

Una revisión de la versión probabilística del método de las componentes demográficas para pronosticar la población de México

A review of the probabilistic population projection for Mexico using the cohort-component method

Víctor Manuel García Guerrero*^o

RESUMEN

En este trabajo se revisa el método de proyección de población de las de componentes demográficas en su versión probabilística. Este método es aplicado a la información de la conciliación demográfica elaborada por la Sociedad Mexicana de Demografía para el periodo 1960-2010. Se analiza y evalúa la proyección para el periodo 2011-2050 comparando sus resultados con información observada o estimada al año 2015 para la población total, la fecundidad y la migración. La proyección de los niveles y tendencias tanto de las variables demográficas y como de la población total por edades y sexo es precisa en el corto plazo; sin embargo, se ubican algunas áreas oportunidad en cuanto a la proyección de estructura por edades de la fecundidad, en particular para el año 2015, y respecto a la amplitud de los intervalos de predicción. La diferencia en el volumen de población masculina proyectada y observada entre los 15 y 50 años para el año 2015 podría estar relacionada con el incremento de la mortalidad en este grupo étnico.

ABSTRACT

In this paper, the probabilistic version of the cohort-component method is reviewed. Such a method is applied to the information of demographic conciliation made by the Mexican Society of Demography for the period 1960-2010. The projected period 2011-2050 is analyzed and assessed comparing these results with information observed or estimated for total population, fertility and migration by 2015. The projection of levels and trends for both the demographic variables and the total population by age and sex is accurate in the short term; however, there are some areas of opportunity in fertility age structure projection, in particular by 2015, and in the amplitude of prediction intervals. The difference between projected and observed male population volume between the ages of 15 and 50 by year 2015 could be related to the increase in mortality in this age group.

Recibido: 5 de noviembre del 2016
Aceptado: 1 de septiembre del 2017

Palabras clave:

Proyecciones de población; métodos demográficos; proyecciones probabilísticas de población; dinámica demográfica.

Keywords:

Population projections; demographic methods, probabilistic population projections, demographic dynamic.

Cómo citar:

García Guerrero, V. M. (2017). Una revisión de la versión probabilística del método de las componentes demográficas para pronosticar la población de México. *Acta Universitaria*, 27(NE-1), 5-16. doi: 10.15174/au.2017.1665

INTRODUCCIÓN

Una proyección de población es un escenario hipotético de lo que ocurriría con la dinámica demográfica de un país o región si determinadas condiciones sociales, económicas y demográficas ocurriesen. Para su cálculo se necesitan dos insumos generales: 1) una estimación inicial tanto de la población total como de cada una de sus componentes demográficas básicas (mortalidad, fecundidad y migración internacional) por sexo y grupos de edad y, 2) un conjunto de supuestos sobre la evolución futura de cada una de dichas componentes.

En México, las proyecciones de población han sido una herramienta utilizada para planear aspectos estratégicos del Estado dirigidos a satisfacer

* Centro de Estudios Demográficos, Urbanos y Ambientales, El Colegio de México A.C. Camino al Ajusco 20, Pedregal de Santa Teresa, Ciudad de México, C.P. 10740. Tel. +52 55 5449 3000 ext. 4265. Correo electrónico: vmgarcia@colmex.mx

^o Autor de correspondencia.

algunas demandas sociales básicas como: vivienda, educación, salud y empleo (Consejo Nacional de Población [Conapo], 2014a). Estas proyecciones son necesarias también para conocer la demanda futura de alimento, agua, energía, transporte y otros servicios. El Programa Nacional de Población (PNP) que se elabora cada sexenio y que se encuentra articulado con el respectivo Plan Nacional de Desarrollo (PND), se nutre de los resultados de la última conciliación y proyección demográfica calculada por el Conapo (2014b). Con las proyecciones también se calculan los ponderadores de las diversas encuestas que levanta el Instituto Nacional de Estadística y Geografía (INEGI) que, entre otras cosas, son utilizadas como insumo para calcular la distribución del presupuesto a las entidades federativas y municipios¹, para medir y establecer programas de combate a la pobreza, para establecer metas de empleo y evaluar los logros en este campo, por citar algunas.

Así, es indiscutible el papel fundamental que tienen las proyecciones de población en la planeación de los distintos niveles de gobierno; sin embargo, están lejos de ofrecer datos ciertos y puntuales sobre el futuro demográfico ya que, como en cualquier extrapolación del futuro, se articulan muchos factores que incrementan su incertidumbre o rango de error (García Guerrero, 2014a). Dichas proyecciones son simulaciones sobre el futuro de la demografía de tipo “qué pasaría si...” ocurriese determinada evolución de la mortalidad, fecundidad y migración (Alho & Spencer, 2005; Lee 1998) Es decir, las proyecciones de población parten de un esquema metodológico determinista.

Bajo este esquema, es prácticamente imposible calcular todos los escenarios posibles o todas las combinaciones de trayectorias posibles de la mortalidad, fecundidad y migración que conlleven a escenarios de alto o bajo crecimiento demográfico. Usualmente se calculan escenarios alternativos donde se simula un alto crecimiento poblacional y uno bajo con el fin de mostrar la plausibilidad de la proyección que será utilizada para la toma de decisiones (Conapo, 2014a; Gómez & Partida, 2001). Sin embargo, se ha demostrado que dichas variantes de proyección son estadísticamente incorrectas en el sentido de que de manera artificial suponen que las variables demográficas relacionadas se correlacionan perfectamente a lo largo del horizonte de proyección (García, 2014a; Lee, 1998).

Para subsanar las deficiencias de las proyecciones de población deterministas se han realizado proyecciones de tipo probabilístico. Estas consisten en determinar el conjunto de escenarios probables que tendría la dinámica demográfica en su conjunto, así como la de las componentes que la conforman a partir de una serie temporal de información estimada (la cual, para el caso de México puede ser la conciliación demográfica). Existen tres tipos de aproximaciones para el cálculo de proyecciones probabilísticas: 1) la bayesiana (Raftery, Alkema & Gerland, 2014), 2) utilizando métodos de simulación estocástica (García, 2014a) y 3) utilizando las técnicas de suavizamiento exponencial y datos funcionales (Hyndman & Booth, 2008).

Para México se han realizado diversas aproximaciones probabilísticas para proyectar población total o tasas de crecimiento en una región (Silva, Guerrero & Peña, 2011; Silva & Ordorica, 2013), para proyectar alguna de las componentes demográficas (García-Guerrero, 2016; García & Ordorica, 2012; Morales-Eslava, Silva-Urrutia & Ordorica-Mellado, 2016) y para proyectar la estructura por edades de la población (García, 2014a; 2014e). En este trabajo se revisan tanto las proyecciones probabilísticas de las componentes demográficas (excepto la mortalidad ya que no hay datos observados a la fecha de elaboración de este documento) y de la estructura por edades de la población del país.

Cualquier tipo de proyecciones de población, ya sean probabilísticas o deterministas, depende principalmente de dos factores: 1) la estimación inicial de la dinámica demográfica y 2) los supuestos que se utilizan para su extrapolación. Las estimaciones iniciales dependen a su vez de la calidad de la información; así como de las técnicas estadísticas y demográficas para realizarlas.

Desde la última década del siglo pasado, después de cada censo y conteo de población en México se realiza un tipo muy particular de estimación demográfica, llamada conciliación demográfica. Con la conciliación se establece la población inicial de cada ejercicio de proyección poblacional. El procedimiento para su cálculo consiste básicamente, en lo siguiente:

1. Analizar la evolución histórica de los niveles y tendencias de la fecundidad, mortalidad, migración interna e internacional a través de distintas fuentes de información.

¹ El Fondo General de Participaciones establecido en la Ley de Coordinación Fiscal señala que, en la ecuación de reparto presupuestal, la variable poblacional que se debe utilizar es la última publicada por el INEGI, por lo que, además de una encuesta también podría ser la de un censo o conteo.

2. Seleccionar las fuentes de datos más adecuadas o consistentes (para ciertos puntos en el tiempo pueden combinarse) para determinar, por un lado, los niveles y tendencias y, por otro, la estructura por edades y sexo.
3. Usar métodos de corrección y suavizamiento.
4. Se proyecta la población (utilizando los datos de las estimaciones de las componentes demográficas) desde distintos puntos en el tiempo hasta el año del censo más reciente y de retroproyección de dicho año hasta distintos puntos en el tiempo en el pasado.
5. Finalmente, se elige la serie histórica que, a juicio del demógrafo, resulte más consistente con la dinámica del pasado y que minimice la diferencia entre los datos conciliados y los censados.

Por otro lado, las técnicas de proyección conforman un cuerpo de conocimiento que integra, en una sola metodología, gran parte de las herramientas metodológicas del análisis demográfico (García, 2014a). Como se mencionó antes, existen dos enfoques metodológicos: el determinista y el estocástico. Desde ambos enfoques se puede proyectar el volumen total de la población a partir de la extrapolación de cohortes por sexo. A esta forma de proyección se le conoce, en español, como método de las componentes demográficas (MCD)².

Fuentes de Información y Metodología

En términos generales, el MCD consiste en proyectar los niveles y tendencias de la mortalidad (utilizando la esperanza de vida al nacimiento), fecundidad (por medio de la tasa global de fecundidad), inmigración y emigración, utilizando algún método de extrapolación; luego se imputa una estructura por edades de alguna otra fuente (por ejemplo, para la mortalidad, en algunos casos se utilizan las tablas modelo de las Naciones Unidas)³.

Para proyectar los sobrevivientes futuros por edades, primero se proyecta la mortalidad y se calculan las probabilidades de supervivencia, mismas que se multiplican por la población a inicio de cada año; después, al resultado, se le agregan los inmigrantes y se le restan los emigrantes para cada edad. Finalmente,

se calculan los nacimientos a partir de la proyección de las tasas específicas de fecundidad y de multiplicarlas a la población femenina proyectada en edades reproductivas. El procedimiento se repite hasta cubrir todo el horizonte de tiempo que se desea proyectar (es decir, cada uno de los años en el futuro que se desean proyectar, también llamado horizonte de proyección).

En su vertiente determinista, este procedimiento se calcula una sola vez. Es decir, cada punto en el tiempo calculado implica un solo punto en el tiempo en el futuro. En cambio, en la versión probabilística, cada punto en el tiempo calculado implica un conjunto de posibles realizaciones del futuro; por ello, usualmente a este tipo de proyecciones se les conoce como pronósticos. Este término hace referencia a que no se calcula un solo escenario futuro, sino que se calculan miles de ellos en función de la incertidumbre asociada (llamadas realizaciones del pronóstico). Dichas realizaciones se agrupan en un intervalo de predicción que lleva consigo una probabilidad de ocurrencia. A partir de estos cálculos es posible extraer un escenario que resume el comportamiento promedio de la dinámica de la variable demográfica en cuestión o de la población total. En este marco, la principal diferencia entre las proyecciones y los pronósticos de población es que en las primeras se construye un escenario y después se evalúa su plausibilidad; en cambio en las estocásticas primero se mide y pronostica su variabilidad y luego se extrae el escenario medio de referencia.

Es decir, en las proyecciones deterministas la incertidumbre no es considerada para su cálculo, aunque al calcular escenarios alternativos se reconoce que juega un rol en la precisión de la proyección. En cambio, los pronósticos de población parten del hecho de que existe incertidumbre en torno a la dinámica demográfica.

Como se mencionó anteriormente, las proyecciones de población tienen dos fuentes principales de incertidumbre: 1) la población inicial o de partida para la proyección y 2) los supuestos sobre la evolución futura de cada una de las variables demográficas. Los primeros no son despreciables y los errores atribuibles a los supuestos incorporan más incertidumbre a las proyecciones en lugar de disminuirla (García, 2011; 2014a.) Además, la demografía mexicana tiene intrínsecamente otra gran fuente de incertidumbre: la información a partir de la que se estima la población base. Si esta fuera perfecta entonces la incertidumbre de las proyecciones solo radicaría en los supuestos.

² En inglés se le conoce como *Cohort-Component Method* (CCM).

³ Las cuales pueden ser consultadas en <https://esa.un.org/unpd/wpp/Download/Other/MLT/>

En este sentido, la información básica utilizada para calcular las proyecciones probabilísticas de población provienen de la conciliación demográfica que realizó la Sociedad Mexicana de Demografía (Somede, 2011). Esta es la fuente de información demográfica armonizada más confiable con la que se cuenta hasta el momento además de que concilia o armoniza la información demográfica desde 1960 hasta la fecha del último censo de población. Existe también la información que ofrece el Consejo Nacional de Población, sin embargo, además de que data de 1990 (es decir, ofrece 30 años menos de información histórica), se encuentra armonizada no solo demográficamente sino políticamente (García, 2014a).

Pronóstico de las componentes demográficas básicas

El principal método para modelar y pronosticar probabilísticamente las componentes demográficas ha sido desarrollado en los trabajos de Lee (1993), Lee & Carter (1992) y Lee & Tuljapurkar (1994). Luego, para proyectar la población total por edades y sexo, se utiliza cada una de las realizaciones de las componentes demográficas y se articulan por medio del método de las componentes. Como se mencionó, el proceso se repite miles de veces considerando los factores de variabilidad de los modelos de series de tiempo con que fueron proyectadas. Con todas esas simulaciones se calculan los percentiles que se desean para cada punto en el tiempo. En este caso se calculan los percentiles 2.5% y 97.5% para determinar el intervalo de 95% de confianza. Para su cálculo se utiliza la función “quantile” del lenguaje de programación R. (R Core Team, 2016). El cálculo de los percentiles o cuantiles está sustentado en el trabajo de Hyndman & Fan (1996) donde se demuestra que todos los cuantiles muestrales pueden ser definidos como promedios ponderados de las estadísticas de orden consecutivas. Es decir, si x_i es la j -ésima estadística de orden, n es el tamaño de la muestra, m es una constante determinada por el tipo de cuantil muestral, γ es una función de $j = [np + m]$, $g = np + m - j$, $(j - m) / n \leq p < (j - m + 1) / n$, e $1 \leq i \leq 9$, entonces el cuantil de tipo i está definido por⁴:

$$Q_i(p) = (1 - \gamma) x_i + \gamma x_{i+1}$$

En este trabajo y con fines de sintetizar la parte técnica de este trabajo, el modelo de Lee-Carter (LC)

es generalizado para proyectar cada una de las componentes demográficas básicas. El modelo de LC parte del supuesto de que existe una relación lineal entre el logaritmo natural de las tasas específicas de los eventos demográficos y dos factores que las determinan: la edad x y el tiempo t .

Sea $\eta_{x,t}$ la tasa específica del evento demográfico (en este caso solo mortalidad o fecundidad) a edad x al tiempo t y sea $\xi_{x,t}$ el volumen del evento demográfico (en este caso sólo inmigración o emigración)⁵ a edad x al tiempo t . De esta manera, si se define a α_x como el patrón por edades de referencia de $\ln(\eta_{x,t})$, β_x como el cambio por edades a lo largo del tiempo (intensidad) y k_t como la tendencia temporal, entonces,

$$\ln(\eta_{x,t}) = \alpha_x + \beta_x k_t + \xi_{x,t} \quad (1)$$

garantizando que tiene solución única restringiéndola a que $\sum_{x=0}^{\omega+} \beta_x = 1$ y $\sum_{t=1}^N k_t = 0$.⁶ Lo anterior conlleva a que α_x sea estimada como el promedio del logaritmo de las tasas a lo largo del tiempo; sin embargo, es posible utilizar algún criterio empírico que mejore la estimación. Por ejemplo, para el caso de la modelación del logaritmo natural de las tasas específicas de fecundidad en México se ha propuesto aproximar este parámetro utilizando la última información histórica o, en su caso, el promedio de los últimos años (García, 2014a).

Los parámetros β_x y k_t son estimados utilizando la descomposición en valores singulares (DVS) de la matriz $\ln(\eta_{x,t}) - \alpha_x$. Es decir, ambos parámetros quedan aproximados por la primera componente principal en dicha descomposición, la cual es la que representa la mayor cantidad de variabilidad asociada. En el texto original de Lee & Carter (1992) sugieren tomar solo dicha primera componente; sin embargo, en Hyndman & Booth (2008) sugieren incorporar todas las componentes involucradas, es decir, estimar todas las β_x^i y k_t^i para $i = \{1, 2, \dots, n\}$, en virtud de que la capacidad computacional lo permite hoy en día. En este trabajo se toma únicamente la primera componente ya que, como se explica más adelante, el incorporar más parámetros incrementaría considerablemente la complejidad del algoritmo de simulación, además que, de la segunda componente en adelante los parámetros carecen de interpretación.

Para la mortalidad se utilizó el método usual de LC aplicado al caso mexicano (Aburto & García-Guerrero,

⁴ Los tipos de cuantil están definidos en Hyndman & Fan (1996). Una explicación muy sintética sobre ellos puede encontrarse en <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/quantile.html>

⁵ En trabajos previos se ha usado el volumen de la migración neta (García, 2014a).

⁶ $\omega +$ es la notación para el último grupo abierto de edades.

2015; García & Ordorica, 2012; García, 2014a; 2014b; 2014c; 2014d) y para la migración se utilizó la aproximación propuesta en García (2016; 2014a; 2014f). Respecto a la migración, en este trabajo se propone desagregar la migración neta en inmigración y emigración. Como se detalla en García (2016) el modelo LC se modifica utilizando la información a nivel y no ajustada logarímicamente. Es decir,

$$\xi_{x,t} = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$$

donde α_x es el promedio en el tiempo del número de inmigrantes o emigrantes para cada edad x , a diferencia de los casos de la fecundidad y mortalidad, β_x y k_t no se estiman utilizando la DVS, sino que

$$k_t = \sum_{x=0}^{\omega} (\xi_{x,t} - \alpha_x)$$

para cada t . Luego, β_x se estima utilizando mínimos cuadrados ordinarios. Finalmente, se estima un parámetro de tendencia que garantiza que la volatilidad de la migración total (inmigración o emigración) no conlleve a valores imposibles de población (por ejemplo, valores del volumen poblacional negativos). Este parámetro es modelado a través de una transformación logística de la migración dirigida total (inmigración o emigración). Es decir,

$$Z_t = \ln \left(\frac{Mig_t - L}{U - Mig_t} \right)$$

donde Mig_t es la migración total en el periodo t y, U y L son las cotas superior e inferior de la migración total que se refieren a los niveles de sostenibilidad del flujo de inmigrantes y emigrantes (García, 2016).

Para realizar el pronóstico, el parámetro k_t para el caso de la fecundidad y mortalidad, y el z_t para el caso de la inmigración y emigración son modelados como procesos autorregresivos integrados de promedios móviles (ARIMA). Es decir, para $W_t = (1-B)^d k_t$ o $W_t = (1-B)^d z_t$ donde B es el operador de rezago tal que $B^d k_t = k_{t-d}$, se tiene que,

$$W_t = \phi_1 W_{t-1} + \dots + \phi_v W_{t-v} + X_t + \theta_1 X_{t-1} + \dots + \theta_a X_{t-a}$$

donde X_t son procesos puramente aleatorios con media cero y varianza constante.

Usualmente, en los eventos demográficos de fecundidad y mortalidad no se necesita más de una diferencia ya que cada diferencia implica un cambio estructural en la dinámica de la variable; ello, en términos

demográficos está directamente relacionado con la transición demográfica. En otras palabras, más de una diferencia implicaría más de una transición demográfica, lo cual hasta el momento no ha ocurrido en algún otro contexto. En ocasiones es necesario adicionar un parámetro de tendencia determinista o deriva, μ , tal que $\phi_0 = \mu + \sum_{i=1}^p \phi_i$. La función “auto.arima” de la librería “forecast” del lenguaje de programación R permite determinar el mejor modelo de series de tiempo que se ajusta a la información (Hyndman, 2016; Hyndman & Khandakar, 2008).

Por último, para cada evento demográfico y cada sexo se calculan 10 000 simulaciones de los parámetros ϕ_j y θ_i modelados de manera multivariada como $\Theta \sim \mathcal{N}_N(M, \Sigma)$ donde M es el conjunto de valores medianos del conjunto de parámetros Θ y Σ es la matriz de varianzas y covarianzas de dicho conjunto de parámetros. Con cada una de esas realizaciones se extrapola cada k_t y z_t y manteniendo constantes los parámetros α_x y β_x respectivos a cada evento demográfico y cada sexo, se obtiene una extrapolación de las tasas específicas o niveles por edad de cada evento demográfico (10 000 por cada año proyectado). Con esas 10 000 realizaciones de las tasas específicas de mortalidad, fecundidad y de los volúmenes por edad de los inmigrantes y emigrantes, se calculan, por un lado, los percentiles correspondientes para calcular los indicadores resumen (por ejemplo, esperanza de vida y tasa global de fecundidad) con intervalos de predicción y utilizando cada uno de ellos, se emplea el método de las componentes demográficas (véase García, 2014a para una explicación más detallada respecto al cálculo del método de las componentes demográficas) para calcular 10 000 realizaciones de la población total por edades y sexo. Después también utilizando el cálculo de los percentiles explicado arriba se calculan intervalos de predicción para la población total por edades y sexo. Es importante señalar que, si se desea calcular algún otro tipo de indicador demográfico como tasas de crecimiento o razones de dependencia, es necesario utilizar todas las realizaciones calculadas y después calcular los respectivos percentiles.

ANÁLISIS Y RESULTADOS

Mortalidad

De esta manera, la figura 1 muestra las esperanzas de vida al nacimiento (EVO) por sexo pronosticadas de acuerdo al método descrito anteriormente para el periodo 2011-2015. Se utiliza la esperanza de vida como

indicador resumen de la mortalidad en virtud de que este indicador resume la experiencia de mortalidad que tiene la población considerando su estructura por edades. Se presentan los intervalos de predicción del 67% y 95% con el fin de que el lector evalúe el incremento en la incertidumbre en el pronóstico de esta variable demográfica.

La mortalidad ha sido una de las variables demográficas menos inciertas a lo largo de la historia de México. Salvo por dos episodios epidémicos de sarampión (a principios de la década de 1970 y de la década de 1990) que frenaron el avance de la EVO, esta recuperó su tendencia casi inmediatamente debido a que dicha epidemia fue atendida con oportunidad. Sin embargo, desde el año 2006 ha habido un descenso sostenido en la EVO debido principalmente al fuerte incremento en homicidios en las edades jóvenes (entre 15 y 45 años de edad) y debido a la mortalidad relacionada con enfermedades como la diabetes en edades posteriores a los 50 años de edad (Aburto, Beltrán-Sánchez, García-Guerrero & Canudas-Romo, 2016; Canudas-Romo, Aburto, García-Guerrero & Beltrán-Sánchez, 2017; Canudas-Romo, García-Guerrero & Echarri-Canovas, 2014).

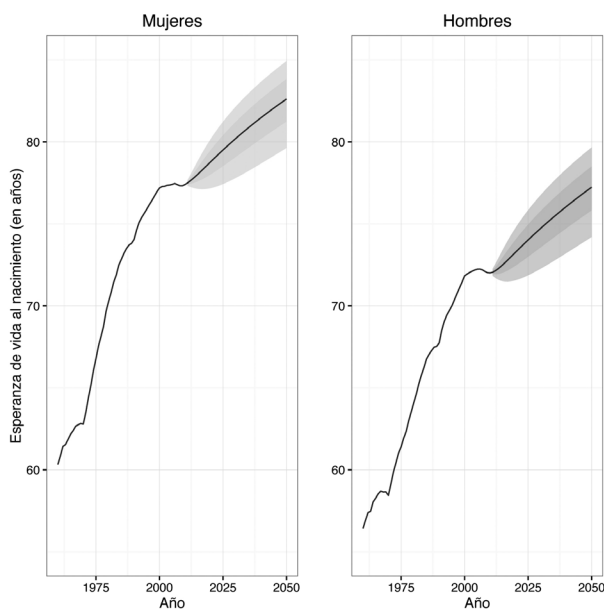


Figura 1. Esperanza de vida al nacimiento por sexo estimada 1950-2010 y pronosticada 2011-2050 con intervalos de predicción del 67% y 95%. Fuente: Cálculos propios con base en Somede (2011).

A la fecha en que se escribe el presente trabajo no se han actualizado las tablas de mortalidad para el país, sin embargo, el volumen de homicidios no presenta una disminución significativa que permita inferir que dicha tendencia ha cambiado. De ser así, al año 2016 se tendría una EVO de 71.4 años para hombres y de 77.1 años para mujeres. Esto es, poco más de 7 meses menos de EVO respecto de la del año 2010 y poco más de 10 meses menos de EVO respecto al año 2006 para el caso de la población masculina. Con relación a la población femenina, dicha disminución representa cerca de 4 meses menos de EVO respecto tanto para el año 2010 como para el 2006.

Como indica el intervalo de predicción inferior, para el caso de la población femenina es posible que dicha disminución alcance un mínimo de 77.14 años alrededor del año 2017. Si ocurriese esta situación, sería hasta el año 2023 que se recupere el nivel de 2010. Pero con un 50% de probabilidad la tendencia creciente se podría recuperar inmediatamente. Para el caso de la población masculina el valor mínimo también se alcanzaría en el mismo año 2017 pero se recuperaría el nivel de 2010 hasta el año 2025. Finalmente, como los intervalos de predicción del 67% y 95% con el fin de que el lector evalúe el incremento en la incertidumbre en el pronóstico de esta variable demográfica.

Fecundidad

Por otra parte, en la figura 2 se muestra el pronóstico actualizado de la Tasa Global de Fecundidad. Este indicador resumen se utiliza para determinar el nivel de la fecundidad de la población controlando el efecto de la estructura por edades de la fecundidad (TGF). Hacia el año 2010 se había estimado un promedio de 2.35 hijos por mujer, de acuerdo con esta proyección, al año 2016 la TGF podría estar oscilando entre 2 y 2.4 hijos por mujer (considerando el intervalo de predicción del 95%). La mediana de dicho indicador para ese año es de 2.2 hijos por mujer en promedio. Considerando a la mediana se esperaría que para el año 2025 se alcance el nivel del reemplazo y después continúe la tendencia por debajo de dicho nivel, pero a un ritmo de decrecimiento cada vez menor.

Ahora bien, para el año 2014 se proyecta una TGF de entre 2.1 y 2.4 hijos por mujer en promedio, con una mediana de 2.27. La Encuesta Nacional de la Dinámica Demográfica (Enadid) de ese año arroja un resultado medio de 2.2 hijos por mujer; esto es, la dinámica de la fecundidad se mantuvo de acuerdo a lo

esperado (considerando la volatilidad histórica del indicador), no hubo algún evento “raro” que permitiera que el valor observado de la TGF saliera de los límites de predicción. El “error” respecto a la mediana del pronóstico fue de sólo 0.07 hijos por mujer. Sin embargo, el calendario de la fecundidad pronosticado no es similar al reportado en la encuesta (figura 3).

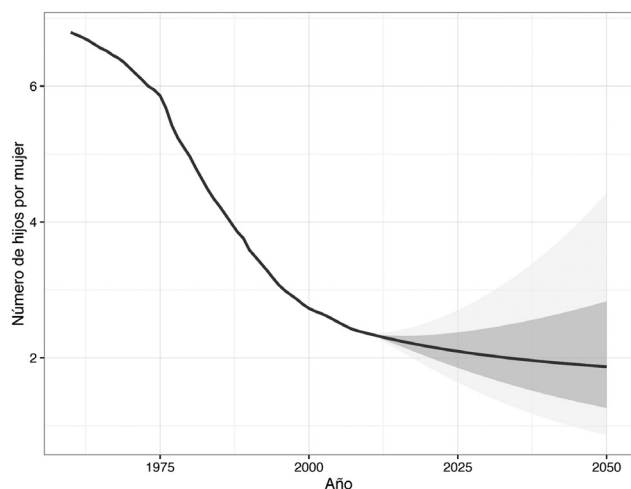


Figura 2. Tasa Global de Fecundidad estimada 1950-2010 y pronosticada 2011-2050 con intervalos de predicción del 67 y 95%.
Fuente: Cálculos propios con base en Somede (2011).

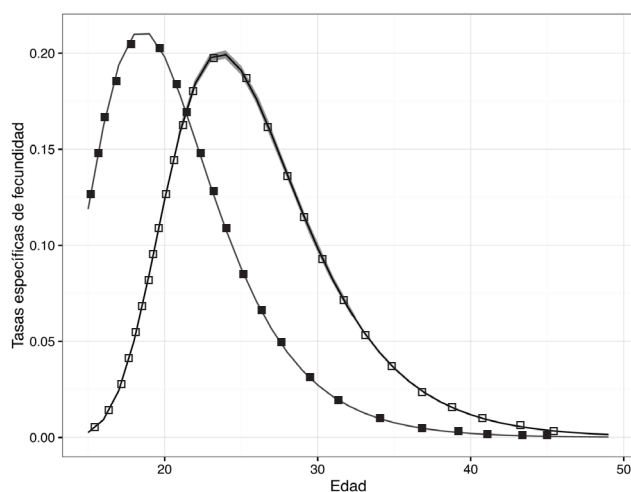


Figura 3. Comparativo entre tasas específicas de fecundidad pronosticadas al año 2014 (-) con intervalos de predicción del 95% (+), respecto a las obtenidas en la Encuesta Nacional de la Dinámica Demográfica 2014.
Fuente: Cálculos propios con base en Somede (2011) y INEGI (2015).

En efecto, como se observa en la figura 3 la distribución por edades reportada en la Enadid se encuentra sesgada a la izquierda. Ello se debe principalmente a que en el ejercicio de conciliación demográfica (Somede, 2011), se privilegió la información de las estadísticas vitales para determinar el calendario de la fecundidad. Es decir, si la diferencia en la estructura por edades es errónea en el pronóstico se debe principalmente a un error en la estimación base de la conciliación demográfica.

Una última cuestión que hay que tomar en cuenta es que el modelo considera probable un repunte acelerado de la fecundidad, como se puede observar en los intervalos de predicción superiores en la figura 2. Estadísticamente esto es cierto, pero demográficamente, aún no se sabe. Esto abre una ventana de oportunidad para futuras investigaciones al respecto, en donde el juicio de expertos sea incorporado para determinar los escenarios o realizaciones “factibles” que arroja el modelo, a partir de las cuales sea posible calcular intervalos de predicción “razonables”, demográficamente hablando.

Migración internacional

La migración internacional es la variable más incierta de la dinámica demográfica ya que depende en gran medida de factores económicos, que son a su vez muy volátiles. Lo anterior se confirma con la evolución histórica de la migración mexicana. Por un lado, se tiene el inesperado incremento de la migración de mexicanos entre 1999 y 2000 y, por otro, una disminución de dicha migración a partir del año 2006. El primer hecho está relacionado con un periodo de auge económico de la economía estadounidense y con una recesión de la economía mexicana, relacionada con la crisis económica de 1995 y, el segundo, está relacionado con la crisis económica mundial en 2008 que afectó a muchos de los sectores laborales en los que se insertan los migrantes mexicanos (Passel, 2011; Passel, Cohn & González-Barrera, 2016).

Así, en las figuras 4 y 5 se muestra la estimación y pronóstico estocástico de los inmigrantes y emigrantes por sexo. Observe que la incertidumbre representada por los intervalos de predicción para estas variables es muy alta; sin embargo, las medianas se aproximan a los resultados observados últimamente. En efecto, de acuerdo a la información presentada en González-Barrera (2015) la migración neta anualizada del periodo 2009-2014 fue de alrededor de 26 000 personas. La migración neta proyectada arroja un saldo de 21 000 personas para el 2014. Sin embargo, habría que hacer un análisis más detallado sobre la estructura por edades de la migración real u observada con la proyectada.

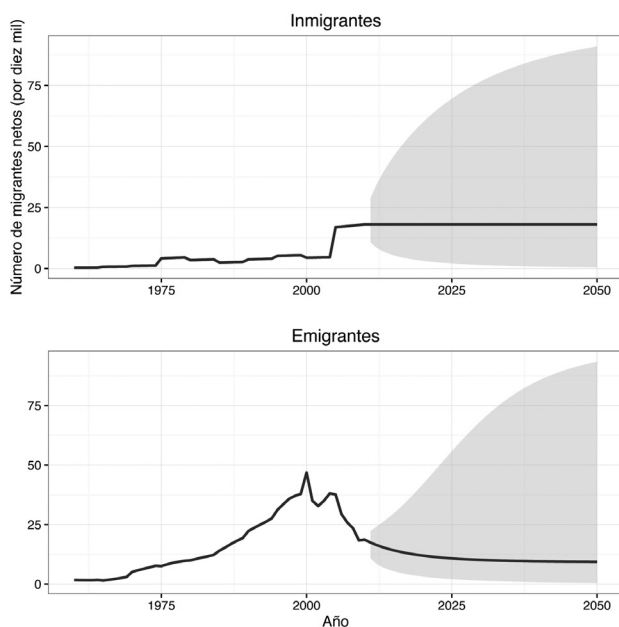


Figura 4. Migración internacional masculina estimada 1950-2010 y pronosticada 2011-2050 (línea negra) con intervalos de predicción del 95% (franja gris). Fuente: Cálculos propios con base en Somedé (2011).

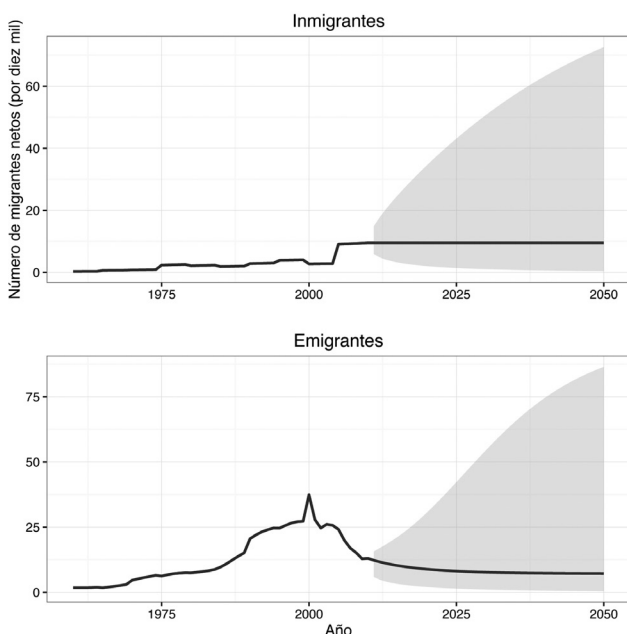


Figura 5. Migración internacional femenina estimada 1950-2010 y pronosticada 2011-2050 (línea negra) con intervalos de predicción del 95% (franja gris). Fuente: Cálculos propios con base en Somedé (2011).

En suma, los resultados sobre la proyección del número de inmigrantes y el de emigrantes no pueden ser muy útiles como insumos puntuales para la elaboración de planes y programas gubernamentales, pero sí para mostrar la gran incertidumbre que rodea a esta variable demográfica.

Población total y estructura por edades

Para proyectar estocásticamente la población total se incorpora la incertidumbre de los modelos que sirven para aproximar las variables demográficas básicas en el método de las componentes. Así, a partir de dicha incorporación se realizan miles de simulaciones de la población por sexo y edades que representan una realización futura de la dinámica demográfica a lo largo de todo el horizonte de proyección si las variables puramente aleatorias tienen un valor determinado. Así, con todas las simulaciones obtenidas se calculan los percentiles deseados (usualmente el del 97.5% y 2.5% para conformar un intervalo del 95% y el de 50% para señalar el valor mediano) que señalan un rango de confiabilidad de que un valor futuro de la proyección se encuentre dentro de él. El método de las componentes es “femenino-dominante” en el sentido de que se proyecta primero la población femenina para que a partir de ella se calculen los nacimientos femeninos y masculinos.

Así, en la figura 6 se muestra la estimación (1960-2010) y proyección (2011-2050) de la población total femenina y masculina con sus respectivos intervalos de predicción. Asimismo, en la figura se muestra el valor observado de la población total en 2015 a partir de la información de la EIC-2015. En general, a partir de la información histórica estimada se espera que ambas poblaciones continúen creciendo, pero después de la década 2030 a un ritmo cada vez menor. Conforme pasa el tiempo la incertidumbre se incrementa considerablemente (tiene una amplitud de cerca de 15 millones de mujeres y 20 millones de hombres para el 2050).

Ahora bien, para el 2015 la Encuesta Intercensal 2015 arroja una población de 61.5 millones de mujeres y de 58 millones de hombres (INEGI, 2016) mientras que el pronóstico aquí presentado indica que dichas poblaciones pueden variar, para ese mismo año, entre 61.4 millones y 62 millones de mujeres y entre 58.4 millones y 60 millones de hombres. En este sentido, se puede decir que la población femenina resultó dentro lo previsto, aunque muy cercana al límite inferior; mientras que la población masculina resultó fuera de lo esperado, aunque no por mucho, la diferencia es de solo 400 000 hombres respecto al límite inferior.

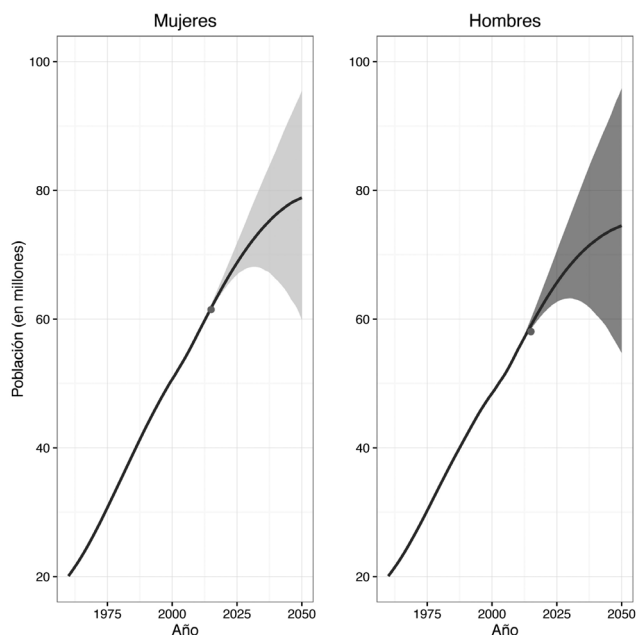


Figura 6. Población total por sexo estimada 1950-2010, pronosticada 2011-2050 (líneas negras) con intervalos de predicción del 95% (franja gris) y estimada al 2015 de acuerdo a la Encuesta Intercensal 2015 (punto gris).
Fuente: Cálculos propios con base en Somede (2011) y INEGI (2016).

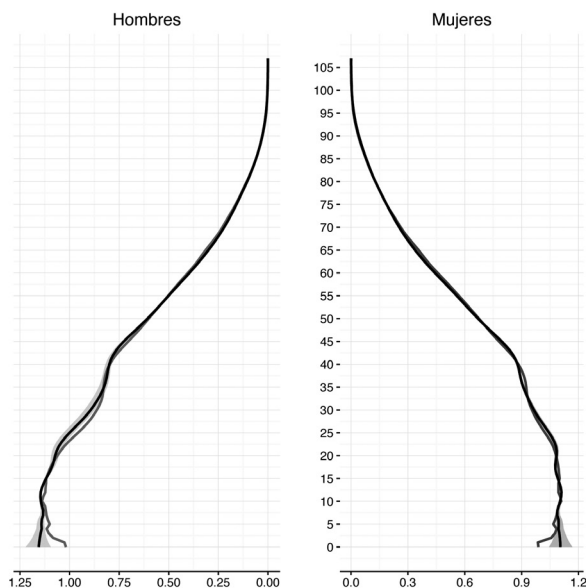


Figura 7. Distribución por edad y sexo para la población estimada 1950-2010, pronosticada 2011-2050 (curvas negras) con intervalos de predicción del 95% (franja gris claro) y estimada al 2015 de acuerdo a la Encuesta Intercensal 2015 (curva gris). En millones.
Fuente: Cálculos propios con base en Somede (2011) y INEGI (2016).

Para analizar esa discrepancia es necesario comparar la estructura por edades de la población. Como se muestra en la figura 7, las principales diferencias en la distribución etárea se encuentran en la población menor a 5 años y la que se encuentra entre los 15 y 40 años. Esa última diferencia es relativamente mayor en la población masculina. Por un lado, la migración proyectada para ese año es relativamente cercana a la que ha sido estimada y la fecundidad, si bien tiene un problema de calendario, no afecta esos grupos por edades. Es importante señalar que parte de la diferencia en el volumen poblacional se puede deber al subregistro de niños menores de 5 años que usualmente ocurre; pero es de mayor magnitud en los grupos jóvenes. En este sentido, se puede especular que hay algún evento que afecta la mortalidad masculina que está fuera del alcance de la información histórica, es por ello que sale de los límites de predicción.

Entre el año 2000 y 2010 ha existido un estancamiento en la esperanza de vida al nacimiento ocasionado principalmente por un importante incremento en el número de homicidios de hombres entre 15 y 40 años así como por la diabetes en los grupos de mayor edad (Aburto *et al.*, 2016; Canudas-Romo *et al.*, 2014; 2016). Es probable que ese problema se haya agudizado en los últimos años, aunque, para corroborarlo, habría que calcular tablas de mortalidad a partir de la información de la EIC-2015 así como corroborar si hay algún faltante poblacional a partir del cálculo de una conciliación demográfica más actualizada.

Razón de dependencia

Lo anterior conlleva a que la razón de dependencia total (figura 8) alcance un valor mínimo a mediados de la década de los 2030s. Lo que significa que, de no suscitarse algún cambio abrupto en la dinámica demográfica en el futuro, aún quedan cerca de 15 años para generar las capacidades necesarias que permitan aprovechar o capitalizar el bono demográfico. La “nueva demografía económica” señala que “[l]as oportunidades que ofrece esa peculiar y transitoria estructura etaria se relacionan [...] con los aspectos productivos y de consumo, pero el alcance de esas oportunidades se [pueden extender] a aspectos de ahorro financiero” (Alba, 2009).

Así, quedan dos retos y oportunidades por atenderse. Por un lado, se encuentra la oportunidad económica que genera el creciente volumen de población en edades laborales; esto más que ser un reto hacia el futuro es una oportunidad que debe ser aprovechada creando las condiciones necesarias y suficientes para

capitalizar esa capacidad productiva. En este sentido, la reducción de la población en edades infantiles provee de la oportunidad de mejorar la calidad de los servicios educativos y generar un círculo virtuoso que permita capitalizar mejor su potencial cuando lleguen a las edades laborales (García Guerrero, 2014b, 2014c). Como señala Alba, el bono o primer dividendo demográfico no es un resultado automático ni inmediato, ya que “la dimensión demográfica no puede ser tratada en forma aislada de las otras dimensiones de la realidad social, del contexto del que dicha dimensión forma parte, como la estabilidad política, las estrategias económicas, las condiciones de los mercados laborales (empleo, desempleo, subempleo), la calificación de los recursos humanos, los niveles salariales los equilibrios y desigualdades sociales y regionales entre otros” (Alba, 2009).

Reflexiones finales

En este trabajo se explicaron y aplicaron los métodos que han sido desarrollados hasta el momento para pronosticar la población mexicana y sus componentes demográficas básicas. Se compararon los resultados de dicho ejercicio prospectivo con la información observada o estimada de aquellas variables de las que ya se dispone de información. Todas las variables demográficas y la población total por edades y sexo son proyectadas razonablemente bien en el corto plazo. Sin embargo, se ubicaron algunas áreas que pueden coadyuvar a la mejora del método de pronóstico aquí propuesto.

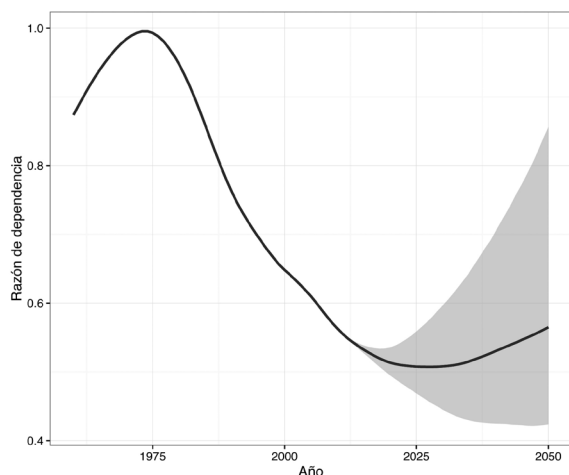


Figura 8. Razón de dependencia total estimada 1950-2010, pronosticada 2011-2050 (curva negra) con intervalos de predicción del 95% (franja gris).
Fuente: Cálculos propios con base en Somede (2011).

La primera de ellas se refiere a que dicho método depende totalmente de la información histórica de cada componente demográfico así como de la población total. Es decir, se capta toda la información que se tiene sobre el comportamiento del nivel, tendencia y estructura de cada componente demográfica. Dicho comportamiento tiene una componente determinista y otra aleatoria. Al realizar el pronóstico se extrapola dicha experiencia, o sea, que sale de la perspectiva del modelo el impacto que tenga algún evento fuera de la experiencia histórica. Esto queda evidente con la diferencia de población masculina que se observa hacia el 2015 respecto a la información captada por la EIC-2015.

Por otra parte, se encontró que la migración internacional fue pronosticada razonablemente y refleja las tendencias que han sido observadas en los últimos años. Sin embargo, queda pendiente el controlar la amplitud de los intervalos de predicción. Estos señalan en gran medida que aún desconocemos muchos aspectos acerca de la naturaleza de la migración, o bien, que hay una gran heterogeneidad dentro de su medición. Una posibilidad es hacer un ejercicio de pronóstico de este fenómeno diferenciando entre migrantes temporales y permanentes, por sexo y edad. Esto en sí mismo ya es un reto en cuanto la captación de la información básica.

Por último, se analizó el ejercicio prospectivo de la fecundidad. En dicho análisis se observó que en términos generales el nivel y tendencia de la fecundidad fue pronosticado adecuadamente. Sin embargo, la estructura por edades es muy diferente de lo que se esperaba. La información de la Enadid-2014 muestra un inesperado rejuvenecimiento de la fecundidad. En este sentido, un primer ejercicio sería la evaluación de la encuesta. Además, como se mencionó, la conciliación demográfica (que es la base de la proyección) privilegia la información de las estadísticas vitales para determinar el calendario de la fecundidad. Habría que comparar la información obtenida en el certificado de nacimientos. Adicionalmente, se señaló que el modelo considera probable un repunte acelerado de la fecundidad. Esto abre una ventana de oportunidad para futuras investigaciones en donde el juicio de expertos sea incorporado para determinar los escenarios o realizaciones “factibles”.

Las proyecciones de corte estocástico que se presentan en este trabajo no serían posibles sin la calidad que ofrece la conciliación demográfica como la que realizó Somede (2011) y que parte del mismo aspecto metodológico sobre el cual se han construido conciliaciones previas que dieron origen a las proyecciones

oficiales de población de la década de los 90 y principios de la década del 2000. Un avance importante al respecto sería incorporar la perspectiva estocástica al proceso de conciliación. Ello permitiría reconocer la incertidumbre que existe en la dinámica demográfica del pasado y daría pie a mejoras técnicas para el cálculo de la dinámica demográfica del futuro.

El pronóstico demográfico es una herramienta que permite comprender de una manera integral la dinámica demográfica de las poblaciones, tanto de manera aislada como conjunta. En los últimos 30 años se ha avanzado en el desarrollo de técnicas y métodos que permitan informar lo que conocemos y, sobre todo, lo que desconocemos de la dinámica demográfica; la perspectiva estocástica lo ha permitido. México tiene una invaluable fuente de información pública y la capacidad técnica para desarrollar y aplicar métodos con dicha perspectiva. Ahora una de las tareas es poder traducir esa información en conocimiento útil y sintético para una eficiente toma de decisiones.

REFERENCIAS

- Aburto, J. M., Beltrán-Sánchez, H., García-Guerrero V. M. & Canudas-Romo, V. (2016). Homicides In Mexico Reversed Life Expectancy Gains For Men And Slowed Them For Women, 2000-10. *Health Affairs*, 35(1), 88-95. doi: 10.1377/hlthaff.2015.0068
- Aburto, J. M., & García Guerrero, V. M. (2015). El modelo aditivo doble multiplicativo. Una aplicación a la mortalidad mexicana. *Papeles de Población, CIEAP-UAEMEX, México*, 21(84), 9-44.
- Alba, F. (2009). El uso político de los "dividendos demográficos". *Este País. Tendencias y Opiniones*, 218, 4-10.
- Alho, J., & Spencer, B. (2005). *Statistical demography and forecasting*. Springer Series in Statistics, Nueva York: Springer.
- Canudas-Romo, V., Aburto J. M., García-Guerrero, V. M., & Beltrán-Sánchez, H. (2017). Mexico's epidemic of violence and its public health significance on average length of life. *Journal of Epidemiology Community Health*, 71(2), 188-193.
- Canudas-Romo, V., García-Guerrero, V. M., & Echarri-Canovas, C. J. (2014). The stagnation of the Mexican male life expectancy in the first decade of the 21st century: the impact of homicides and diabetes mellitus. *Journal of Epidemiology & Community Health*, 69(1), 28-34, doi:10.1136/jech-2014-204237
- Consejo Nacional de Población (Conapo). (2014a). *Proyecciones de la población de México. 2010-2050*. Documento Metodológico, Consejo Nacional de Población, México. Recuperado el 26 de enero del 2016 de <http://www.Conapo.gob.mx/es/Conapo/Proyecciones>
- Consejo Nacional de Población (Conapo). (2014b). *Programa Nacional de Población 2014-2018*. Consejo Nacional de Población, México. Recuperado el 26 de enero del 2016 de http://www.Conapo.gob.mx/es/Conapo/Documento_Completo_PNP_2014_2018
- García Guerrero, V. M. (2011). Un análisis de las diferencias entre las proyecciones de población 2006-2050 y el censo de población 2010. *Coyuntura Demográfica. Revista sobre los procesos demográficos en México hoy*, Sociedad Mexicana de Demografía, 1(1), 62-65.
- García Guerrero, V. M. (2014a). *Proyecciones y Políticas de Población en México*, Centro de Estudios Demográficos, Urbanos y Ambientales, México: El Colegio de México.
- García Guerrero, V. M. (2014b). El futuro demográfico de México y las proyecciones de población. Uso de los pronósticos probabilísticos en la planeación nacional. En S. Giorguli y V. Ugalde (coords.). *Gobierno, Territorio y Población: Las políticas públicas en la mira*. (pp. 71-120). Centro de Estudios Demográficos, Urbanos y Ambientales, México: El Colegio de México.
- García Guerrero, V. M. (2014c). Las proyecciones de la población de México. En C. Rabell (Coord.). *Los mexicanos: Un balance del cambio demográfico*, (pp. 857-900). México: Fondo de Cultura Económica.
- García Guerrero, V. M. (2014d). Las estimaciones y proyecciones demográficas como herramienta para el diseño de políticas de población. En L. M. Valdés (Coord.). *Hacia una nueva Ley General de Población en México*. (pp. 259-278). Instituto de Investigaciones Jurídicas, México: UNAM.
- García Guerrero, V. M. (2014e). Pronóstico estocástico de la población mexicana utilizando modelos de datos funcionales. Realidad, Datos y Espacio. *Revista Internacional de Estadística y Geografía*. Instituto Nacional de Estadística y Geografía (INEGI), México, 5(3), 18-33.
- García Guerrero, V. M. (2014f). Análisis de los supuestos sobre la migración internacional en las proyecciones de población de México 2001-2050 y 2006-2050. *Estudios Demográficos y Urbanos*, Centro de Estudios Demográficos, Urbanos y Ambientales, El Colegio de México, 29(2), 229-256.
- García-Guerrero, V. M. (2016). A Probabilistic Method to Forecast the International Migration of Mexico by Age and Sex. *Papeles de Población*, 22(88), 113-140.
- García Guerrero, V. M., & Ordorica, M. (2012). Proyección estocástica de la mortalidad mexicana por medio del método de Lee-Carter. *Estudios Demográficos y Urbanos*, Centro de Estudios Demográficos, Urbanos y Ambientales, El Colegio de México, 27(2), 409-448.
- Gómez de León, J., & Partida, V. (2001). Proyecciones de la población para el nuevo siglo: el proceso de envejecimiento de la población mexicana. En J. Gómez de León y C. Rabell (Coords.). *La Población de México. Tendencias y Perspectivas Sociodemográficas hacia el Siglo XXI*. (pp. 991-1022). México: Conapo, Fondo de Cultura Económica.
- González-Barrera, A. (2015). More Mexicans Leaving than coming to the U.S. Net loss of 140 000 from 2009 to 2014; Family reunification top reasons for return. Washington D.C. USA: Pew Research Center.
- Hyndman, R. J. (2016). Forecast: Forecasting functions for time series and linear models. R package version 7.1. Recuperado el 26 de enero del 2016 de <http://github.com/robjhyndman/forecast>.
- Hyndman, R. J., & Booth, H. (2008). Stochastic population forecasts using functional data models for mortality, fertility and migration. *International Journal of Forecasting*, 24(3), 323-342.
- Hyndman, R. J., & Fan, Y. (1996). Sample quantiles in statistical packages. *American Statistician*, 50, 361-365.

- Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3), 1-22.
- Instituto Nacional de Estadística y Geografía, (INEGI). (2015). Encuesta Nacional de la Dinámica Demográfica 2014, recuperado el 15 de enero de 2016 en www.inegi.org.mx
- Instituto Nacional de Estadística y Geografía (INEGI). (2016). Encuesta Intercensal 2015, INEGI, México. Recuperado el 15 de enero de 2016 de <http://www.inegi.gob.mx>
- Lee, R. (1993). Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level, *International Journal of Forecasting*, 9, 187-202.
- Lee, R. (1998). Probabilistic Approaches to Population Forecasting en Lutz, Vaupel y Ahlburg (edits.). *Frontiers of Population Forecasting*, suplemento de *Population and Development Review*, 24, Nueva York, E.U.A.: Population Council.
- Lee, R., & Carter, L. (1992). Modeling and Forecasting U.S. Mortality, Supplement: Aging, Health, and Public Policy, *Journal of the American Statistical Association*, 87(419), 659-671.
- Lee, R., & Tuljapurkar, S. (1994). Stochastic population forecast for the United States: Beyond high, medium and low, *Journal of the American Statistical Association*, 89(428), 1175-1189.
- Morales-Eslava, A. J., Silva-Urrutia, E., & Ordorica-Mellado, M. (2016). Pronósticos de la fecundidad en México: una aplicación de modelos multivariados de series de tiempo. *Papeles de Población*, 22(89), 99-131.
- Passel, J. (2011). Mexico-U.S. Migration Flows 1990-2010: Preliminary Assessment Based on U.S. Sources, Coyuntura Demográfica. *Revista sobre los procesos demográficos en México hoy*. Sociedad Mexicana de Demografía, 1(1), 62-65.
- Passel, J., Cohn, D., & González-Barrera, A. (2012). Net Migration from Mexico Falls to Zero – and Perhaps Less. Pew Hispanic Center, Washington D.C., E.U.A. Recuperado el 26 de enero del 2016 de http://www.pewhispanic.org/files/2012/04/Mexican-migrants-report_final.pdf
- R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Recuperado el 26 de enero del 2016 de <https://www.R-project.org/>.
- Raftery, A. E., Alkema, L., & Gerland, P. (2014). Bayesian Population Projections for the United Nations. *Statistical Science. A Review Journal of the Institute of Mathematical Statistics*, 29(1), 58-68. <http://doi.org/10.1214/13-STS419>
- Silva, E., Guerrero, V. M., & Peña, D. (2011). Temporal disaggregation and restricted forecasting of multiple population time series, *Journal of Applied Statistics*, 38, 799-815.
- Silva, E., & Ordorica, M. (2013). Pronósticos multivariados de poblaciones con series de tiempo: el caso de la ZMCM contrastado con los datos del Censo 2010. *Estudios Demográficos y Urbanos*, 28, 167-188.
- Sociedad Mexicana de Demografía A.C. (Somede). (2011). *Conciliación demográfica de México y entidades federativas 1990-2010*. Inédito.