacieron vivos(as), de...

¿Cuántos están vivos actualmente?

¿Cuántos? ¿Cuántos hombres?

¿Cuántas mujeres?

No sabe

3. ¿Cuál es el mes y año del último hijo e hija nacido vivo o viva?

Las estimaciones se realizaron con las observaciones válidas de estas tres preguntas. Es decir, aquellas que tenían la información completa. Por tanto, las mujeres que respondieron que tenían hijos en la primera pregunta, pero respondieron que no sabían cuántos de ellos estaban vivos en la segunda no fueron incluidas. Esta es una de las advertencias que hace TDE, ya que la inclusión de estas mujeres sobreestimaría la probabilidad de morir. En el caso colombiano esta sobreestimación sería de alrededor del 0,02 o 2 puntos porcentuales para cada año. Para estimar la probabilidad de morir durante el primer año de vida, ${}_1q_0$, se necesitan los siguientes insumos de acuerdo con TDE:

${}_5N_x$: número de mujeres agrupadas en quinquenios de edad.

${}_5CEB_x$: número de niños nacidos vivos para cada grupo quinquenal.

${}_5CS_x$: número de niños nacidos vivos que han sobrevivido hasta el día del Censo para cada grupo quinquenal.

${}_5f_x$: tasa específica de fecundidad a la edad x.

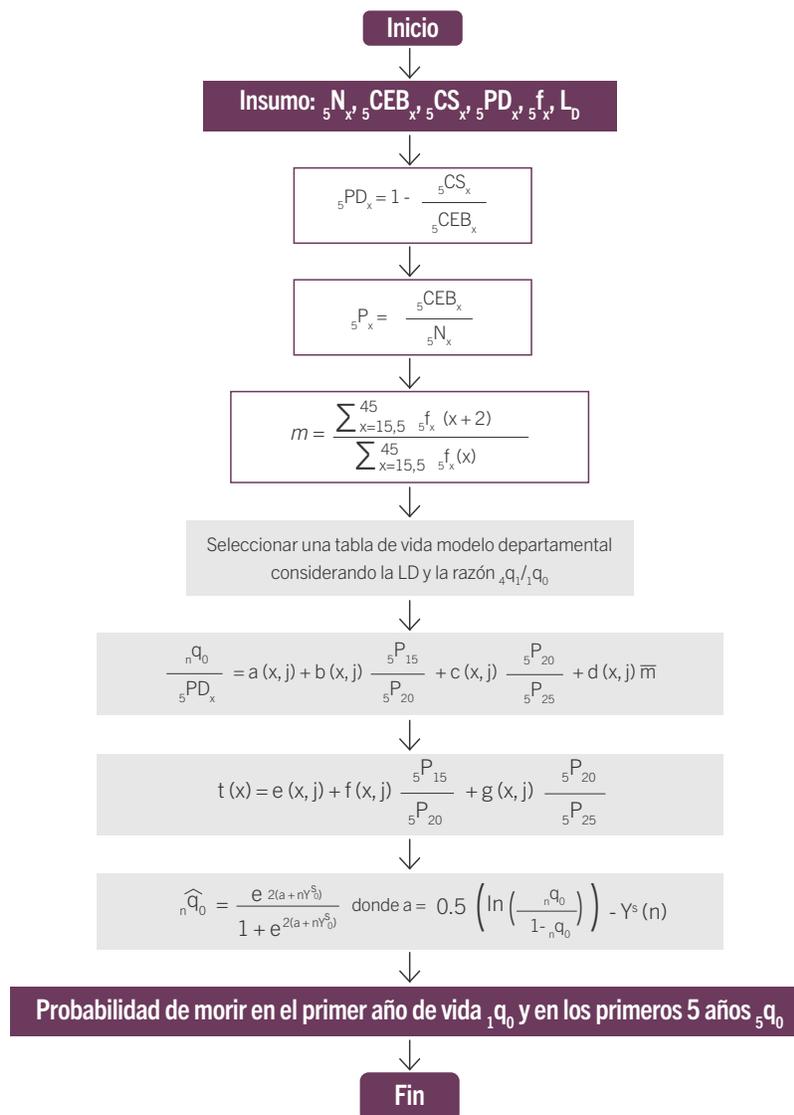
Adicionalmente, para escoger la tabla de vida modelo por departamento, se usaron como insumos el indicador de la lactancia materna continúa al año de vida, L_D , y las ${}_1q_0$ y ${}_4q_1$ reportadas por las Encuestas Nacionales de Demografía y Salud - ENDS 2010 y 2015. Sin embargo, por la consistencia de la información, el principal criterio fue la lactancia materna continúa al año de vida. Este indicador es la proporción de niños y niñas de 12 a 15 meses de edad que fueron amamantados durante el día anterior a la encuesta, y se calculó con la Encuesta Nacional de Situación Nutricional - ENSIN 2015. El capítulo 31 de TDE explica que en las poblaciones donde la lactancia materna es una práctica común y el destete ocurre a una edad tardía (12 meses o más), se espera que la mortalidad en la niñez sea relativamente alta en comparación a la infantil. Esto se debe a que la leche materna previene la desnutrición y las enfermedades infecciosas. Por tanto, después del destete el niño está menos protegido y es más probable que muera.

La Figura 5 ilustra el algoritmo del cálculo de la probabilidad de morir en el primer año y en los primeros 5 años de vida. La información de entrada son los insumos que se explicaron anteriormente: ${}_5N_x$, ${}_5CEB_x$, ${}_5CS_x$, ${}_5PD_x$, ${}_5f_x$ y L_D . El primer paso es el cálculo de la proporción de niños muertos para cada grupo quinquenal, ${}_5PD_x$. Luego se calcula el número promedio de niños nacidos por mujer en cada grupo

quinquenal, ${}_5P_x$. Después se estima la edad media de la fecundidad, m . Posteriormente se selecciona la tabla de vida a partir de la L_D y la razón ${}_4q_1/{}_1q_0$; este paso se explicará en detalle en la sección siguiente. Cada tabla de vida tiene unos parámetros asociados para calcular el ${}_nq_0$ para cada ${}_5PD_x$ y para estimar el tiempo de referencia $t(x)$. Los parámetros del primer

cálculo son $a(x,j)$, $b(x,j)$, $c(x,j)$ y $d(x,j)$. Por su parte, los parámetros para el tiempo de referencia son $e(x,j)$, $f(x,j)$ y $g(x,j)$. Por último, con base en $Y^s(1)$ y $Y^s(5)$ que son la transformaciones logit de las tablas de vida estándar, se convierte cada estimación de ${}_nq_0$ en una estimación de ${}_1q_0$ y de ${}_5q_0$.

Figura 5. | Algoritmo para calcular la probabilidad de morir en el primer año ($n=1$) y en los primeros 5 años ($n=5$) de vida



Fuente: Dirección de Censos y Demografía basados en el Capítulo 16 de TDE

Nota: ${}_nq_0$: probabilidad de morir durante los primeros n años de vida, ${}_5N_x$: número de mujeres agrupadas en quinquenios de edad, ${}_5CEB_x$: número de niños nacidos vivos para cada grupo quinquenal, ${}_5CS_x$: número de niños nacidos vivos que han sobrevivido hasta el día del Censo para cada grupo quinquenal, ${}_5PD_x$: proporción de niños nacidos vivos que han muerto por grupo quinquenal, $Y^s(n)$ son transformaciones logit de las tablas de vida estándar y L_D : lactancia materna continúa al año de vida.

Las **tablas de vida** modelo proveen una descripción sucinta de los aspectos más relevantes de la mortalidad, mostrando los cambios en la probabilidad de morir en función de la edad. Las tablas usadas en las proyecciones y retroproyecciones 2020 son las referenciadas en TDE. Estas son las de Princeton (Norte, Sur, Este y Oeste) y las de las Naciones Unidas (General, América Latina, Chile, Sur de Asia y Lejano Oriente).

Según TDE, y como se explicó antes, en una población donde la lactancia es una práctica común y donde el destete ocurre a una edad relativamente tardía, 12 meses o más, es razonable esperar que la ${}_4q_1$ sea relativamente alta en comparación con la ${}_1q_0$. En estos casos, es probable que la

mortalidad en la infancia esté bien representada por el estándar de Princeton North o General de las Naciones Unidas. En cambio, en poblaciones donde la lactancia ha sido abandonada por una proporción alta de mujeres se espera que la ${}_1q_0$ sea relativamente alta en comparación a la ${}_4q_1$. En estas condiciones el modelo Este de Princeton o el de Chile de las Naciones Unidas pueden representar la mortalidad infantil. Considerando estas sugerencias, se establecieron rangos para el indicador de lactancia materna continúa al año de vida para cada tabla modelo, ver Tabla 1. Estos rangos no fueron una condición estricta para escoger la tabla de vida modelo, pero sirvieron para saber aproximadamente donde estaba ubicado cada departamento.

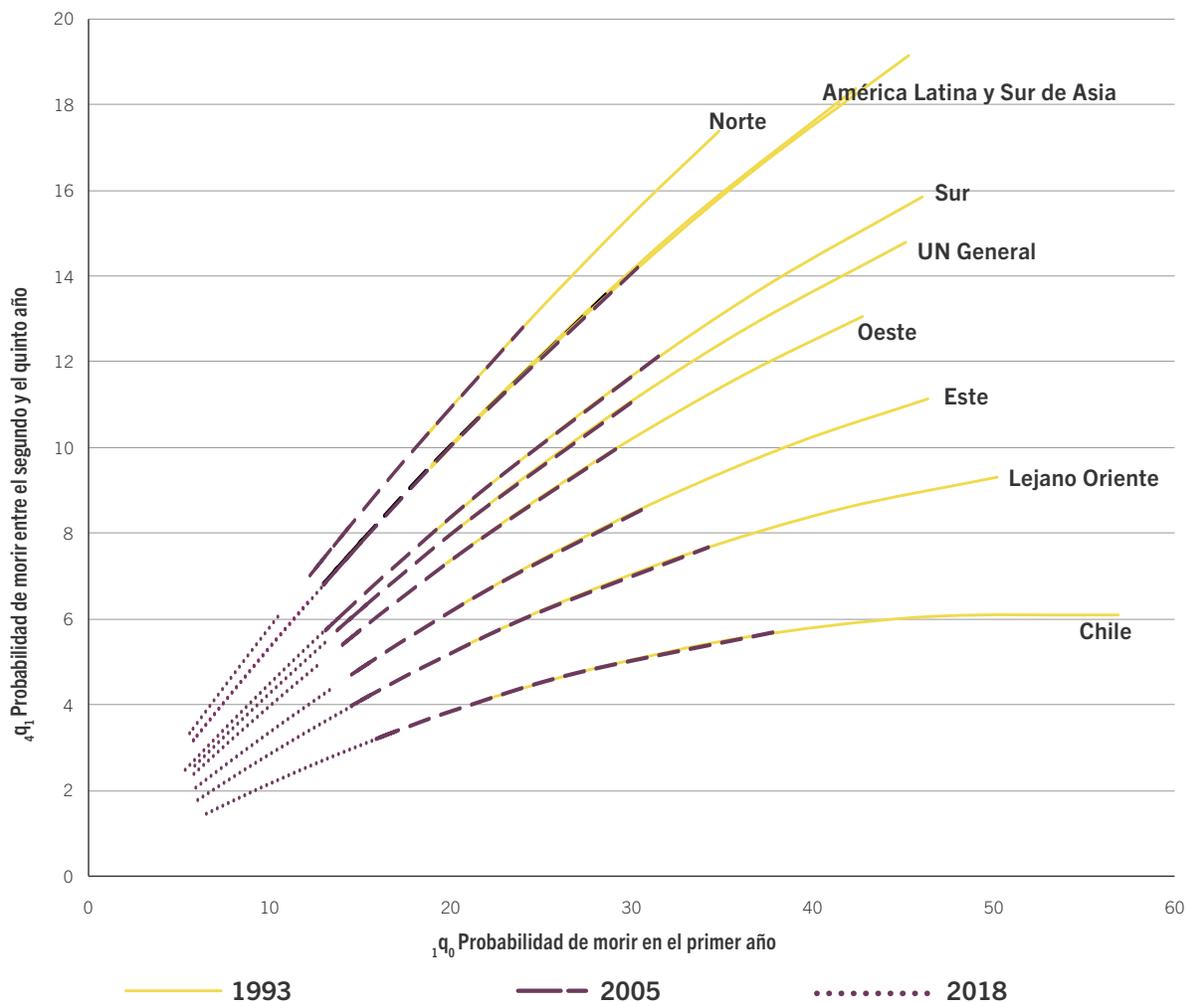
Tabla 1. | Criterios para escoger las tablas de vida modelo

Familia	Lactancia continua al primer año	${}_4q_1/{}_1q_0$
Princeton North	[80-100]	(0,46- 1)
Latin America or South Asia	[70-80)	(0,38 - 0,46]
UN General or South	[60-70)	(0,32 -0,38]
West	[40-60)	(0,27 - 0,32]
Far East	[30-40)	(0,21 - 0,27]
East	[20-30)	(0,14 - 0,21]
Chile	[0-20)	[0 - 0,14]

Fuente: Dirección de Censos y Demografía, DANE

Aunque, TDE indica que las condiciones higiénicas también son importantes, no se cuenta con un indicador para estas. Por este motivo, y para tener más criterios de elección, se establecieron unos rangos para la relación ${}_4q_1/{}_1q_0$ a partir de los resultados nacionales, que se presentan el Gráfico 1. Este gráfico muestra la ${}_1q_4$ versus la ${}_1q_0$ estimadas con los últimos tres Censos (1993, 2005 y 2018) para el total nacional utilizando el método indirecto de Brass-Coale (1968). Como se observa en el gráfico, la razón ${}_1q_4/{}_1q_0$ más alta corresponde a la de la familia Princeton-North, mientras la más baja a la de las Naciones Unidas-Chile. Considerando las pendientes de esta gráfica se establecieron los rangos que se presentan en la tercera columna de la Tabla 1, y que se utilizaron para clasificar a los departamentos de acuerdo con las ${}_4q_1$ y ${}_1q_0$ reportadas por las ENDS 2010 y 2015.

Gráfico 1. | Estimaciones del ${}_1q_0$ y ${}_4q_1$ con los Censos de 1993, 2005 y 2018 - Colombia

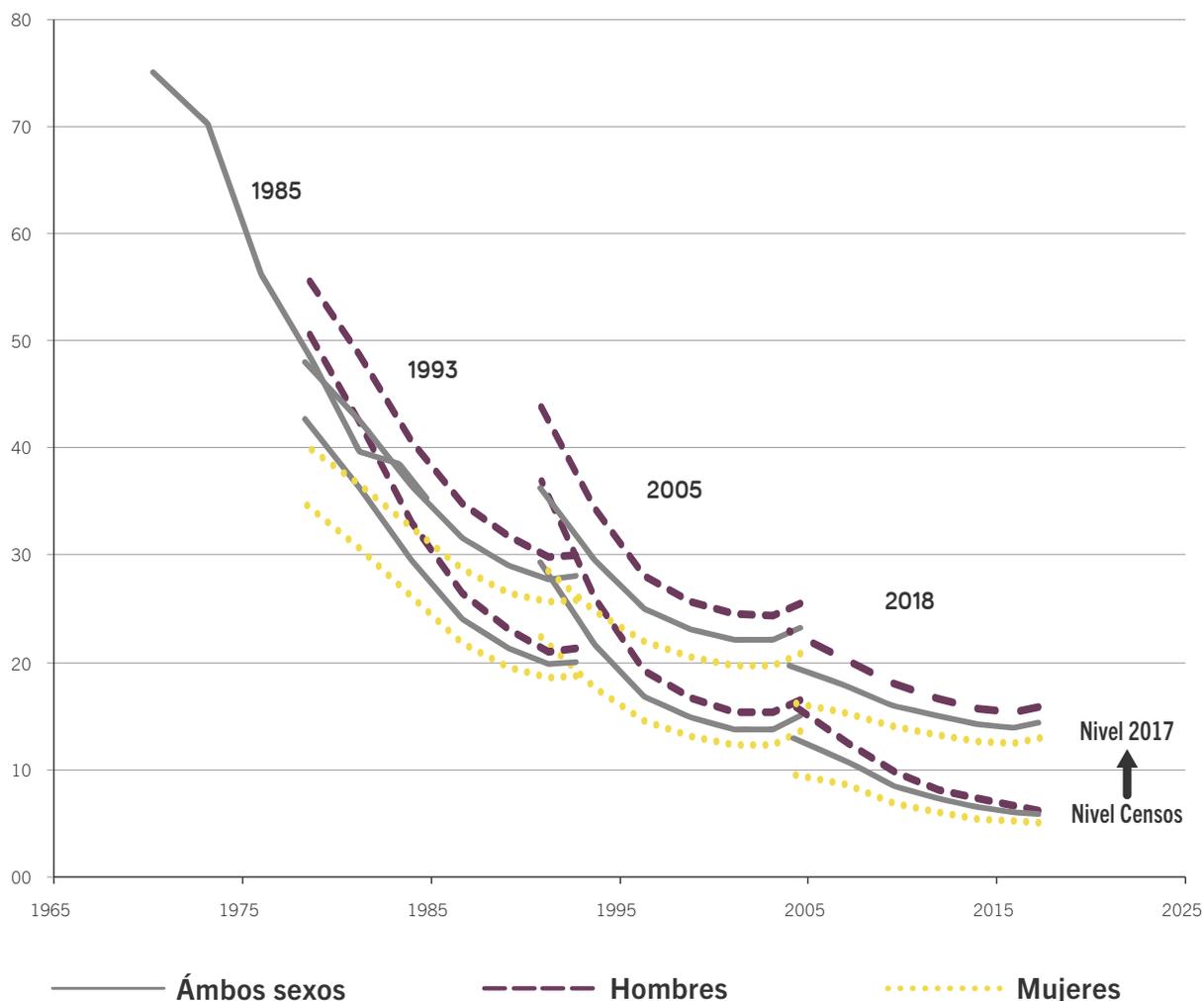


Fuente: Censos de 1993, 2005 y 2018, Dirección de Censos y Demografía, DANE

El paso 3 para la estimación de la mortalidad infantil en el nivel departamental – cabecera y resto (ver Figura 4) se hizo a través del método indirecto de Brass-Coale (1968) ajustado a los niveles del método directo corregido. Lo primero que se hizo en este paso fue ajustar el nivel nacional del método de Brass-Coale al nivel que se obtuvo con el método directo, el cual uso los factores de corrección del Censo. Este ajuste se observa en el Gráfico 2. A partir del nuevo nivel nacional,

se estimó el número de defunciones necesarias para que el método de Brass-Coale alcanzara este nivel de mortalidad infantil. Posteriormente, estas defunciones se distribuyeron entre los grupos quinquenales y departamentos de manera proporcional al número de hijos muertos para cada desagregación. Con este nuevo nivel de hijos muertos se volvió a realizar el algoritmo de la Figura 5 para todos los departamentos y ambos sexos.

Gráfico 2. | Mortalidad infantil nacional estimada con el Método de Brass-Coale (1968) utilizando los últimos cuatro censos y ajuste al nivel de 2017



Fuente: Censos de 1985, 1993, 2005 y 2018, Dirección de Censos y Demografía, DANE

Según la Figura 4, el paso 4 es la elección de la ${}_1q_0$ entre el valor observado corregido con el Censo (Paso 1) y el Método de Brass-Coale ajustado (Paso 3). Para este fin, se tuvo en cuenta los rankings de este indicador a nivel departamental estimados por el DANE para 1985, 1993 y 2005. A partir de los rankings se calculó un índice igual a 0.25 posición del departamento en 1985 + 0.25 posición en 1993 + 0.5 posición en 2005.

Con base en este índice, se decidió el nivel más adecuado del ${}_1q_0$ por departamento.

El escoger para unos departamentos el resultado del paso 1 y para otros el del paso 3 afectó el valor nacional. Por tanto, fue necesario ajustar levemente las estimaciones departamentales escogidas en este paso para que al agregarlas dieran el valor nacional.

El procedimiento del paso 5 de la Figura 4 para el cálculo del ${}_1q_0(\text{resto})$ usando el ${}_1q_0(\text{cabecera})$ se explica en detalle en la siguiente sección.

2.2 | Estimación de las curvas de mortalidad para los restos departamentales

Los factores de corrección del Censo 2018 para el área resto no dieron resultados consistentes; para algunos departamentos, estos factores fueron considerablemente bajos, obteniéndose menores probabilidades de morir que en la cabecera, o muy altos, haciendo que estas probabilidades fueran extremadamente elevadas. Lo que se deduce de este resultado es que una proporción importante de personas en el resto no conocen la existencia o no tienen claro el uso del certificado de defunción. Sin embargo, lo que ellos sí saben es si algún miembro del hogar murió en 2017. Partiendo de esta premisa se decidió estimar la probabilidad de morir a la edad x con el Censo, con el fin de obtener la razón $q_x(\text{resto})/q_x(\text{cabecera})$ por sexo y departamento. Al tener esta razón y las curvas ${}_1q_x$ de cabecera, la estimación de las curvas de resto fue directa.

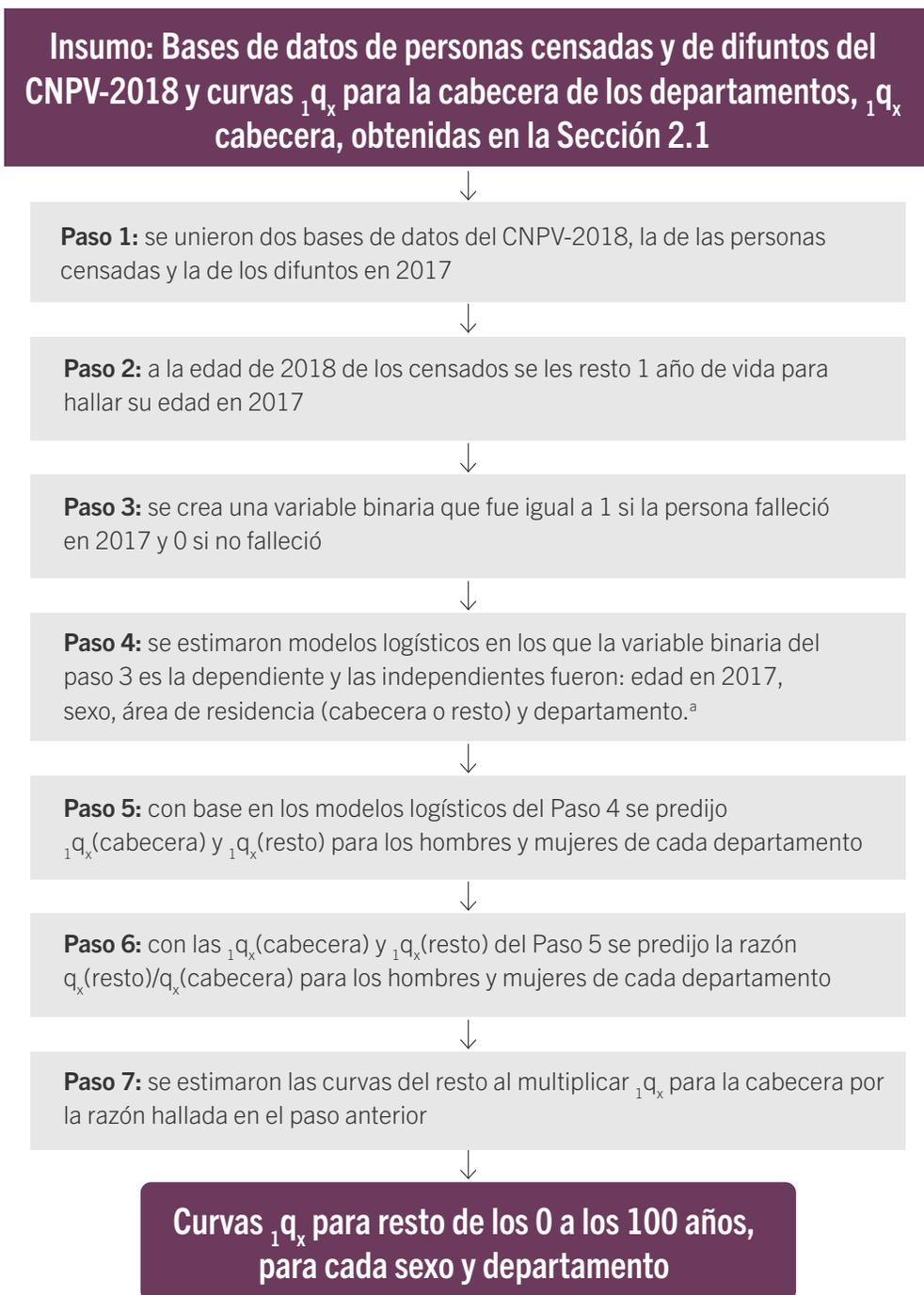
La Figura 6 muestra el procedimiento de la estimación de las curvas del área resto de los departamentos. Para realizar la estimación $q_x(\text{resto})/q_x(\text{cabecera})$, se unieron dos bases: la que contenía la información de las personas censadas en 2018 y la de los difuntos reportados en el Censo. El primer paso fue hallar la edad en 2017 de las personas censadas en 2018 para que fuera comparable a la de los difuntos; es decir, se les resto un año de vida. Posteriormente, se creó una variable binaria que era igual a 1 si la persona falleció en 2017 y 0 si no falleció. Esta fue la variable dependiente en un modelo logístico que consideraba las siguientes variables

independientes: edad en 2017, sexo, área de residencia (cabecera o resto) y departamento. A partir del modelo logístico se predijo la probabilidad de morir a la edad x , ${}_1q_x$, por área, cabecera, ${}_1q_x(\text{cabecera})$, y resto, ${}_1q_x(\text{resto})$, y para cada sexo y departamento. Así, fue posible estimar la razón $q_x(\text{resto})/q_x(\text{cabecera})$ para los hombres y las mujeres de todos los departamentos. Al tener esta razón para todas las edades, para las mujeres de un departamento, por ejemplo, y la curva de mortalidad de cabecera deducida en la Sección 2.1, la estimación de la ${}_1q_x$ para resto fue directa:

$${}_1q_{x \text{ resto}} = {}_1q_{x \text{ cabecera}} \cdot q_x(\text{resto})/q_x(\text{cabecera})$$

Las estimaciones de los modelos logísticos se hicieron sobre grupos de departamentos considerando similitudes en las condiciones de vida y la diferencia de estas condiciones entre cabecera y resto, ya que estas son determinantes importantes de la mortalidad. Para este fin, se hicieron clusters de departamentos a partir del NBI total y la diferencia del NBI del resto y el de la cabecera. Adicionalmente, los modelos se estimaron por rangos de edad: 0, 1 a 4, 5 a 14, 15 a 29, 30 a 64, y 65 años y más. Se asumió población cerrada debido a que no se preguntó el departamento de residencia del difunto al momento de morir; por tanto, en el caso de los hogares migrantes no se sabe si el difunto murió en el lugar de residencia de hace un año o en el actual. Es así que, se recomienda incluir esta pregunta en futuros censos.

Figura 6. | Algoritmo para estimar las curvas de mortalidad de los restos de los departamentos



Fuente: Dirección de Censos y Demografía, DANE

Nota: ^a Los modelos logísticos fueron estimados por clusters de departamentos a partir del NBI total y la diferencia del NBI del resto y el de la cabecera; además, se estimaron modelos separados para los siguientes rangos de edad: 0, 1 a 4, 5 a 14, 15 a 29, 30 a 64, y 65 años y más.

3.

Resultados



Resultados

Esta sección presenta los resultados para La Guajira, uno de los departamentos con mayor nivel de subregistro de muertes en los Registros de Estadísticas Vitales.

3.1 | Estimación de las curvas de mortalidad en las zonas urbanas departamentales

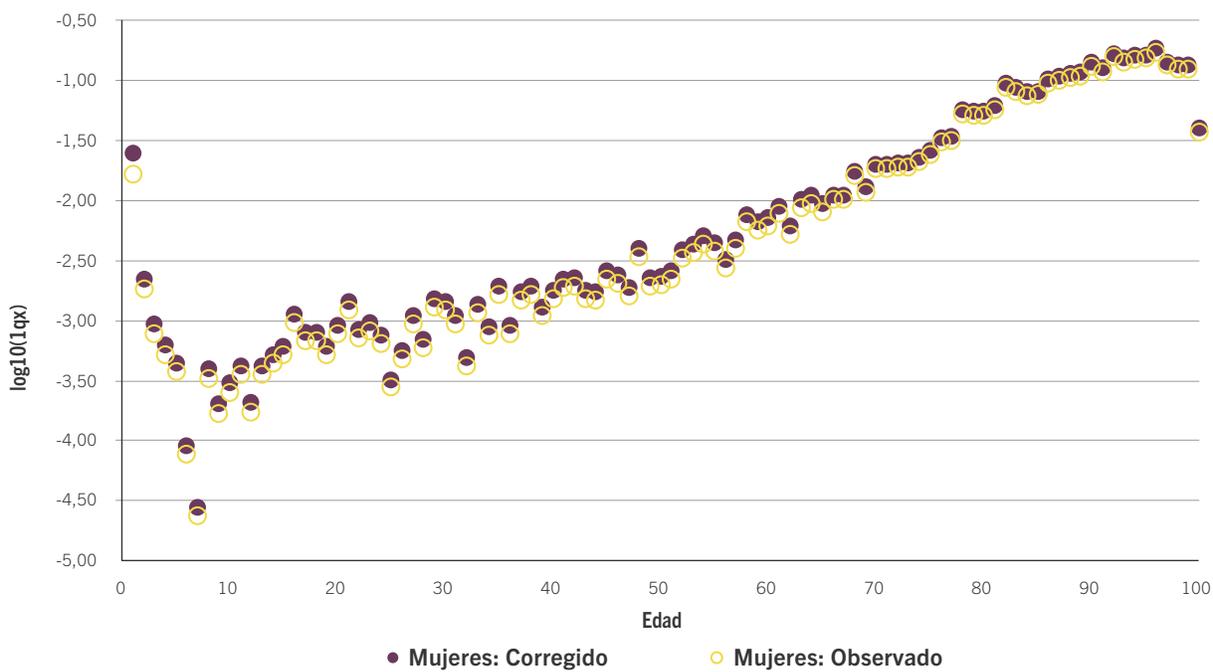
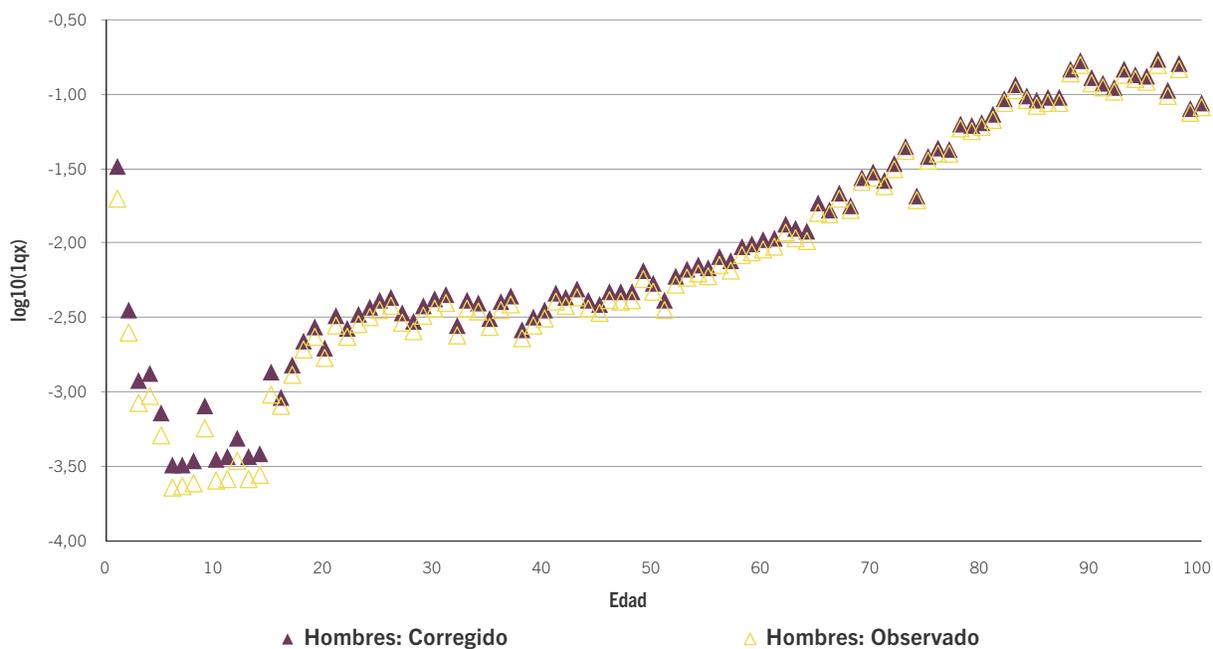
Esta parte muestra los resultados de los tres pasos explicados en la sección 2.1, los cuales se utilizaron para estimar las curvas de mortalidad en la cabecera a nivel departamental. Estos pasos son: **a.** uso del factor de corrección del Censo 2018 para ajustar el nivel de las curvas de mortalidad, **b.** suavizamiento de las curvas u obtención del patrón y **c.** estimación de la mortalidad infantil.

a. Estimación de las tablas de vida a nivel departamental a partir de las defunciones corregidas

En el Gráfico 3 se observa las ${}_1q_x$ estimadas para la cabecera de La Guajira, considerando y sin considerar los factores de corrección del Censo. Para este departamento, estos factores fueron calculados por separado para las áreas de cabecera y resto, y diferentes rangos de edad. En el caso de las cabeceras los factores fueron: 1,47 para el grupo de edad de 0 años, 1,19 para el de 1 a 14, 1,17 para el de 15 a 64 y 1,09 para los de 65 años y más. Por lo tanto, en el Gráfico 3 los marcadores púrpura son iguales a los amarillos multiplicados por factor de corrección.



Gráfico 3. | Probabilidad de morir a la edad x , q_x , para la cabecera de La Guajira estimado con las defunciones sin corregir y corregidas con el factor del CNPV- 2018

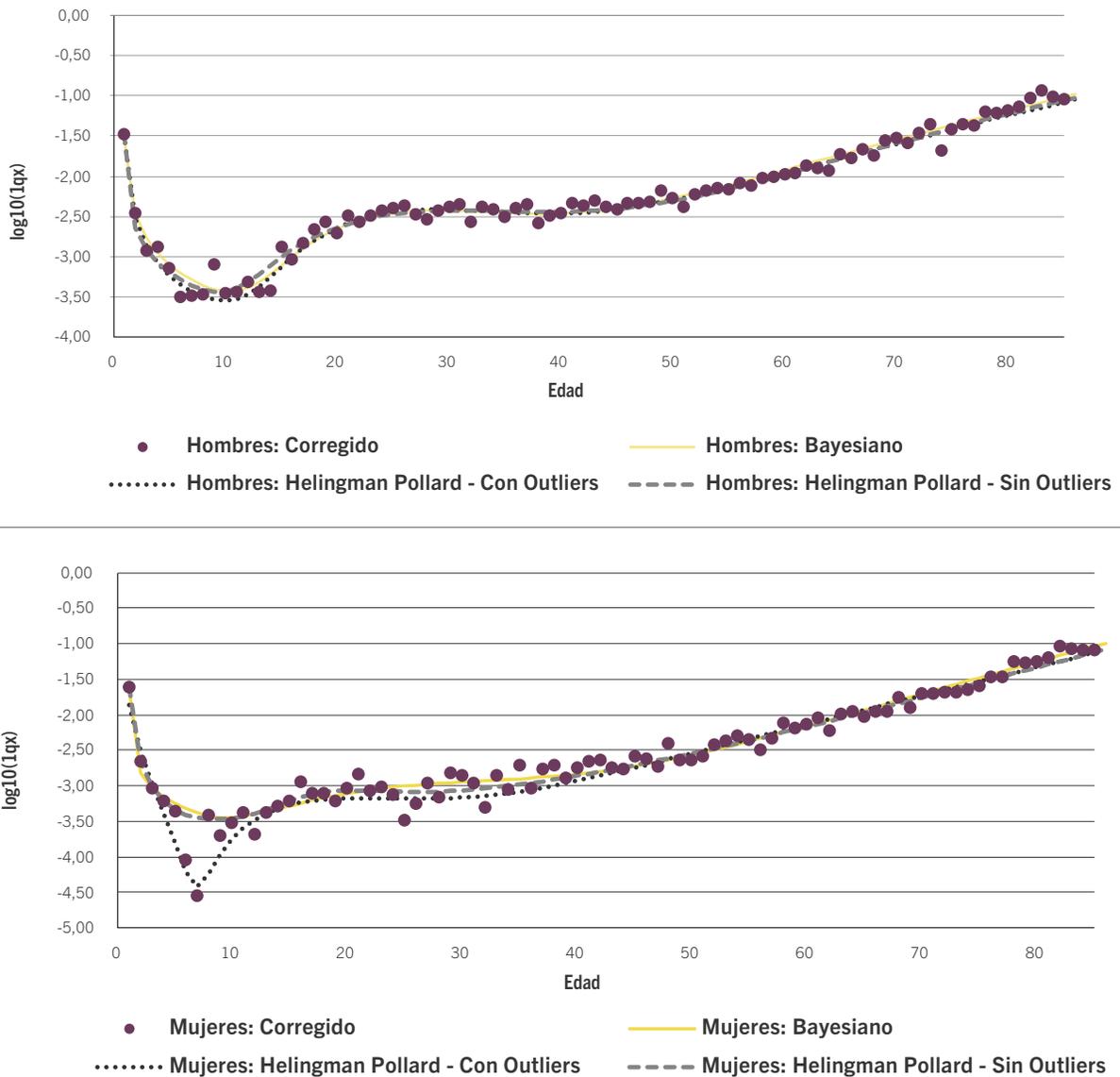


Fuente: EEV 2016-2018 y CNPV-2018, Dirección de Censos y Demografía, DANE

b. Suavizamiento o estimación del patrón

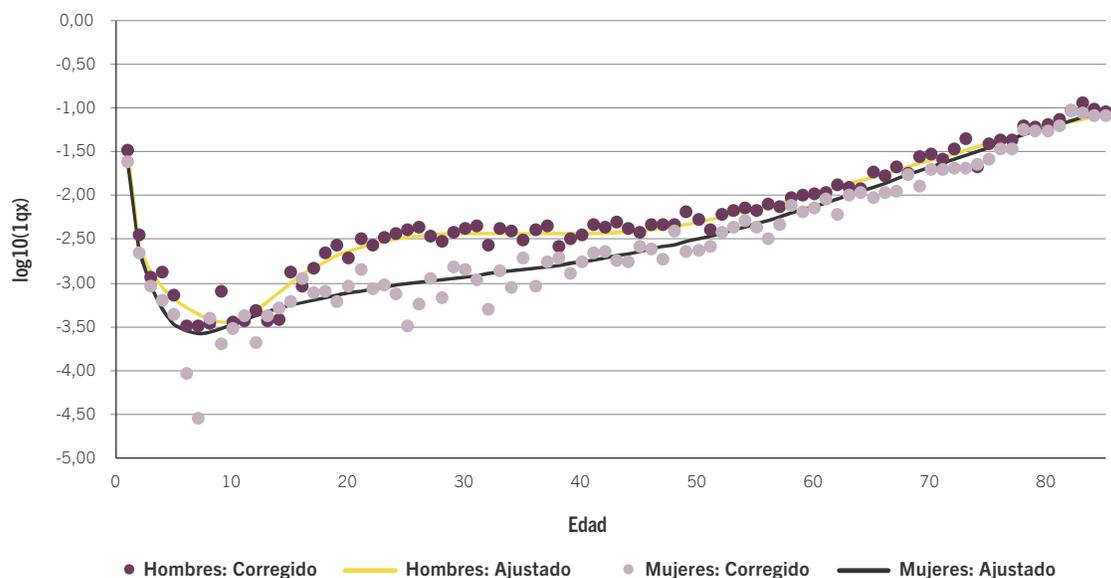
El Gráfico 4 presenta el ajuste de las tres curvas para La Guajira: Heligman Pollard considerando valores atípicos, Heligman Pollard sin considerar estos valores y el ajuste Bayesiano de Alexander et al. El Gráfico 5 muestra la curva seleccionada para cada sexo. En el caso de La Guajira, el suavizamiento vía Heligman Pollard sin considerar valores atípicos se eligió para los hombres, y el ajuste Bayesiano para la curva de las mujeres.

Gráfico 4. | Métodos de suavizamiento de la curva de ${}_1q_x$,
 La Guajira - cabecera



Fuente: EEV 2016-2018 y CNPV-2018, Dirección de Censos y Demografía, DANE

Gráfico 5. | Curva q_x seleccionada para la cabecera de La Guajira – cabecera



Fuente: EEVV 2016-2018 y CNPV-2018, Dirección de Censos y Demografía, DANE

c. Estimación de la mortalidad infantil en 2017

Aunque, el indicador principal para escoger la tabla de vida modelo fue el de lactancia materna, también se consideraron criterios de contexto. Por ejemplo, a varios departamentos del Amazonas y de la Orinoquía, según la Tabla 1, les corresponde Princeton North. Sin embargo, esta familia describe el comportamiento de tres países nórdicos: Noruega, Suecia e Islandia. Por considerarlos contextualmente diferentes a la Amazonía y Orinoquía colombiana, se escogió la tabla modelo más cercana a la realidad de estas regiones, pero que a su vez la mortalidad postinfantil fuera relativamente alta en comparación a la mortalidad infantil, tal como ocurre con la familia Princeton North. Por tanto, para los departamentos de la Amazonía y la Orinoquía se escogió entre América Latina, Sur y Sur Asía.

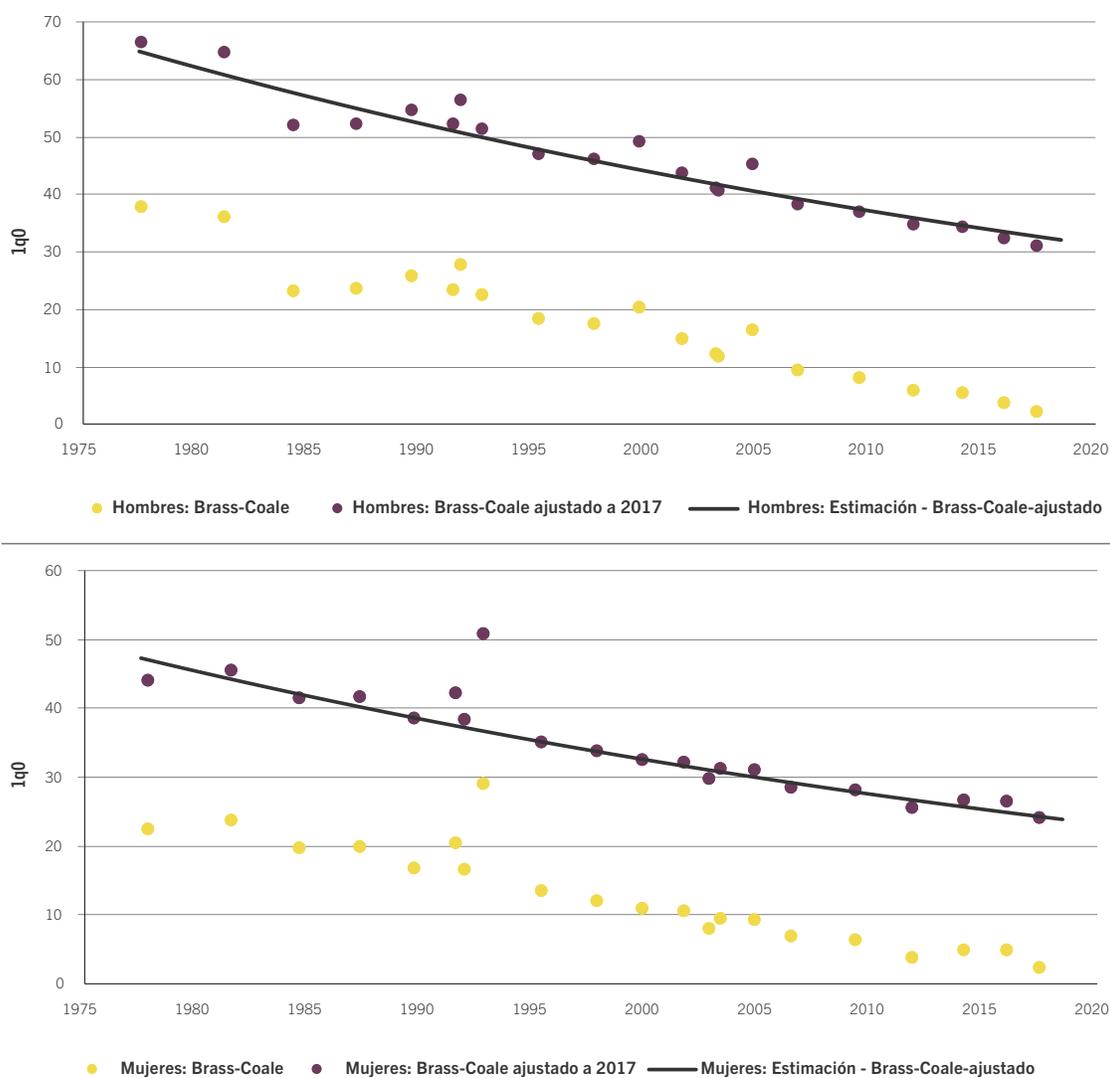
Por otra parte, se observó que la tabla de vida modelo que se escogió a partir del indicador de lactancia materna es la más consistente con las

condiciones sociales, demográficas y ambientales de los departamentos. En cambio, las tablas elegidas considerando la relación ${}_4q_1/q_0$ no presentan tanta consistencia. Por ejemplo, Caquetá es un departamento principalmente rural, que pertenece a la región amazónica, tiene temperaturas superiores a los 25°C, y cuyo indicador de lactancia materna es 80,3%. Por tanto, se espera que la mortalidad infantil sea relativamente baja en comparación a la mortalidad postinfantil. Sin embargo, las ${}_4q_1/q_0$ de 2010 y 2015, 0.11 (2/19) y 0.27 (4/15) respectivamente, señalaron las tablas de vida modelo de Chile y del Lejano Oriente como las apropiadas, siendo que estas tablas de vida son las que presentan la menor razón ${}_4q_1/q_0$ según el Gráfico 1.

El Gráfico 6 muestra los resultados para La Guajira del método de Brass-Coale antes y después de ajustar el nivel departamental de la mortalidad infantil al nivel nacional de 2017. Para aplicar el método de

Brass-Coale para La Guajira, se utilizó la familia de tablas de vida de Sur; según TDE para las regiones menos desarrolladas, las tablas de vida Sur y Sur de Asia son las más apropiadas. Los marcadores identificados por “Brass-Coale” son las estimaciones con los datos originales de los censos; en cambio, los identificados por “Brass-Coale ajustado a 2017” están ajustados a la mortalidad infantil nacional de este año. En el Gráfico 6 se observa que los tres censos indican claramente el patrón de la mortalidad infantil; por esta razón, en la retroproyección se respetó el patrón de los Censos.

Gráfico 6. Mortalidad infantil de la Guajira estimada con el Método de Brass-Coale (1968) utilizando los últimos tres censos (1993, 2005 y 2018) y ajuste del método al nivel nacional de 2017



Nota: La estimación del Brass-Coale ajustado, línea continua, se realizó teniendo en cuenta las recomendaciones de UNICEF, WHO, The World Bank & UN Population Division (2007).

Fuente: Censos de 1993, 2005 y 2018, Dirección de Censos y Demografía, DANE.

3.2 | Estimación de las curvas de mortalidad en el área resto de los departamentos

Como se explica en el capítulo 2.2, las curvas del resto se estimaron utilizando un método simple e innovador, que funcionó muy bien: la relación $q_x(\text{resto})/q_x(\text{cabecera})$ se estimó por medio de la probabilidad de morir a la edad x , la cual se aproximó a través de un modelo logístico estimado con la información del Censo 2018. Así, con esta relación, la estimación de las curvas para el área resto fue directa, debido a que ya se conocía la curva ${}_1q_x$ de

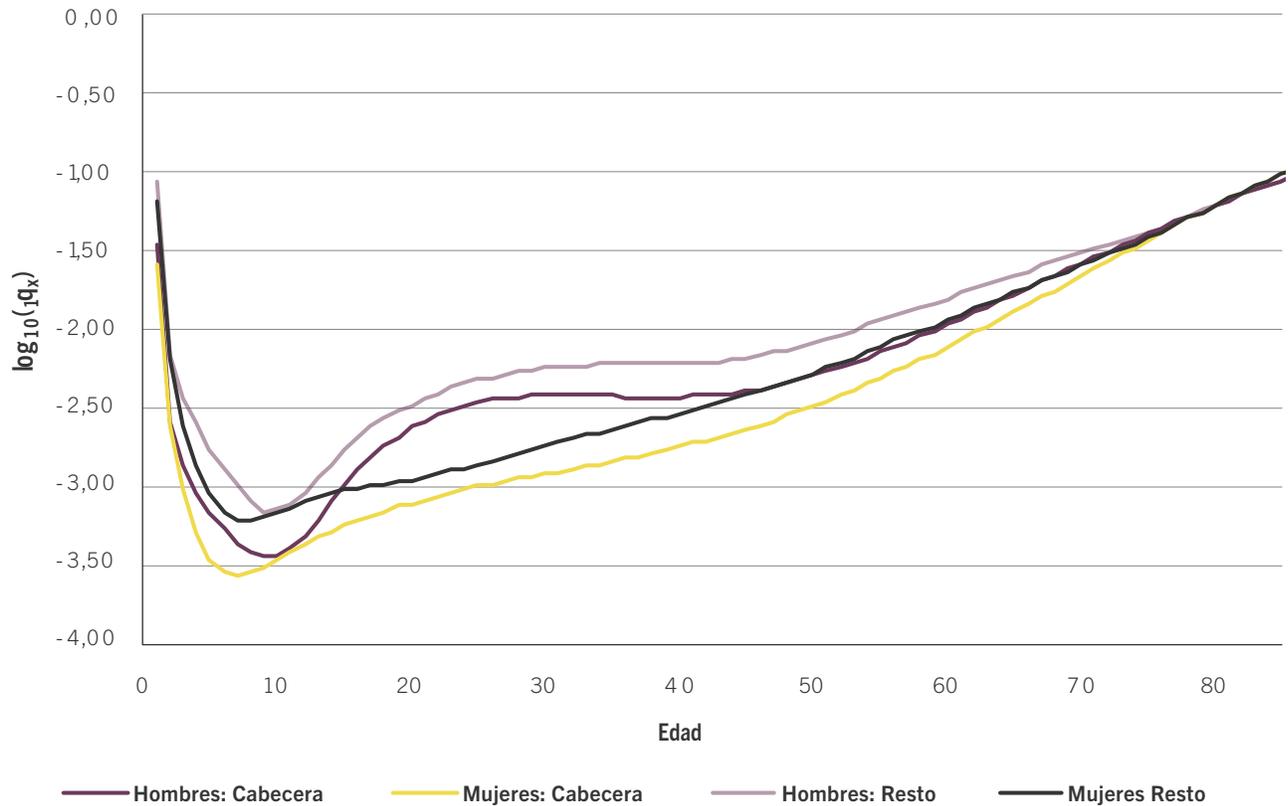
la cabecera (ver Gráfico 5). Las estimaciones del modelo se realizaron por grupos de departamentos teniendo en cuenta dos criterios: el Índice de Necesidades Básicas Insatisfechas -NBI y la diferencia entre el NBI del resto y de la cabecera. Los grupos aparecen en la Tabla 2; además, se estimaron modelos separados para los siguientes rangos de edad: 0, 1 a 4, 5 a 14, 15 a 29, 30 a 64, y 65 años y más.

Tabla 2. | Clusters de departamentos para hacer los modelos probabilísticos

Departamento	Cluster
Guainía, Vichada, La Guajira, Vaupés, Chocó	1
Tolima, Putumayo, Antioquia, Nariño, Cesar, Meta, Casanare, Norte de Santander, Cauca	2
Cundinamarca, Risaralda, Archipiélago de San Andrés, Providencia y Santa Catalina, Atlántico, Huila, Quindío, Boyacá, Caldas, Santander, Valle del Cauca, Bogotá, D.C.	3
Córdoba, Arauca, Caquetá, Sucre, Magdalena, Guaviare, Amazonas, Bolívar	4

Las curvas de las ${}_1q_x$ para cabecera y resto por sexo para La Guajira se presentan en el Gráfico 7. En este gráfico se observa que la probabilidad de morir a cualquier edad por sexo es mayor en el resto que en la cabecera.

Gráfico 7. | Curvas ${}_1q_x$ para cabecera y resto por sexo para La Guajira



Fuente: EEV 2016-2018 y CNPV-2018, Dirección de Censos y Demografía, DANE

Por último, es importante aclarar que las curvas ${}_1q_x$ – 2017, cuyo procedimiento se explica en este documento, son las curvas insumo que se introducen al RUP (Rural Urban Projections). Este es el programa que interacciona los tres componentes demográficos y hace las proyecciones de población.³

³ <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacio>

4.

Comparación de los resultados con otros métodos



Comparación de los resultados con otros métodos

Las curvas estimadas con el anterior procedimiento se compararon con las estimadas por dos métodos demográficos: el método de la ecuación del equilibrio del crecimiento de Brass (The Brass Growth Balance Method) y el método del equilibrio del crecimiento generalizado (The Generalized Growth Balance Method).

El método del equilibrio del crecimiento generalizado no se pudo implementar inicialmente porque dependía de la estimación de la migración neta, y esta, a su vez, de la mortalidad y la fecundidad. En el procedimiento general de las proyecciones, primero se calibró la fecundidad y la mortalidad, asumiendo una población cerrada, y posteriormente se introdujo la migración. No obstante, cuando se cerró el círculo fecundidad – mortalidad – migración, el método generalizado se utilizó para evaluar los resultados de la mortalidad.

En general, la evaluación fue satisfactoria, ya que se obtuvo que las curvas estimadas a partir del método generalizado, el cual tiene en cuenta la migración, fueron similares a las estimadas por el método implementado por el DANE. El Gráfico 8 muestra esta comparación para La Guajira. Este gráfico resalta la importancia de tener en cuenta la migración. La curva del método de la ecuación del equilibrio del crecimiento de Brass, que no la tiene en cuenta, pasa considerablemente por encima de las curvas estimadas por los otros dos métodos.

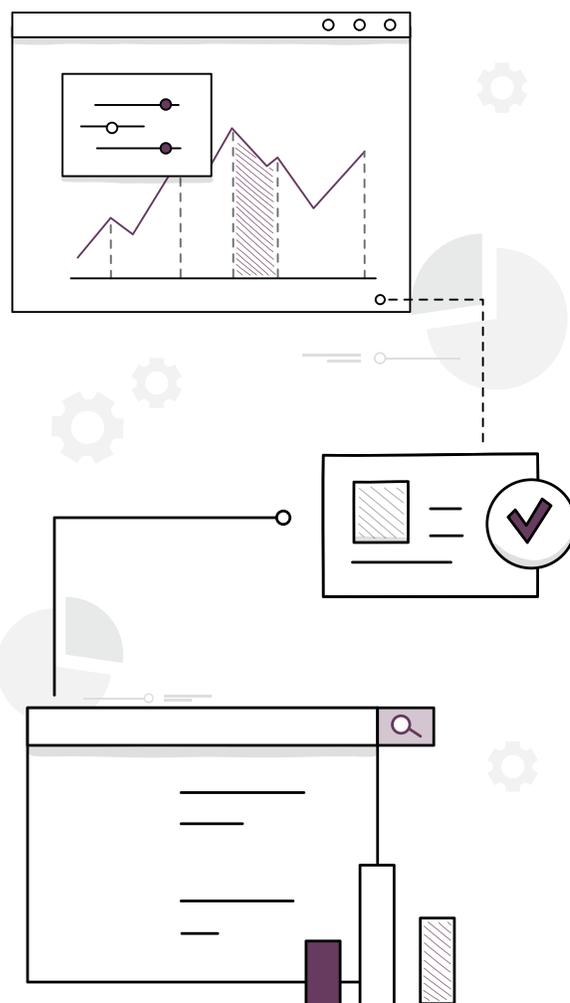
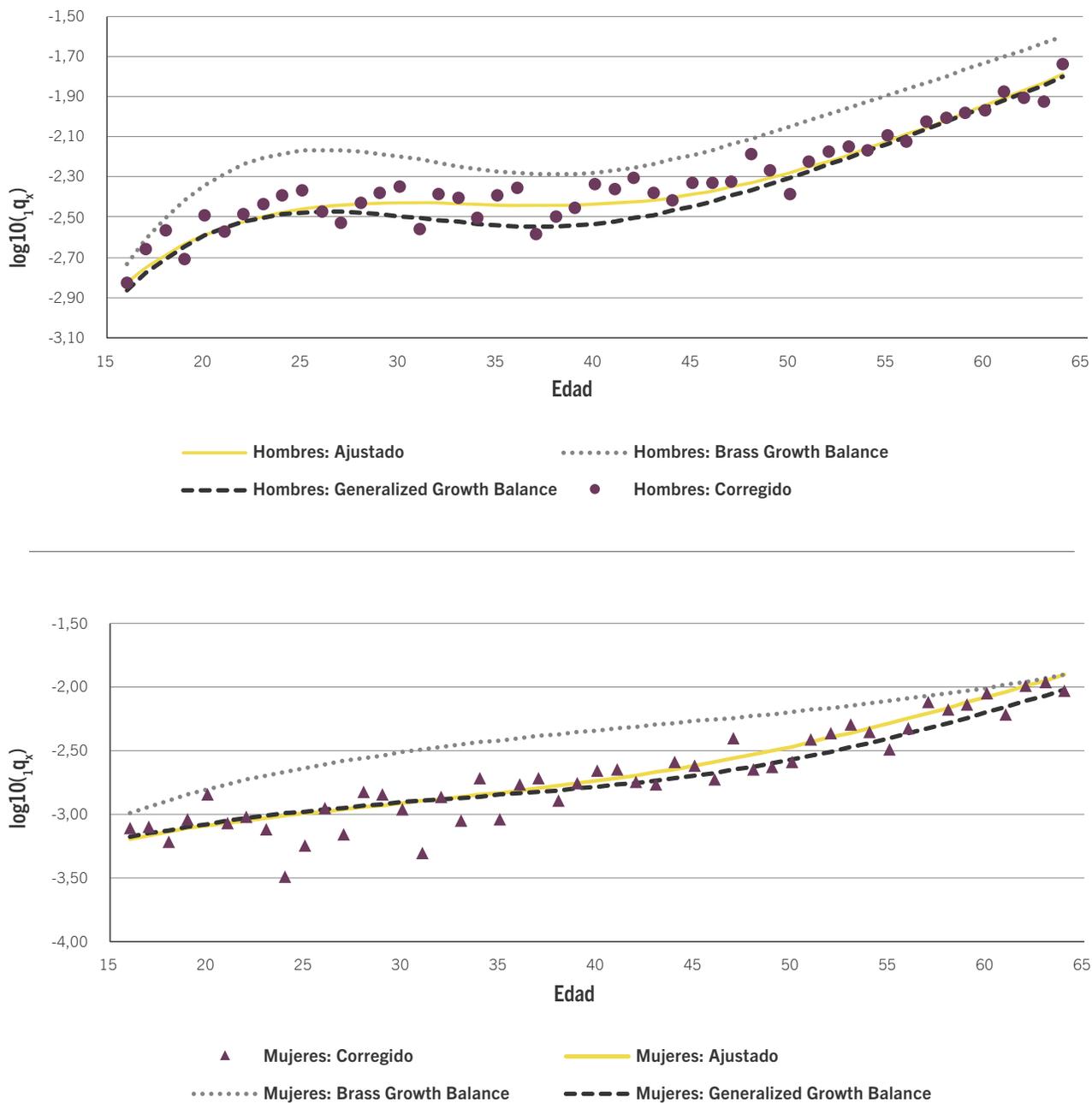


Gráfico 8. Comparación de las curvas de mortalidad estimada para La Guajira con la de otros métodos



Fuente: EEVV 2016-2018 y CNPV-2018, Dirección de Censos y Demografía, DANE

5.

Conclusiones



Conclusiones

Los factores de corrección del Censo, estimados con la pregunta sobre la expedición o no del certificado de defunción, son una herramienta útil y confiable para corregir las muertes que no se reportan en los Registros de Estadísticas Vitales, al menos para las cabeceras o áreas urbanas. En las zonas rurales o en el resto, los resultados no fueron satisfactorios, ya que en algunos departamentos el número corregido de muertes no fue suficiente, y en otros, sobrepasó lo esperado.

Para superar el obstáculo de estimar las curvas de mortalidad para el resto, se propuso un método simple e innovador, cuyos resultados fueron consistentes: estimar la probabilidad de morir a la edad x con el fin de determinar la relación $q_x(\text{resto})/q_x(\text{cabecera})$. La probabilidad de morir a la edad x se aproximó a través de un modelo logístico que se estimó usando las bases de personas censadas y difuntas del CNPV-2018. Utilizando la relación $q_x(\text{resto})/q_x(\text{cabecera})$ y las curvas de mortalidad de la cabecera, la aproximación de las curvas del resto fue directa. Este método se basa en el siguiente supuesto: las personas en los restos pueden no conocer la existencia de un certificado de defunción; sin embargo, lo que sí saben es si algún miembro del hogar murió el año anterior. Las estimaciones de la relación $q_x(\text{resto})/q_x(\text{cabecera})$ se hizo por grupos de edad y de departamentos teniendo en cuenta el NBI departamental y la brecha de este indicador entre cabecera y resto.

En cuanto a la mortalidad infantil, para elegir la tabla de vida modelo por departamento, se utilizaron dos indicadores: el indicador de lactancia materna continúa al año de vida y la relación ${}_4q_1/{}_1q_0$ de acuerdo con las Encuestas Nacionales de Demografía y Salud 2010 y 2015. Este último indicador presentó inconsistencias al considerar las condiciones de vida y las características de la población que habita en los departamentos; en contraste, el de lactancia materna diferenció mejor entre los departamentos urbanos y más desarrollados y los departamentos más rurales.

Al comparar los dos métodos de suavizamiento se concluyó que tienen lógicas similares. Heligman-Pollard estima el patrón a partir de tres funciones matemáticas, mientras que, el Bayesiano de Alexander et al. (2017) lo hace con base en tres componentes principales que se derivan de curvas de mortalidad de referencia.

Finalmente, la evaluación que se hizo del método implementado por el DANE al compararlo con dos métodos clásicos (el método de la ecuación del equilibrio del crecimiento de Brass y el método del equilibrio del crecimiento generalizado) fue satisfactoria. Los resultados del método del DANE son similares a los obtenidos por el método generalizado, el cual considera la migración.

Bibliografía

- Alexander, M., Zagheni, E & Barbieri, M. (2017). A Flexible Bayesian Model for Estimating Subnational Mortality. *Demography* (2017) 54:2025–2041. DOI 10.1007/s13524-017-0618-7
- Arriaga, E. (2011). *Análisis Demográfico de la Mortalidad*. Universidad Nacional de Córdoba.
- Heligman, L. & Pollard, J. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, 107(1), 49-80. <https://doi.org/10.1017/S0020268100040257>
- Kostaki, A. (1991). The Heligman-Pollard formula as a tool for expanding an abridged life table. *Journal of Official Statistics*, 7(3), 311-323. <http://www.jos.nu/Articles/abstract.asp?article=73311>
- Moultrie, T., Dorrington, RE., Hill, A., Hill, K., Timaeus, I., & Zaba, Basia. (2013). *Tools for Demographic Estimation*.
- UNICEF, WHO, The World Bank & UN Population Division (2007). *Levels and Trends of Child Mortality in 2006: Estimates developed by the Inter-agency Group for Child Mortality Estimation*, New York.

DANE

INFORMACIÓN PARA TODOS



@DANE_Colombia



/DANEColombia

www.dane.gov.co



/DANEColombia



@DANEColombia

