

DIRECCIÓN DE CENSOS Y DEMOGRAFÍA -DCD-

COORDINACIÓN DE PROYECCIONES Y ANÁLISIS DEMOGRÁFICO

NOTA METODOLÓGICA ACTUALIZACIÓN GENERAL DE PROYECCIONES DE POBLACIÓN Y ESTIMACIONES DEMOGRÁFICAS (POST COVID19)

MARZO de 2023

CONTENIDO

Introducción	3
1. El método de componentes demográficos	4
1.1 Población base por edad y sexo.	5
2. Metodología de proyecciones para mortalidad	6
3. Metodología de proyecciones para fecundidad.....	10
4. Metodología de proyecciones para migración.....	12
4.1 Consideraciones preliminares.....	12
4.2 Evaluación de las fuentes.....	14
4.3 Estimación y modelación de estructuras.....	15
5. Métodos de proyecciones para áreas pequeñas (desagregaciones por área y a nivel de municipios)	19
Bibliografía	22

Introducción

Proveer estadísticas oficiales sobre población con un elevado grado de confiabilidad y mayores niveles de desagregación, es uno de los retos que enfrentan actualmente los Institutos Nacionales de Estadística. Este ejercicio requiere el uso de herramientas especializadas y capacidades técnicas, enfocadas en métodos y materiales de análisis demográfico.

Las proyecciones de población son uno de los productos de información demográfica que brindan insumos a entidades públicas, privadas y organismos internacionales, para prever y cuantificar demandas de bienes y servicios. Además, proporcionan insumos fundamentales para una amplia gama de actividades gubernamentales tales como la planeación de políticas públicas y ordenamiento del territorio, la toma de decisiones con base en las características de su población, entre otras actividades, considerando a la población no solo como demandante de recursos sino también como sujeto productivo que propicia el desarrollo de una sociedad.

Las proyecciones demográficas se pueden definir como una estimación sobre la probable evolución de la población, que toma como base la información histórica reciente de los componentes demográficos: fecundidad, mortalidad y migración. En el periodo de la proyección se pueden analizar aspectos relacionados con el tamaño, composición y localización poblacionales. Por ejemplo, se pueden obtener y analizar indicadores sobre la población en edades avanzadas (65 años y más), la proporción de mujeres en edad reproductiva (desde los 10 hasta los 49 años), y el peso relativo de la población de cada área subnacional dentro del total país, entre otros indicadores.

Este documento describe los principales aspectos metodológicos implementados por la Dirección de Censos y Demografía del DANE, para la producción de proyecciones de población actualizadas de acuerdo a las dinámicas poblacionales observadas post pandemia por COVID-19 y sus medidas de contención y prevención de contagio, a partir de información demográfica proveniente del Censo Nacional de Población y Vivienda – CNPV 2018, en conjunto con otras principales fuentes de información como los registros de estadísticas vitales, las encuestas especializadas y el aprovechamiento estadístico de registros administrativos.

1. El método de componentes demográficos

La necesidad de disponer de proyecciones desagregadas implica el uso de un método como el de componentes demográficos, que aporta resultados por edad, sexo y componente demográfico. Dicho método considera el cambio poblacional en función de los niveles y tendencias de la fecundidad, la mortalidad y la migración, sus interrelaciones y las de estos factores con las estructuras poblacionales por edad y sexo. Resulta relevante destacar que la elaboración de estimaciones y proyecciones por edad es clave para la planificación de la economía y las políticas públicas, dado que indicadores relacionados con los niveles de bienestar, las necesidades, los patrones de productividad y el comportamiento del consumo, entre otros, muestran resultados diferenciales según esta variable. Así mismo, la distribución por sexo es importante dado que los desbalances entre el número de mujeres y hombres pueden afectar los patrones de fecundidad y nupcialidad, la participación laboral y los roles de género dentro de la sociedad.

El modelo de la ecuación compensadora que define el cambio poblacional entre dos momentos puede ser expresado de la siguiente manera:

$$P(T) = P(0) + N(0, T) - D(0, T) + I(0, T) + E(0, T)$$

Donde

$P(T)$ es el número de personas en la población en el momento T .

$P(0)$ es el número de personas en la población en el momento 0 .

$N(0, T)$ es el número de nacimientos en la población entre los tiempos 0 y T .

$D(0, T)$ es el número de defunciones en la población entre los tiempos 0 y T .

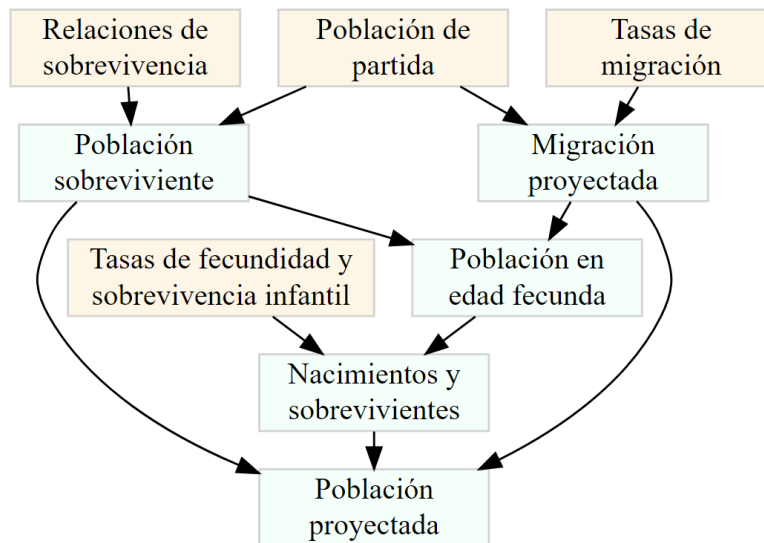
$I(0, T)$ es el número de inmigrantes entre los tiempos 0 y T .

$E(0, T)$ es el número de emigrantes entre los tiempos 0 y T .

En el siguiente diagrama se resume de manera sucinta el esquema de trabajo utilizado para la implementación del método de componentes demográficos para la elaboración de proyecciones. Dado que este método considera de manera explícita los cambios por edades, generaciones de individuos y sexo, garantiza un nivel adecuado de consistencia interna de los resultados, que no puede obtenerse mediante otros métodos de proyección. No obstante, dicho nivel de detalle tiene un costo: la demanda

de información detallada de los componentes demográficos y la población de base para cada unidad de análisis. La elaboración de cálculos para proyectar la población por edad y sexo es desarrollada de manera iterativa hasta que se llegue al año correspondiente al horizonte de proyección.

Figura 1. La proyección por el método de componentes



Fuente: DANE, Proyecciones de población y estudios demográficos.

1.1 Población base por edad y sexo.

El establecimiento de una línea de base que se aproxime lo mejor posible a la realidad es clave para que los resultados de las proyecciones sean consistentes en el corto y mediano plazo y muestren tendencias plausibles a largo plazo, dada la información de la que se dispone actualmente. El método que fue utilizado para determinar la población base en el marco de la elaboración de este escenario de proyecciones es el conocido como conciliación censal, el cual tiene por objetivo que dicha población sea coherente con lo encontrado en censos y las características de la evolución demográfica de la población.

Algunos elementos que se tuvieron en cuenta para la estimación de la estructura de la población base:

- Las tasas de omisión censal presentan niveles altos entre 0 y 9 años, principalmente en las primeras edades. La falta de cobertura en las primeras edades se tiene registrada desde mitad del siglo XX en Estados Unidos, Europa continental y Reino Unido y también es relacionada en manuales de la CEPAL. Entre otras causas, los niños faltantes tienden a estar altamente concentrados en los hogares y viviendas que son más difíciles de enumerar (difícil acceso, pobreza, alta movilidad, viviendas precarias). Así mismo, algunos informantes pueden omitir el reporte de los niños debido a que no ven la importancia del mismo. Para la estimación de la omisión censal en las primeras edades (0 a 9 años) se utiliza un método basado en el análisis demográfico, debido a que tiene en cuenta datos de las estadísticas vitales
- La omisión de censos y registros en edades intermedias tiende a ser superior en los hombres que en las mujeres, especialmente en el grupo de edades 15 a 29. Los hombres, en el caso de Colombia, tienden a tener mayor omisión censal debido a que tienen una mayor frecuencia estadística dentro de grupos armados al margen de la ley y en el caso de los hogares unipersonales pueden mostrar una mayor movilidad residencial.
- Existe una sobreestimación de los efectivos poblacionales de 60 años y más tanto en censos como en encuestas y registros. Esta tendencia tiende a ser más marcada en edades muy avanzadas -80 años y más-. También puede haber transferencia de efectivos de una edad a otra superior debido a problemas de memoria del informante idóneo, que con frecuencia tiende a exagerar sus edades.
- Cabe mencionar que se utilizaron métodos matemáticos para el suavizamiento de las estructuras poblacionales de los censos corregidas por la falta de cobertura. Esto permite una mejor representación de los patrones principales de dichas estructuras.

2. Metodología de proyecciones para mortalidad

Para obtener medidas acerca de cuantos individuos mueren y cuantos sobreviven durante un periodo, según sus edades, fue necesario elaborar tablas de vida, las cuales se pueden definir como representaciones matriciales asociadas a modelos teóricos, que resumen el proceso de extinción de las poblaciones a través de su ciclo de vida.

Los principales indicadores resultantes de la construcción de tablas de vida son la esperanza de vida al nacer y la tasa de mortalidad infantil. El primero es el indicador de referencia para el análisis del nivel de la mortalidad de una población; mientras que la tasa de mortalidad infantil suele ser visto no solo como un indicador demográfico sino como un indicador de desarrollo de departamentos y municipios, al estar ampliamente correlacionado con otros indicadores como el de necesidades básicas insatisfechas y de logro educativo.

En el ejercicio desarrollado, las tablas de vida fueron derivadas de las tasas específicas de mortalidad m_x . En Colombia, no se dispone de información completa acerca de las defunciones por generaciones. Por tanto, dado que se conocen las defunciones por edad, se utilizan las tasas de tipo edad – periodo, las cuales combinan la mortalidad de un año calendario, de una edad y dos generaciones, perspectiva que se considera adecuada para el análisis a largo plazo y de las fluctuaciones coyunturales. Las tasas específicas por edad fueron calculadas de la siguiente manera:

$$m_x^t = \frac{D_x^t}{0,5(P_x^t + P_x^{t+1})}$$

Los cocientes de mortalidad q_x se construyeron a través de la siguiente fórmula, suponiendo que la función de sobrevivencia varía en forma lineal en cada intervalo $x, x+1$.

$$q_x = \frac{2m_x}{2+m_x}$$

La serie de cocientes de mortalidad permitieron calcular la evolución de una generación de $l_0 = 100$ mil individuos, obteniendo una serie de supervivientes a edades exactas sucesivas l_x :

$$l_{x+1} = l_x - l_x * q_x$$

El número de años persona vividos entre las edades x y $x+1$ es definido por la siguiente ecuación:

$$L_x = a_x l_x + (1 - a_x) l_{x+1}$$

Donde a_x es el número medio de años persona vividos entre las edades x y $x+1$ para aquellos que fallecieron antes de completar el intervalo entero.

Finalmente, la probabilidad perspectiva de paso es definida mediante la siguiente expresión:

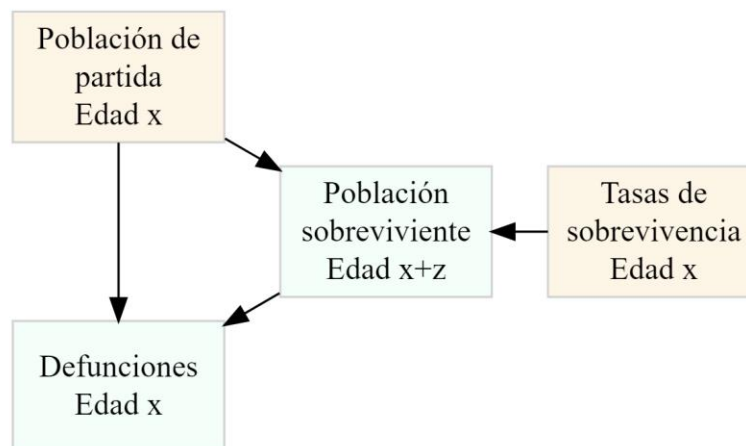
$$Z_x = \frac{L_{x+1}}{L_x}; Z_{nac} = \frac{L_0}{l_0}$$

Dicho indicador, nos sirve de base para el cálculo de sobrevivientes, tomando en cuenta solo el efecto de la mortalidad.

$$NS_{x+1}^{t+1} = P_x^t * Z_x$$

En la metodología implementada fue necesario contar con una tabla de mortalidad inicial y una tabla de mortalidad límite para proyectar la mortalidad, lo cual reduce el problema de la proyección a una interpolación entre dos tablas. En general, se ha identificado que las tasas de mortalidad tienden a cambiar a través del tiempo, de forma más lenta y consistente que las de fecundidad y migración, por lo cual su proyección tiende a ser más precisa. En cualquier caso, se debe tener en cuenta que los límites de la vida humana no han sido fijados hasta ahora de manera unívoca, debido a los continuos avances en medicina y científicos, como los avances en la tecnología biomédica, la efectividad de los sistemas de salud para la atención médica, cambios comportamentales, las enfermedades emergentes y reemergentes, cambios ambientales, las crisis políticas y económicas de los países, entre otros, que dificultan el establecimiento de una edad límite para la vida humana.

Figura 2 La proyección por el método de componentes. El módulo de mortalidad



Fuente: DANE, Proyecciones de población y estudios demográficos.

El desarrollo de métodos indirectos para la estimación de la mortalidad ha tenido gran auge en los países en desarrollo, debido a que tanto los sistemas de estadísticas vitales como los censos de

población no tienen una cobertura perfecta en ellos. Las proporciones de hijos nacidos vivos que han fallecido fueron utilizadas para la estimación de indicadores de mortalidad en la niñez; mientras que en el caso de la mortalidad adulta se realizó una estimación del grado de cobertura del registro de defunciones según edad, con el fin de ajustar las tasas de mortalidad que sirvieron de base para la elaboración de las tablas de vida.

Para la estimación de la cobertura del registro de defunciones por edad existen múltiples alternativas. En la aplicación desarrollada se hizo uso de la información recopilada mediante la pregunta censal sobre defunciones no certificadas en los hogares. Mientras que en el modelamiento y suavizamiento de la mortalidad por edades se utilizó el método de Heligman Pollard.

Ilustración 1. Pregunta utilizada en los censos de Colombia 2005 y 2018.

26. ¿Cuántas personas que eran miembros de este hogar fallecieron en el 2017?

Total (Si es 0, continúe con la pregunta 27)
(Si es 1 o más, relaciónelos en la siguiente tabla)

Número de orden	Sexo		Edad al morir (para menores de 1 año escriba 0)	¿Se expidió certificado de defunción?		
	1. Hombre	2. Mujer		1. Sí	2. No	3. No sabe
<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Si en el hogar fallecieron más de 6 personas en el 2017, continúe el registro en un cuestionario adicional

(Relacione en la tabla si es mayor a 00)
↓

Sexo	Edad al morir	¿Se hizo certificado de defunción?
1 <input type="radio"/> Hombre 2 <input type="radio"/> Mujer	<input type="text"/>	1 <input type="radio"/> Sí 2 <input type="radio"/> No
1 <input type="radio"/> Hombre 2 <input type="radio"/> Mujer	<input type="text"/>	1 <input type="radio"/> Sí 2 <input type="radio"/> No
1 <input type="radio"/> Hombre 2 <input type="radio"/> Mujer	<input type="text"/>	1 <input type="radio"/> Sí 2 <input type="radio"/> No
1 <input type="radio"/> Hombre 2 <input type="radio"/> Mujer	<input type="text"/>	1 <input type="radio"/> Sí 2 <input type="radio"/> No
1 <input type="radio"/> Hombre 2 <input type="radio"/> Mujer	<input type="text"/>	1 <input type="radio"/> Sí 2 <input type="radio"/> No

Fuente: DANE, cuestionarios de los censos de población de Colombia 2005 y 2018.

Si bien el subregistro de las defunciones no permitió la estimación directa de la mortalidad, como es la práctica en los países desarrollados, los registros de estadísticas vitales permitieron obtener estimaciones sobre la estructura de la mortalidad por edad y sexo que no pudieron ser estimadas de manera consistente a partir de otras fuentes de información. En este sentido, el nivel de desagregación de los resultados, que permite el registro, es mucho mayor que el de las encuestas por muestreo y permite la elaboración de comparaciones por periodos interanuales. Por otro lado, la información de la mortalidad en edades avanzadas fue ajustada para arrojar resultados consistentes en el software RUP, en particular para la población de 90 años y más.

Así mismo, al igual que en países de mayor avance en sus procesos de transición de la mortalidad, en las proyecciones subnacionales de Colombia fue necesario formular hipótesis diferenciadas acerca de la probable convergencia o divergencia del comportamiento de mortalidad para grupos de departamentos, fijando un escenario límite que realiza una partición en dos grupos: departamentos líderes y departamentos rezagados en el proceso de transición.

3. Metodología de proyecciones para fecundidad

En cuanto al componente de la fecundidad, es importante destacar que en este caso se partió de diversas variables de entrada. Entre estas, son clave las que ayudaron a consolidar una estimación de la población de mujeres en edades reproductivas, que está conformada no solo por las poblaciones nativas sino por las inmigrantes. Este conjunto de información, junto con las tasas específicas de fecundidad, permiten obtener una estimación de los nacimientos, que es distribuida por sexo y proyectada teniendo en cuenta las tasas de sobrevivencia en las primeras edades. El indicador del nivel de fecundidad de una población utilizado fue la tasa global de fecundidad, también conocido como índice sintético de fecundidad. Mientras que los indicadores de estructura de la fecundidad utilizados fueron las tasas específicas de fecundidad por edades.

El procedimiento utilizado para la proyección de los nacimientos se basó en las tasas de fecundidad por edades del periodo. En el caso colombiano, el cambio poblacional futuro está todavía fuertemente determinado por la transición de la fecundidad, de ahí la importancia de validar de manera extensa los resultados de este componente demográfico. En referencia a la tasa global de fecundidad, un valor de referencia que hasta ahora fue frecuentemente tomado en cuenta en las proyecciones es el denominado "nivel de reemplazo", de 2.1 hijos por mujer. No obstante, el avance de la transición de la fecundidad en el país ha hecho que esto sea revaluado como valor de referencia, y en vez de este, el patrón de comparación comienza a ser una TGF de 1.5. Algunos de los determinantes de este proceso acelerado

de descenso de la fecundidad en el país han sido la expansión educativa, el incremento de la igualdad de género, la participación de la mujer en la fuerza laboral, la urbanización, la globalización, el uso de métodos modernos de anticoncepción, entre otros factores. En este sentido, el descenso de la fecundidad ha sido un determinante importante del proceso de envejecimiento de la estructura poblacional en Colombia.

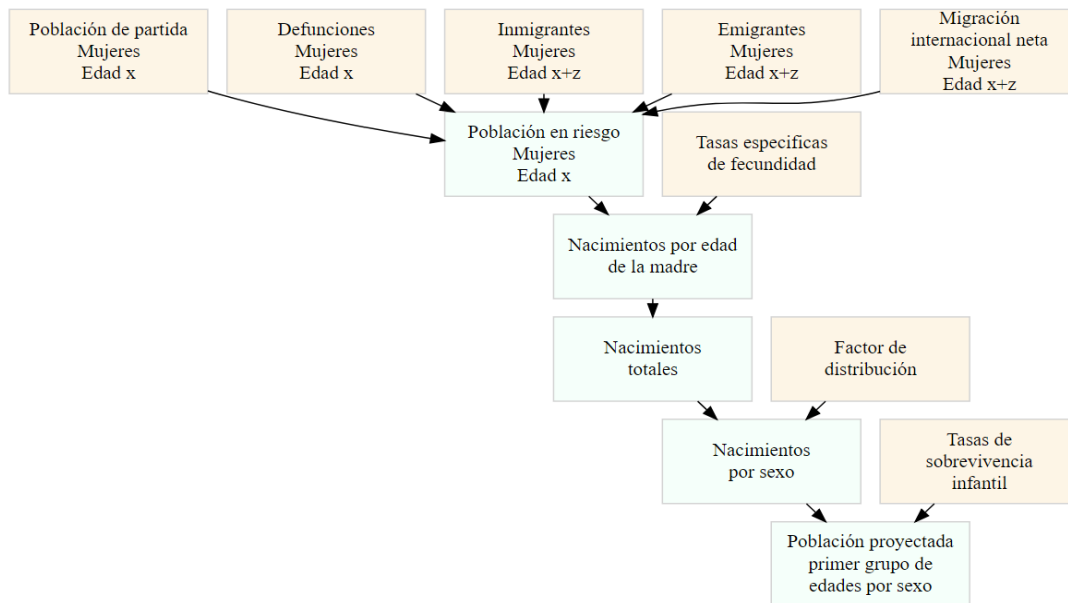
El número de nacimientos de ambos sexos en el periodo calendario es una función de, la población femenina en edades reproductivas y las tasas específicas de fecundidad por edad, y es calculado mediante la siguiente expresión:

$$B_t = \frac{1}{2} \sum_x f_{x,t} (P_{x,t}^f + P_{x,t+1}^f)$$

Posteriormente se distribuyen los nacimientos según sexo, generalmente asumiendo que la razón de masculinidad al nacer es de 105 hombres por cada 100 mujeres.

$$B_t^h = B_t * \frac{105}{205} \text{ y } B_t^m = B_t * \frac{100}{205}$$

Figura 3 La proyección por el método de componentes. El módulo de fecundidad



Fuente: DANE, Proyecciones de población y estudios demográficos.

Las estimaciones de fecundidad, al igual que las de mortalidad, fueron realizadas a partir de métodos indirectos. En el caso de la fecundidad, se utilizó el modelo de Gompertz, debido a que el patrón de fecundidad estándar es aplicable en general a regímenes de fecundidad alta o media, además que no necesita el supuesto de fecundidad constante que tiene la formulación original del método P/F de Brass.

Por otra parte, al igual que en el componente de mortalidad, el subregistro de los nacimientos es un tema de interés en la elaboración de proyecciones. Se partió del supuesto de que en Colombia al igual que en otros países, la omisión implícita porcentual de los nacimientos tiende a ser menor que la omisión implícita porcentual de las defunciones.

Para proyectar la tendencia de las proyecciones de fecundidad, a nivel nacional, se utilizó un modelo logístico, utilizando una TGF límite de 1.43. En cuanto a la proyección de la estructura de la fecundidad, se utilizó el patrón de Brasil estimado por *World Population Prospects*, considerando que este país se encuentra en un estado más avanzado de transición demográfica y tiene una estructura y tendencia de la fecundidad similar a la de Colombia.

4. Metodología de proyecciones para migración

4.1 Consideraciones preliminares.

El componente migratorio fue dividido en migración internacional y migración interna. La metodología para la proyección de la migración interna e internacional siguió el enfoque top-down, que enmarca las proyecciones en general. En el caso específico de la migración, este enfoque permitió una coherencia entre los sistemas de migración en la escala nacional y subnacional.

La proyección de migración consiste en la definición precisa de un conjunto de tasas de migración por sexo y edad, dentro de un sistema (interno o internacional) y una tendencia a futuro (el año de 2070 en este caso). La formulación de las hipótesis de migración interna e internacional se fundamenta en un modelo teórico, que incluye consideraciones respecto al efecto de la pandemia por COVID-19.

La literatura reportó un descenso de la intensidad de las migraciones internas debido a las restricciones a la movilidad durante la pandemia. Adicionalmente, hubo una ruralización de las migraciones principalmente desde las grandes áreas metropolitanas hacia sus zonas rurales aledañas. No obstante, este fue un fenómeno coyuntural y transitorio.

"[...] Los resultados muestran alteraciones reseñables en 2020, que se plasmaron principalmente en un descenso de los movimientos hacia las ciudades y en un aumento en las salidas desde ellas, especialmente hacia municipios rurales. Estos flujos de personas tuvieron un notable impacto demográfico en las áreas rurales. No obstante, a pesar de los cambios observados, la mayor parte de las migraciones internas fueron mayoritariamente entre ciudades, al igual que antes de la pandemia, y todo apunta a que se trató de un hecho coyuntural" (González et al, 2022).

La pandemia de coronavirus, en general, provocó una tendencia inversa de las migraciones, desde los lugares "destino" a los de "origen". Cientos de miles de trabajadores y estudiantes retornaron a sus hogares en busca de calidez, empatía y seguridad (Dandekar & Ghai, 2020). El COVID junto con el deterioro de la economía y las restricciones a la movilidad por la cuarentena tuvieron un impacto directo en la migración interna. A través de una encuesta de expertos, realizada en Australia, se concluyó que el escenario más plausible para 2020-21 era una disminución del 15 al 30% en las tasas de migración interna (Bernard et al, 2020).

Además del impacto de la pandemia por COVID 19, se incorporó también a la tendencia proyectada el modelo de transición de la migración, planteado inicialmente por Zelinsky (Skeldon, 2012). Este modelo prevé una disminución generalizada de las migraciones internas e internacionales, concomitante con el aumento de adopción de tecnología de comunicación, transporte e información, que facilitan el trabajo a distancia y la pendularidad. Además de que una migración intensa es insustentable a largo plazo.

Así, considerando el modelo de transición de la migración, se proyectó una disminución substancial de las migraciones internacionales e interdepartamentales, para 2070. No obstante, se estableció que los departamentos con mejores indicadores socioeconómicos, a pesar de la disminución, continuarán siendo destinos más atractivos. Respecto a la migración internacional, el modelo nos permitió asumir que la corriente de inmigración desde Venezuela es un fenómeno transitorio, que tenderá a disminuir paulatinamente en los próximos años, y que a futuro el saldo migratorio internacional neto tenderá a cero.

De esta manera, los modelos teóricos, en los periodos de pandemia y postpandemia, permitieron diseñar la trayectoria de la intensidad de la emigración para cada una de las áreas consideradas en la proyección. Las estimaciones del volumen e intensidad de los flujos migratorios se realizaron en tres

fases metodológicas consecutivas: i) Evaluación de las fuentes; ii) Estimación y modelación de estructuras; iii) Proyección de las tendencias a 2070.

4.2 Evaluación de las fuentes

Inicialmente se evaluaron las fuentes que informan sobre el nivel, la estructura y la tendencia de las migraciones internas e internacionales.

Para las migraciones internas se consideró el CNPV 2018, la GEIH (2018-2021), la ECV (2018-2021) y el REBP (2018-2020). A continuación, se presenta un resumen de la evaluación de la calidad de estas fuentes.

Tabla 1. Comparación de las fuentes para medir migración interna en cuanto a su representatividad; presencia de información en el periodo de pandemia; calidad del registro por sexo y edad; y coherencia con la información del CNPV 2018

Fuente	Representatividad de los migrantes internos	¿Presenta información durante la pandemia?	Calidad de la información de los migrantes por sexo y edad	Coherencia del nivel (intensidad de la migración) respecto al CNPV
CNPV 2018	Buena	No	Buena	--
GEIH (2018-2021)	No tiene	Sí	Solo para algunos departamentos	Regular. En pandemia no hay consistencia con lo esperado
ECV (2018-2021)	No tiene	Sí	Mala	Poca
REBP (2018-2020)	Permite medir cambios de residencia, hay subregistro	Sí	Sobrerrepresentación de adultos mayores	Alta

Fuente: DANE, Proyecciones de población y estudios demográficos.

La evaluación de las fuentes evidenció que el CNPV 2018 permite estimar con mejor precisión el nivel inicial, bien como las tasas de inmigración y emigración interdepartamentales por sexo y edad. Y, por otro lado, que la fuente que mejor describe la disminución prevista de los flujos interdepartamentales, durante 2020 (año de cuarentena por COVID 19), fue el REBP, que reveló una disminución de la intensidad de los flujos interdepartamentales de 30,3%.

En el caso de la inmigración internacional, el análisis reveló que la fuente para medir la tendencia, intensidad (nivel) inicial de la inmigración internacional fue la GEIH (2014-2018). Mientras que, para estimar la estructura (calendario) de la inmigración y su distribución departamental, se usó el CNPV-2018. La GEIH permite entender mejor el fenómeno de alta inmigración desde Venezuela en los últimos años. El acumulado de flujos anuales provenientes desde ese país, que se obtiene a través de la GEIH, se aproxima al stock reportado por diferentes fuentes. Además, se contrastó la proyección de los flujos con las estimaciones realizadas por GIFMM a 2024. Esta última fuente permitió incorporar el flujo anual estimado de venezolanos que retornan a su país de origen.

Finalmente, para la emigración internacional se aprovechó la información de entradas de salidas de Migración Colombia (Registro Estadístico de Migración Internacional – REMI). Esta fuente reveló también una caída substancial de las migraciones internacionales durante el periodo de pandemia, lo cual se adecúa a lo esperado.

4.3 Estimación y modelación de estructuras

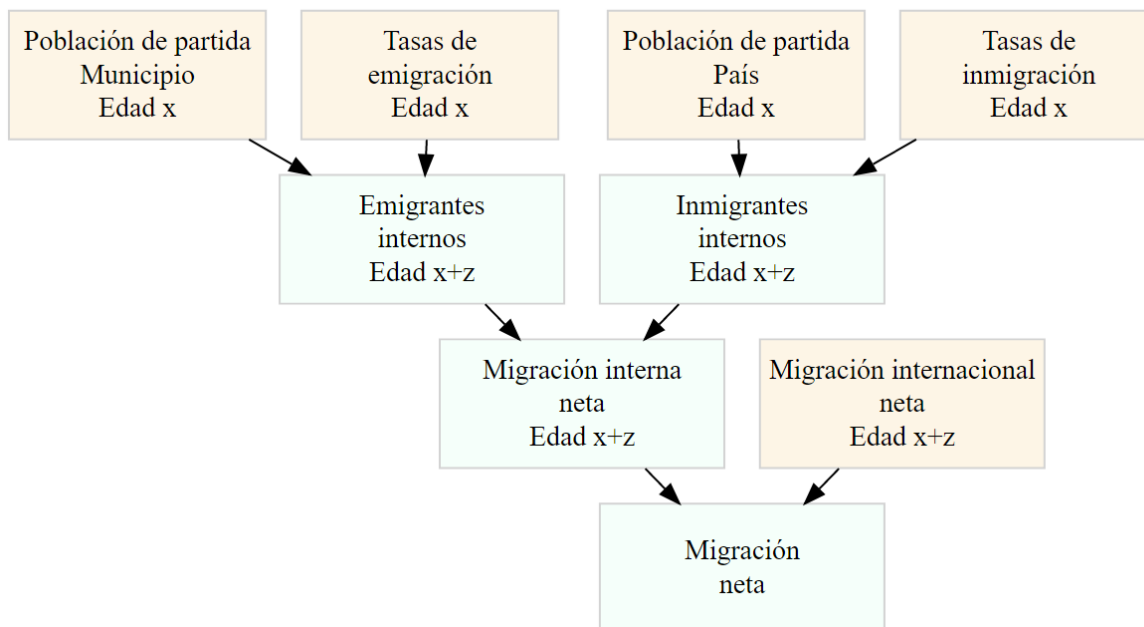
A partir de la información censal se estimaron tasas específicas de emigración e inmigración, interna e internacional, por sexo y edades simples. Para modelar y suavizar las estructuras de migración que se obtienen de dichas tasas, se usó el modelo de Rogers y Castro (1981). Este método modela y suaviza las tasas de migración por edad. Para esto se utiliza un proceso de ajuste, minimizando el error por mínimos cuadrados. El resultado permite obtener el calendario parametrizado de las migraciones internas e internacionales para cada departamento, desagregado en edades simples y sexo.

Una vez se obtiene la estructura de las migraciones, y habiendo estimado el nivel de estas al inicio de la migración, se calcula índice sintético de migración (ISM) como indicador de la intensidad migratoria. Posteriormente, se extrapolan los ISM usando un modelo logarítmico.

Para la distribución de la diferencia anual entre inmigrantes e inmigrantes internos, se usó un método matemático, que no generó diferencias significativas con el método “pool migration” (Wilson & Bell, 2004).¹

Por otra parte, cabe señalar que para la implementación del módulo de migración en el software RUP fue necesario utilizar como indicador de referencia para el nivel de migración al saldo neto migratorio y como indicadores de estructura a los saldos migratorios por edades.

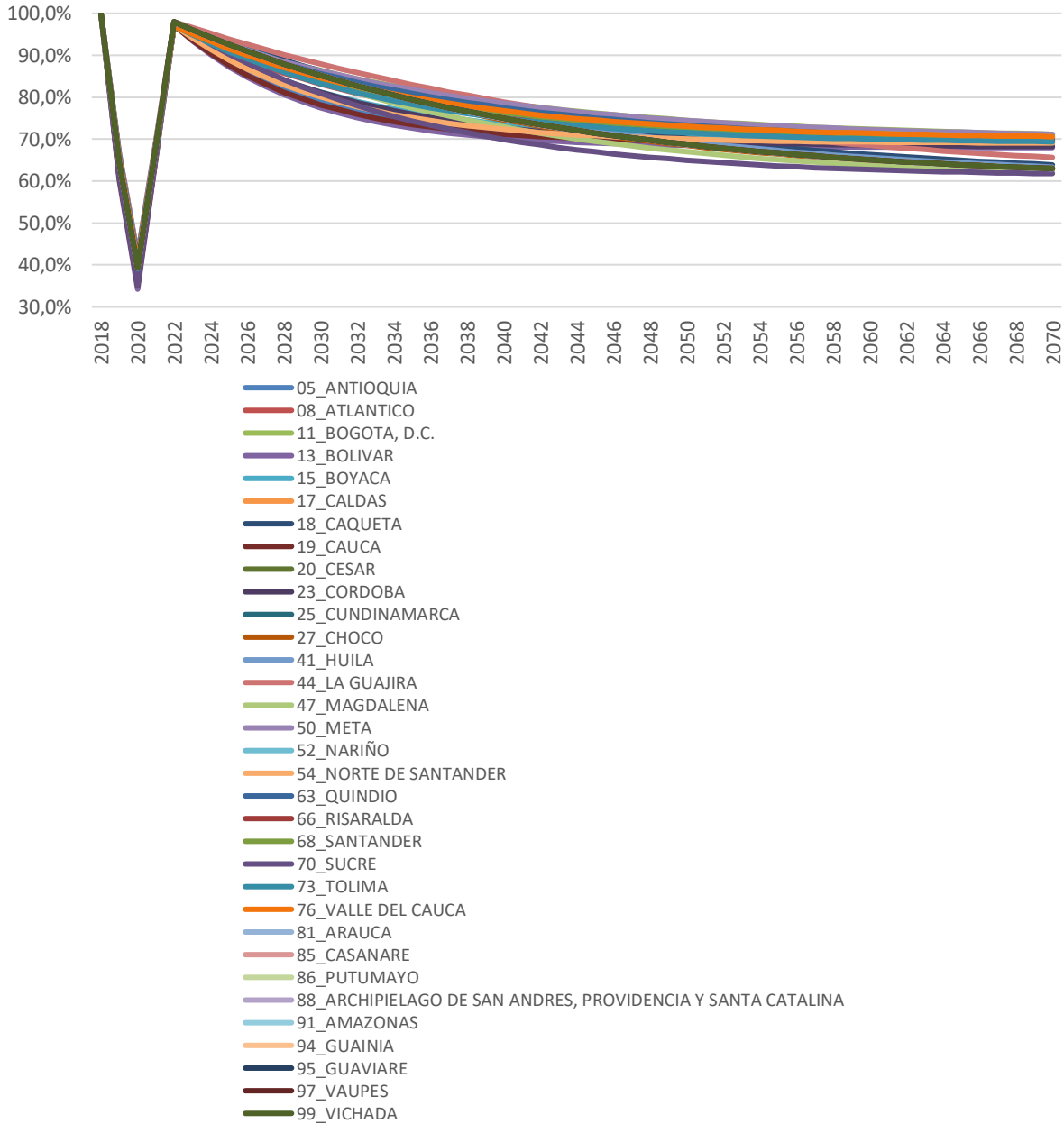
La proyección por el método de componentes. El módulo de migración



Fuente: DANE, Proyecciones de población y estudios demográficos.

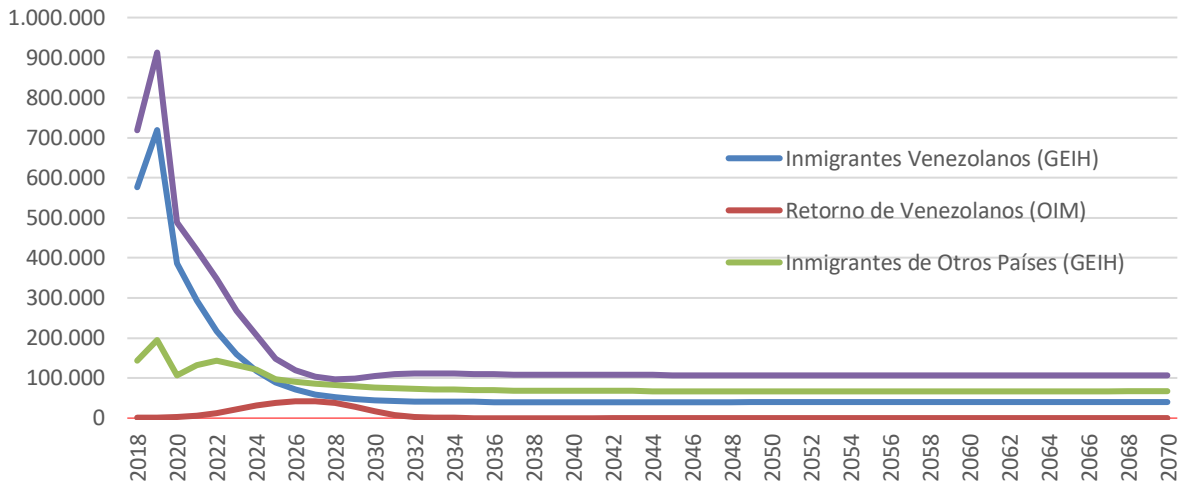
¹ Es un ejemplo de modelo dividido, trata la migración en dos etapas, en primer lugar, se proyecta el número de emigraciones en cada región y colocando a los migrantes en un “pool” común y en segundo lugar, asignando esos flujos migratorios a los destinos. El modelo usa coeficientes de reparto, por edades quinquenales, sexo y áreas geográficas, para distribuir la población migrante. La caja negra (pool), permite que, dado que la suma del total de migrantes debería ser cero, se garantice que al final del procedimiento la bolsa quede vacía.

Gráfico 1. Proyección de la tendencia del ISM interdepartamental 2018-2070



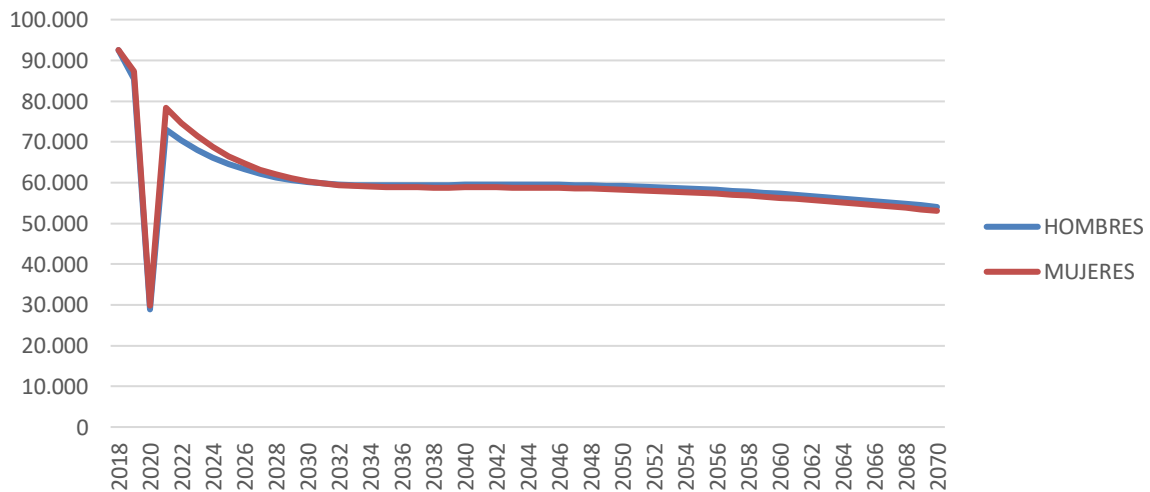
Fuente: DANE, Proyecciones de población y estudios demográficos.

Gráfico 2. Proyección de la tendencia de flujo de inmigrantes internacionales 2018-2070



Fuente: DANE, Proyecciones de población y estudios demográficos.

Gráfico 3. Proyección de la tendencia de flujo de inmigrantes internacionales 2018-2070



Fuente: DANE, Proyecciones de población y estudios demográficos.

5. Métodos de proyecciones para áreas pequeñas (desagregaciones por área y a nivel de municipios)

Existe una creciente demanda de proyecciones desagregadas a escala subnacional. No obstante, el desarrollo de estos métodos de proyección desagregada presenta importantes rezagos respecto a los desarrollos realizados para proyecciones nacionales. En la elaboración de las nuevas proyecciones poblacionales, se tuvo en cuenta que, al contrario que en poblaciones de gran volumen, en áreas menores la influencia del crecimiento natural de la población es frecuentemente sobrepasada por el efecto de la migración. Por otro lado, la elaboración de este tipo de proyecciones con un alto nivel de detalle geográfico es importante para la atención de subpoblaciones con demandas especiales, el monitoreo de igualdad de oportunidades en el mercado laboral y de vivienda, y la planeación de ciudades que se están despoblando. De igual manera, se tuvo en cuenta que factores como el tiempo de desplazamiento a las capitales departamentales y a los municipios pertenecientes a aglomeraciones urbanas restringen en alguna medida las posibilidades de crecimiento poblacional de algunos municipios y áreas.

En Colombia, cabe esperar que como consecuencia del alto impacto de los patrones de migración se haga más visible la existencia de fuertes desequilibrios en la composición por edades de las poblaciones. Los cambios de ritmo en la dinámica de crecimiento y la llegada a los municipios de importantes flujos de población en periodos relativamente cortos de tiempo van a producir cambios en la distribución de la población por edad, sexo y localización geográfica.

En términos de las metodologías a aplicar, se optó por realizar una extensa revisión de literatura sobre el tema de proyecciones en áreas menores. En este caso, al igual que en otros países del contexto latinoamericano se aplicaron modelos de base estadístico – matemática para proyectar áreas pequeñas, dado que la calidad de las estadísticas en estas poblaciones no permitió la utilización del método de componentes. En particular para los totales se optó por utilizar aproximaciones compromiso entre los resultados de métodos de series de tiempo para el volumen y crecimiento poblacional y la información auxiliar de la que se disponía en el momento. Estos resultados fueron desagregados por áreas urbanas y rurales buscando guardar consistencia con estimaciones realizadas en el penúltimo ejercicio de proyecciones poblacionales a escala municipal.

Control de dos o más dimensiones para la consistencia entre áreas geográficas

El método utilizado para controlar dos o más dimensiones de las proyecciones fue el Iterative Procedure Fitting (IPF), conocido también como método de las tablas cuadradas. En el caso de la implementación para las proyecciones poblacionales en Colombia, la tabla de contingencia ayudó a respetar el total departamental de cada edad simple como la suma de la misma edad de todos los municipios del departamento. El resultado final fue la estructura de población de cada municipio, donde la suma de los municipios garantiza ser consistente con la estructura departamental.

En el caso de dos dimensiones, los totales de fila o de columna pueden diferir de distribuciones marginales de las cuales se disponía de estimaciones o proyecciones previas. El ajuste de los datos de la tabla de contingencia puede ser alcanzado mediante el ajuste proporcional iterativo de filas y columnas. Esto es iniciar con el ajuste de filas, luego sobre el resultado ajustar columnas, posteriormente filas, luego columnas y seguir el proceso hasta que el margen de diferencia entre una tabla y la anterior sea mínimo. La convergencia de este proceso es generalmente alcanzada en el segundo ciclo -el cuarto ajuste proporcional- (Siegel y Swanson, 2004).

El primer ajuste proporcional puede ser expresado de la siguiente manera:

$$N_{ij}^{(1)} = N_{ij}^{(0)} * \frac{N_i}{n_i}$$

En este caso se ajusta el total de cada edad simple del departamento, con la suma de cada edad simple de los municipios. Para ello se calcula una relación de ajuste del total de población de cada edad simple del departamento y la suma de la misma edad simple de los municipios, entonces se crea un factor de ajuste por edad, el cual es aplicado en todos ellos.

El segundo ajuste proporcional se expresa mediante la siguiente fórmula:

$$N_{ij}^{(2)} = N_{ij}^{(1)} * \frac{N_j}{n_j}$$

Se busca respetar la población total por sexo a nivel departamental para cada año, teniendo en cuenta la suma de la población para cada uno de los sexos a nivel municipal. Este ajuste se realiza de la misma manera, es decir calculando la relación entre el total departamental y la suma de la distribución municipal resultante del paso anterior.

Los pasos 1 y 2 corresponden a un ciclo y, para continuar un proceso que garantizara convergencia, fue necesario desarrollar cinco ciclos.

Siguiendo el proceso detallado, el tercer ajuste proporcional puede ser expresado de la siguiente manera:

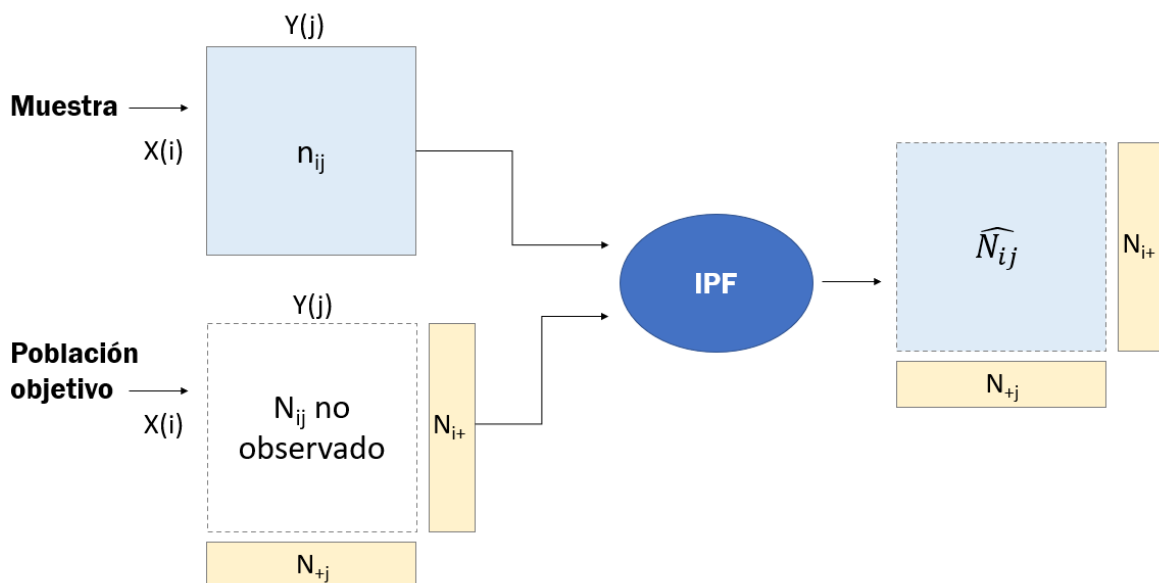
$$N_{ij}^{(3)} = N_{ij}^{(2)} * \frac{N_{i.}}{n_{i.}}$$

El cuarto ajuste proporcional se expresa mediante la siguiente formula:

$$N_{ij}^{(4)} = N_{ij}^{(3)} * \frac{N_{.j}}{n_{.j}}$$

Y se continúa el proceso hasta la convergencia de resultados.

Diagrama 1. Mecanismo de ajuste proporcional para dos dimensiones IPF para datos muestrales y marginales conocidas.



Fuente: DANE, Proyecciones de población y estudios demográficos.

El método utilizado para el ajuste de las estimaciones a números enteros partió de reconocer que en el caso de utilizar métodos de estimación como el IPF (tabla cuadrada) los resultados no corresponden con números enteros, como se requieren en proyecciones o estimaciones poblacionales. La alternativa

que fue tomada fue la planteada por Lovelace y Ballas (2013): truncar números enteros, replicar y muestrear (en inglés TRS: truncate, replicate, sample).

Por último, cabe destacar que el estudio de los errores de proyección de ejercicios postcensales anteriores sirvió como insumo para identificar algunos de los principales problemas que podían corregirse en la elaboración de nuevos conjuntos de proyecciones poblacionales, en particular en la consistencia de las poblaciones de menor dimensión y las que hacen parte de conglomerados urbanos.

Bibliografía

Bernard, A., Charles-Edwards, E., Alvarez, M., Wohland, P., Loginova, J., & Kalemba, S. (2020). Anticipating the impact of COVID-19 on internal migration.

Borges, G. M. (2018). Consistent population estimates: an application to Brazil. University of California, Berkeley.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International journal of forecasting*, 22(3), 547-581.

Chackiel, J. (1984). Proyección de la fecundidad: Criterios y procedimientos utilizados en CELADE. En: *Métodos para proyecciones demográficas - LC/DEM/CR/G. 5-1984-p. 43-85.*

Dandekar, A., & Ghai, R. (2020). Migration and reverse migration in the age of COVID-19. *Economic and Political Weekly*, 55(19), 28-31.

Del Popolo, F. (2000). Los problemas en la declaración de la edad de la población adulta mayor en los censos. *Serie población y desarrollo*, No. 8. Santiago de Chile: Centro Latinoamericano de Demografía (CEPAL).

Felipe, M. y Guillén, M. (1999). *Evolución y Predicción de las Tablas de Mortalidad Dinámicas para la Población Española*. Cuadernos de la Fundación. Madrid: Fundación Mapfre Estudios.

Heligman, L. and J.H. Pollard (1980), The age pattern of mortality. *Journal of the Institute of Actuaries* 107, pp. 49-80.

Gonzales, L., López-Gay, A., Valverde, J. R., & Rowe, F. Cambios de residencia en tiempos de COVID-19: un poco de oxígeno para el despoblamiento rural. *Perspectives Demographiques*.

Lovelace, R., Dumont, M., Ellison, R., & Založnik, M. (2017). Spatial microsimulation with R. Chapman and Hall/CRC.

Lutz, W., Goldstein, J.R., y Prinz, C. (1996a). Alternative approaches to population projections. In Lutz, W. (ed) *The Future Population of The World: What Can We Assume Today?* Earthscan Publication Ltd., London. pp. 14-44.

Moultrie, T. A., Dorrington, R. E., Hill, A. G., Hill, K., Timæus, I. M., y Zaba, B. (2013). Tools for demographic estimation. International Union for the Scientific Study of Population.

O'Hare, W. P. (2019). Differential undercounts in the US census: who is missed? (p. 167). Springer Nature.

Rincón, M. J. (1984). Conciliación censal y determinación de la población base. En: *Métodos para proyecciones demográficas - LC/DEM/CR/G. 5-1984-p. 13-42.*

Rogers, A., and Castro, L. J. (1981). Model migration schedules. IIASA Research Report, 81, 1 to 160.b

Skeldon, R. (2012). Migration transitions revisited: Their continued relevance for the development of migration theory. *Population, Space and place*, 18(2), 154-166.

Smith, S. K., Tayman, J., y Swanson, D. A. (2013). *A practitioner's guide to state and local population projections.* Springer Netherlands.

Swanson, D. A., y Tayman, J. (2012). *Subnational population estimates (Vol. 31).* Springer Science & Business Media.

Wilson, T., & Bell, M. (2004). Comparative empirical evaluations of internal migration models in subnational population projections. *Journal of Population Research*, 21(2), 127-160.