

Máster Universitario en Ciencias Actuariales y Financieras
Curso académico 2020-2021

Trabajo Fin de Máster

“Recalibración del shock de longevidad en Solvencia II: Índice Europeo de Longevidad”

Victoria Muñoz Bartolomé

Tutores:

José Miguel Rodríguez-Pardo del Castillo
Jesús Ramón Simón del Potro

Madrid, junio de 2021



Esta obra se encuentra sujeta a la licencia Creative Commons
Reconocimiento - No Comercial - Sin Obra Derivada

RESUMEN

Gracias a las mejoras de las condiciones higiénicas, sanitarias, nutricionales y sociales, la esperanza de vida tiene una tendencia alcista desde hace mucho y parece que continuará en los próximos años. Situaciones catastróficas puntuales pueden alterar a corto plazo esta serie temporal. Las nuevas generaciones disfrutan de todos los avances acumulados por la humanidad, y cabe esperar que un mayor porcentaje alcance edades avanzadas.

Esto afecta de forma directa a las compañías aseguradoras, que aplican en sus cálculos distintos "*shocks*" o impactos, analizan mediante test de *stress* los distintos escenarios a los que se pueden tener que enfrentar según circunstancias poco favorables y, en consecuencia, reservan provisiones en cuantías suficientes para garantizar su solvencia.

En el presente trabajo se estudia el riesgo de longevidad y cómo se aplica. Solvencia II utiliza un porcentaje fijo. Sin embargo, diversos autores han planteado que eso podría no ser lo más conveniente. Las tablas de mortalidad no recogen actualmente el ajuste generacional que parece necesario aplicar sobre todo a los más jóvenes.

Para ello se han utilizado los modelos de Lee-Carter, Cairns-Blake-Dowd y P-splines. Como resultado obtenemos unos factores correctivos que tienen en cuenta la edad del asegurado y la duración del seguro. Este Índice Europeo de Longevidad se puede utilizar como instrumento de medida de la tendencia de la longevidad y la transferencia del riesgo.

Palabras clave: ciencias actuariales, longevidad, mortalidad, impacto, shock, Solvencia II, EIOPA, tendencia, Lee Carter, Cairns-Blake-Dowd, P-splines, Índice Europeo de Longevidad.

ABSTRACT

Due to improvements in hygienic, sanitary, nutritional and social conditions, life expectancy has been on an upward trend for a long time and will keep on in next years. Sporadic eventual catastrophic events can shift this trend in the short term. New generations enjoy all the advances achieved by humanity, and it seems most of them will reach old ages.

This directly affects insurance companies. They apply a variety of shocks in their calculations, use stress tests to different scenarios forecasting unfavorable circumstances and, accordingly, reserve enough provisions to guarantee solvency.

This work deals about longevity risk and how it should be applied. Solvency II sets a fixed percentage. However, several authors have argued that this might not be the most convenient way to proceed. Mortality tables currently do not include the generational adjustment needed, especially when younger people is involved.

With this goal in mind, Lee-Carter, Cairns-Blake-Dowd and P-splines models have been used. The outcome we have obtained are corrective factors that take into account the age of the insured and the duration of the insurance. This European Longevity Index can be used as an instrument for measuring longevity trend and to assess risk transfer.

Palabras clave: actuarial science, longevity, mortality, shock, Solvency II, EIOPA, trend, Lee Carter, Cairns-Blake-Dowd, P-splines, European Longevity Index.

DEDICATORIA

A mi familia, por su apoyo incondicional.

Siempre.

ÍNDICE GENERAL

1. INTRODUCCIÓN.	1
2. EL RIESGO DE LONGEVIDAD	4
2.1. Concepto	4
2.2. Análisis de la esperanza de vida	5
2.3. Rectificación	11
3. MODELIZACIÓN	14
3.1. Introducción: Las Tablas de Mortalidad	14
3.2. Modelos a utilizar	16
3.2.1. Lee-Carter	18
3.2.2. Cairns-Blake-Dowd	19
3.2.3. P-splines	20
3.3. Backtesting	22
4. ANÁLISIS DE LA MORTALIDAD.	23
4.1. Obtención de los datos.	23
4.2. La longevidad en España	26
4.3. La longevidad en Europa	34
5. PROPUESTA DE SHOCK DE LONGEVIDAD	39
5.1. Solvencia II y el <i>shock</i> de longevidad	39
5.2. Aplicación de los modelos	40
5.2.1. Lee-Carter	42
5.2.2. CBD	46
5.2.3. P-splines	48
5.3. Resultados	50
5.4. Índice Europeo de Longevidad	56
6. CONCLUSIONES	65
BIBLIOGRAFÍA	67

ÍNDICE DE FIGURAS

1.1	Pilares de Solvencia II	2
1.2	Estructura general de la fórmula estándar	2
2.1	Tasa de fallecimiento en España desde 1908	7
2.2	Impacto de la Segunda Guerra Mundial	8
2.3	Tasa de fallecimiento en España finales S XX	8
2.4	Tasa de fallecimiento Italia y Suiza a finales S XX	9
2.5	Tasa de fallecimiento Bélgica a finales S XX	9
2.6	Histórico tasa de fallecimiento en neonatos	10
2.7	Tasa de fallecimiento en neonatos Europa desde 1990	10
2.8	Rectangularización en España	12
4.1	Matriz de datos	24
4.2	Número de fallecidos en España desde 1908	26
4.3	Número de fallecidos en España desde 1978	28
4.4	Número de personas expuestas en España desde 1978	29
4.5	Tasa de fallecimiento en España desde 1978	30
4.6	Esperanza de vida al nacer en España	32
4.7	Decesos en España	33
4.8	Esperanza de vida al nacer en Europa	34
4.9	Tasa de fallecimiento a los 55 años en Europa	35
4.10	Pirámide poblacional europea en 2017	36
4.11	Fallecimientos en Europa en 2017	37
4.12	Decesos en Europa	38
5.1	Estructura del paquete “StMoMo”	42
5.2	Parámetros Lee-Carter	43
5.3	Forecast Lee-Carter	44
5.4	Residuos Lee-Carter	44

5.5	Simulación Lee-Carter	45
5.6	Parámetros CBD	46
5.7	Forecast CBD	47
5.8	Residuos CBD	47
5.9	Simulación CBD	48
5.10	Suavizado P-spines	49
5.11	Proyección P-spines	49
5.12	Residuos P-spines	50
5.13	Comparación de los 3 modelos	51
5.14	Proyecciones de los modelos	52
5.15	Errores relativos	53
5.16	Mejora de los modelos	55
5.17	Evolución de la tasa de mortalidad	60
5.18	Sección de la tasa de mortalidad	60
5.19	Índice Europeo de Longevidad	63

ÍNDICE DE TABLAS

4.1	Esperanza de vida al nacer en España	32
5.1	Mejora de la mortalidad	58
5.2	Mejora de la mortalidad, media	59
5.3	Índice Europeo de Longevidad	62

1. INTRODUCCIÓN

En la Unión Europea encontramos el Sistema Europeo de Supervisión Financiera (ESFS), compuesto por tres órganos o autoridades independientes cuya función es garantizar una uniformidad y una supervisión eficiente de los principales sectores financieros dentro de la Unión. Estos órganos son los siguientes:

- **Autoridad Bancaria Europea (EBA):** se encarga de que la regulación internacional del sistema bancario resulte eficiente y coherente, a la vez que supervisa la correcta ejecución de ello de forma prudente en todo el sector bancario de la unión.
- **Autoridad Europea de Valores y Mercados (ESMA):** su función asegurar la estabilidad en el sistema financiero, regularizando los mercados para que sean estables y ordenados, lo cual aporta protección a los inversores.
- **Autoridad Europea en Seguros y Pensiones de Jubilación (EIOPA):** su principal objetivo es la protección de los intereses públicos, procurando garantizar tanto estabilidad como efectividad al sistema financiero a nivel económico, empresarial e individual.

Centrándonos en EIOPA, encontramos sus principales objetivos en su página web [1]:

1. Mejorar la protección de los clientes y reconstruir la confianza de éstos en el sistema financiero.
2. Asegurar una regulación y una supervisión fuertes, efectivas y consistentes, teniendo en cuenta el interés de todos los estados miembros y las diferentes naturalezas de las instituciones financieras.
3. Conseguir una aplicación armonizada y coherente de las normas para las instituciones y los mercados financieros en toda la Unión Europea.
4. Favorecer la supervisión de los grupos transfronterizos.
5. Promover respuestas supervisoras coordinadas de la Unión Europea.

Para ello, desde EIOPA crearon un sistema normativo con el objetivo de ser utilizado por todas las aseguradoras europeas, llamado Solvencia II. Este reglamento entró en vigor el 1 de enero de 2016. Con ello incrementaron el control de los riesgos principales a los que se exponen las aseguradoras, como pueden ser los riesgos de mercado, de vida, operacional... Establece requerimientos tanto cuantitativos como cualitativos y de transparencia de mercado, organizados por niveles o “pilares”, que son los que muestra la figura 1.1 :

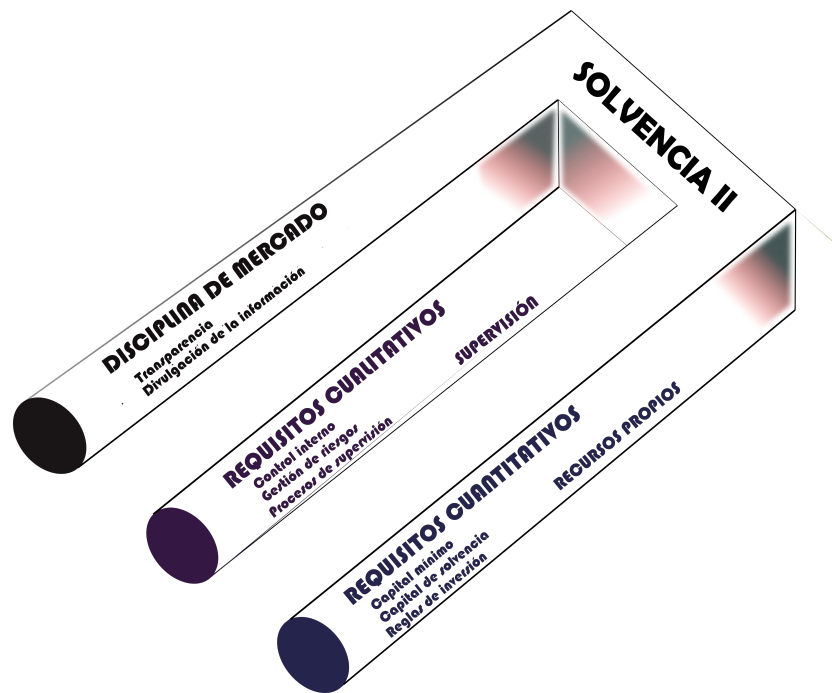


Figura 1.1. Pilares de Solvencia II. Fuente: Elaboración propia

Además, también fija el capital de solvencia obligatorio (SCR), que son los fondos propios que las aseguradoras deben tener en su balance, es decir, regula los recursos que debe tener la entidad para poder hacer frente a sus obligaciones en caso de ser necesario.

A la hora de calcular el SCR básico, encontramos los diferentes riesgos clasificados en cinco submódulos: suscripción distinto del de vida, suscripción de vida, suscripción de enfermedad, mercado e incumplimiento de la contra-parte, según la figura 1.2.

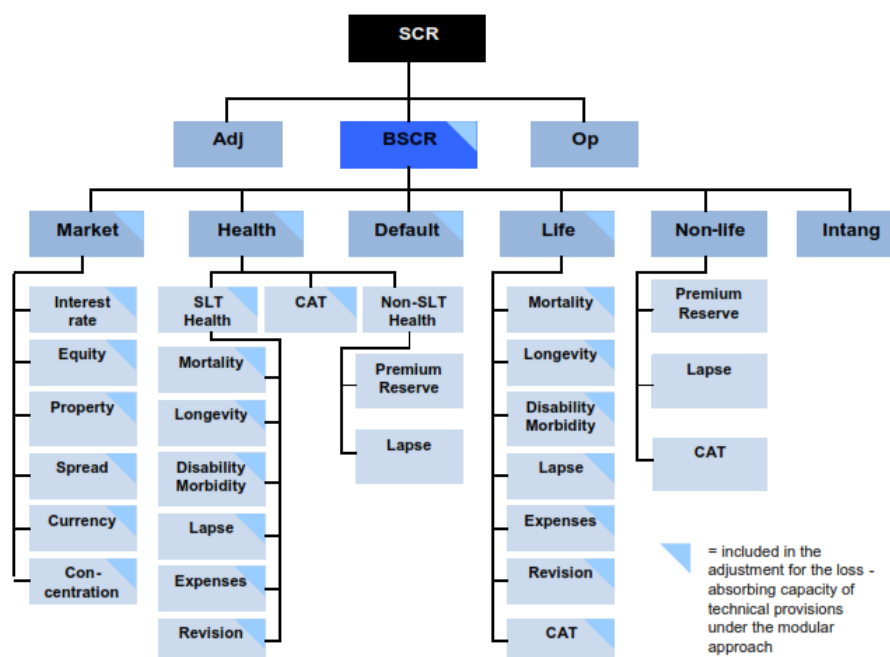


Figura 1.2. Estructura general de la fórmula estándar. Fuente: EIOPA

Centrándonos en el riesgo de subscripción de vida, encontramos que está a su vez formado por siete submódulos, que son [2]:

- **Riesgo de mortalidad:** cuando un incremento en las tasas de mortalidad ocasiona un aumento de las obligaciones de la compañía aseguradora.
- **Riesgo de longevidad:** caso contrario, un descenso en las tasas de mortalidad ocasiona un aumento de las obligaciones.
- **Riesgo de discapacidad y morbilidad:** cuando una variación de las tasas de invalidez, enfermedad o morbilidad provocan un incremento de las obligaciones.
- **Riesgo de gastos del seguro:** cuando la variación se produce en los gastos de ejecución de un seguro, con la consecuente pérdida o el incremento de valor de las obligaciones pertinentes.
- **Riesgo de revisión:** cuando por cambio legislativo o por deterioro en la salud del asegurado ocasionan una variación en las tasas de variación aplicables a las rentas concedidas.
- **Riesgo de reducción o caducidad:** cuando variaciones producidas en la volatilidad de las tasas de discontinuidad, cancelación, renovación o rescate ocasionan un incremento de las obligaciones.
- **Riesgo de catástrofe:** cuando la incertidumbre, debida a sucesos extraordinarios, a la hora de tarificar o constituir provisiones lleva a un incremento de las obligaciones.

En el presente trabajo nos centramos en los shocks relativos a los riesgos de mortalidad y longevidad, principalmente debido a la relación intrínseca y natural que encontramos entre ellos. El estudio realizado en las próximas páginas pretende ser una extensión del trabajo escrito sobre la adecuación de estos shocks establecidos por EIOPA, el cual podemos encontrar en la referencia [3]. De igual modo, los resultados obtenidos para estos impactos serán diferentes, considerándose adecuado el shock aplicado al riesgo de mortalidad y proponiéndose uno alternativo en el caso del riesgo de longevidad.

EIOPA establece el shock de mortalidad fijo de un 15 %, el cual, como comentaremos más adelante, consideraremos prudente y adecuado. Por otro lado, establece el shock de longevidad fijo de un 20 %. Numerosos estudios actuariales nos muestran que, debido al riesgo de tendencia, este impacto debería variar dependiendo de la edad del asegurado (ver por ejemplo [4], [5]).

2. EL RIESGO DE LONGEVIDAD

2.1. Concepto

En el último siglo, gracias a los avances científicos y tecnológicos que van llegando a toda la población, la esperanza de vida ha aumentado notoriamente. La llegada del llamado “estado de bienestar”, el decremento en fallecimientos neonatales e infantiles, el menor efecto de enfermedades agudas (aunque este último se ha visto sustituido por el aumento de enfermedades crónicas) y el cambio en los hábitos alimenticios son algunas de las causas más importantes que han conducido a las mejoras en la mortalidad.

Según el Instituto Nacional de Estadística (INE), la esperanza de vida se define como la media de años que una persona puede esperar vivir en caso de que las tasas de mortalidad observadas por edad no varíen [6]. La muerte es un evento seguro, pero el momento exacto en el que ocurrirá es una incógnita, por lo que utilizaremos probabilidades para estudiar este fenómeno.

Según la Organización Mundial de la Salud, la esperanza de vida a nivel global es, de media, 22 años mayor que a principios del siglo pasado.

En España, a principios del siglo XX la esperanza de vida era aproximadamente de 41 años, mientras que al inicio del presente siglo es de unos 81 años. En 2019, la esperanza de vida al nacer era de 83,483 según el Banco Mundial [7]. Es decir, en los últimos 100 años se ha doblado la edad que, en condiciones normales, alcanzan los ciudadanos. Debido a este incremento y al descenso en la natalidad, la pirámide poblacional española se está invirtiendo, causando que la financiación pública mediante el sistema de pensiones actual pudiera no ser suficiente. Una solución a este fenómeno sería contratar productos con pagos en forma de rentas vitalicias. Por ello, el estudio de la longevidad se ha vuelto tan relevante en los últimos tiempos para las aseguradoras del ramo de vida. Cabe esperar que este riesgo de longevidad se vea desplazado del sector público al privado.

Por otro lado, en la documentación técnica de la normativa de Solvencia II encontramos la definición de “riesgo de longevidad” y “riesgo de mortalidad”. El riesgo de longevidad recoge la probabilidad de experimentar una mortalidad más baja de la esperada mientras que el riesgo de mortalidad representa la probabilidad de que se produzca una mortalidad más alta de la esperada (en ambos casos es independiente de si se debe a una fluctuación arbitraria o sistemática).

Para un mejor análisis del impacto que el riesgo de longevidad puede tener sobre el sector asegurador, lo podemos desglosar principalmente en cuatro subriesgos:

1. **Riesgo de nivel:** se refiere al hecho de que se haya elegido de forma errónea el modelo utilizado, debido a que éste no capture de forma estructurada o adecuada las dinámicas a estudiar. Para contrastar este riesgo, se utilizan criterios de “juicio experto”, escogiendo el modelo que se considere más oportuno de acuerdo a hipótesis de máxima prudencia.
2. **Riesgo de catástrofe:** contempla posibles variaciones en la mortalidad debido a acontecimientos puntuales en un año concreto, como pueda ser una pandemia, una guerra o algunos desastres naturales (terremotos, inundaciones, temporales...). En este caso no se pueden contrarrestar dichos sucesos, simplemente se han de tener en cuenta dichos escenarios a la hora de realizar estudios de solvencia.
3. **Volatilidad:** tiene en cuenta la desviación aleatoria que pueden tener las hipótesis actuariales respecto al riesgo estudiado, establecidas en las bases técnicas demográficas del objeto de estudio. Para contrarrestar este riesgo se puede aumentar el tamaño muestral de exposición, añadir nuevas variables respecto al estilo de vida o contrarrestarlo mediante la contratación de reaseguros.
4. **Tendencia:** hace referencia a la dinámica temporal que puede tener la biométrica humana. Este riesgo elimina el principio de estacionariedad en el tiempo, el cual se ha utilizado en múltiples estudios actuariales. Debido a que los estudios se hacen mediante datos históricos, la tendencia futura a largo plazo es muy difícil de determinar con suficiente fiabilidad. Resulta fácil caer en subestimaciones de los valores esperados de supervivencia que establecen las bases técnicas actuariales originales.

Por último, cabe destacar que es recomendable realizar un estudio a conciencia de la probabilidad de fallecimiento (q_x) y de la de supervivencia (p_x), dado que afectan de forma directa a la hora de medir de manera acertada dicho riesgo. Esto afecta tanto a la tarificación del seguro como a las reservas o provisiones que la aseguradora deberá tener para hacer frente a sus obligaciones futuras.

2.2. Análisis de la esperanza de vida

Una posible forma de enfocar el análisis de la esperanza de vida es mediante la dinámica de sistemas. Ésta es la disciplina matemática que analiza las interacciones entre los múltiples elementos que componen cualquier sistema. El ejemplo más simple e intuitivo de sistema (que es comúnmente utilizado a nivel docente) es la relación entre el número de zorros y el de conejos en un cierto ecosistema. En este caso encontramos que la población de zorros se ve afectada positivamente cuando los conejos proliferan dado que tendrán menor escasez de comida mientras que, en el caso contrario, las causas son negativas, dado que a más depredadores, la población de presas disminuirá.

A este fenómeno se le conoce como realimentación. Matemáticamente se analiza mediante ecuaciones diferenciales, y puede tener tanto efectos de realimentación negativa

(contrarios y reguladores; cuanto más crece una magnitud, menor es el valor de otra) como positiva (cuanto más valor alcanza una variable, mayor es la respuesta obtenida). Un ejemplo de realimentación positiva sería, por ejemplo, cuando un micrófono se acopla a un altavoz y se produce el típico sonido estridente.

Ahora bien, centrándonos en la longevidad humana y la investigación, podemos encontrar tanto realimentaciones positivas como negativas. En primer lugar, a mayor longevidad, mayor probabilidad de que se tenga oportunidad de disfrutar de los avances. Cuantos más avances, mayor longevidad. En segundo lugar, a mayor edad de la población, más recursos se precisarán para sus cuidados, lo que podría llevar a disponer de menos recursos para dedicarlos a la investigación.

Un avance científico puede cambiar para bien las condiciones de la población sin que ello implique un mecanismo infinito de realimentación sobre el agente causante. Un ejemplo de ello serían las mejoras en higiene, que han supuesto una variación significativa en la esperanza de vida en el siglo XX por su extensión a un mayor porcentaje de la población, pero llegado a un punto, los cambios son inapreciables: efectos asintóticos.

Se emplea, por ejemplo, el análisis de cohortes, fraccionando la población en grupos con alguna característica en común. También se utiliza el análisis de los datos, el cual nos permite estudiar lo que ya ocurrió, pero no anticipar lo que va a ocurrir. En estos momentos sería muy interesante poder prever las consecuencias de la pandemia de COVID-19, pero ni siquiera podemos compararla con la provocada por la “gripe española” porque ésta afectó a toda la población y COVID-19 ha afectado principalmente a los más mayores así que no podemos extrapolar lo observado entonces.

Otra disciplina matemática a la que podríamos acudir serían los sistemas lineales invariantes en el tiempo [8], que son aquellos cuya respuesta a dos entradas es igual a la suma de las respuestas de cada una. Además, deben comportarse de la misma forma en cualquier momento del tiempo. En esas circunstancias es posible analizar la respuesta a un impulso o delta de Dirac $\delta(t)$ y utilizar la integral de convolución para caracterizar como el paso por dicho sistema afecta a una determinada función que represente la población, la tasa de mortalidad... Los sistemas sociales ni tienen porqué cumplir las propiedades de linealidad e invarianza temporal, ni se prestan fácilmente a este tipo de análisis.

Sin el rigor matemático que sería de desear, sí hay una similitud entre los efectos que tienen ciertas circunstancias sobre la población y el efecto de un filtro electrónico sobre una señal, como una convolución donde los electrones de la señal fuesen humanos y el equipo electrónico que lo transforma, el suceso. Por ejemplo, si nos centramos en los datos de la población española observamos los siguientes gráficos:

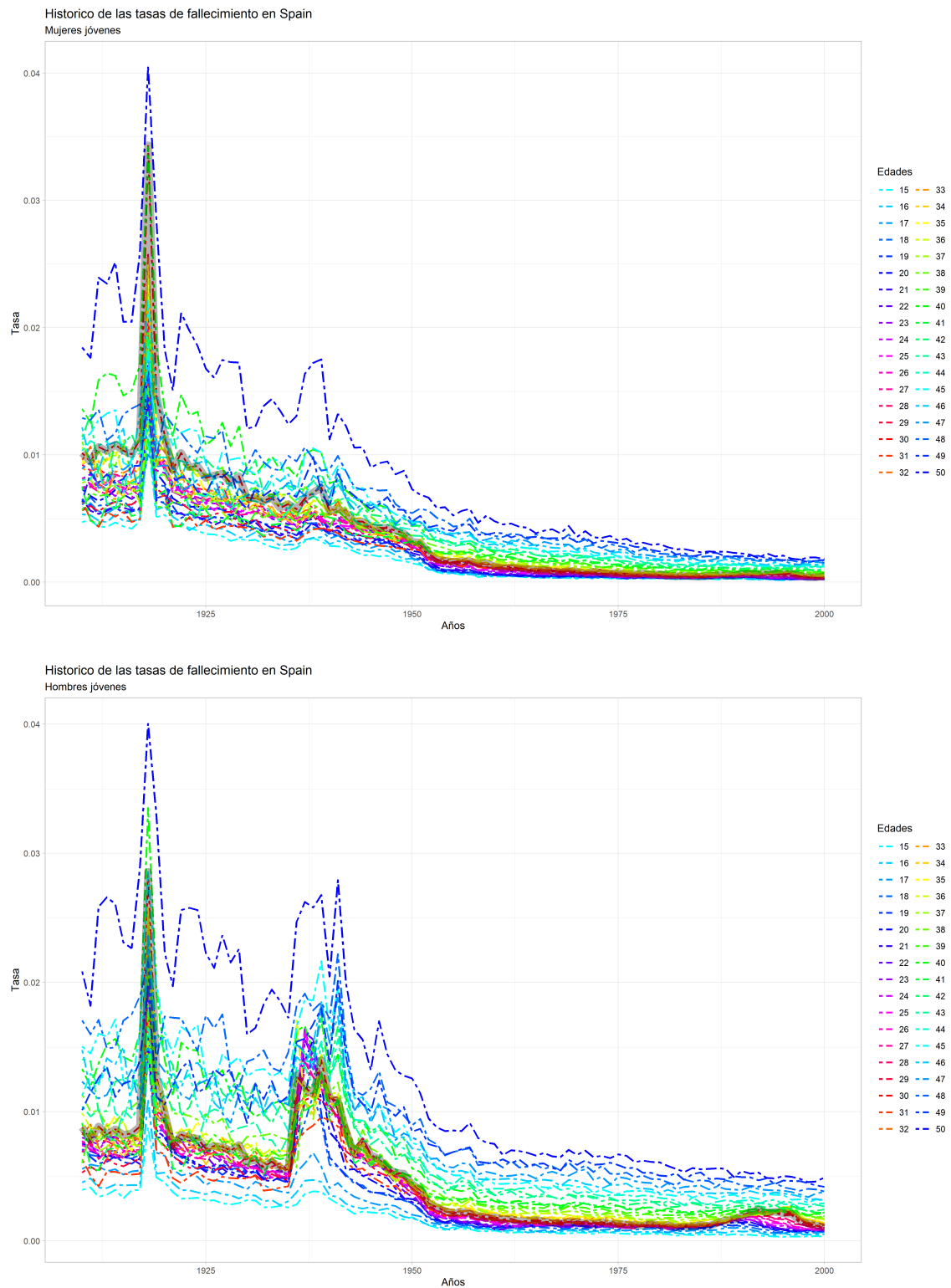


Figura 2.1. Tasa de fallecimiento en España entre 1908 y 2018. Fuente: Elaboración propia

Observamos que en 1918 se produce la gripe española, que afecta por igual a hombres y mujeres de todas las edades. Su influencia en el tiempo reducida: pico de corta duración. Por otro lado en 1936 comienza la guerra civil, que dura 3 años pero tiene un efecto acumulado de más duración (convolución), superior en varones que en mujeres.

Un resultado peculiar que encontramos en este gráfico es que a la gripe española no afectó, o lo hizo a muy poca escala, a los países nórdicos.

Si nos centramos en el periodo de la Segunda Guerra Mundial, se puede ver cómo los países que participaron en dicho suceso se ven mucho más afectados que España en ese periodo:

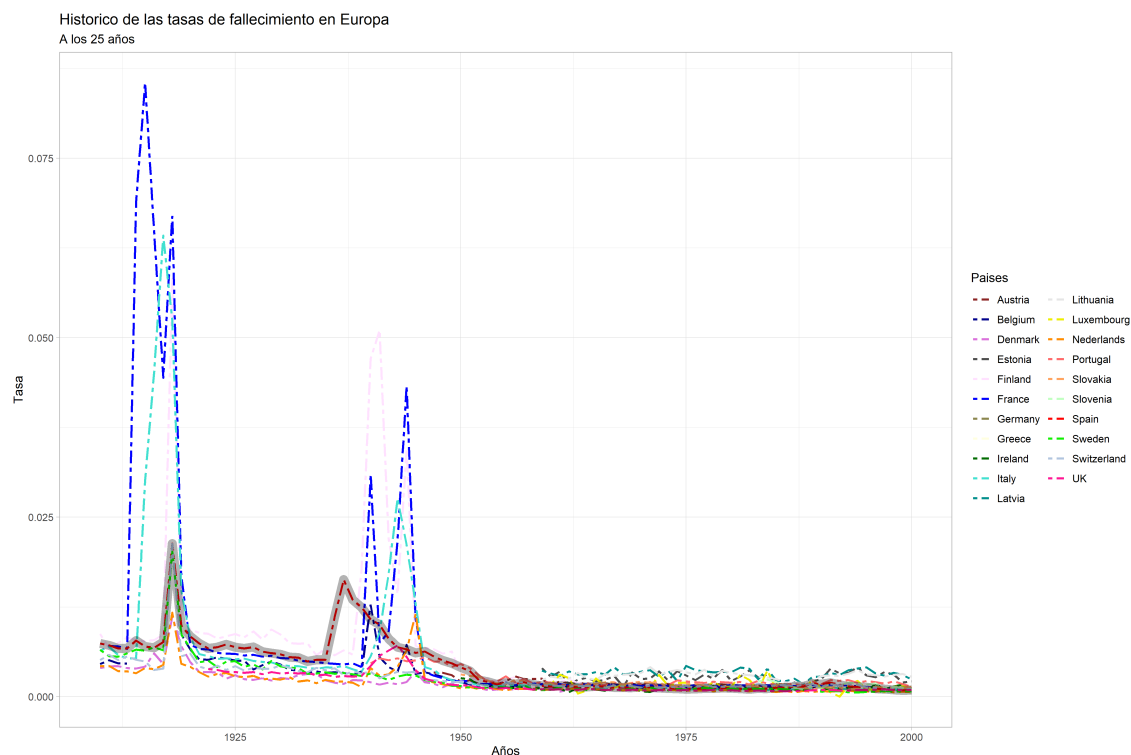


Figura 2.2. Impacto de la Segunda Guerra Mundial. Fuente: Elaboración propia

Este gráfico se ve muy penalizado dado que en la *Human Mortality Database* sólo se dispone de datos de Alemania a partir de 1990, y este país fue uno de los más afectados en Europa. Por otro lado, Francia e Italia, dos de las grandes potencias del momento que participaron en esta Guerra, sí se vieron sumamente afectadas.

Otros fenómenos no son tan evidentes, por lo que hay que ajustar la escala para poder apreciarlos.

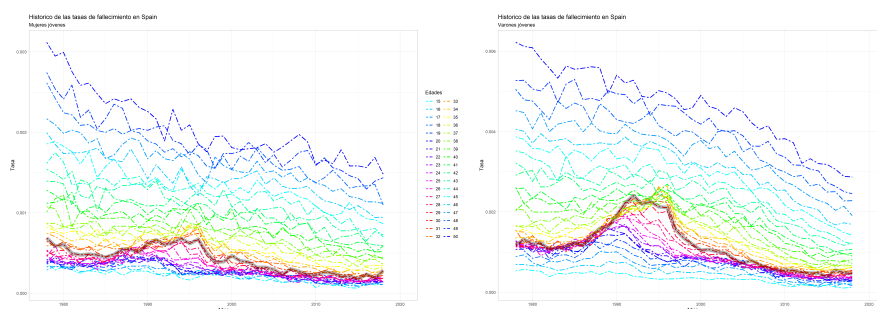


Figura 2.3. Tasa de fallecimiento en España a finales del siglo XX. Fuente: Elaboración propia

En España a finales de los 80 se produjo una subida significativa en la mortalidad de la población joven. Este cambio fue considerablemente más significativo en los varones que en las mujeres. Se puede deber, como trataremos más adelante en profundidad, a la llegada de ciertas drogas y, a partir de los años 90, al SIDA.

Este fenómeno se observa, a distinta escala, en otros países del entorno:



Figura 2.4. Tasa de fallecimiento en Italia y Suiza a finales del siglo XX. Fuente: Elaboración propia

Mientras que en otros países europeos esto no se produce.

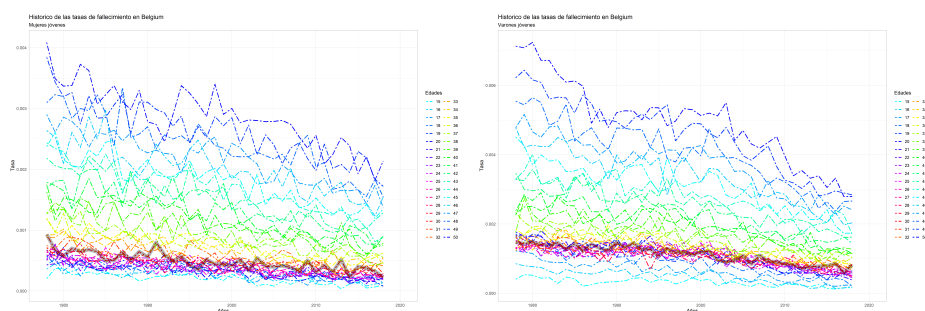


Figura 2.5. Tasa de fallecimiento en Bélgica a finales del siglo XX. Fuente: Elaboración propia

Dentro de los numerosos progresos alcanzados en el siglo pasado, destaca el decremento en la tasa de fallecimiento de niños recién nacidos (menores de un año). Como podemos apreciar a continuación en el gráfico, esta tasa ha llegado a estar casi en 0,3 mientras que a día de hoy toma valores cercanos a 0. En el gráfico siguiente hemos destacado los datos correspondientes a España.

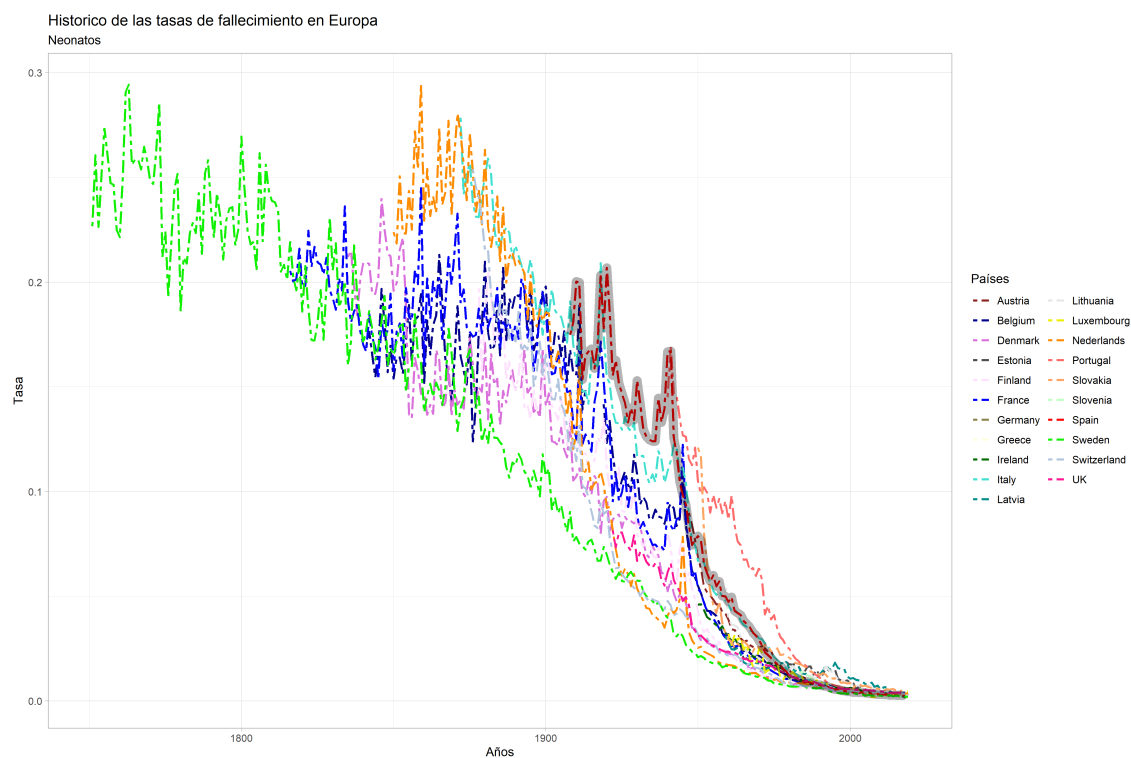


Figura 2.6. Histórico de la tasa de fallecimiento en personas neonatas. Fuente: Elaboración propia

Y, ajustando la escala a Europa desde 1990:

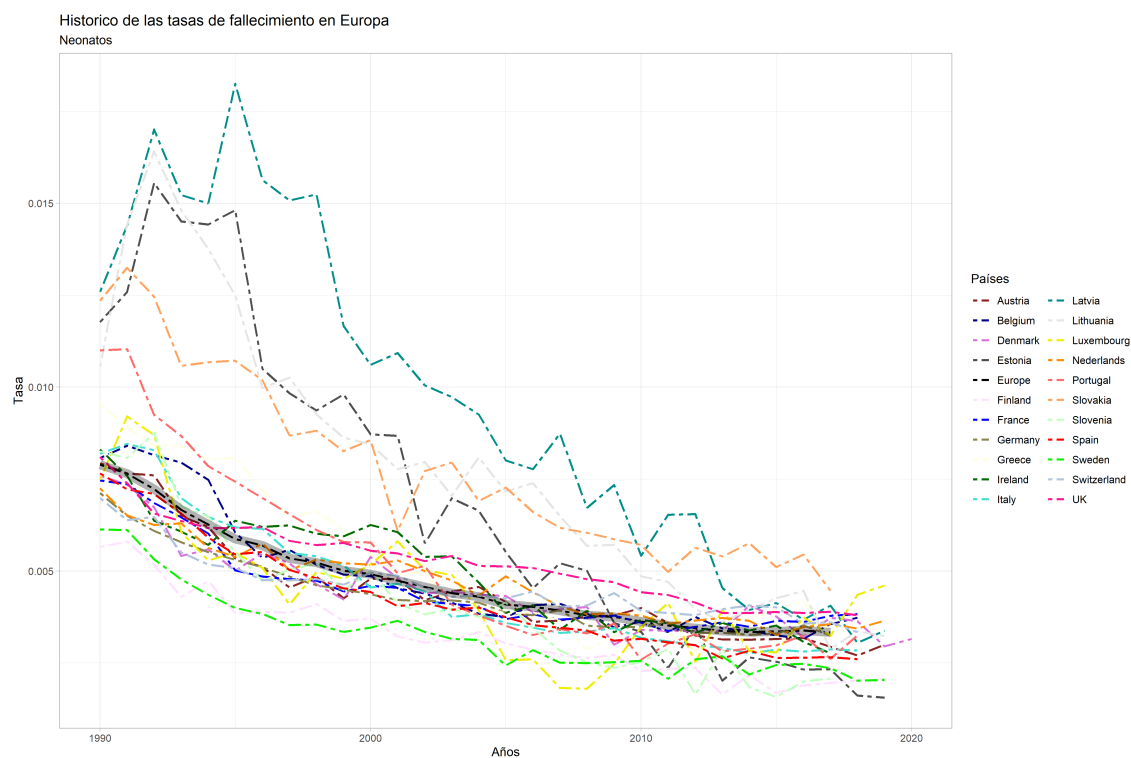


Figura 2.7. Tasa de fallecimiento en personas neonatas en Europa desde 1990. Fuente: Elaboración propia

Esta tendencia no se observa con tanta contundencia a ninguna otra edad. El descenso se debe a la generalización de la sanidad a toda la población. Los medios disponibles, en distintos años según el país, conducen asintóticamente a unas tasas que reflejan únicamente la probabilidad de tener un accidente o sufrir una enfermedad aún incurable.

Para poder analizar estos efectos y extraer conclusiones, utilizaremos muestras grandes. Es decir, aplicando la ley de los grandes números, podemos prever el comportamiento medio a futuro de una población, pero no el de un individuo o una muestra pequeña. Pese a la variedad de rangos a la hora de obtener los datos, al analizarlos se pueden identificar ciertos cambios que sí han alterado la serie histórica, y si se dispusiese de datos con mayor número de detalle, seríamos capaces de apreciar, en especial en ciertos colectivos, los efectos provocados.

Proyectaremos a futuro los resultados medios obtenidos. Para ello, utilizaremos los datos históricos y aplicaremos algoritmos matemáticos a un futuro cercano. El objetivo es conseguir que los resultados se encuentren dentro de un intervalo de confianza establecido, acotando los posibles errores cometidos.

2.3. Rectificación

Pese al aumento en la esperanza de vida, podemos observar ciertos límites sobre la edad que un ser humano puede alcanzar. Observando los datos verificados de las personas que han llegado o superado los 115 años, a los que llamaremos “supersupercentenarios”, el registro más antiguo lo encontramos a partir de 1871. Desde entonces hasta hoy, únicamente se tiene constancia de que 55 personas han alcanzado o superado dicha edad [9].

Por ello, dado que la gente vive hasta edades más avanzadas pero a cierta edad la mortalidad aumenta en gran medida, la curva de supervivencia en los últimos tiempos ha tendido a incrementar enormemente su pendiente a partir de los 80 años de edad. Es decir, los fallecimientos se producen cercanos a la moda. Por ello, la curva de supervivencia se parece cada vez más a un rectángulo, como podemos apreciar en el próximo gráfico. A este fenómeno se le denomina “rectangularización”.

Rectangularización: España desde 1908, sexo: ambos

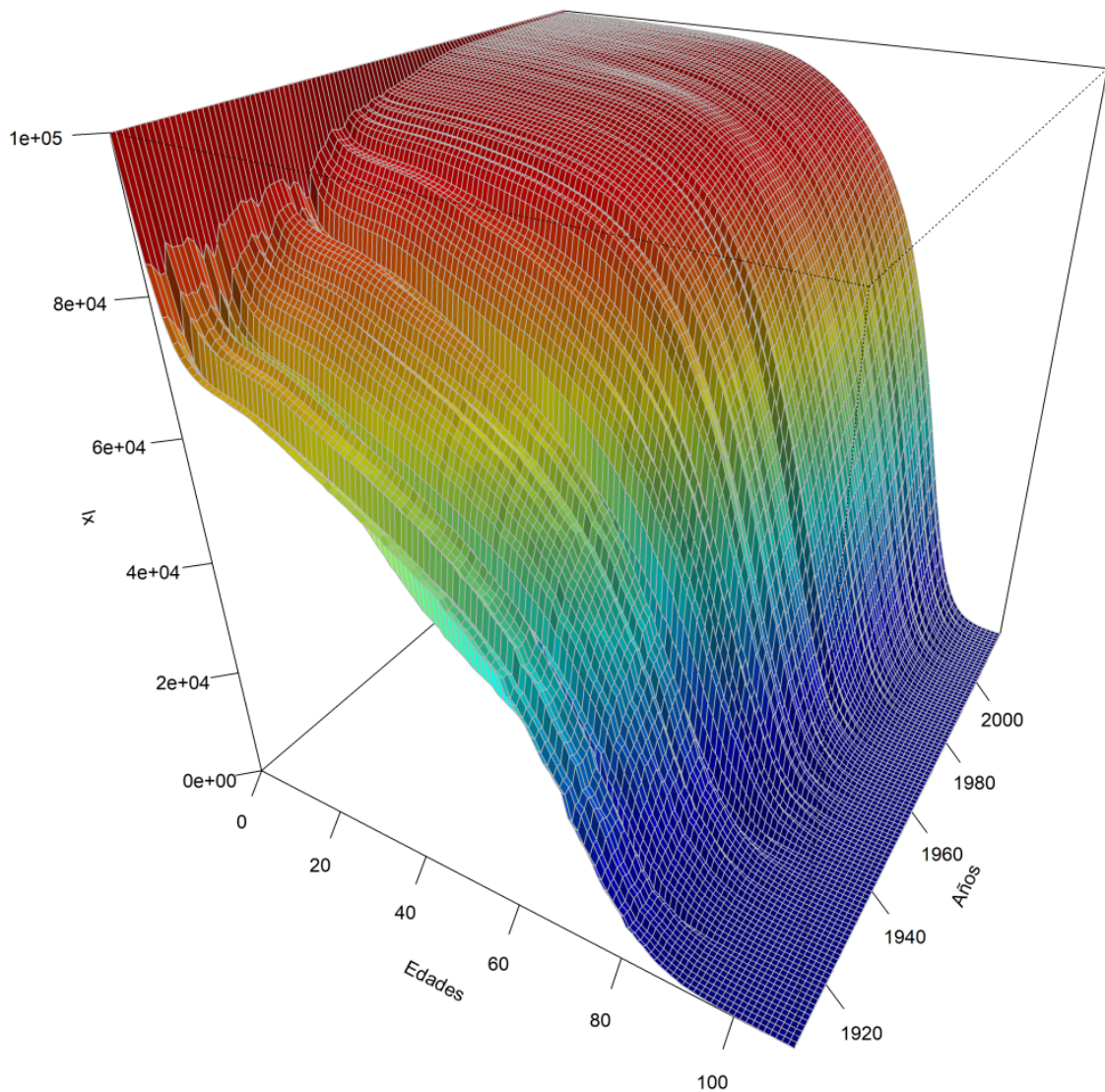


Figura 2.8. Rectangularización en España. Fuente: Elaboración propia

Se puede observar claramente como a inicios del siglo pasado, el número de personas vivas descendía de forma suave, casi lineal, mientras que a finales de siglo, el número de fallecimientos se concentro enormemente a edades avanzadas. Esto hace que la población descienda bruscamente manteniéndose casi horizontal hasta edades cercanas a los 90 años (ver figura 4.7).

Además, al hecho de que cada vez la población sea más longeva, es decir, la moda de los fallecimientos se alcance a edades más avanzadas, se le denomina “extensión” o “expansión” de la función de supervivencia.

Pero, ¿por qué, pese a todas estas mejoras, el número de supersupercentenarios no se ha incrementado? ¿A qué se debe este límite en la edad que puede vivir una persona?

Una posible respuesta la podríamos encontrar al estudiar la Teoría de Weismann sobre la muerte programada [10]. La explicación que encontramos en dicha teoría consiste en que, al faltar los individuos más antiguos de una población, el resto de elementos dispondrán de un mayor número de recursos, por lo que será más probable que la especie sobreviva.

Otra posibilidad, que seguramente encajaría mejor con las poblaciones de seres humanos, sería la Teoría del Soma Desechable. La explicación que encontramos en este caso es que el mantenimiento somático y la reparación de tejidos presentan ciertas limitaciones. A lo largo de la vida de un individuo, sus células se van dañando. Este deterioro se acumula con el paso de los años, lo que contribuye al envejecimiento del organismo. Los genes que determinan la tasa de los cambios producidos a lo largo del tiempo son los genes cargados de la reparación y el mantenimiento del organismo. Esto nos conduce a una longevidad programada. El fallecimiento del organismo por vejez se puede retrasar hasta ciertos límites consiguiendo una baja exposición a riesgos que causen daños a las células o mejorando el mantenimiento corporal.

La senescencia celular es un término que hace referencia a los cambios producidos a lo largo del tiempo en una célula o en un conjunto de células. Este proceso se produce por un acortamiento de los telómeros. Los telómeros son secuencias de ADN no codificadas que se encuentran al final de los cromosomas y se encargan de proteger de posibles daños a la información genética contenida en dicho cromosoma. Durante el proceso de división celular, los telómeros se acortan poco a poco, por lo que, con el tiempo, la célula deja de ser capaz de dividirse [11].

Uno de los eventos que causan daño en las células es la oxidación. Este proceso se da cuando a una molécula le falta la pareja de un electrón en su orbital más externo, es decir, está en estado inestable, busca obtener un electrón adicional. Algunos procesos comunes que son causantes de este proceso de oxidación son la polución, los rayos ultravioleta, el humo del tabaco, el estrés o la falta de sueño. Estos “robos” de electrones causan no sólo daño a las células, sino también mutaciones en las mismas.

Por tanto, por mucho que mejoren las condiciones de vida o la sanidad, seguiremos encontrándonos con un límite en la longevidad. Simplemente dicho límite cada vez se encuentra en edades más avanzadas.

3. MODELIZACIÓN

3.1. Introducción: Las Tablas de Mortalidad

A la hora de estudiar la longevidad (o la mortalidad) de un colectivo, se utilizan las llamadas tablas de mortalidad. En ellas se recoge el número de fallecimientos por cada edad que hay en una población concreta. Gracias a este modelo teórico podemos estudiar las probabilidades de fallecimiento o supervivencia en función de la edad de la persona.

En estas tablas podemos observar el número esperado de personas vivas a una cierta edad, representado por el símbolo l_x donde x es la edad observada en tiempo discreto ($x = 0, 1, 2, 3 \dots$). Para la estimación de l_x se suelen utilizar datos históricos de una cohorte de individuos. Además, estas tablas también recogen la función de supervivencia, representada por $S(x)$, es decir, la probabilidad de que una persona recién nacida haya sobrevivido hasta una cierta edad x , con $x > 0$. Los resultados representados en estas tablas son producto de un estudio estadístico basado en observaciones históricas de la mortalidad de una región. A la probabilidad anual de muerte se la representa por q_x . Una vez obtenida dicha probabilidad, se estudia la fuerza de la mortalidad, denotado por μ_x .

Estos estudios, mediante los cuales se obtienen tanto q_x como μ_x , se realizan observando las frecuencias de fallecimiento a una cierta edad durante varios años. Por esta razón, se suele considerar que l_x y $S(x)$ no varían a lo largo del tiempo. Como hemos visto en el capítulo anterior, esta suposición es incorrecta, dado que con el tiempo la curva de supervivencia ha ido variando. Para paliar dicho inconveniente, en los modelos se incluye, además de la edad x , el año del calendario t . Además, las tablas generacionales aplican a su vez un factor de mejora, λ . Un ejemplo de este tipo de ajustes lo encontramos en [12], en los anexos, donde para realizar el ajuste deberemos utilizar la fórmula:

$$q_x^T = q_x^{base2012} \times e^{-\lambda \times (T-2012)}$$

Antes de pasar a explicar los distintos modelos que utilizamos en el presente estudio, definamos la notación a utilizar y las hipótesis que serán necesarias para el desarrollo de la siguiente sección.

- t : periodo de tiempo entre el inicio del año t y el inicio del año $t + 1$.
- ${}_tq_x$: probabilidad de que en el periodo t fallezca un individuo de edad x .
- ${}_tp_x$: probabilidad de que un individuo de edad x sobreviva al periodo t (${}_tp_x = 1 - {}_tq_x$).
- $D(x, t)$: número de fallecidos de edad x en el periodo t .
- $E(x, t)$: población de edad x durante el periodo t .

- $m(x, t)$: tasa bruta de mortalidad.

$$m(x, t) = \frac{D(x, t)}{E(x, t)}$$

- $\mu(x, t)$: fuerza de mortalidad o tasa instantánea de fallecimiento.

$$\mu(x, t) = \frac{\text{función de probabilidad / densidad}}{\text{función de supervivencia}}$$

- Φ : conjunto de parámetros que se van a estimar. Normalmente los más utilizados son:

- $\beta_x^{(i)}$: efectos debidos a la edad.
- $\kappa_t^{(i)}$: efectos debidos al periodo.
- $\gamma_c^{(i)}$: efectos debidos a la cohorte.

Por otro lado, consideraremos que la tasa instantánea de fallecimiento es constante para cada periodo t y edad x , es decir:

$$m(x + r, t + s) = \mu(x, t) \quad \text{con } r < 1 \text{ y } s \geq 0$$

Además, supondremos que el tamaño de la población permanece constante a lo largo de los años, por lo tanto tenemos:

$${}_tq_x = 1 - e^{-\mu(x, t)} \iff \mu(x, t) = -\ln(1 - {}_tq_x)$$

Supondremos que $D(x, t)$ sigue un proceso de Poisson cuyo parámetro es igual al producto de $E(x, t)$ con $m(x, t)$:

$$D(x, t) \approx P[E(x, t) \times m(x, t)]$$

Para comparar los distintos modelos entre sí, utilizaremos su verosimilitud, la cual depende de $m(x, t)$, que es igual a $\mu(x, t)$ cuando $r, s = 0$. Dado que todos los modelos siguen una distribución de Poisson, en primer lugar veamos la función de verosimilitud de una distribución de Poisson genérica y después de la distribución de Poisson que necesitamos.

Sea X una variable que sigue una Poisson con media λ , entonces:

$$P(X = x) = \frac{e^{-\lambda} \times \lambda^x}{x!}$$

El siguiente paso es aplicar el productorio a la función anterior:

$$L(\lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \times \lambda^{x_i}}{x_i!} = \frac{e^{-n \times \lambda} \times \lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

Finalmente, aplicando logaritmos obtenemos:

$$\begin{aligned} \ln[L(\lambda)] &= -n \times \lambda + \sum_{i=1}^n x_i \times \ln(\lambda) - \sum_{i=1}^n \ln(x_i!) \\ &= \sum_{i=1}^n (\lambda + x_i \times \ln(\lambda) - \ln(x_i!)) \end{aligned}$$

Aplicándolo a nuestro caso particular, en el que $D(x, t) \approx P[E(x, t) \times m(x, t)]$, para cada parámetro $\phi \in \Phi$ que queramos estimar tenemos:

$$\ln[L(D, E, \phi)] = \sum_{i=1}^n (E(x, t) \times m(x, t) + D(x, t) \times \ln[E(x, t) \times m(x, t)] - \ln(D(x, t)!))$$

3.2. Modelos a utilizar

Debido a las mejoras en la mortalidad en los últimos años, podemos deducir que la naturaleza de los procesos subyacentes es estocástica. Los procesos estocásticos se definen matemáticamente como la sucesión de distintas variables aleatorias que, en función de otra variable, evolucionan. Por ejemplo, el número de ventas de una tienda varía de manera aleatoria cada día en función de la hora y de la cantidad de clientes que entren en la tienda con intención de comprar.

En nuestro caso, la esperanza de vida de una persona al nacer dependerá del momento en el que haya nacido (por ejemplo una persona nacida durante la guerra civil española tendrá menor esperanza de vida que una que nació a finales del siglo XX), de los avances a los que tenga acceso (médicos, tecnológicos...), etcétera. Por ello, los modelos que utilizamos en este estudio son de carácter estocástico. Además podremos suponer, sin pérdida de generalidad, que el número de fallecidos sigue una distribución de Poisson.

Existen principalmente dos tipos de modelos, los paramétricos y los no paramétricos. A su vez, los modelos paramétricos se pueden dividir en modelos estructurales o no estructurales, dependiendo de si tienen en cuenta el paso del tiempo sobre a los parámetros que utilizan o si lo consideran como una variable más. Por otro lado, los modelos no paramétricos no dependen de la edad a estudiar. Además, cabe destacar que los modelos que se utilizan en este estudio se pueden expresar modelizándolos como modelos lineales generalizados (GLM) o modelos no lineales generalizados (GNM).

Los ocho modelos que más han sido desarrollados en estudios actuariales son los siguientes [5]:

1. Modelo de Lee-Carter
2. Modelo de Renshaw y Haberman
3. Modelo Edad-Periodo-Cohorte (cuyas siglas son APC: Age-Period-Cohort)
4. Modelo con P-splines
5. Modelo de Cairns, Blake y Dowd (en adelante modelo CBD)
6. Modelo CBD con efecto cohorte
7. Modelo CBD con efecto cohorte y componente cuadrático
8. Modelo CBD con efecto cohorte que decrece con el paso del tiempo

Estos modelos se basan en el análisis de los datos históricos y, en referencia al pasado más cercano al momento de estimación, presuponen que el futuro cercano a dicha fecha se comportará según esa última tendencia. Las principales variables que estos modelos tienen en cuenta para realizar la estimación son la edad, el sexo de la persona y el periodo en el que nos encontremos.

Para reducir el riesgo de modelización, estudiaremos los datos utilizando tres de los modelos citados anteriormente y haremos un promedio de los resultados obtenidos. Los modelos utilizados son el modelo de Lee-Carter, el modelo CBD y el modelo con P-splines.

El modelo de Lee-Carter es el más conocido de todos los modelos citados anteriormente. Se trata de un modelo paramétrico que nos devuelve buenos resultados a la hora de estimar cómo variarán los datos en el futuro (*forecast*). Por otro lado, a la hora de estimar los datos ya conocidos no arroja resultados tan satisfactorios, dado que no ajusta a los datos con especial precisión. Estos resultados eran de esperar dado que es un modelo que se diseñó con el objetivo de predecir observaciones futuras.

El modelo CBD es una versión suavizada del modelo de Lee-Carter, por lo que también es un modelo paramétrico. Tanto los puntos a favor como los defectos de este modelo coinciden con los citados en el párrafo anterior.

Por último, utilizaremos el modelo con P-splines, el cual es un modelo no paramétrico. Este modelo, al contrario que los anteriores, no predice de forma demasiado satisfactoria los cambios futuros, pero modela los datos ya conocidos de manera muy satisfactoria.

No hemos seleccionado ningún modelo que incluya efectos cohorte dado que son modelos que han de utilizarse con gran cautela, en especial a la hora de realizar predicciones, debido a su elevada relación de dependencia.

3.2.1. Lee-Carter

Este modelo fue desarrollado por Ronald D. Lee y por Lawrence R. Carter en 1992. Su principal propósito fue modelizar y predecir la mortalidad estadounidense, evaluando posteriormente la precisión de la estimación futura obtenida. Para este proyecto utilizaron datos históricos locales entre los años 1933 y 1987. Algunas de sus principales ventajas son la sencillez, la fácil interpretación de los parámetros que utiliza y su parsimonia (da prioridad a las explicaciones sencillas frente a todas las existentes). El modelo que propusieron para los tantos centrales de mortalidad es el siguiente:

$$\ln[m(x, t)] = \alpha_x + \beta_x \times \kappa_t$$

donde α_x es el efecto de la edad x , β_x es la velocidad con la que la fuerza de la mortalidad varía a lo largo del tiempo y κ_t es el efecto del tiempo.

El modelo de Lee-Carter, para asegurar que ${}_tq_x \leq 1$ modela las tasas de mortalidad como:

$$\text{logit}({}_tq_x) = \ln\left(\frac{{}_tq_x}{1 - {}_tq_x}\right) = \ln\left(\frac{{}_tq_x}{{}_tp_x}\right) = \alpha_x + \beta_x \times \kappa_t$$

y, despejando ${}_tq_x$, obtenemos:

$${}_tq_x = \frac{e^{\alpha_x + \beta_x \times \kappa_t}}{1 + e^{\alpha_x + \beta_x \times \kappa_t}}$$

Esta estructura es invariante frente a las siguientes transformaciones en los parámetros:

$$\begin{aligned}\tilde{\alpha}_x &= \alpha_x + b \times \beta_x \\ \tilde{\beta}_x &= \frac{\beta_x}{a} \\ \tilde{\kappa}_t &= a \times (\kappa_t - b)\end{aligned}$$

con a, b constantes. Esto deriva en problemas de identificabilidad en los parámetros. Por tanto, para solucionar este problema y obtener una única solución del sistema, añadimos las siguientes restricciones:

$$\begin{aligned}\sum_t \beta_x &= 1 \\ \sum_t \kappa_t &= 0\end{aligned}$$

Estas hipótesis suponen que el estimador de α_x es aproximadamente igual a la media sobre t del $\ln[m(x, t)]$ para cada edad x y que la variable explicativa κ_t está repartida a lo largo de todas las edades.

Cuando los valores de los parámetros α_x, β_x y κ_t son estimados exitosamente, se puede utilizar un modelo ARIMA, mediante la aplicación de la metodología de Box-Jenkins, para predecir los valores futuros del índice κ_t .

Bajo la hipótesis de que α_x y β_x se mantienen constantes frente al paso del tiempo, diversos estudios concluyen que un modelo ARIMA(1,2,1) es el que mejor predice la variación de la mortalidad. Si tenemos en cuenta que θ es un parámetro de deriva y ϵ_t representa los errores independientes e idénticamente distribuidos que siguen una normal $N(0, \sigma^2)$, entonces obtenemos la siguiente estimación:

$$\kappa_t = \kappa_{t-1} + \theta + \epsilon_t$$

3.2.2. Cairns-Blake-Dowd

Este modelo fue formulado por Cairns, Blake y Dowd en el año 2009. Su principal objetivo consiste en estimar las probabilidades de fallecimiento. Para ello, tomamos como supuesto:

$$\begin{aligned}\beta_x^{(1)} &= 1 \\ \beta_x^{(2)} &= x - \bar{x}\end{aligned}$$

donde \bar{x} es la media muestral de las edades incluidas en el modelo.

Por tanto, el modelo queda de la siguiente manera:

$$\ln\left(\frac{{}_tq_x}{{}_tp_x}\right) = \beta_x^{(1)} \times \kappa_t^{(1)} + \beta_x^{(2)} \times \kappa_t^{(2)} + \epsilon_t = \kappa_t^{(1)} + (x - \bar{x}) \times \kappa_t^{(2)} + \epsilon_t$$

donde nuevamente ϵ_t representa los errores independientes e idénticamente distribuidos que siguen una normal $N(0, \sigma^2)$.

Por un lado, el índice temporal $\kappa_t^{(1)}$ es el intercepto del tiempo y afecta de igual manera a todas las edades, representa el nivel de mortalidad en el tiempo t . Si este factor disminuye con el paso del tiempo, podemos deducir que la tasa de mortalidad también se ha reducido en dicho periodo temporal. Por otro lado, el índice temporal $\kappa_t^{(2)}$ representa la pendiente del modelo, que es independiente para cada edad considerada.

Este modelo paramétrico realmente es una reparametrización del modelo CBD de dos factores de mortalidad que propusieron originalmente estos autores en 2006. Ese primer modelo que idearon intenta tener en cuenta el hecho de que la mortalidad no presenta una buena correlación en todas las edades. Por lo tanto, podríamos considerar que tanto ese primer modelo CBD como su reparametrización, son versiones modificadas del modelo de Lee-Carter (el cual sólo incluye un parámetro).

En este caso, no observamos problemas con la identificabilidad en los parámetros, a diferencia de lo que ocurría en el modelo de Lee-Carter.

3.2.3. P-splines

Los splines se definen matemáticamente como curvas diferenciables definidas a intervalos mediante el uso de polinomios. A la hora de realizar interpolaciones resultan de gran utilidad dado que los polinomios que definen estas curvas son de grados bajos. Este hecho evita los problemas de oscilaciones que generan polinomios de grados superiores. Por otro lado, si el objetivo consiste en ajustar curvas, los splines permiten aproximar los resultados a formas de gran complejidad.

Los splines con penalizaciones combinan los puntos fuertes que presentan los dos siguientes enfoques clásicos:

- splines de suavizado: el número de parámetros a utilizar depende de la cantidad de observaciones;
- splines de regresión: se selecciona el número de nodos y se ajustan mediante mínimos cuadrados.

En cambio, mediante los splines con penalización, no utilizamos tantos parámetros como en el modelo de suavizado y a la hora de realizar la selección de los nodos, no resulta tan relevante como en el modelo que utiliza los splines de regresión. Además, como mucho podremos elegir 40 nodos, por lo que el proceso resulta más sencillo y eficiente.

El modelo de suavizado que utiliza splines con penalizaciones fue desarrollado por Eilers y Marx en el año 1996. El método que decidieron se puede dividir en dos fases:

1. A la hora de realizar la regresión, aplicar una base adecuada.
2. Introducir penalizaciones, basadas en la diferencia existente entre distintos coeficientes adyacentes entre sí, a la función de verosimilitud.

En el supuesto caso en el que las penalizaciones aplicadas no tuvieran influencia en el modelo, estaríamos ante un caso de Basic splines (en adelante B-splines).

Para determinar los B-splines, se precisa establecer las dimensiones a aplicar en el modelo. Para ello, hay que conocer el lugar que ocupan los puntos nodales en el espacio. En caso de considerar una única dimensión, suponiendo que el rango de los datos sea de t_0 a t_n , los nodos del B-spline podrán estar en un rango mayor, por ejemplo entre t_{-2} y t_{n+2} . En este supuesto entonces tendríamos, para $k = -2, \dots, n - 1$, el spline de orden k :

$$B_k(x) = \sum_{i=k}^{k+3} \prod_{j \neq i}^{k+3} \frac{(x - t_i)_+^2}{t_j - t_i}$$

con $(x - t_i)_+^2 = \max \{x - t_i, 0\}$.

Por lo tanto, cuando x se encuentra fuera del rango entre t_k y t_{k+3} , este spline de orden k vale cero, pero tiene en toda la recta real tanto la primera como la segunda derivada continuas.

Sea $\delta = t_{k+1} - t_k$, entonces bajo la hipótesis de que los nodos están distribuidos de manera uniforme tendríamos:

$$B_k(x) = \delta^{-3} \left[\frac{(x - t_k)_+^2}{6} - \frac{(x - t_{k+1})_+^2}{2} + \frac{(x - t_{k+2})_+^2}{2} - \frac{(x - t_{k+3})_+^2}{6} \right]$$

Si consideramos más dimensiones, el procedimiento es análogo. En el caso de dos dimensiones, por ejemplo sea “ x ” la edad, con nodos entre t_{-3}^e y $t_{n_e+3}^e$, y sea “ y ” el año del calendario, con nodos entre t_{-3}^a y $t_{n_a+3}^a$. En este caso tendremos dos bases, $B_k^e(x)$ para $k = -3, \dots, n_e - 1$ y $B_l^a(t)$ para $k = -3, \dots, n_a - 1$. Los B-splines que obtendríamos en este caso serían de la forma:

$$B_{k,l}(x, t) = B_k^e(x) \times B_l^a(t)$$

Para el cálculo de la base a usar en la regresión, encontramos principalmente dos metodologías a utilizar. En la primera se obtiene mediante el uso de las bases de polinomios truncados. En la segunda se utilizan las bases B-splines. Para una información más sobre estas y otras metodologías, consultar [13].

Los P-splines se utilizan cuando para estimar los B-splines se requiere un conjunto de datos en que, para ajustar la verosimilitud, se usa una función de penalización con el objetivo de conseguir un suavizado de los resultados.

Sea B una base con K nodos. La función objetivo $S(a, y)$, bajo el supuesto de que para ajustar el modelo utilizamos mínimos cuadrados sería:

$$S(a, y) = (y - B \times a)' \times (y - B \times a) \rightarrow \tilde{a} = (B' \times B)^{-1} \times B' \times y$$

Dependiendo del tamaño de la base, tendremos una curva ajustada u otra. El tamaño de la base y la suavidad de la curva tienen una relación indirecta, es decir, cuanto mayor sea la base utilizada menor será la suavidad de la curva. En el caso en que el número de nodos coincidiera con la cantidad de datos, el resultado sería una curva que interpola la información utilizada.

Para evitar este fenómeno, una de las penalizaciones más comunes sería aplicar la integral de la segunda derivada de la curva, elevándolo al cuadrado. Así, nuestra función objetivo quedaría de la forma:

$$S(a, y, \lambda) = (y - B \times a)' \times (y - B \times a) + \lambda \int_x (B'' \times a)^2 dx$$

la novedad introducida por los P-splines se encuentra en que se penalizan directamente los coeficientes, no la curva, por lo que se simplifica el problema.

El tipo de penalización a utilizar dependerá de la base que necesitemos aplicar. Para más información ver [13].

La expresión de este modelo sería:

$$\ln [m(x, t)] = \sum_{i,j} \theta_{i,j} B_{i,j}^{e,a}(x, t)$$

con $\theta_{i,j}$ suavizado en función de la edad y del tiempo.

3.3. Backtesting

Para comprobar si las predicciones futuras obtenidas con los distintos modelos son correctas o incorrectas, utilizaremos datos históricos para evaluar la calidad de las estimaciones obtenidas, es decir, haremos un ejercicio de “*backtesting*”. Los pasos que seguimos para realizar este proceso son:

1. Elegir un periodo histórico que queramos estudiar dejando fuera los últimos años.
2. Obtener una predicción para los datos que hemos dejado fuera en el paso anterior.
3. Contrastar la predicción realizada con los datos reales observados en dicho periodo.

En el mercado asegurador, los cambios inesperados en la tendencia podrían provocar errores en las estimaciones obtenidas en los modelos. Por ello, desde la Directiva de Solvencia II, comunican que es necesario realizar un ejercicio de estrés para la correcta gestión de los riesgos (en este caso es del 99,5 %). De esta manera, se calcula el capital mínimo que es necesario para afrontar los posibles cambios que se pueden producir en las obligaciones tanto de tomadores como de beneficiarios.

4. ANÁLISIS DE LA MORTALIDAD

4.1. Obtención de los datos

Los datos utilizados en el presente trabajo son los recogidos en la “*Human Mortality Database*” [14] (en adelante *HMD*). Esta base de datos es utilizada en numerosos estudios actuariales debido a la calidad y la riqueza de la información que contiene. Con la información que nos aporta esta página, analizaremos la evolución histórica de la supervivencia en Europa y, en particular, en España.

Los datos más antiguos y con continuidad los encontramos en Suecia, cuyo primer registro de la población se produjo en el año 1751. Por otro lado, países como Grecia y Alemania no presentan un registro demasiado amplio de información, dado que recogen datos desde 1981 en el caso griego y desde 1990 en el caso alemán.

El lenguaje de programación utilizado a lo largo de este trabajo es “*R*”, y el interfaz de programación “*R-studio*”. En primer lugar, descargamos los datos directamente de la página web [14], para lo que necesitaremos crearnos una cuenta. Continuaremos con una limpieza general de los datos, dado que presentan caracteres que nos pueden llevar a error (por ejemplo, que ponga edad 110+ en lugar de edad 110). A continuación, organizaremos dichos datos en una matriz que contenga diferentes matrices de datos organizadas según cada país contenido en el estudio. Este proceso lo podemos encontrar en el anexo, en concreto en el [código 1](#).

A esta estructura de datos la he denominado “miHMDdata”, extendiendo el original “HMDdata” del paquete “*Mortality Smooth*” con más países y datos más actualizados. Su primer escalón está compuesto por cada uno de los países escogidos. Dentro de cada país, en primer lugar tenemos los datos genéricos que lo definen, como son el nombre del mismo, el tipo de datos con los que estamos trabajando, los años de calendario incluidos en la base de datos y las edades recogidas. A continuación, tenemos el siguiente escalón, que son los datos propiamente dichos. En cada país hay cuatro subgrupos de datos, que son:

- *Población (pop)*: Número de personas vivas en una cierta población a 1 de enero de cada año recogido.
- *Fallecimientos (death)*: Numero de fallecimientos registrados durante un año.
- *Exposición al riesgo (expousure)*: Estimación de la población expuesta al riesgo de fallecimiento en cada intervalo de tiempo. Se calcula a partir de la población registrada, ajustando el valor con los fallecimientos producidos en dicho intervalo.
- *Tasa de fallecimiento (rate)*: Número de fallecimientos entre las personas expuestas al riesgo.

Esta matriz de datos tiene el siguiente formato:

miHMDdata	list [21]	List of length 21
Austria	list [8]	List of length 8
Belgium	list [8]	List of length 8
Denmark	list [8]	List of length 8
France	list [8]	List of length 8
Germany	list [8]	List of length 8
Greece	list [8]	List of length 8
Estonia	list [8]	List of length 8
Finland	list [8]	List of length 8
Italy	list [8]	List of length 8
Ireland	list [8]	List of length 8
Latvia	list [8]	List of length 8
Lithuania	list [8]	List of length 8
Luxembourg	list [8]	List of length 8
Netherlands	list [8]	List of length 8
Portugal	list [8]	List of length 8
Slovakia	list [8]	List of length 8
Slovenia	list [8]	List of length 8
Spain	list [8]	List of length 8
country	character [1]	'Spain'
type	character [1]	'mortality'
year	integer [112]	1908 1909 1910 1911 1912 1913 ...
age	integer [111]	0 1 2 3 4 5 ...
pop	list [3]	List of length 3
female	double [111 x 112]	293128 265371 247725 248121 243290
male	double [111 x 112]	303916 268001 254761 249421 245973
total	double [111 x 112]	597044 533372 502486 497542 489264
death	list [3]	List of length 3
exposure	list [3]	List of length 3
rate	list [3]	List of length 3
Sweden	list [8]	List of length 8
Switzerland	list [8]	List of length 8
UK	list [8]	List of length 8

Figura 4.1. Matriz de datos. Fuente: Elaboración propia

A su vez, cada subgrupo de datos está dividido en tres matrices, que contienen los datos de las mujeres, de los hombres, y del conjunto de ambos sexos. Estas matrices recogen en cada columna un año de calendario y en cada fila la edad.

Para construir los datos de Europa he extendido esa misma estructura con un “país” adicional, cuyas poblaciones, exposiciones y fallecimientos se calculan mediante la suma del conjunto de países citados anteriormente y siendo la tasa (“*rate*”) el cociente entre los fallecimientos (“*death*”) y las exposiciones (“*expousure*”). Este proceso se desarrolla en el [código 5](#).

Debido a la utilización de esta técnica, únicamente disponemos de datos de Europa a partir del momento en que todos los países tienen datos, es decir, desde 1990 tras la unificación de Alemania.

Se han seleccionado todos los países de la eurozona menos Chipre y Malta, dado que en la *HMD* no se dispone de datos suyos [15]. Además, se ha añadido Suiza, Reino Unido y Suecia, y se mantiene Dinamarca pese a sus cláusulas de exclusión.

Además, he descargado otros dos conjuntos de datos:

- **Esperanza de vida al nacer:** en base al mismo conjunto de países y que en *HMD* corresponde al fichero *E0per.txt*, que se utilizará en diversos datos pero no en los modelos de simulación (ver [código 2](#)).
- **Tablas de vida:** dentro de las tablas de vida de *HMD*, los ficheros *flptper_1x1.txt*, *mltper_1x1.txt* y *bltper_1x1.txt* para los gráficos que muestran la rectangularización y que tampoco se emplean a la hora de modelizar (ver [código 3](#)).

4.2. La longevidad en España

En primer lugar, analizaremos el número de fallecidos de mujeres y de hombres que hubo en España entre 1908 y 2018. Cada década se representa con un color distinto, para facilitar su identificación visual.

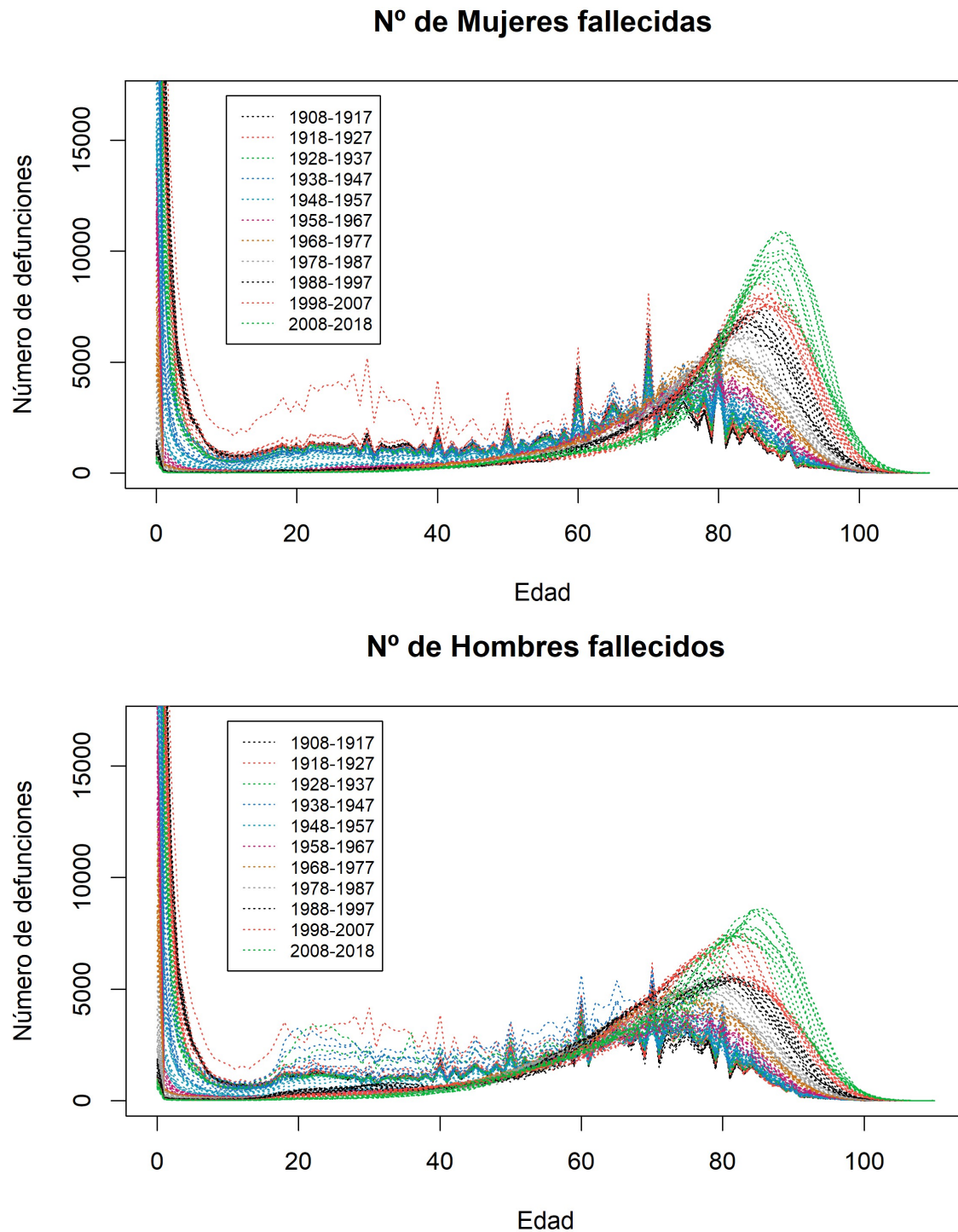


Figura 4.2. Número de fallecidos en España entre 1908 y 2018. Fuente: Elaboración propia

A primera vista se puede observar que, en general, los gráficos cumplen con lo esperado. La mortalidad en personas recién nacidas, de hasta un año de edad, era elevada hasta mediados del siglo XX, cuando cae asintóticamente a cero, igualándose a las tasas del resto de niños. De igual modo y principalmente en el gráfico que representa los fallecimientos en hombres, podemos observar que, al llegar a la mayoría de edad y durante el periodo que incluye tanto juventud como inicio de la etapa adulta (entre los 18 y los 35 años), el número de fallecimientos vuelve a ascender, probablemente debido a que es una etapa en la que las personas se sienten con confianza y energía, y asumen más riesgos. Después, antes de llegar a la década de los cuarenta el número de fallecidos vuelve a descender, dado que las personas tienden a tener en esos instantes una vida más asentada, y se estabiliza durante un tiempo para finalmente ascender hasta alcanzar su máximo a edades avanzadas. Este valor máximo se encuentra cerca de los 90 años en las mujeres y cerca de los 80 en los hombres. Además, también se observa claramente que el número de fallecimientos en personas de 100 años o mayores es muy pequeño en todas las décadas dado que, como vimos anteriormente, son pocas las personas que llegan a esas edades.

Las diferencias apreciadas entre hombres y mujeres se pueden deber a factores socioculturales. Por ejemplo, históricamente los hombres se han dedicado a trabajos más arduos físicamente hablando, lo que les puede llevar a un mayor deterioro físico o a vivir situaciones más arriesgadas. También puede deberse a factores biológicos. Algunos estudios defienden que las mujeres tienen un sistema inmunológico más fuerte que el del género masculino, generan mayor número de anticuerpos y con mayor rapidez, por lo que se ven menos afectadas frente a enfermedades infecciosas [16].

En las primeras décadas en las que se recogieron datos (entre 1908 y 1937), podemos observar picos, probablemente debido a una menor consistencia de la información.

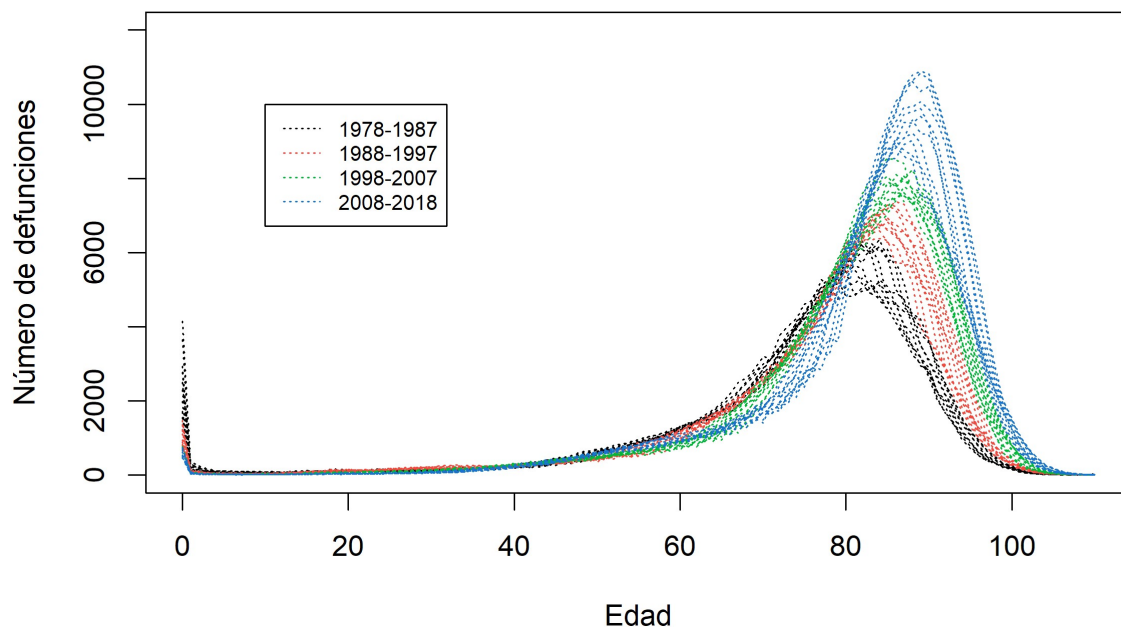
Entre 1918 y 1920, se aprecia un mayor número de fallecimientos, principalmente en gente joven o adultos saludables menores de 40 años. Esto es consecuencia de la mal llamada “gripe española”. Para más información, consultar [17].

En la década de 1930, podemos observar un mayor número de fallecimientos en gente joven, principalmente en hombres. Esto se explica porque fue cuando ocurrió uno de los momentos históricos más duros de nuestro país, la Guerra Civil Española. En esta época falleció mucha gente, principalmente entre la población joven, que eran los que iban al frente, por lo que la cantidad de personas en dicha generación se redujo.

A partir de los años 60 las perturbaciones son menores, y se observa más claramente un patrón hacia el envejecimiento de la población. En estas últimas décadas se aprecia un ligero aumento en el número de fallecidos a partir de los 18 años. Dado que no se produjo ningún evento bélico en este periodo, una explicación de este fenómeno podría ser que, al llegar a la mayoría de edad, los jóvenes obtienen el permiso de conducir pero aún no tienen experiencia al volante, ocasionando un mayor número de accidentes de tráfico. Otra explicación podría ser la expansión de las drogas, principalmente heroína en los 70 y cocaína en los 80, y al SIDA en los años 90, que afectó sobre todo a los más jóvenes.

Para apreciar mejor la evolución de la mortalidad, continuaremos realizando el estudio con datos de la España constitucional, es decir, a partir de 1978.

Nº de Mujeres fallecidas



Nº de Hombres fallecidos

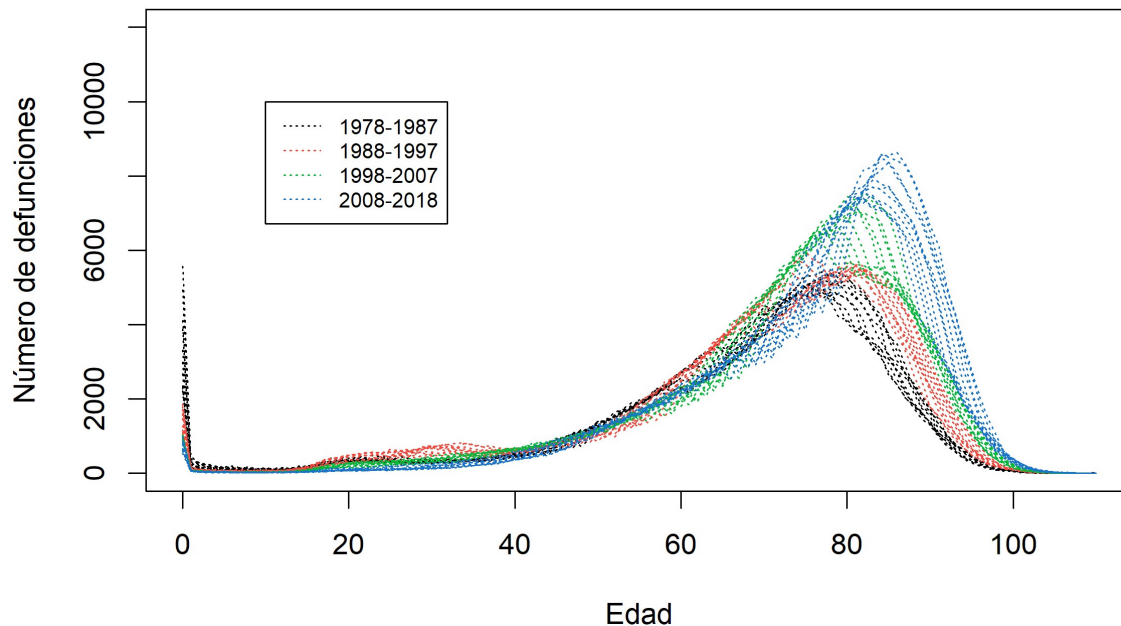


Figura 4.3. Número de fallecidos en España entre 1978 y 2018. Fuente: Elaboración propia

De este modo se aprecia con mayor claridad cómo, con el paso de las décadas, el número de fallecimientos se ha ido concentrando cada vez a edades más avanzadas. En mujeres se encuentra más agrupado a edades un poco mayores que en hombres. Aún así, en ambos gráficos podemos observar cómo cada vez la población fallece a mayor edad.

Los gráficos de exposición al riesgo de fallecimiento recogen la misma tendencia.

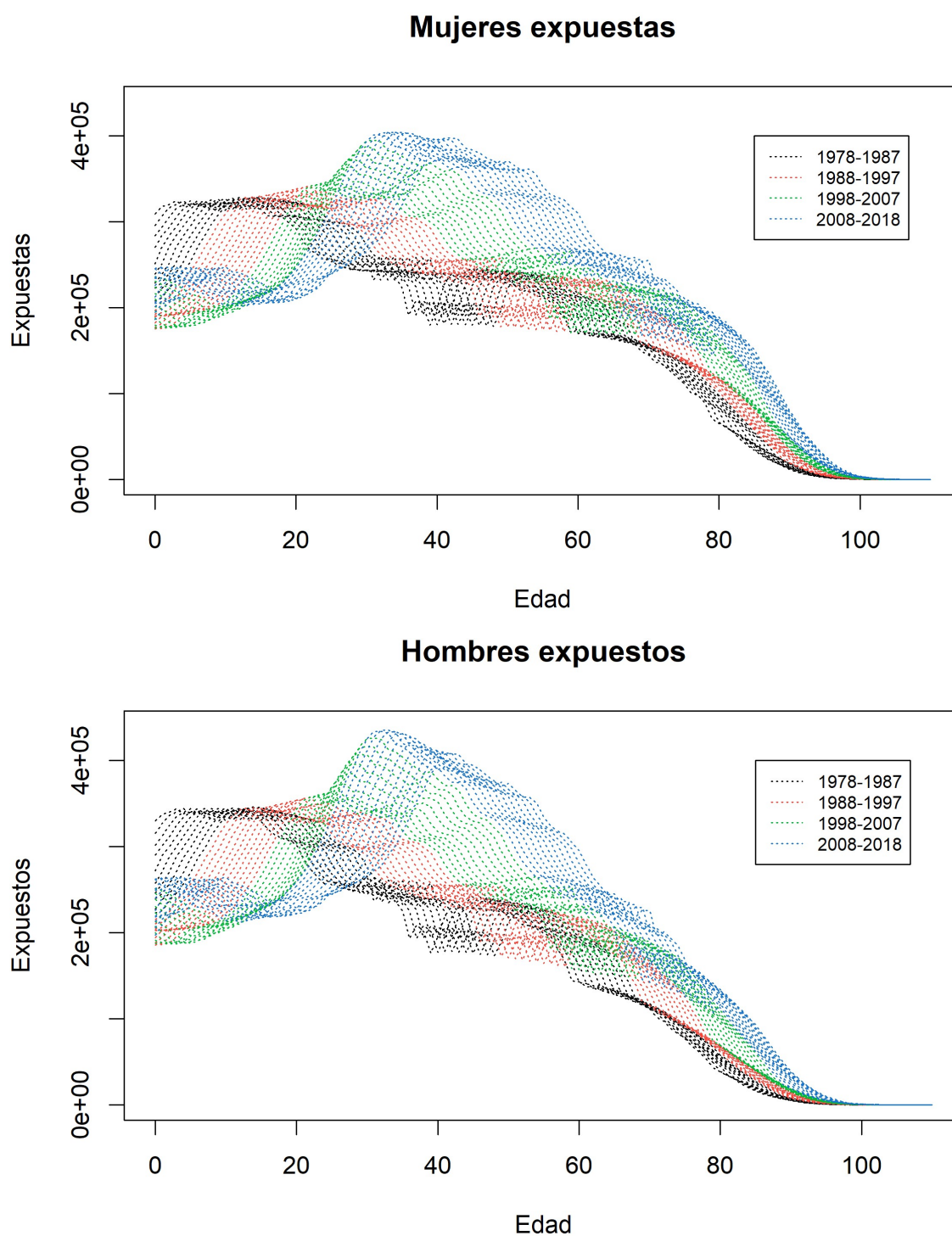


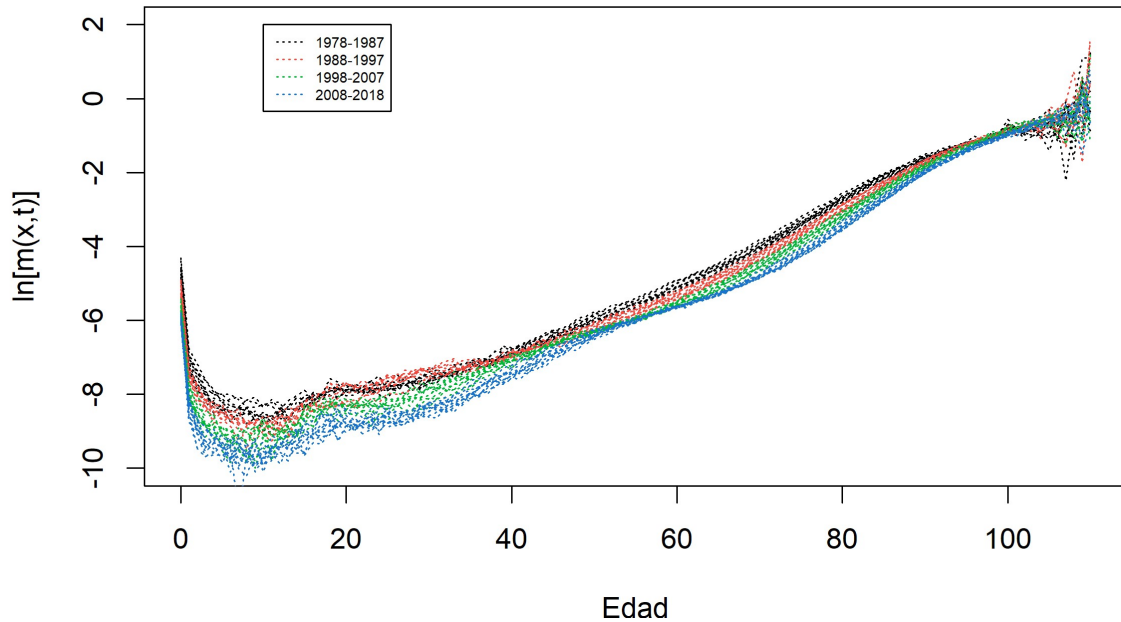
Figura 4.4. Número de personas expuestas en España entre 1978 y 2018. Fuente: Elaboración propia

Los resultados obtenidos en estos gráficos vuelven a ser los esperados. Según van pasando los años, hay más personas expuestas a edades más avanzadas. En esta ocasión también podemos observar cómo el número de nacimientos ha descendido. Esto puede ocasionar un serio problema demográfico dado que las personas cada vez llegan a edades

más avanzadas pero no hay jóvenes suficientes para mantener el actual sistema de pensiones. Aunque es un tema de gran interés, está fuera del objetivo del presente trabajo. Existen numerosos artículos que tratan sobre este tema, como por ejemplo [18].

A continuación tenemos las gráficas relativas a la tasa de mortalidad.

$\ln[m(x,t)]$ en Mujeres



$\ln[m(x,t)]$ en Hombres

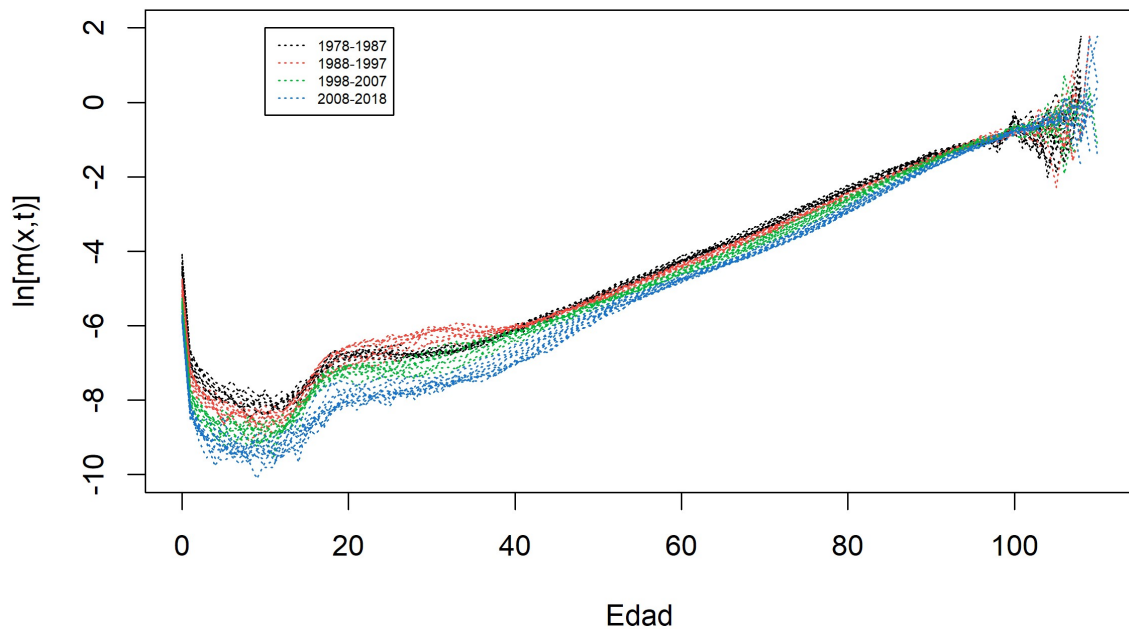


Figura 4.5. Tasa de fallecimiento en España entre 1978 y 2018. Fuente: Elaboración propia

En este gráfico hemos representado el logaritmo de los valores de las tasas brutas de mortalidad. Al aplicar el logaritmo, lo que obtenemos es una sucesión bastante regular,

que no presenta puntos de inflexión debido al periodo que está representado, en el que la calidad de los datos es elevada y no ha ocurrido ningún suceso histórico relevante en términos de mortalidad de la población.

De nuevo podemos observar cómo a lo largo de las décadas la tasa de mortalidad va disminuyendo debido al aumento en la calidad de vida. Gracias a estos gráficos, podemos apreciar más notoriamente ese incremento de la tasa de mortalidad en los jóvenes y la estabilización que hay hasta poco antes de los 40 años de edad, a partir de donde se incrementa linealmente. Además, al superarse los 100 años observamos mucha distorsión en los datos, dado que las personas que llegan a edades tan avanzadas no siguen un patrón de mortalidad que se sepa modelizar.

A la hora de aplicar los modelos expuestos anteriormente, utilizaremos los datos recogidos entre 1950 y 2018. Esto se debe a que entre 1908 y 1950 aproximadamente los datos recogidos no son tan fieles a la realidad como los datos posteriores. Por lo tanto no consideraremos dichos espacios temporales como representativos a la hora de estudiar la evolución de la longevidad de un país. Se podría considerar que son datos atípicos o “*outliers*”, dado que rompen con la tendencia de la serie. El periodo escogido es lo suficientemente amplio como para poder aplicar los modelos de manera precisa y los datos tienen una alta calidad.

Las ciencias actuariales intentan equilibrar los datos que se estiman aplicando los modelos con los datos reales. Para ello, suponemos como cierta la hipótesis de que las irregularidades obtenidas en los valores de la función se deben a errores cometidos en las observaciones realizadas. Además, supondremos que el efecto de dichos fallos se pueden subsanar tomando como resultados finales los correspondientes a una curva que se halle en un intervalo de confianza respecto de los valores experimentales.

Volviendo a la tasa de mortalidad, para evitar problemas ante posibles irregularidades entre las distintas edades, se realizan algunos ajustes. Para estudiar el comportamiento histórico de la mortalidad, es decir, la tendencia que tiene, nos serán útiles las definiciones de “tasa de mejora” y “factor de reducción de la mortalidad”.

Siguiendo la notación anterior, el factor de mejora de la mortalidad, al que denominamos por ${}_t\lambda_x$, se define como:

$${}_t\lambda_x = 1 - \frac{{}_tq_x}{{}_{t-1}q_x}$$

y el factor de reducción de la mortalidad, al que representaremos por ${}_tFR_x^0$, está definido por:

$${}_tFR_x^0 = 1 + \frac{{}_tq_x - {}_tq_0}{{}_tq_0}$$

donde ${}_tq_0$ es la mortalidad en el periodo de referencia.

Estos dos factores nos permiten valorar la mortalidad experimentada por la población objeto de análisis con gran nivel de detalle. Al realizar un estudio previo de la información, es más probable encontrar posibles errores que pasarían desapercibidos. De esta manera, evitamos asumir mejoras recurrentes o utilizar periodos de tiempo incorrectos.

Cuando no disponemos de datos suavizados de la mortalidad, entonces el factor de reducción nos permite realizar un análisis más preciso de la evolución de la mortalidad que el factor de mejora. Esto se debe a que el factor de reducción se ve menos afectado por la volatilidad presente en las observaciones brutas entre periodos. Por otro lado, para poder utilizar este factor es necesario marcar un periodo de referencia base, de donde cogemos ${}_tq_0$, mientras que para utilizar el factor de mejora de la mortalidad, únicamente necesitamos las probabilidades de fallecimiento para una persona de edad x en un periodo y en el anterior.

Para finalizar esta sección, veamos la evolución que ha tenido la esperanza de vida al nacer en España en el último siglo. A continuación expondremos la esperanza de vida diferenciando sexo y la esperanza de vida conjunta de la población.

Año	Mujer	Hombre	Total
1908-1909	42.27	40.26	41.25
1910-1914	43.12	41.11	42.10
1915-1919	40.70	38.93	39.80
1920-1924	44.76	42.25	43.48
1925-1929	49.93	46.82	48.35
1930-1934	52.61	49.04	50.80
1935-1939	53.34	45.29	49.13
1940-1944	55.30	48.48	51.82
1945-1949	62.08	56.60	59.39
1950-1954	66.39	61.69	64.15
1955-1959	69.81	65.11	67.52
1960-1964	72.24	67.15	69.76
1965-1969	73.93	68.47	71.25
1970-1974	75.30	69.66	72.53
1975-1979	77.16	71.18	74.22
1980-1984	79.10	72.85	76.02
1985-1989	80.12	73.43	76.80
1990-1995	81.13	73.89	77.51
1996-1999	82.22	75.10	78.66
2000-2004	83.20	76.42	79.82
2005-2009	84.21	77.87	81.05
2010-2014	85.25	79.51	82.42
2015-2018	85.66	80.23	82.98

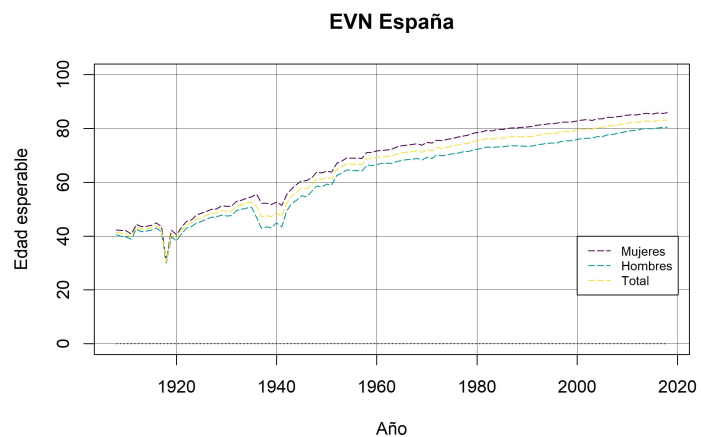


Tabla 4.1. Esperanza vida

Figura 4.6. Esperanza vida entre 1908 y 2018

Fuente: Elaboración propia con datos de “*The Human Mortality Database*”

La tendencia de la esperanza de vida, como era de esperar, es positiva. Cuanto más tarde nace una persona, mejores son las condiciones de las que disfruta y, por tanto, vive un mayor número de años. En este gráfico podemos observar dos momentos en los que la esperanza de vida disminuyó de manera puntual. El primero se produjo antes de 1920, en los años en los que la pandemia de la gripe española afectó de forma más directa a la población. La segunda bajada en la esperanza de vida fue causada por la Guerra Civil Española, que trajo consigo no solo la muerte de los soldados principalmente jóvenes, sino también hambre y la pobreza para con sus familias.

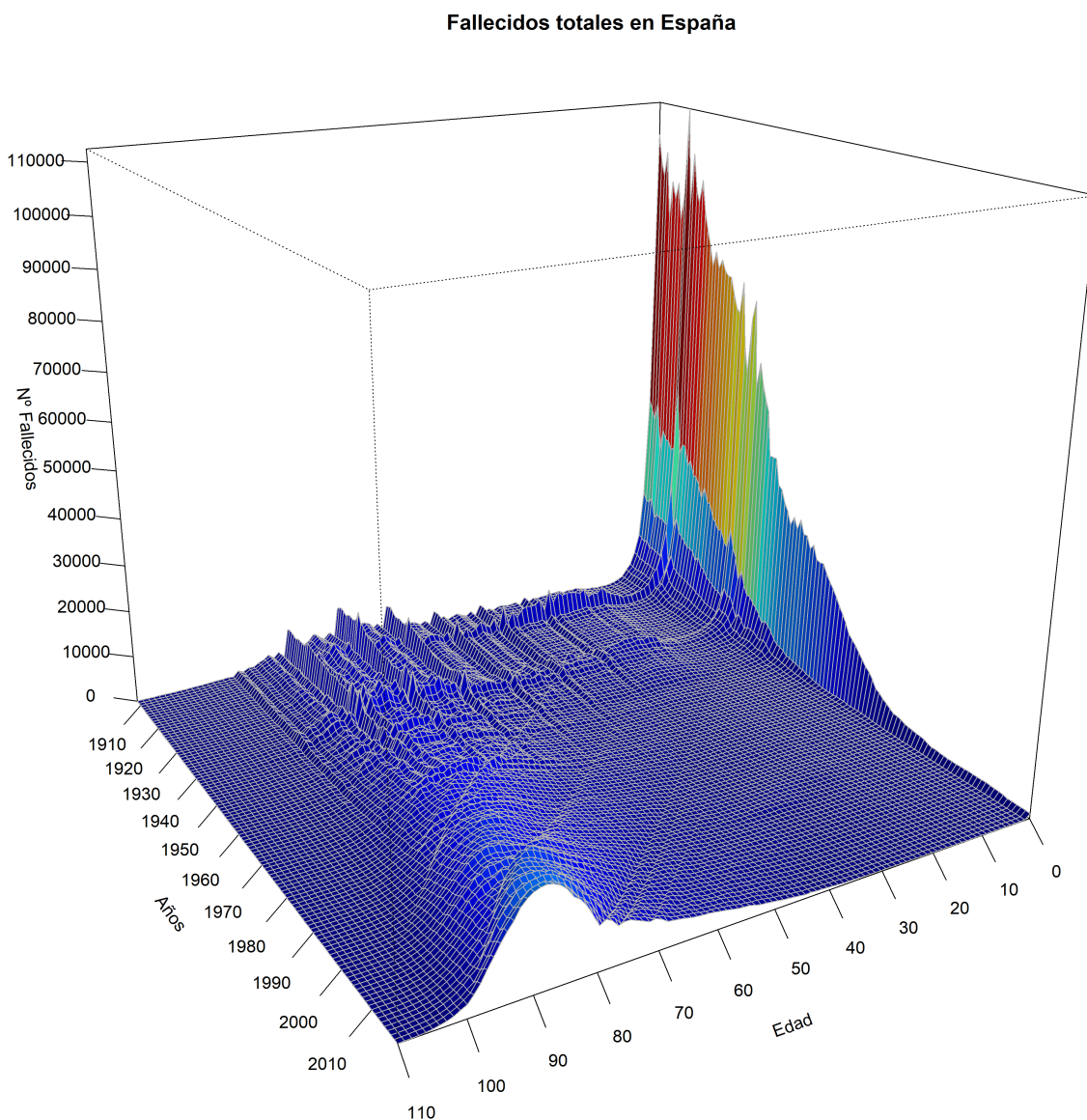


Figura 4.7. Decesos en España desde 1908. Fuente: Elaboración propia

El gráfico anterior nos puede ayudar a recapitular todo el análisis realizado anteriormente.

En primer lugar, los datos de principios del siglo XX tienen “ruido” debido a su menor calidad respecto a los datos actuales. Este fenómeno es apreciable en los picos que encontramos en el fondo del gráfico.

Por otro lado, la cantidad de fallecimientos en personas neonatas ha descendido de forma drástica, siendo el cambio más notorio respecto a la mortalidad.

También cabe destacar que las muertes se producen cada vez a edades más avanzadas. Esto se aprecia en forma de “ola” en los últimos años del gráfico. A comienzos del siglo XX, los fallecimientos son similares en todas las edades, por lo que la superficie es casi plana, mientras que en los últimos tiempos esto se ha concentrado entre las edades de 80 y 90 años.

4.3. La longevidad en Europa

En Europa cabe esperar un comportamiento análogo al observado en el caso de España. Los datos de la mayoría de los países seleccionados incluyen el siglo XX. Durante ese siglo cabe destacar, al igual que en el caso de España, la pandemia de la gripe española. También tuvieron lugar tanto la Primera como la Segunda Guerra Mundial.

En la Primera Guerra Mundial, por primera vez en la historia, participaron países de los cinco continentes. Millones de personas lucharon y fallecieron durante esta guerra de trincheras, tras soportar condiciones infrahumanas. La mortalidad creció no sólo a causa de las víctimas mortales durante las batallas o las provocadas a la hora de arrasar ciudades, también se sufrió la escasez de alimentos y una gran crisis económica. No se conoce el número de muertos exactos pero se estima que hubo alrededor de 20 millones.

La Segunda Guerra Mundial se llevó a cabo con grandes avances armamentísticos, por lo que el número de muertos fue aún mayor que durante la Primera Guerra Mundial, rondando los 55 millones de fallecidos.

En el siguiente gráfico podemos observar la esperanza de vida de cada país.

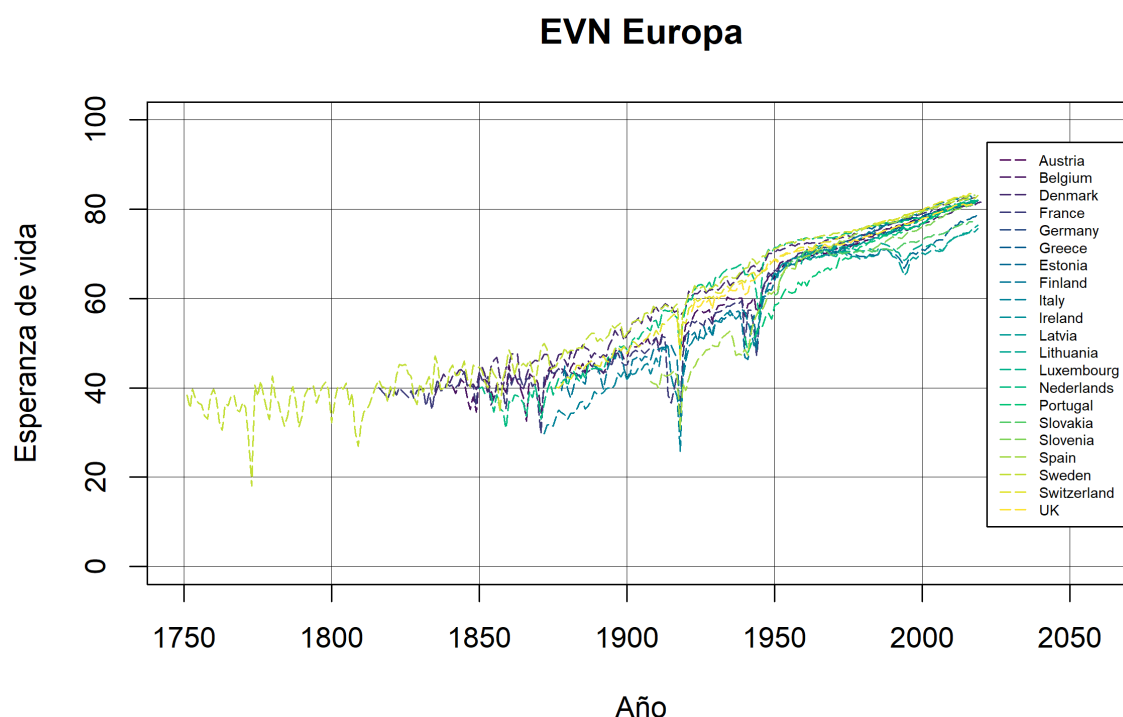


Figura 4.8. Esperanza de vida al nacer en Europa. Fuente: Elaboración propia

En todos los países seleccionados se aprecia una tendencia creciente salvo en momentos puntuales. Además, el gráfico representa esta esperanza desde el momento en el que cada país empezó a tener datos. Suecia es el país que tiene datos desde hace más tiempo (1751), mientras que Alemania recoge datos desde hace escasas décadas (1990), condicionando los datos que utilizamos en las proyecciones.

Pese a existir momentos puntuales en los que la esperanza de vida disminuyó, en general la tendencia en todos los países es creciente. Es un hecho que cada vez se vive más y mejor. Aún así, las dos guerras mundiales son claramente apreciables, la esperanza de vida de la gente nacida durante dichos periodos es claramente inferior, llegando incluso a igualar valores vividos a mediados del siglo XIX.

Veamos ahora, por ejemplo, la tasa de fallecimiento a los 55 años para los países de nuestra base de datos

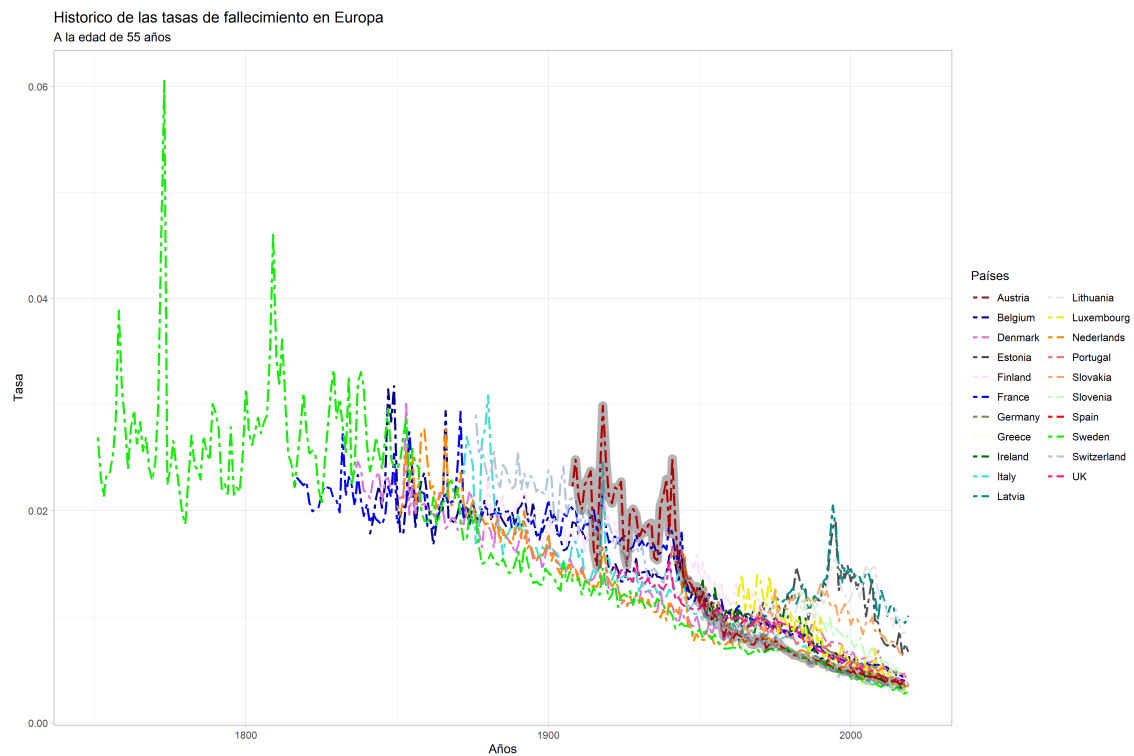


Figura 4.9. Tasa de fallecimiento a los 55 años en Europa. Fuente: Elaboración propia

En esta ocasión, nuevamente podemos observar los picos correspondientes a los eventos citados anteriormente. Y, como era de esperar, esta tasa tiene una tendencia decreciente, es decir, la esperanza de vida aumenta con el paso del tiempo.

Para dar otro enfoque al problema que tiene el aumento de la longevidad, veamos la pirámide poblacional de, por ejemplo, el año 2015. Además, también resulta interesante comparar las tasas de mortalidad de España y las de Europa.

Pirámide de población: España vs Europa en 2017

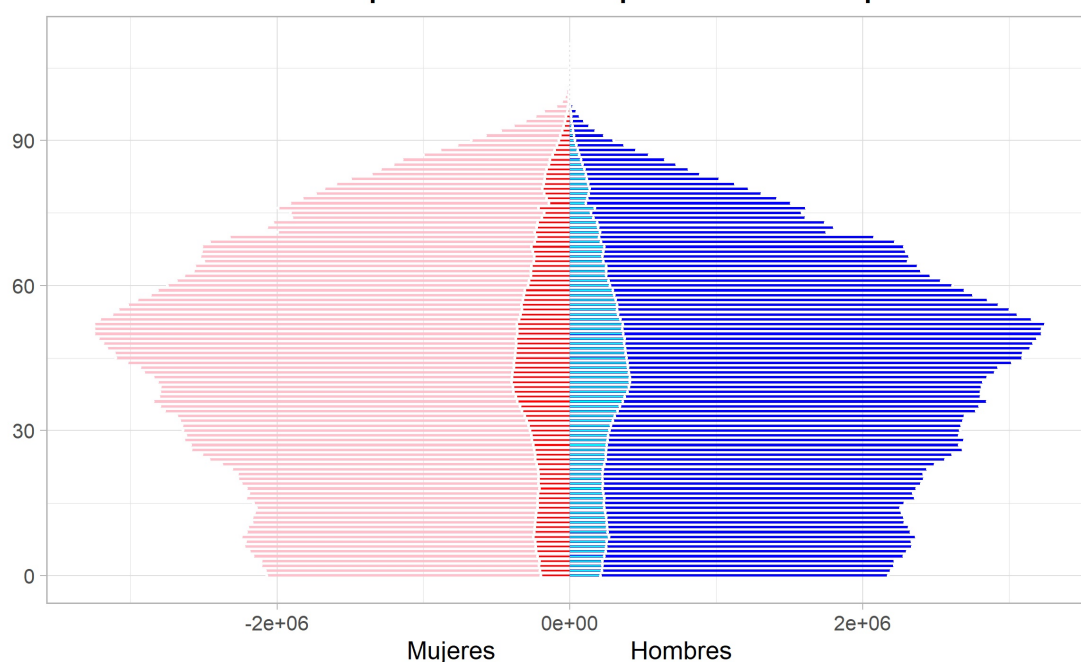


Figura 4.10. Comparación entre la pirámide poblacional europea y la española en 2017. Fuente: Elaboración propia

En el gráfico anterior podemos apreciar en tonalidad rojiza la población de mujeres y en color azulado la población de hombres. La pirámide corresponde a los datos de Europa, y las dos franjas interiores son los datos correspondientes a la población española. Lo habitual era que, como su propio nombre indica, tuviera forma de pirámide, dado que los fallecimientos se distribuían, como dijimos anteriormente, a lo largo de todas las edades. Podemos observar que la mayor parte de la población se concentra en los 50 años aproximadamente, es decir, gente nacida durante el llamado “baby boom”. La pirámide refleja más los efectos de las variaciones en la natalidad (conflictos, crisis...) que la propia mortalidad.

Dado que la esperanza de vida ha aumentado, y cada vez más porcentaje de personas alcanzan edades avanzadas, en unos años nos encontraremos con un problema demográfico importante, dado que habrá mucha gente jubilada y pocos jóvenes cubriendo los puestos de trabajo, por lo que el sistema de pensiones se podría ver desbordado. He aquí la razón por la que cada vez será mas importante tener un plan privado de jubilación.

Por otro lado, la forma de la pirámide es similar tanto para hombres como para mujeres, sobre todo a edades tempranas. Según van pasando los años, observamos que las mujeres tienden a vivir más años que los hombres y fallecen menos a edades jóvenes.

Centrándonos en España, también podemos apreciar cómo la pirámide poblacional se está invirtiendo. El número de nacimientos lleva descendiendo desde 2009. El número medio de hijos por mujer se sitúa en 2019 en 1,23. Si entramos en detalle respecto al dato

anterior, el número medio de mujeres de origen español es 1,17 mientras que de mujeres de origen extranjero alcanza el 1,59. Este hecho, junto con la gran cantidad de personas extranjeras principalmente jóvenes que vienen a España ha ocasionado que la pirámide no envejeciera tan rápido como podría haber sido. Aún así, dado que el número de nacimientos en España ha disminuido notablemente y hay mucha gente de edad avanzada, en 2019 hubo un crecimiento vegetativo poblacional negativo de 57.146. [19].

Pirámide de fallecidos en España vs Europa en 2017

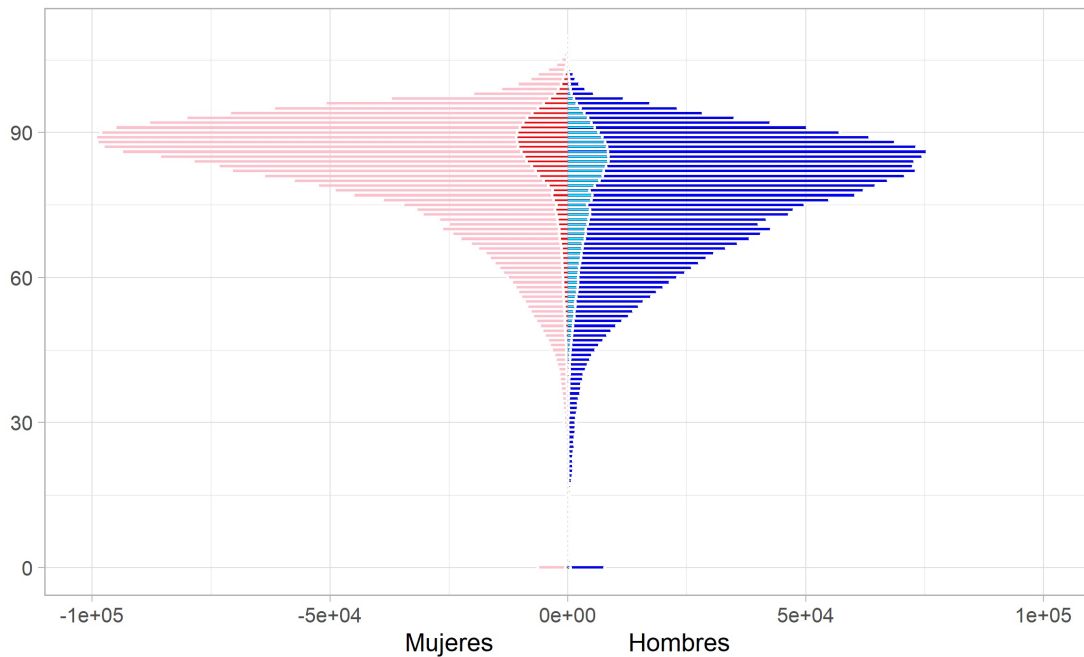


Figura 4.11. Comparación entre los fallecimientos en Europa y en España por edades en 2017.

Fuente: Elaboración propia

La tasa de fallecimiento en niños y adolescentes, exceptuando los recién nacidos, es despreciable. Por otro lado, en este gráfico se observa claramente como la mayor tasa de mortalidad la alcanzan los hombres a edades más tempranas que las mujeres, tanto en el caso español como en el europeo.

De todo el estudio anterior podemos concluir que existe una tendencia en la longevidad que lleva a la población a vivir cada vez más años (hasta cierto umbral) no sólo en España, sino en toda Europa. Ya sea antes o después se ha alcanzado un estado de bienestar en todos los países de nuestro continente. Debido a esta tendencia, como veremos en el siguiente capítulo, la aplicación de un factor constante para “estresar” la longevidad a la hora de calcular el capital de solvencia obligatorio no concuerda con lo observado en el presente estudio. Por otro lado, aunque los fallecimientos cada vez se produzcan más tarde, seguimos encontrando una barrera al llegar a los 110 años.

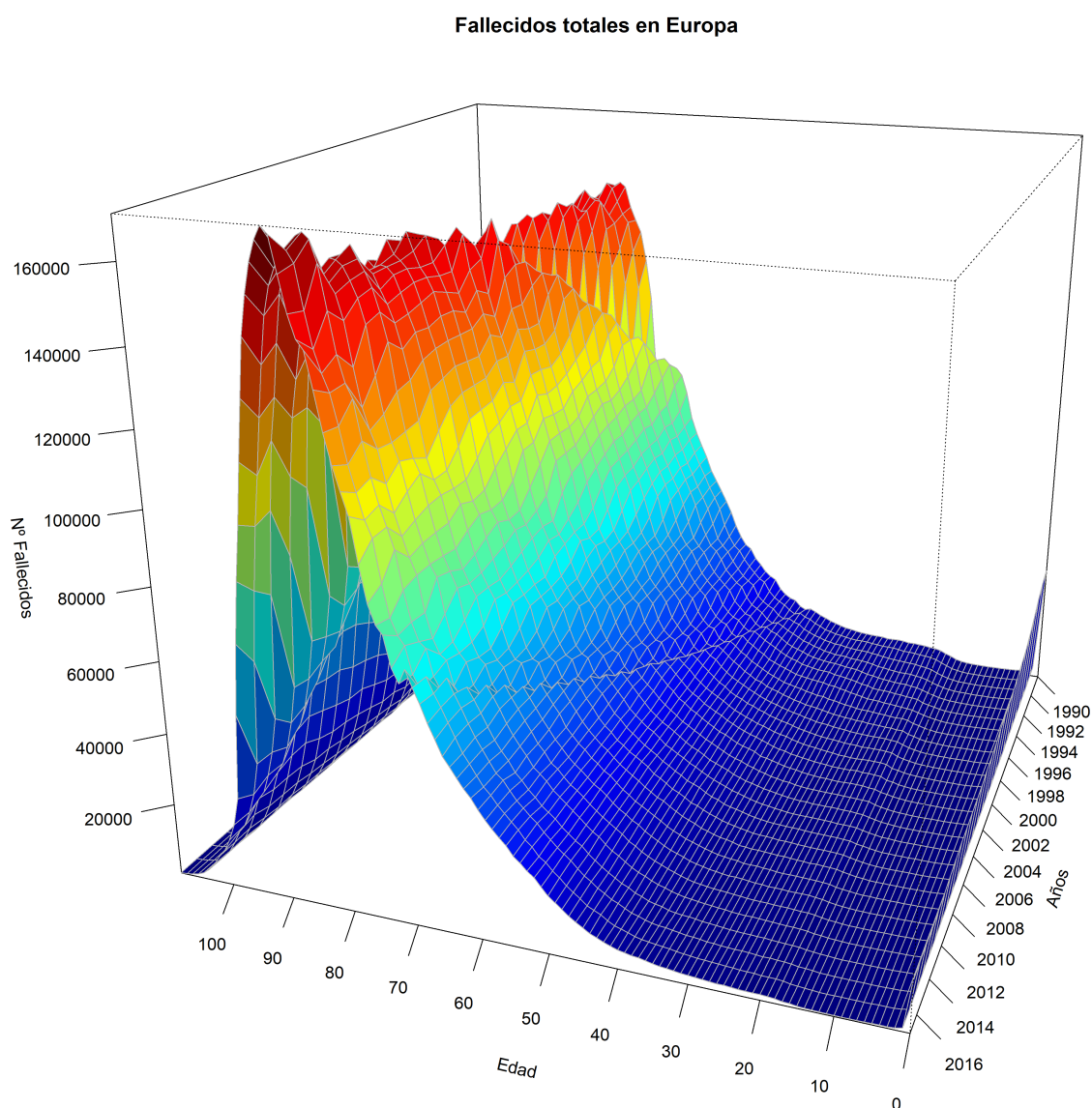


Figura 4.12. Decesos en Europa desde 1990. Fuente: Elaboración propia

Dado que de Europa únicamente tenemos datos a partir de 1990, si observamos el número de fallecidos, encontramos que es muy uniforme. La mortalidad en neonatos ya se ha reducido enormemente y los picos de mortalidad se producen cerca de los 90 años. Las grandes mejoras en la mortalidad se produjeron entre los años 70 y los 90, aunque a día de hoy sigamos apreciando la misma tendencia en la longevidad.

Hay que tener en cuenta la integración de países del este que incorporan con cierto retraso los efectos de la universalización de la sanidad avanzada.

5. PROPUESTA DE SHOCK DE LONGEVIDAD

5.1. Solvencia II y el *shock* de longevidad

A la hora de calcular la mortalidad base, el método más comúnmente utilizado por las compañías aseguradoras es mediante los datos basados en su propia experiencia. Este procedimiento no siempre es recomendable. Por ejemplo, si una cartera está compuesta principalmente por nuevos asegurados, su esperanza de vida no tiene por qué coincidir con la observada anteriormente en la empresa.

Existen diversas estrategias para el cálculo de la mejora de la mortalidad. Uno de los procedimientos consiste en adoptar tablas generacionales publicadas. Si se desea ser aún más conservador, se puede incorporar un factor de mejora extra sobre dichas tablas. Otra opción consiste en realizar un estudio de la mortalidad similar al realizado en el capítulo anterior. Además, siempre se puede recurrir a un “juicio experto”.

Para evaluar los riesgos presentes en una compañía de seguros y calcular sus necesidades de capital, se puede utilizar un modelo interno o la fórmula estándar proporcionada por la Comisión Europea. El desarrollo de un modelo interno es muy costoso debido a su gran sofisticación, pero los resultados alcanzados presentan una gran precisión. La fórmula estándar, mucho más sencilla de implementar, sirve para darnos una idea aproximada del negocio. Esta fórmula busca un equilibrio entre la fácil implementación para las entidades aseguradoras y la precisión de los resultados obtenidos. El enfoque que se dio a la fórmula estándar es conservador, de este modo cubre la exposición al riesgo de forma general para todos los tipos de cartera de cualquier aseguradora.

Antes de continuar explicando las fórmulas anteriores, conviene definir el término “Valor en Riesgo” (en adelante VaR, en inglés “*Value at Risk*”). El VaR es una técnica estadística que nos indica con qué probabilidad sufriremos una determinada pérdida en un periodo temporal concreto. Determina la pérdida máxima que se puede experimentar en dicho intervalo de tiempo.

Centrándonos en la fórmula estándar, ésta consiste en aproximar el VaR con un nivel de confianza del 99,5 % y un horizonte temporal de un año. A continuación, se combinan en una matriz de correlación y se obtiene la cantidad final de capital que la empresa deberá tener para poder hacer frente a sus obligaciones.

Para medir el riesgo de longevidad, el VaR se aproxima mediante la variación en el patrimonio neto que se produce al aplicar un *shock* constante de reducción de un 20 % a las tablas de mortalidad esperada, sin importar la edad, el género ni la duración del seguro.

Para calcular el SCR de longevidad en primer lugar debemos cuantificar la incertidumbre de la mortalidad, que dependerá del tamaño de la cartera a estudiar. Para ello, utilizaremos técnicas estándar de intervalos de confianza.

A continuación se cuantifica el riesgo de tendencia de la longevidad con intervalo temporal superior al año. Para ello, se utilizan modelos estocásticos y se asume la hipótesis de que existe cierta correlación entre el riesgo de mortalidad y el de longevidad.

A continuación estudiaremos si el *shock* de longevidad de la fórmula estándar está fundamentado actuarialmente o es una simplificación conservadora de lo que debiera ser. A priori, tras haber estudiado la evolución histórica de la longevidad, parece que este impacto constante del 20 % no es la mejor de las opciones. La longevidad va cambiando de forma progresiva, por lo que el salto aplicado para el cálculo de este SCR a las tasas de mortalidad no concuerda con la evolución histórica de los datos.

Además, existen documentos publicados desde EIOPA que defienden que un cambio gradual en las tasas de mortalidad, dependiendo de la edad del asegurado y del periodo que cubra su seguro, sería más correcto que el *shock* constante que utiliza la fórmula estándar. Esta imprecisión de la fórmula estándar podría ocasionar errores a la hora de valorar la carga del SCR. Además, el riesgo de no cumplir con los niveles de capital necesarios podría ser significativamente mayor al establecido en Solvencia II.

Por tanto, veremos si el margen que proporciona el impacto constante de la fórmula estándar respecto a la longevidad es insuficiente, apropiado, conservador o muy conservador. Además, no debemos olvidar la importancia de que se pueda aplicar de forma sencilla este *shock*.

Pasando ahora a los modelos internos, tener uno no exime a la compañía aseguradora de utilizar la fórmula estándar. Debido a ello los modelos internos se desarrollan de tal forma que son capaces de producir resultados adicionales utilizando las bases de la fórmula estándar. Dado que la fórmula estándar no parece recoger de forma adecuada el *shock* de longevidad, sería recomendable para las empresas utilizar modelos internos.

En este trabajo se sugiere utilizar un *shock* que dependa de la edad y del sexo del asegurado, así como de la duración residual del contrato de seguro analizado. Por ello, estudiaremos el impacto de la longevidad en la fórmula estándar.

Las entidades aseguradoras se enfrentan al riesgo de longevidad cuando ofertan seguros de capital diferido o contratos de rentas vitalicias. En caso de ofertar dichos productos, la asignación de capital que deberá realizar la entidad deberá tener muy en cuenta la posible desviación de la longevidad.

5.2. Aplicación de los modelos

En primer lugar, agruparemos los datos de todos los países estudiados anteriormente en una nueva base de datos. Esto nos permitirá trabajar de forma más cómoda a la hora de modelar la longevidad en Europa. Ver [código 5](#).

Para utilizar los modelos nos servimos de los datos históricos desde 1990 (dado que es el primer año en el que todos los países tienen datos) hasta 2011 correspondientes a

Europa. Mediante esta base de datos que contiene todos los países europeos, podemos proyectar la tasa mortalidad futura hasta 2018 y obtener el índice europeo de mortalidad. Esta tasa de mortalidad dependerá del sexo de la persona, de su edad y del año de calendario en el que se encuentre el individuo. Es decir, las suposiciones hechas a la hora de realizar los cálculos están basadas en observaciones históricas de la mortalidad.

A la hora de estudiar los modelos, se pueden coger los datos de distinta manera. Por ejemplo podríamos coger datos de personas aseguradas en toda Europa, o datos de una cartera concreta de asegurados. En este trabajo utilizaremos la experiencia recogida en los datos de la *HMD*, sin pérdida de generalidad, para obtener mejoras futuras de la mortalidad que después aplicaremos a la probabilidad de fallecimiento base. Es decir, utilizaremos el conjunto de toda la población, ya sean personas con seguros contratados o no.

Dado que las personas son cada vez más longevas, se ha incrementado en las aseguradoras la preocupación por el riesgo que conlleva un aumento en la esperanza de vida de los asegurados. Una manera eficaz de estudiar y gestionar este riesgo es utilizando un índice de proyección de longevidad, herramienta muy útil a la hora de mitigar la tendencia del riesgo de longevidad. Para un mercado concreto, como pueda ser el mercado de la población europea, se puede estandarizar una métrica de consumo de capital, medido sobre valores poblacionales. Dicho de otra manera, el uso de los datos poblacionales de la *HMD* no nos quita precisión a la hora de obtener resultados realistas sobre posibles personas que contraten seguros de vida.

Por ello podemos decir que un índice que mida la tendencia de supervivencia de una población es equivalente a un índice de longevidad aplicado a un grupo de personas aseguradas. Mediante la edad y el sexo de las personas podremos explicar las tasas de mortalidad, las de supervivencia o la esperanza de vida al nacer.

Por todo esto, en este capítulo nos centraremos en el desarrollo de un índice de longevidad europeo que mida la evolución del riesgo longevidad teniendo en cuenta la tendencia creciente que ha tenido en los últimos tiempos en los diversos países que componen Europa. Para ello, nos serviremos de los factores de mejora que obtenemos con cada uno de los modelos.

Estos sofisticados modelos nos proporcionan datos precisos y robustos, con los que podremos estimar el índice europeo de longevidad (al que en adelante llamaremos “*ELI*”, *European Longevity Index*). Valoraremos la mortalidad y obtendremos un modelo interno para evaluar el impacto que puede tener un incremento de la longevidad en las entidades aseguradoras que se pueda utilizar como alternativa a la fórmula estándar establecida en Solvencia II. Lo haremos con un nivel de confianza del 99,5 %, acorde con el nivel de precisión requerido en la normativa vigente. Además, utilizaremos un escenario sin estresar y buscaremos que se pueda realizar una gestión eficaz del riesgo de supervivencia y de la transferencia del mismo a la hora de trabajar con carteras de seguros de rentas vitalicias.

A continuación tenemos el estudio realizado con los tres modelos seleccionados. Para la aplicación del modelo de Lee-Carter y para el CBD utilizamos la librería de R “St-MoMo” [20] mientras que para el modelo de P-splines utilizamos la librería “Mortality Smooth” [21].

La librería StMoMo tiene la siguiente estructura:

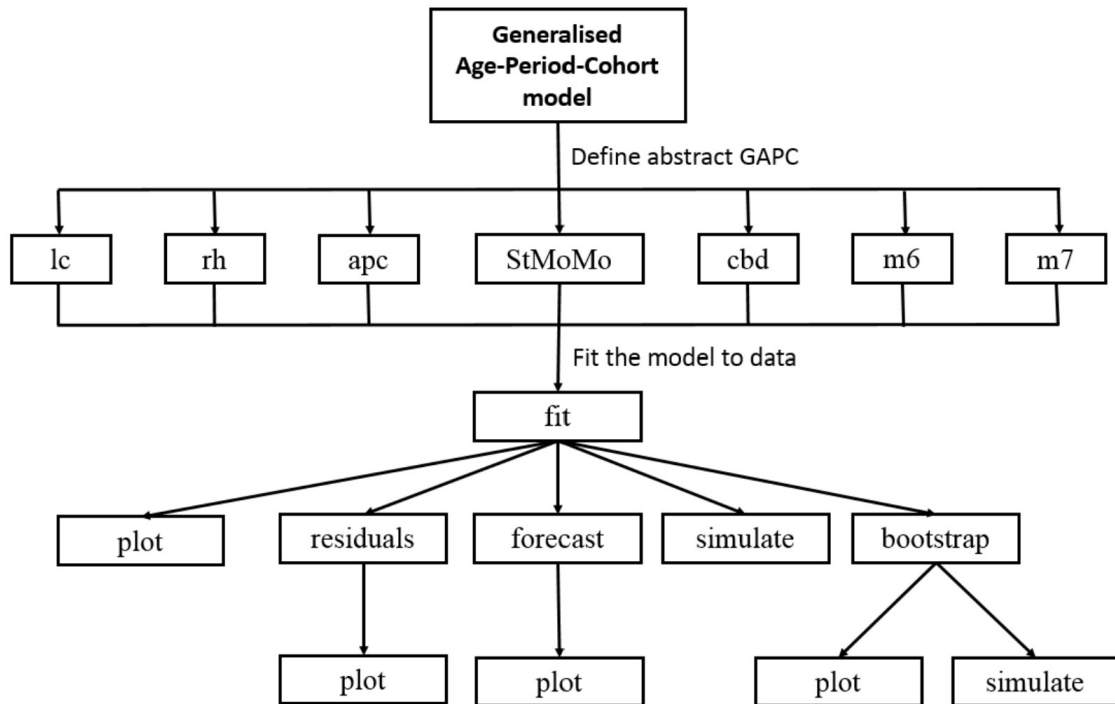


Figura 5.1. Estructura del paquete “StMoMo”. Fuente: [22]

Es decir, definiremos los GAPC como “lc” y “cbd” respectivamente. Además, gracias a este paquete es muy cómodo realizar simulaciones.

Para ver la forma de aplicar estos modelos y el modelo de suavizado por P-splines, ver [código 12](#).

5.2.1. Lee-Carter

Como vimos en el capítulo 3, el estudio se ha realizado bajo la hipótesis de que el número de fallecidos de una cierta edad x en un cierto año t sigue un proceso de Poisson.

Recordemos que el modelo Lee Carter se define de la siguiente manera:

$$\ln[m(x, t)] = \alpha_x + \beta_x \times \kappa_t$$

Aplicando este modelo a los datos de Europa obtenemos los siguientes parámetros.

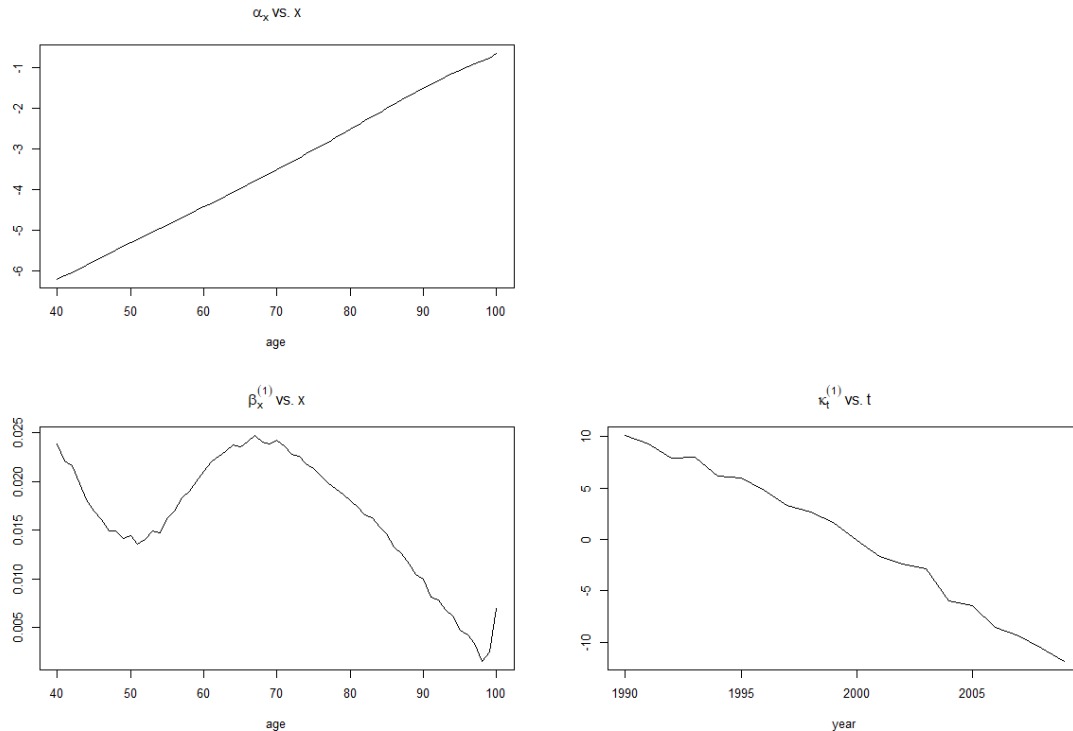


Figura 5.2. Parámetros de Lee-Carter para Europa. Fuente: Elaboración propia

En este gráfico se puede observar lo siguiente:

- El parámetro α_x nos indica la forma que tiene nuestro modelo. Se han seleccionado edades de personas adultas, entre los 40 y los 100 años, dado que la mayoría de los productos de seguros de vida se orientan a personas de más de 40 años. El objetivo es estudiar el impacto en las pensiones, el comportamiento de este parámetro al aplicar una escala logarítmica es lineal.
- El parámetro β_x nos indica a qué edades la mortalidad ha mejorado con mayor rapidez. En este caso podemos observar que las mayores mejoras se han producido a los 40 años y entre los 60 y los 70 años.
- El parámetro κ_t representa el descenso en la mortalidad. Como se puede apreciar, también nos da un resultado bastante lineal. A partir de este parámetro podemos, proyectar la tendencia de la mortalidad (o dicho de otro modo, hacer un “forecast”).

A continuación, tenemos la proyección obtenida sobre la tendencia de la mortalidad. Como era de esperar, esta tendencia continuará siendo descendiente con el paso de los años.

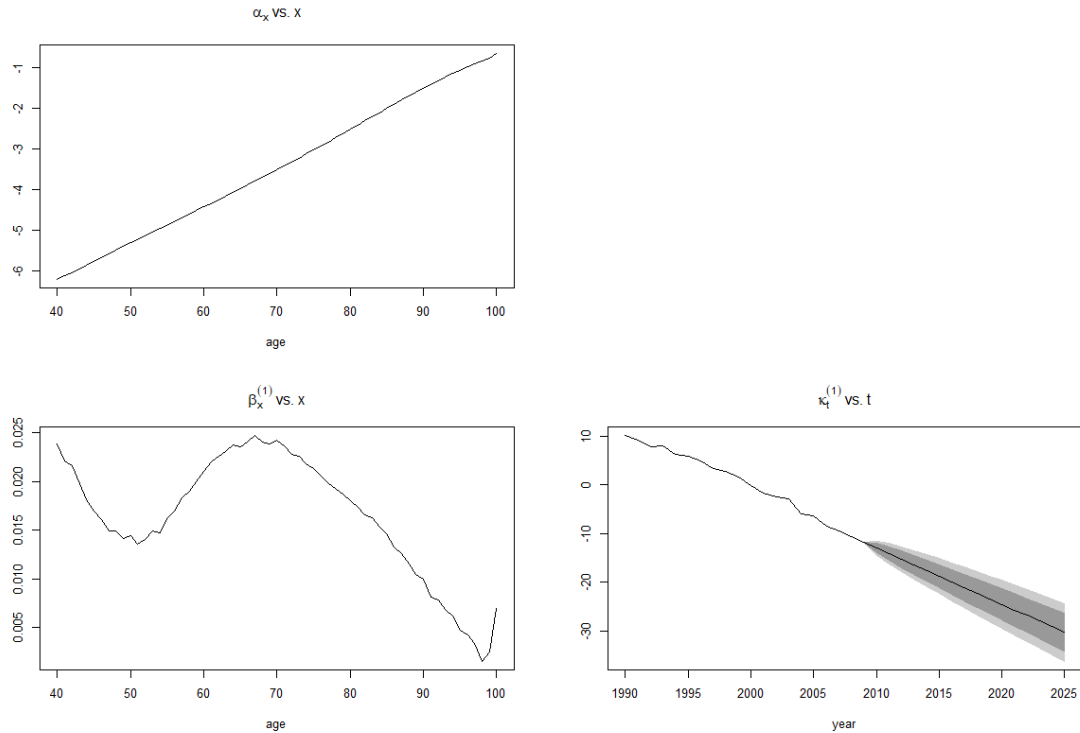


Figura 5.3. Forecast del modelo Lee-Carter para Europa. Fuente: Elaboración propia

Para hacer un análisis exploratorio del ajuste del modelo podemos observar los residuos (posibles errores de estimación).

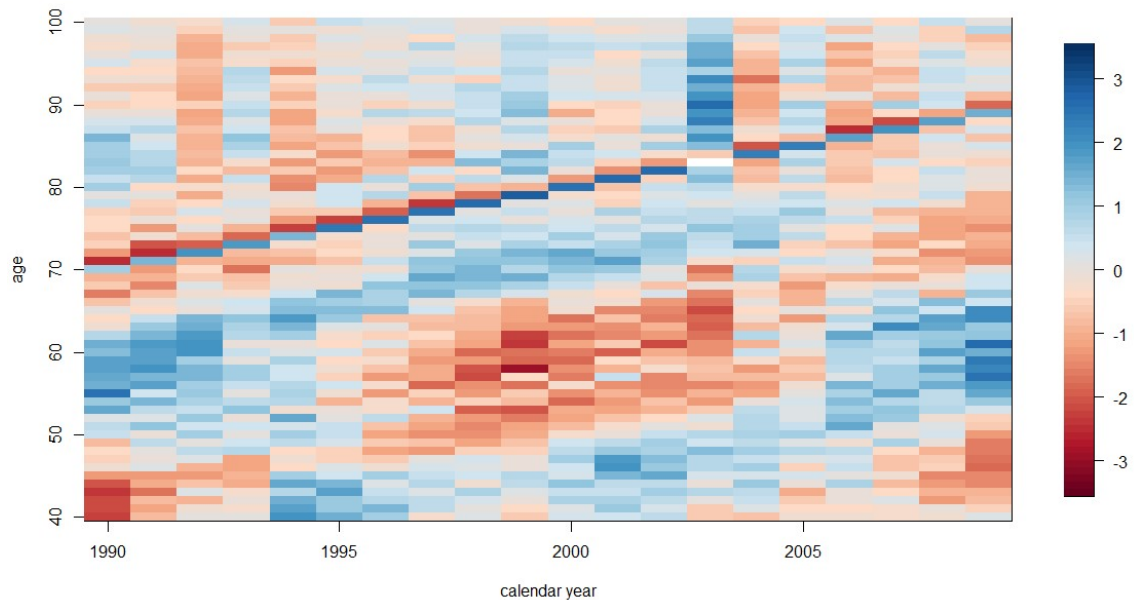


Figura 5.4. Residuos de Lee-Carter para hombres en Europa. Fuente: Elaboración propia

En este caso apreciamos agrupaciones, por ejemplo, de valores negativos entre los años 1996 y 2003 y edades entre 50 y 60 años. Dado que las concentraciones de valores no son especialmente notorias, a primera vista no descartaríamos un buen ajuste del modelo.

Para hacernos una mejor idea de lo que nos podemos encontrar en el futuro, hemos realizado una serie de proyecciones de escenarios para distintas edades (1.000 escenarios por simulación). De este modo podemos observar cómo se puede comportar la mortalidad en un futuro cercano y actuar en consecuencia.

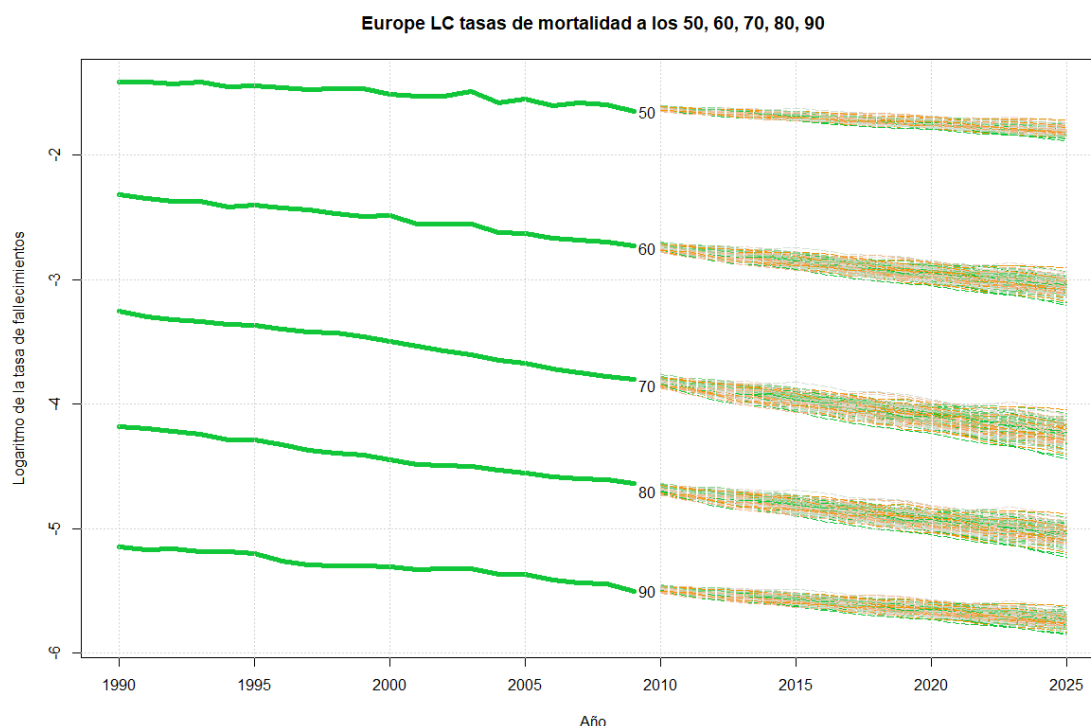


Figura 5.5. Simulación de Lee-Carter para hombres en Europa. Fuente: Elaboración propia

Como se puede apreciar en el gráfico anterior, la tendencia se mantiene igual que en los últimos años, es decir, la mortalidad cada vez disminuyó más y, por lo tanto, la población cada vez alcanzará edades más avanzadas.

En resumen, hemos utilizado un Modelo Lineal Generalizado para definir una familia Edad-Periodo-Cohorte Generalizada de modelos. En nuestro caso no hemos tenido en cuenta la cohorte, únicamente hemos modelado la mortalidad en base a la edad de la persona y al periodo en el que nos encontremos.

Gracias al paquete “StMoMo” hemos podido ajustar el modelo a los datos históricos, comprobar la calidad de dicho ajuste, proyectar a futuro y realizar simulaciones de posibles escenarios que vendrán. Todo ello nos permite obtener los factores de mejora para la creación del “ELI”, como veremos más adelante.

5.2.2. CBD

En primer lugar, recordemos que el modelo CBD se define como:

$$\ln\left(\frac{{}_tq_x}{{}_tp_x}\right) = \beta_x^{(1)} \times \kappa_t^{(1)} + \beta_x^{(2)} \times \kappa_t^{(2)}$$

Al aplicar el modelo a nuestra base de datos obtenemos los siguientes parámetros:

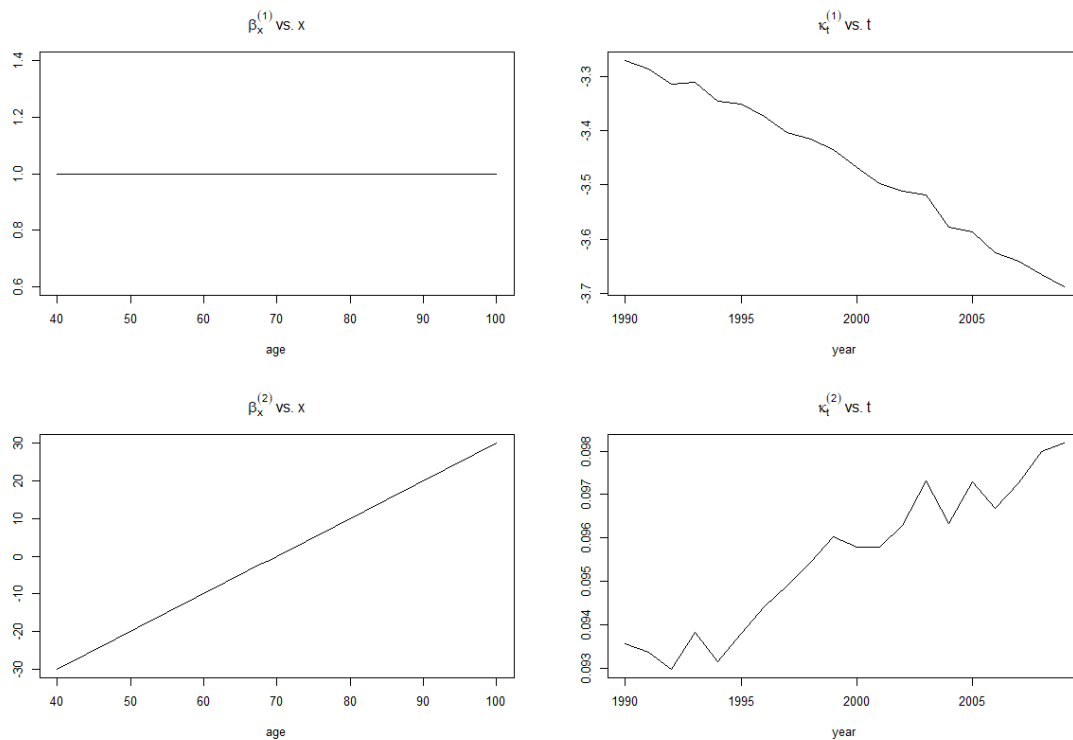


Figura 5.6. Parámetros CBD para Europa. Fuente: Elaboración propia

Estos parámetros nos proporcionan la siguiente información:

- $\beta_x^{(1)}$: tiene valor constante 1, como dijimos en el capítulo 3
- $\beta_x^{(2)}$: es el resultado de, a la edad x , restarle la media entre la edad más joven considerada y la más antigua. Esto nos devuelve, como era de esperar, una función lineal creciente.
- $\kappa_t^{(1)}$: este parámetro se puede interpretar como el “nivel” de la mortalidad. Tiene una tendencia decreciente debido a que refleja la mejora en las tasas de la mortalidad a lo largo del tiempo.
- $\kappa_t^{(2)}$: observamos un crecimiento gradual de este parámetro, llamado coeficiente de la pendiente. Su tendencia se debe a que refleja el hecho de que, históricamente, la mejora de la mortalidad se aprecia de forma más notoria a edades más bajas. La tasa de mejora de la mortalidad a edades avanzadas es escasa.

Procediendo de igual manera, obtenemos las siguientes proyecciones de los parámetros κ . Como era de esperar, la tendencia de estos parámetros se mantiene.

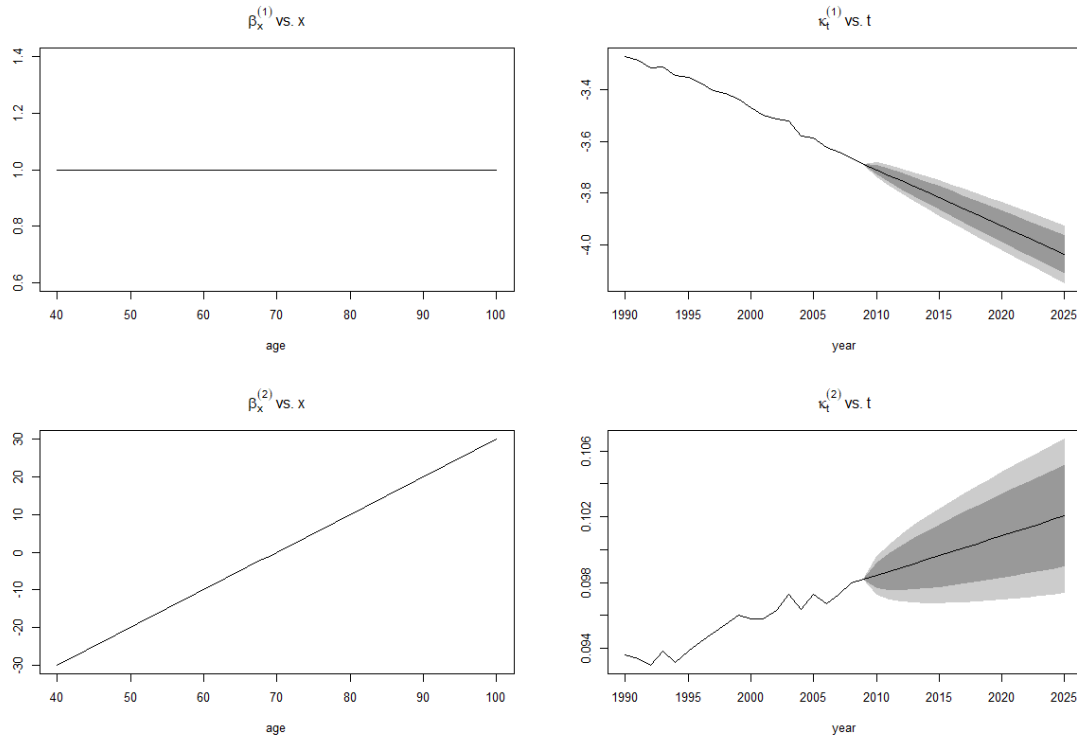


Figura 5.7. Forecast del modelo CBD para Europa. Fuente: Elaboración propia

Al obtener los residuos observamos mayores agrupaciones de valores que en el caso de Lee Carter. Esto nos induce a pensar que el modelo de Lee Carter ajusta mejor que el modelo CBD. Este hecho se observa mejor en 5.15

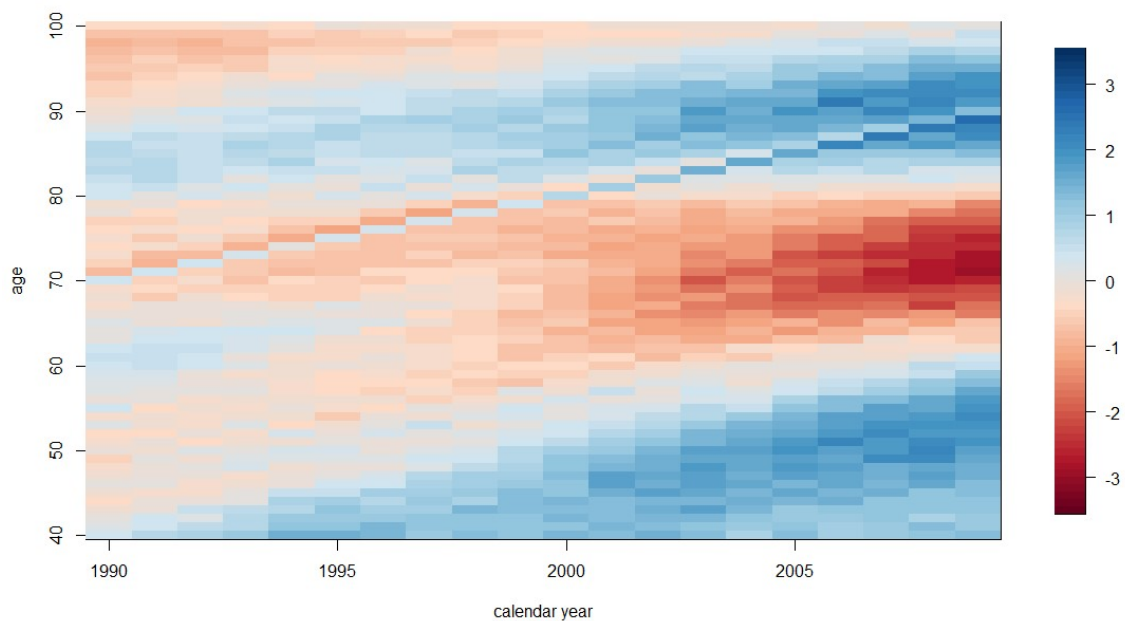


Figura 5.8. Residuos de CBD para hombres en Europa. Fuente: Elaboración propia

Por último, veamos algunos escenarios obtenidos mediante simulaciones estocásticas (1.000 escenarios por caso).

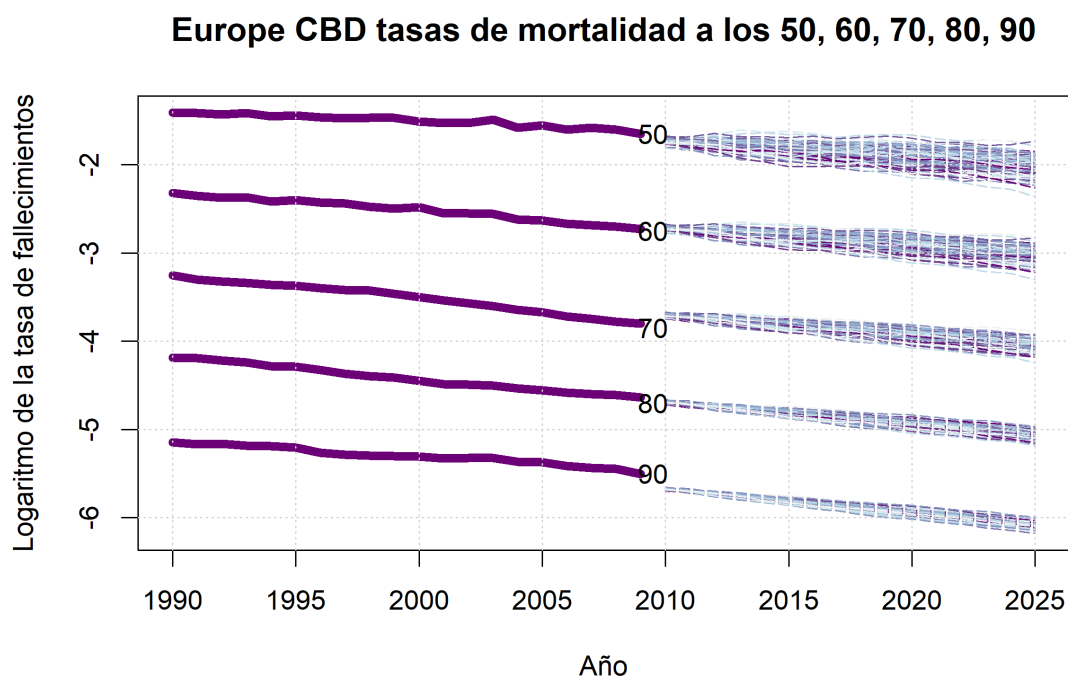


Figura 5.9. Simulación de CBD para hombres en Europa. Fuente: Elaboración propia

Nuevamente, la tendencia es alcanzar edades cada vez más avanzadas. Por otro lado, cabe destacar que la personas de más edad, por mucho que mejore la mortalidad, lo hace de forma mucho más suavizada que a edades tempranas.

5.2.3. P-splines

En primer lugar, veamos el suavizado de los datos obtenido mediante este modelo. En la izquierda tendríamos los datos recogidos directamente de la base de datos y a la derecha el suavizado realizado.

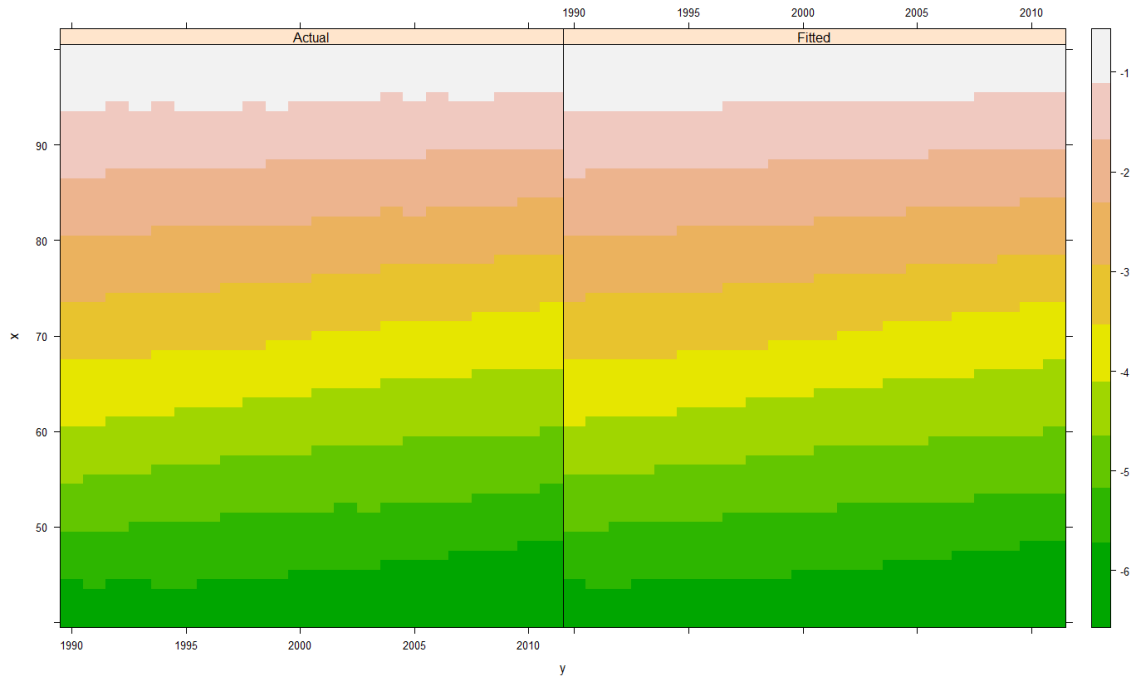


Figura 5.10. Suavizado P-spines. Fuente: Elaboración propia

La proyección obtenida con este modelo, comparando con los datos reales sería:

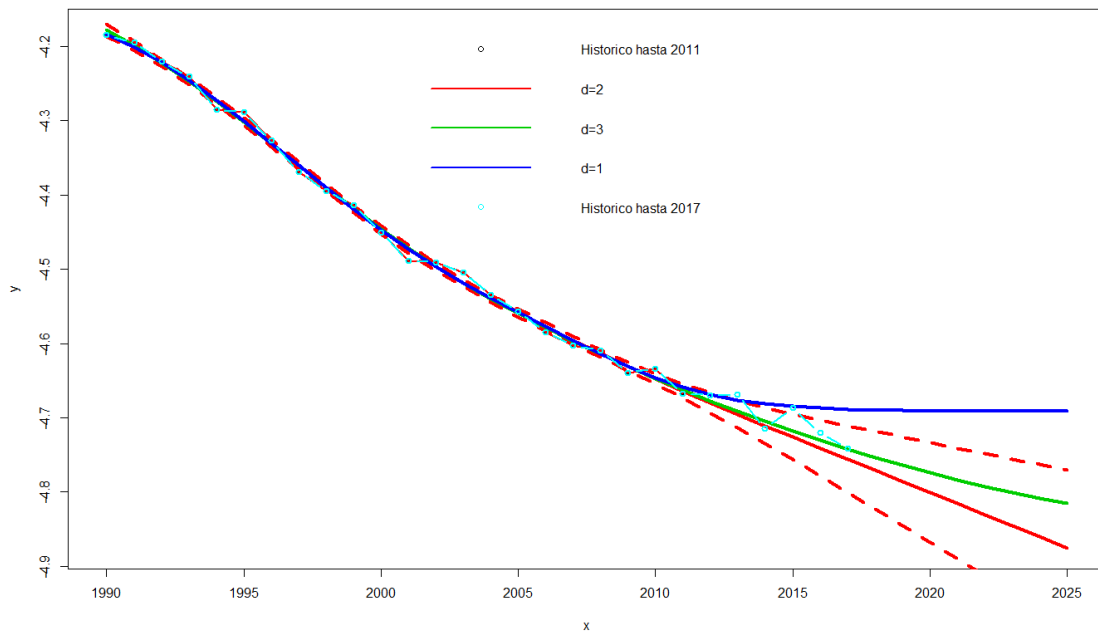


Figura 5.11. Proyección P-spines. Fuente: Elaboración propia

En el caso de $d = 1$ la proyección no se ajusta a los datos reales. Por otro lado, tanto para $d = 2$ como para $d = 3$ el resultado se encuentra dentro del intervalo de confianza, siendo el caso de $d = 3$ el que mejor ajusta a los datos reales.

Si observamos los residuos:

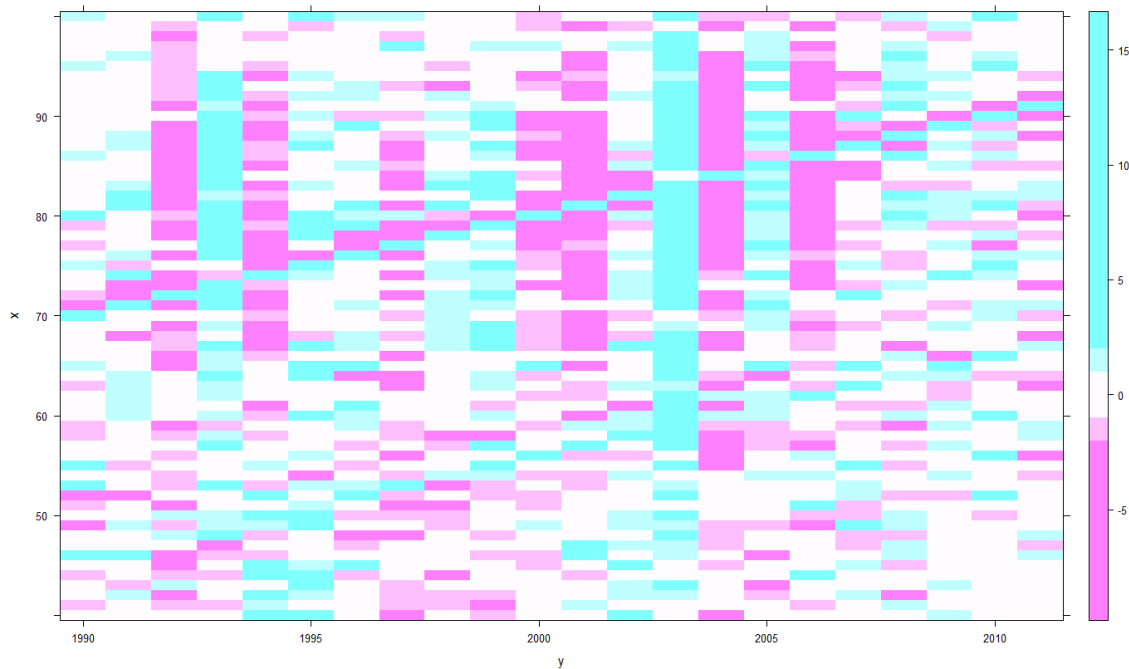


Figura 5.12. Residuos P-spines. Fuente: Elaboración propia

No hay grandes agrupaciones de datos, por lo que podemos decir que el ajuste del modelo es satisfactorio.

5.3. Resultados

Se han utilizado como datos reales los de miHMDdata para Europa desde 1990 hasta 2011, ambos inclusive. Existen datos reales hasta 2017 (en HMD hay, en el momento de realizar este trabajo, datos más recientes de algunos países, pero no de todos), y se utilizarán para comparar los resultados de las simulaciones y valorar cuanto se desvían de la realidad. La proyección se realiza hasta 2025: entre 2018 y 2025 únicamente podremos comparar entre sí los modelos, pero nos permitirá tener mayor perspectiva de cómo las proyecciones divergen al alejarnos de los años con datos reales.

La comparativa comienza seleccionando una edad y representado los datos reales y datos de las simulaciones, en la misma escala. En los siguientes gráficos tenemos este resultado de las proyecciones de los tres modelos para la edad de 84 años, edad que corresponde con la esperanza de vida al nacer en 2019 en España.

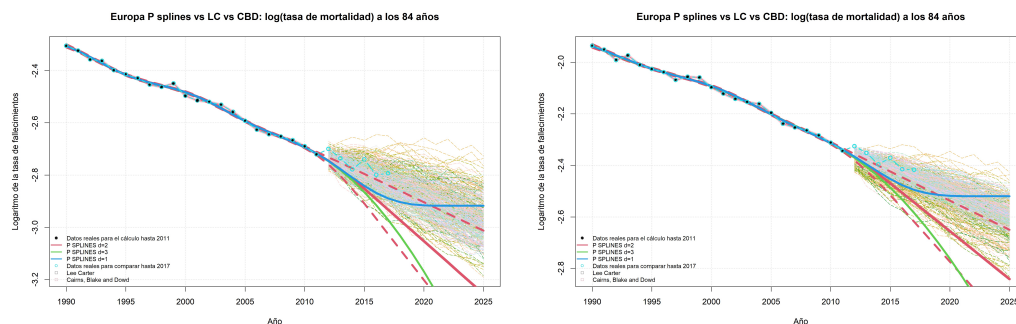


Figura 5.13. Comparación de los 3 modelos para la edad de 84 años. Mujeres a la izquierda y hombres a la derecha. Fuente: Elaboración propia

Un primer análisis de los datos induce a pensar que el modelo de P-splines es el que mejor ajusta a los datos reales, seguido del modelo de Lee Carter.

Efectivamente el modelo de P-splines parece ajustar con gran precisión en los casos de $d = 1$ y $d = 3$. Por su parte, las simulaciones obtenidas con el modelo CBD son demasiado optimistas respecto a la longevidad que se podría alcanzar o, dicho de otro modo, la mortalidad descende mucho más que lo experimentado realmente. Por su parte, el modelo de Lee Carter quedaría entre ambos modelos, siendo mejor que el CBD pero no ajustando con tanta precisión como el P-splines para esta edad.

Este análisis sería correcto para la edad mostrada en el gráfico. Mientras que el modelo de Lee Carter y el de CBD se comportan de manera homogénea en todas las edades, el modelo de P-splines no resulta demasiado estable; para algunas edades, en función de los años que se hayan cogido como datos para la simulación, los resultados se alejan de la realidad.

Pasemos a analizar los 4 métodos con los que estudiamos la longevidad. Estos son:

1. Método tradicional: Solvencia II define la aplicación del 20 % de mejora respecto del último año de datos reales.
2. Método de Lee Carter: Aplicando este modelo, mediante las funciones de la librería StMoMo se realizan simulaciones y se obtienen las tasas de fallecimiento, resultado de calcular la media de los valores obtenidos en las 1000 iteraciones empleadas en la simulación.
3. Método de CBD: procedimiento análogo al caso anterior.
4. Método de P-splines: Aplicando el suavizado de P-splines mediante las funciones de la librería MortalitySmooth, en concreto la función (*predict*).

Los siguientes gráficos representan las proyecciones obtenidas hasta 2025.

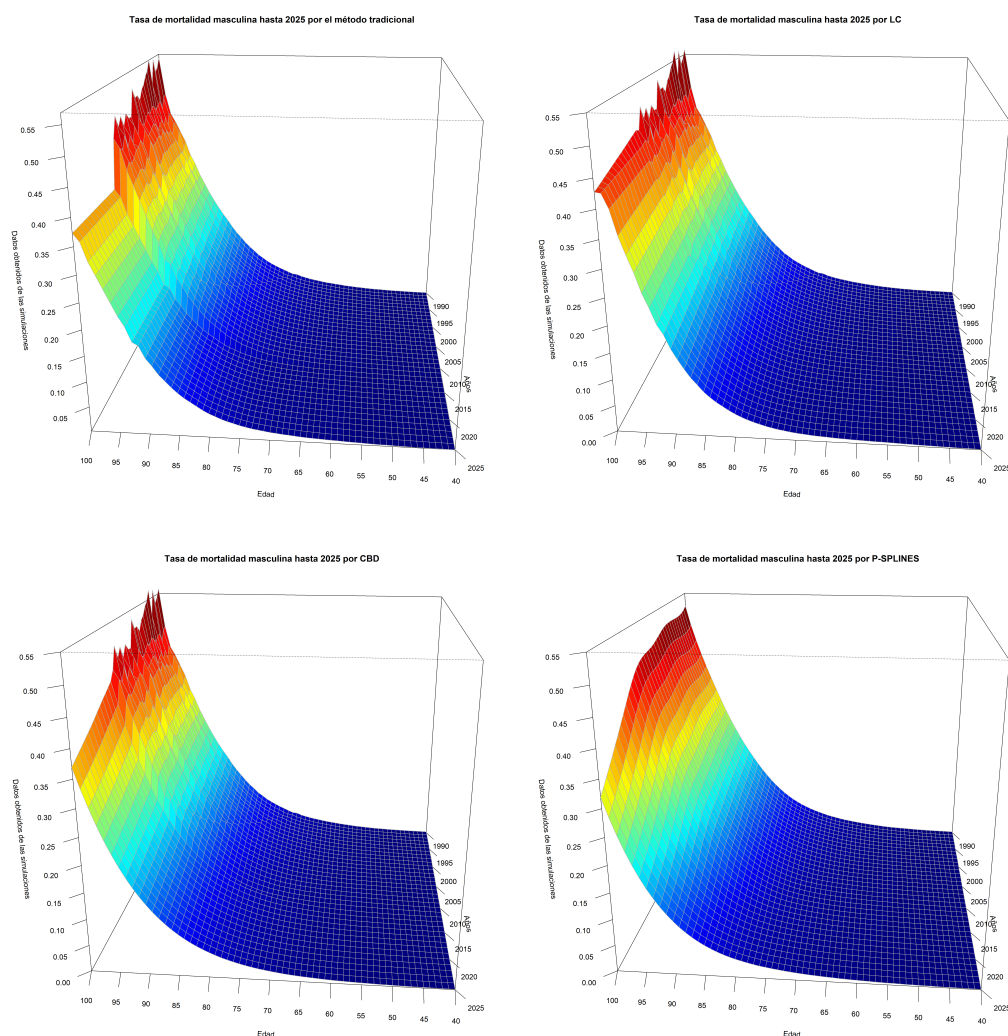


Figura 5.14. Proyecciones de los 4 modelos. Fuente: Elaboración propia

En el método tradicional podemos observar un salto significativo cuando pasamos del último año con datos reales al primero de los proyectados, y a partir de ahí el valor se mantiene constante, que es justamente la crítica de partida a este modelo.

En el caso de Lee Carter apreciamos continuidad entre los datos históricos y los proyectados. La tendencia está dentro de lo esperado, por lo que el modelo es satisfactorio.

El modelo de CBD manifiesta un salto entre los datos reales y la proyección, similar a lo que ocurriría en el modelo tradicional, aunque dicha variación es menos notoria que en el primer caso, y únicamente se aprecia en edades avanzadas.

Por último, mediante el modelo de suavizado con P-splines desaparecen los picos que se observaban en el resto a la edad de 100 años. Esto se debe a la interpolación que realiza el propio modelo. La superficie que tenemos es uniforme, si bien se observan ondulaciones incluso en los primeros años.

Resulta significativo que las variaciones se deban más al avance en el “eje edad” que al de los “años transcurridos”. Hay que fijarse en las edades de 80 a 95 años de 2025 para apreciar que la concavidad es más o menos aguda según el modelo.

A continuación vamos a estudiar la calidad de los modelos calculando los errores cometidos a la hora de proyectar los datos. Los errores absolutos se calculan haciendo la diferencia entre los datos reales y los datos simulados (entre 2012 y 2017), mientras que los errores relativos son los que se obtienen al dividir el error absoluto entre el valor real. A la hora de comparar unos con otros será mucho más significativo estudiar los errores relativos dado que, además de tener en cuenta la diferencia de valores, tienen en consideración la magnitud que se está comparando:

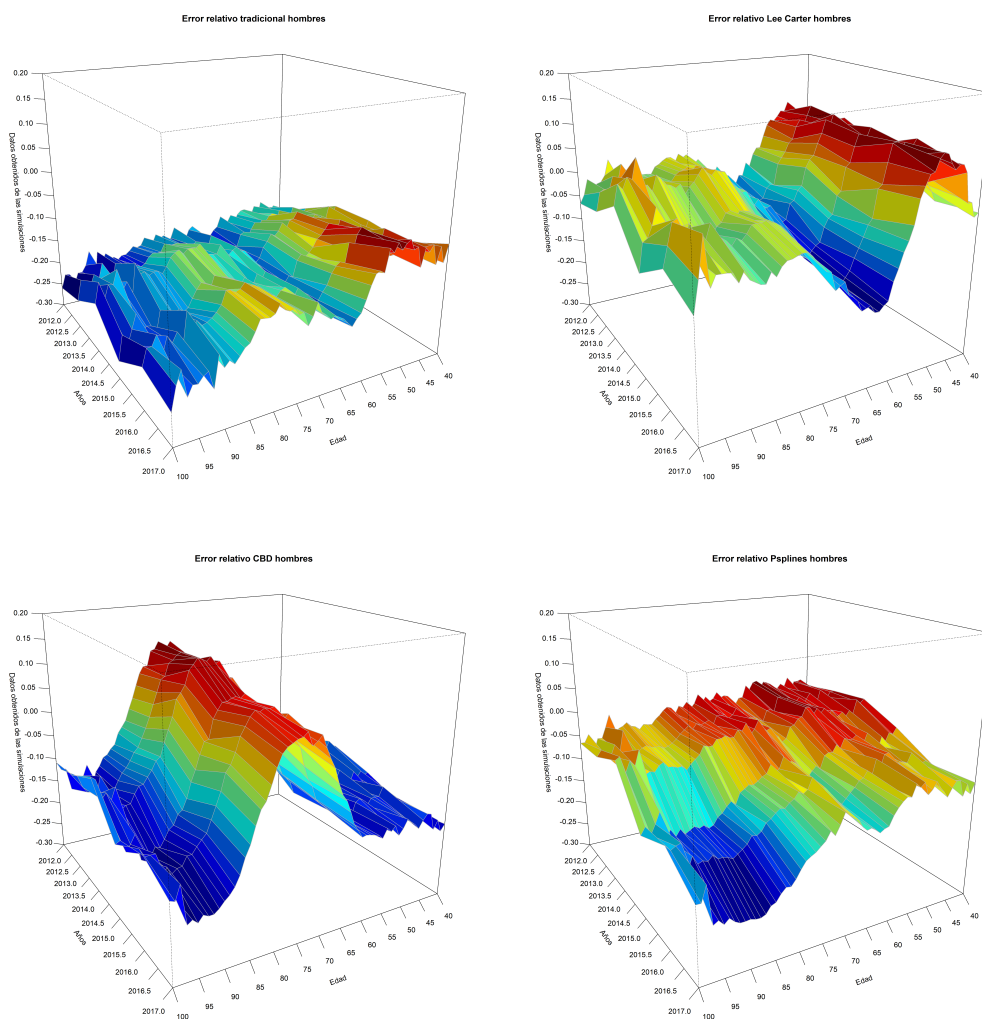


Figura 5.15. Errores relativos (hombres). Fuente: Elaboración propia

Aplicando el método de Solvencia II, obtenemos unos errores relativos negativos respecto a los datos reales cercanos al -20 %. No obstante, es significativo que el error se reduzca cuanto más nos alejamos de los datos reales y crece en función de la edad, aunque no de forma monótona.

Por su parte el modelo de Lee Carter comete errores relativos en el intervalo entre el +10 % y el -10 %. Comete errores por exceso en las primeras décadas de edades consideradas y errores por defecto alrededor de los 60 años. En las edades más avanzadas presenta los mínimos errores.

Respecto al modelo de CBD los errores se encuentran entre +15 % y -15 %, alcanzándose el máximo entre los 65 y los 75 años de edad y el mínimo antes de los 60 y a partir de los 85. Una vez más el año de proyección no afecta significativamente, mientras que la edad es sumamente relevante.

Por último, al usar P-splines el signo del error relativo siempre es negativo, en torno al -5 %, y alcanza su mínimo puntualmente en un -15 %. En este caso tanto la edad como el año de proyección son significativos, alcanzando su peor estimación a los 80 años de edad y en los últimos años proyectados.

A continuación, compararemos los modelos entre ellos. Calculando su diferencia y relativizándola dividiendo por el modelo tomado como referente.

En los gráficos de la página siguiente, los de la columna de la izquierda corresponden a la comparación del método establecido por Solvencia II con el modelo Lee Carter, CBD y P-splines respectivamente. A la derecha, de arriba a abajo, comparamos el modelo de Lee Carter con el CBD, el Lee Carter con el P-splines y el CBD con el P-splines.

LC vs Tradicional oscila entre $\pm 20\%$ de forma, siendo positivo en 2012 y negativo según nos acercamos a 2025. En este caso la edad influye poco.

CBD vs Tradicional oscila con la edad, alcanzando máx a los 75 y mínimos a los 55. De nuevo alcanza valores por encima de 0 % en 2012, llegando casi al 30 %, y por debajo de 0 % cerca de 2025, rozando el 15 %

P-splines vs tradicional, prácticamente plano en 2012 y oscila en 2025 con máximo en 70 años y mínimo en 55 y 85 años. En 2012 el error supera el 20 % y en 2025 está en el -30 %.

Respecto a la columna de la derecha, LC vs CBD forman un plano ondulado que oscila entre $\pm 20\%$, alcanzándose el máximo a los 70 años.

LC vs P-splines oscila de forma similar al anterior, si bien el error apenas supera el 3 %, mientras que a edades avanzadas llega a bajar por debajo del 30 %.

Por último el CBD vs P-splines oscila con mayor suavidad, de nuevo entre $\pm 20\%$

Los modelos parecen más precisos en el centro que en los extremos, por lo que al comparar con el tradicional es donde se encuentran los errores. Además, según nos alejamos de los datos históricos, la diferencia toma signo negativo, es decir, el tradicional tiene valores mayores al modelo.

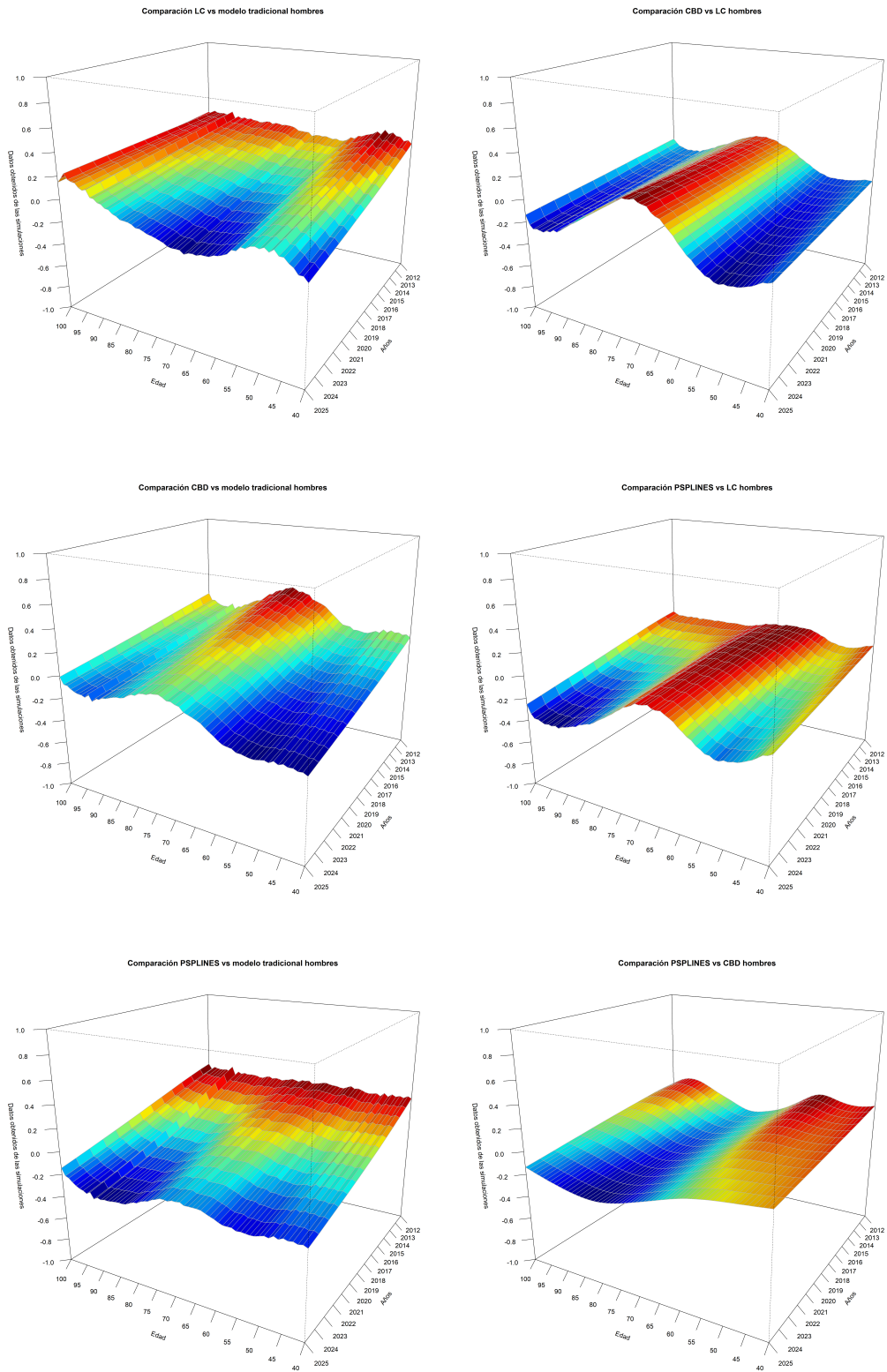


Figura 5.16. Mejora de cada uno de los modelos. Fuente: Elaboración propia

5.4. Índice Europeo de Longevidad

A la hora de determinar para una cierta edad un valor concreto de la tendencia a largo plazo de la longevidad que le sería de aplicación, nos encontramos ante una elección que no debería ser subjetiva. Para intentar tomar un criterio lo más objetivo posible, hacemos simulaciones de caminos estocásticos, es decir, estudiamos diversos escenarios futuros posibles que tienen en cuenta distintas causas de fallecimiento (accidental, por enfermedad o natural). Este procedimiento resulta de gran utilidad a la hora de estimar la tendencia de forma ajustada.

La proyección la hemos realizado en base a los datos recogidos en la *HMD* y almacenados en nuestra base de datos `miHMDdata[["Europe"]]` entre 1990 y 2011. Gracias a esto podemos obtener, con una confianza del 99,5 %, los valores de 2012 a 2025. Con estos resultados podemos calcular el factor de mejora de cada modelo. El objetivo es comparar dichos resultados y valorar si al aplicarlos ponderadamente por igual se obtiene un índice mejor o peor que el tradicionalmente establecido en Solvencia II, analizando en qué casos tiene un signo u otro.

Recordemos que el factor de mejora ${}_t\lambda_x$ se define como:

$${}_t\lambda_x = 1 - \frac{{}_tq_x}{{}_{t-1}q_x}$$

Y mediante estos valores obtenidos, podemos establecer un factor de mejora medio que sea creciente geométricamente para los próximos años. Así pues, para obtener la mortalidad en el último periodo, $\hat{\mu}(x, f)$, aplicamos la siguiente fórmula:

$$\hat{\mu}(x, f) = \hat{\mu}(x, t) \times (1 + {}_t\lambda_x)^{f-t}$$

Siendo f el último periodo observado, es decir, 2025, y t el año de referencia, es decir, el último dato histórico utilizado antes de realizar la proyección, en nuestro caso 2011.

Despejando el valor ${}_t\lambda_x$ de la ecuación anterior obtenemos:

$${}_t\lambda_x = \left(\frac{\hat{\mu}(x, f)}{\hat{\mu}(x, t)} \right)^{\frac{1}{f-t}} - 1$$

Por tanto, aplicándolo a los resultados obtenidos, los factores de mejora que alcanzamos con cada modelo dependiendo del sexo de la persona para edades desde 40 hasta 100 años son los siguientes:

	Lee-Carter		CBD		P-splines	
Edad	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
40	2.25 %	2.73 %	6.43 %	3.86 %	5.57 %	3.57 %
41	2.09 %	2.42 %	6.39 %	3.68 %	4.74 %	3.43 %
42	1.87 %	2.43 %	6.23 %	3.73 %	3.68 %	3.50 %
43	1.99 %	2.09 %	6.51 %	3.59 %	3.08 %	3.40 %
44	1.64 %	1.93 %	6.16 %	3.63 %	1.97 %	3.46 %
45	1.62 %	1.67 %	6.20 %	3.50 %	1.54 %	3.36 %
46	1.56 %	1.67 %	6.13 %	3.64 %	1.29 %	3.53 %
47	1.34 %	1.39 %	5.87 %	3.48 %	1.12 %	3.39 %
48	1.33 %	1.59 %	5.74 %	3.69 %	1.30 %	3.63 %
49	1.27 %	1.41 %	5.73 %	3.60 %	1.78 %	3.56 %
50	1.14 %	1.55 %	5.47 %	3.68 %	2.12 %	3.67 %
51	1.32 %	1.48 %	5.53 %	3.75 %	2.83 %	3.77 %
52	1.38 %	1.54 %	5.30 %	3.75 %	3.20 %	3.80 %
53	1.31 %	1.72 %	4.95 %	3.69 %	3.28 %	3.77 %
54	1.39 %	1.70 %	4.78 %	3.62 %	3.28 %	3.74 %
55	1.56 %	1.98 %	4.51 %	3.62 %	2.82 %	3.78 %
56	1.62 %	1.99 %	4.26 %	3.41 %	2.02 %	3.62 %
57	1.58 %	2.31 %	3.95 %	3.37 %	0.98 %	3.63 %
58	1.75 %	2.45 %	3.60 %	3.32 %	-0.04 %	3.63 %
59	2.04 %	2.57 %	3.49 %	3.13 %	-0.56 %	3.51 %
60	2.06 %	2.61 %	3.28 %	2.90 %	-0.68 %	3.35 %
61	2.10 %	2.77 %	2.85 %	2.72 %	-0.43 %	3.25 %
62	2.16 %	2.79 %	2.50 %	2.51 %	0.31 %	3.13 %
63	2.43 %	2.80 %	2.35 %	2.29 %	1.45 %	3.00 %
64	2.52 %	2.92 %	2.06 %	2.16 %	2.42 %	2.97 %
65	2.55 %	2.88 %	1.70 %	2.07 %	3.08 %	2.97 %
66	2.75 %	2.92 %	1.66 %	1.97 %	3.69 %	2.97 %
67	2.75 %	3.08 %	1.33 %	1.89 %	3.68 %	2.99 %
68	2.82 %	2.86 %	1.15 %	1.63 %	3.56 %	2.84 %
69	2.88 %	2.86 %	1.01 %	1.58 %	3.29 %	2.89 %
70	2.88 %	2.88 %	0.71 %	1.47 %	2.75 %	2.88 %
71	2.79 %	2.83 %	0.62 %	1.49 %	2.39 %	3.00 %
72	2.79 %	2.66 %	0.64 %	1.35 %	2.17 %	2.96 %
73	2.74 %	2.58 %	0.49 %	1.25 %	1.90 %	2.96 %
74	2.63 %	2.49 %	0.50 %	1.29 %	2.01 %	3.10 %
75	2.55 %	2.41 %	0.53 %	1.31 %	2.42 %	3.20 %
76	2.55 %	2.37 %	0.68 %	1.42 %	3.28 %	3.39 %
77	2.44 %	2.26 %	0.67 %	1.42 %	4.20 %	3.47 %

	Lee-Carter		CBD		P-splines	
Edad	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
78	2.34 %	2.17 %	0.74 %	1.47 %	5.20 %	3.58 %
79	2.23 %	2.18 %	0.85 %	1.63 %	6.11 %	3.78 %
80	2.09 %	2.02 %	1.00 %	1.63 %	6.73 %	3.82 %
81	2.02 %	1.94 %	1.20 %	1.77 %	6.95 %	3.96 %
82	2.00 %	1.92 %	1.40 %	1.94 %	6.78 %	4.14 %
83	1.91 %	1.91 %	1.55 %	2.02 %	6.31 %	4.20 %
84	1.86 %	1.75 %	1.68 %	2.03 %	5.73 %	4.19 %
85	1.73 %	1.67 %	1.84 %	2.10 %	5.25 %	4.22 %
86	1.65 %	1.47 %	1.93 %	2.19 %	4.90 %	4.26 %
87	1.44 %	1.45 %	1.99 %	2.30 %	4.74 %	4.32 %
88	1.35 %	1.27 %	2.07 %	2.34 %	4.79 %	4.29 %
89	1.27 %	1.23 %	2.20 %	2.50 %	5.01 %	4.39 %
90	1.18 %	1.05 %	2.19 %	2.47 %	5.17 %	4.29 %
91	1.29 %	1.04 %	2.53 %	2.76 %	5.68 %	4.51 %
92	0.86 %	0.66 %	2.07 %	2.30 %	5.40 %	3.98 %
93	0.81 %	0.74 %	2.18 %	2.60 %	5.59 %	4.19 %
94	0.76 %	0.64 %	2.17 %	2.49 %	5.52 %	4.01 %
95	0.70 %	0.44 %	2.09 %	2.48 %	5.20 %	3.91 %
96	0.65 %	0.42 %	2.04 %	2.40 %	4.67 %	3.76 %
97	0.49 %	0.35 %	1.75 %	2.28 %	3.68 %	3.54 %
98	0.35 %	0.08 %	1.58 %	2.10 %	2.64 %	3.28 %
99	0.47 %	0.36 %	1.40 %	2.09 %	1.50 %	3.18 %
100	0.52 %	0.61 %	1.14 %	1.65 %	0.29 %	2.66 %

Tabla 5.1. Mejora de la mortalidad. Fuente: Elaboración propia

Con el objetivo de mitigar el riesgo de modelo, realizamos la media de estos valores por edad. Gracias a esto obtenemos el factor de mejora promedio o conjunto de manera conservadora, sin comprometer los resultados por la utilización de un único modelo.

Por tanto, los factores de mejora esperados, con horizonte en 2025, son:

Edad	Media	
	Mujeres	Hombres
40	3.35 %	2.25 %
41	3.11 %	2.08 %
42	2.73 %	2.16 %
43	2.75 %	2.00 %
44	2.20 %	2.02 %
45	2.10 %	1.88 %
46	2.01 %	2.02 %
47	1.84 %	1.85 %
48	1.89 %	2.09 %
49	2.09 %	1.99 %
50	2.09 %	2.11 %
51	2.45 %	2.19 %
52	2.51 %	2.22 %
53	2.43 %	2.24 %
54	2.43 %	2.22 %
55	2.26 %	2.31 %
56	1.98 %	2.19 %
57	1.53 %	2.26 %
58	1.10 %	2.28 %
59	0.97 %	2.20 %
60	0.87 %	2.05 %
61	0.83 %	1.99 %
62	0.98 %	1.87 %
63	1.37 %	1.74 %
64	1.57 %	1.71 %
65	1.59 %	1.67 %
66	1.81 %	1.65 %
67	1.65 %	1.66 %
68	1.56 %	1.46 %
69	1.44 %	1.46 %
70	1.15 %	1.41 %

Edad	Media	
	Mujeres	Hombres
71	1.01 %	1.46 %
72	0.96 %	1.34 %
73	0.80 %	1.29 %
74	0.81 %	1.33 %
75	0.93 %	1.36 %
76	1.22 %	1.46 %
77	1.40 %	1.47 %
78	1.62 %	1.50 %
79	1.82 %	1.64 %
80	1.97 %	1.61 %
81	2.10 %	1.70 %
82	2.17 %	1.83 %
83	2.13 %	1.88 %
84	2.07 %	1.85 %
85	2.02 %	1.87 %
86	1.97 %	1.88 %
87	1.91 %	1.95 %
88	1.94 %	1.93 %
89	2.05 %	2.04 %
90	2.06 %	1.96 %
91	2.40 %	2.18 %
92	2.00 %	1.73 %
93	2.12 %	1.98 %
94	2.12 %	1.88 %
95	2.04 %	1.84 %
96	1.93 %	1.79 %
97	1.56 %	1.69 %
98	1.24 %	1.53 %
99	0.92 %	1.61 %
100	0.45 %	1.27 %

Tabla 5.2. Mejora de la mortalidad, media. Fuente: Elaboración propia

Con el fin de aplicar este índice en los distintos seguros la mortalidad sería, teniendo en cuenta la mejora calculada:

$${}_tq'_x = {}_tq_{x+1} \times (1 - {}_t\lambda_x)$$

Por tanto, si queremos saber la probabilidad de fallecimiento de una persona de edad $x + 1$, entonces:

$${}_tq'_{x+1} = {}_tq_{x+1} \times (1 - {}_t\lambda_x) \times (1 - {}_t\lambda_{x+1})$$

Y, de forma general, si la madurez del contrato se alcanza en el periodo n , tenemos:

$${}_tq'_{x+n} = {}_tq_{x+n} \times (1 - {}_t\lambda_x) \times (1 - {}_t\lambda_{x+1}) \times \dots \times (1 - {}_t\lambda_{x+n})$$

Si analizamos cómo ha variado la probabilidad de fallecimiento proyectada respecto la real con la misma escala con que se representaban los 4 modelos 5.14, obtenemos lo siguiente:

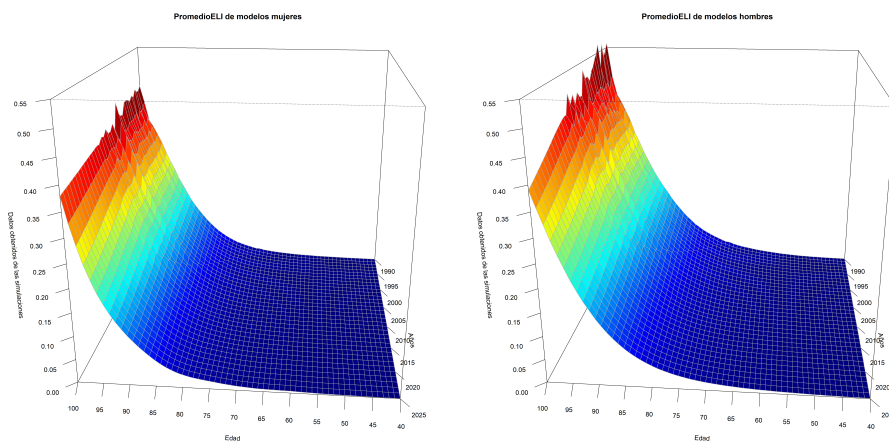


Figura 5.17. Evolución de la tasa de mortalidad. Fuente: Elaboración propia

Si representamos las secciones de dicho índice en distintos periodos de tiempo podemos compararlo con el resultado obtenido mediante la aplicación del 20 % constante de Solvencia II. Centrándonos en los datos masculinos esto sería:

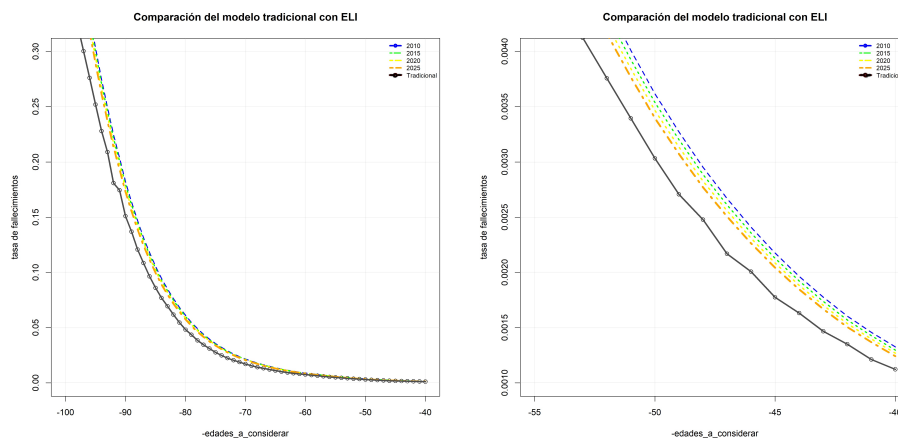


Figura 5.18. Sección de la evolución de la tasa de mortalidad. Fuente: Elaboración propia

En estas secciones transversales sobre la figura 5.17 se muestran las diferencias entre nuestro *ELI* y el 20 % de Solvencia II. Dado que para edades entre 40 y 55 esta variación se aprecia poco debido a la escala utilizada, en el gráfico de la derecha focalizamos los resultados a dichos años y se observa que el comportamiento es similar, y coincide con el esperado: el modelo tradicional plantea un escalón que tiene un efecto significativo el primer año de la proyección respecto del último año con datos reales. En el modelo promediado se mantiene la continuidad en ese momento, representado mediante la curva más distante. Según nos alejamos en el tiempo respecto de los datos reales, el modelo promediado de las estimaciones obtenidas mediante Lee-Carter, CBD y P-splines se va aproximando al planteado por el método tradicional.

Recordemos que, bajo el marco de Solvencia II, el impacto que se aplica sobre la longevidad consiste en una reducción del 20 % de la mortalidad esperada respecto de la mortalidad base. Dicho de otra forma, se compara el número de fallecidos al final del periodo de la mortalidad histórica sobre la población esperada:

$$Shock_longevidad = \frac{{}^tq_{x+n}}{{}^tq'_{x+n}} - 1$$

Dependiendo de la duración del seguro y la edad del asegurado, el *shock* que deberemos aplicar será:

Edad	1 año		5 años		10 años		Vitalicia	
	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
40	3.46 %	2.30 %	10.16 %	7.05 %	17.84 %	13.71 %	70.47 %	61.02 %
41	3.21 %	2.13 %	9.24 %	6.80 %	16.41 %	13.48 %	65.78 %	58.28 %
42	2.81 %	2.20 %	8.35 %	6.73 %	15.30 %	13.44 %	61.59 %	55.81 %
43	2.83 %	2.04 %	7.65 %	6.51 %	14.62 %	13.34 %	58.09 %	53.34 %
44	2.25 %	2.06 %	6.82 %	6.50 %	13.92 %	13.42 %	54.69 %	51.14 %
45	2.15 %	1.92 %	6.59 %	6.45 %	13.89 %	13.48 %	52.09 %	48.96 %
46	2.06 %	2.06 %	6.45 %	6.59 %	13.98 %	13.73 %	49.67 %	47.01 %
47	1.87 %	1.88 %	6.57 %	6.64 %	14.15 %	13.82 %	47.42 %	44.94 %
48	1.93 %	2.14 %	6.99 %	6.93 %	14.43 %	14.12 %	45.43 %	43.13 %
49	2.13 %	2.03 %	7.41 %	6.99 %	14.50 %	14.20 %	43.42 %	41.09 %
50	2.13 %	2.15 %	7.72 %	7.18 %	14.24 %	14.37 %	41.26 %	39.23 %
51	2.51 %	2.24 %	8.03 %	7.31 %	13.80 %	14.38 %	39.16 %	37.32 %
52	2.58 %	2.27 %	7.88 %	7.33 %	12.78 %	14.28 %	36.74 %	35.38 %
53	2.49 %	2.29 %	7.42 %	7.33 %	11.53 %	14.09 %	34.30 %	33.46 %
54	2.49 %	2.27 %	6.73 %	7.34 %	10.31 %	13.81 %	32.01 %	31.57 %
55	2.31 %	2.36 %	5.76 %	7.34 %	9.07 %	13.52 %	29.81 %	29.75 %
56	2.02 %	2.24 %	4.70 %	7.19 %	7.98 %	13.09 %	27.86 %	27.93 %
57	1.55 %	2.31 %	3.71 %	7.11 %	7.27 %	12.72 %	26.21 %	26.26 %
58	1.12 %	2.34 %	3.09 %	6.88 %	7.05 %	12.24 %	25.01 %	24.58 %
59	0.98 %	2.25 %	2.99 %	6.52 %	7.34 %	11.64 %	24.22 %	22.92 %
60	0.88 %	2.10 %	3.18 %	6.16 %	7.83 %	11.07 %	23.57 %	21.40 %

	1 año		5 años		10 años		Vitalicia	
Edad	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
61	0.83 %	2.03 %	3.63 %	5.89 %	8.41 %	10.58 %	23.03 %	20.04 %
62	0.99 %	1.91 %	4.33 %	5.61 %	9.02 %	10.13 %	22.54 %	18.75 %
63	1.39 %	1.78 %	4.98 %	5.42 %	9.43 %	9.72 %	21.93 %	17.59 %
64	1.59 %	1.74 %	5.27 %	5.27 %	9.34 %	9.39 %	21.02 %	16.59 %
65	1.62 %	1.70 %	5.31 %	5.10 %	8.95 %	9.06 %	19.97 %	15.64 %
66	1.84 %	1.68 %	5.19 %	4.93 %	8.45 %	8.75 %	18.91 %	14.74 %
67	1.68 %	1.69 %	4.70 %	4.76 %	7.72 %	8.45 %	17.70 %	13.87 %
68	1.59 %	1.48 %	4.24 %	4.51 %	7.15 %	8.14 %	16.64 %	13.06 %
69	1.46 %	1.48 %	3.71 %	4.41 %	6.73 %	8.01 %	15.72 %	12.40 %
70	1.17 %	1.43 %	3.19 %	4.30 %	6.52 %	7.91 %	14.91 %	11.76 %
71	1.02 %	1.48 %	2.93 %	4.23 %	6.70 %	7.84 %	14.33 %	11.16 %
72	0.97 %	1.36 %	2.88 %	4.13 %	7.14 %	7.75 %	13.88 %	10.60 %
73	0.81 %	1.30 %	2.99 %	4.16 %	7.73 %	7.79 %	13.49 %	10.12 %
74	0.82 %	1.35 %	3.41 %	4.27 %	8.54 %	7.88 %	13.24 %	9.70 %
75	0.93 %	1.38 %	4.03 %	4.40 %	9.36 %	7.92 %	13.03 %	9.27 %
76	1.23 %	1.48 %	4.73 %	4.53 %	10.07 %	7.92 %	12.77 %	8.87 %
77	1.42 %	1.49 %	5.32 %	4.61 %	10.46 %	7.82 %	12.36 %	8.46 %
78	1.64 %	1.53 %	5.86 %	4.74 %	10.60 %	7.70 %	11.87 %	8.06 %
79	1.86 %	1.67 %	6.23 %	4.89 %	10.47 %	7.51 %	11.26 %	7.66 %
80	2.01 %	1.63 %	6.40 %	4.92 %	10.11 %	7.21 %	10.55 %	7.23 %
81	2.14 %	1.73 %	6.41 %	5.00 %	9.56 %	6.87 %	9.77 %	6.87 %
82	2.22 %	1.86 %	6.24 %	5.02 %	8.93 %	6.45 %	8.96 %	6.47 %
83	2.17 %	1.91 %	5.94 %	4.95 %	8.18 %	5.90 %	8.18 %	6.07 %
84	2.11 %	1.88 %	5.64 %	4.82 %	7.43 %	5.30 %	7.44 %	5.66 %
85	2.06 %	1.91 %	5.41 %	4.73 %	6.71 %	4.68 %	6.84 %	5.28 %
86	2.01 %	1.92 %	5.21 %	4.59 %	5.99 %	4.05 %	6.29 %	4.92 %
87	1.95 %	1.99 %	5.14 %	4.48 %	5.27 %	3.42 %	5.83 %	4.60 %
88	1.98 %	1.96 %	5.03 %	4.21 %	4.54 %	2.80 %	5.47 %	4.27 %
89	2.10 %	2.08 %	4.91 %	3.98 %	3.79 %	2.22 %	5.11 %	3.98 %
90	2.11 %	2.00 %	4.69 %	3.65 %	3.03 %	1.69 %	4.76 %	3.65 %
91	2.46 %	2.23 %	4.42 %	3.33 %			4.42 %	3.37 %
92	2.04 %	1.76 %	3.92 %	2.89 %			3.92 %	3.00 %
93	2.17 %	2.02 %	3.53 %	2.62 %			3.63 %	2.84 %
94	2.17 %	1.92 %	3.01 %	2.23 %			3.24 %	2.61 %
95	2.08 %	1.87 %	2.41 %	1.87 %			2.84 %	2.38 %
96	1.96 %	1.82 %					2.36 %	2.14 %
97	1.59 %	1.72 %					1.83 %	1.94 %
98	1.26 %	1.56 %					1.38 %	1.79 %
99	0.93 %	1.63 %					0.93 %	1.63 %

Tabla 5.3. Índice Europeo de Longevidad. Fuente: Elaboración propia

Al representar gráficamente los resultados se facilita la comparación:

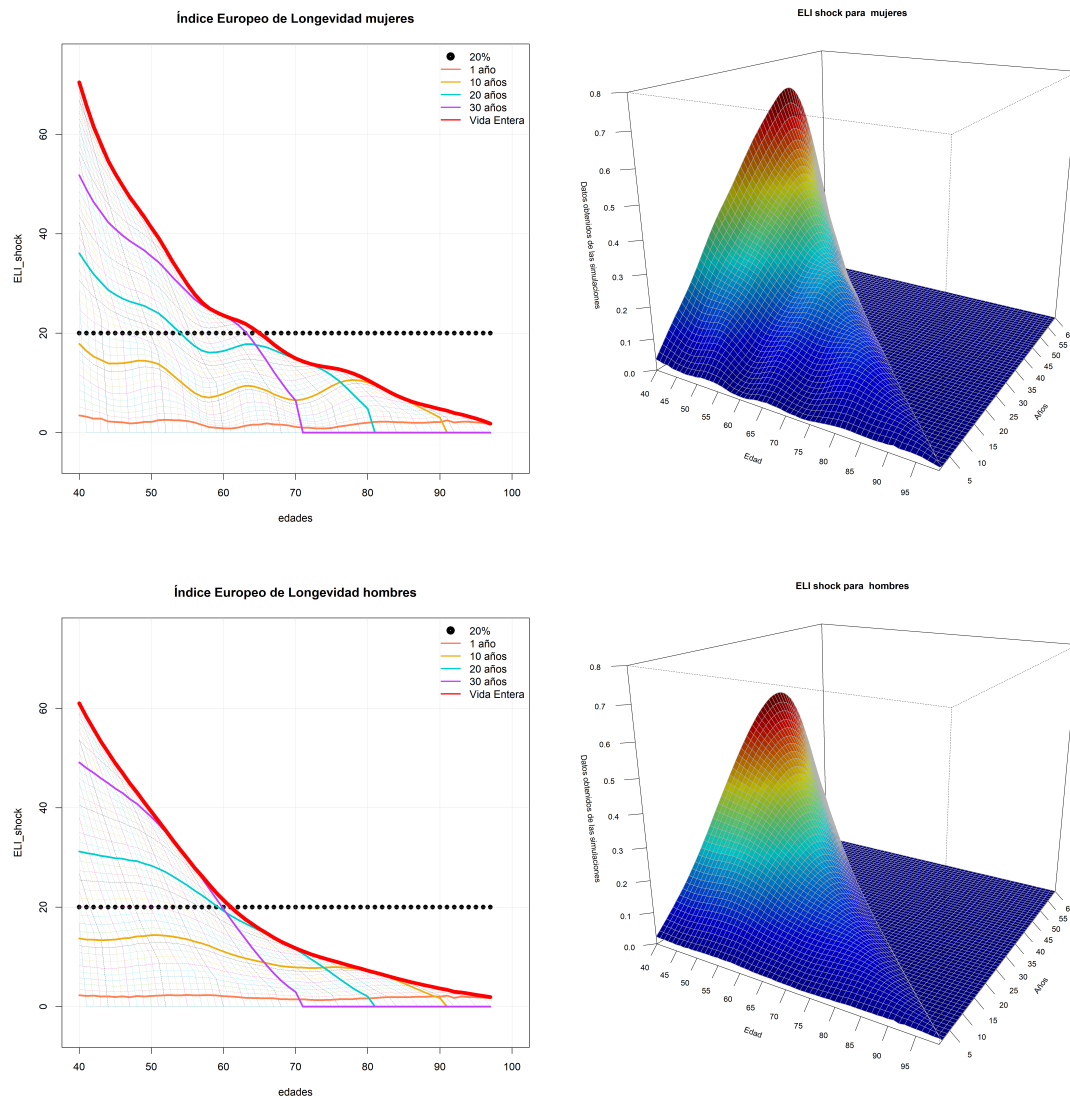


Figura 5.19. Índice Europeo de Longevidad. Fuente: Elaboración propia

El seguro vida entera se ha calculado, sin pérdida de generalidad, acotando la edad máxima que puede vivir un ser humano a 100 años.

Mediante una línea horizontal de puntos negros hemos representado el modelo que corresponde al *shock* propuesto en Solvencia II. De este modo podemos comparar los resultados obtenidos con el valor recomendado en el reglamento. A primera vista es claro que tanto la edad como la duración del seguro son factores determinantes a la hora de aplicar un impacto u otro, siendo este 20 % insuficiente en algunos casos y excesivo en otros. Los resultados también varían considerablemente dependiendo del sexo del asegurado, dado que la esperanza de vida es más elevada en mujeres que en hombres.

En el caso de las mujeres observamos unas ondulaciones significativas en los resultados. Este hecho se debe a que los datos sufrieron pequeñas variaciones en los últimos años antes de realizar la proyección. Por ello, el modelo P-splines, que modeliza fielmente los datos, a la hora de proyectar también mantiene dichas ondulaciones presentes en sus estimaciones, incluso acrecentándolas.

Centrándonos en los resultados obtenidos para las mujeres, si el seguro se contrata alrededor de los 65 años entonces el *shock* aplicable será el estimado en Solvencia II. Para mujeres de 40 años, este resultado no sería suficiente, dado que su esperanza de vida cabe esperar que mejore considerablemente, por lo que la cantidad aplicable debería ser de un 70 %. Para edades superiores el impacto que deberemos aplicar será mucho menor que el utilizado actualmente.

Por su parte, en el caso de los hombres en los seguros vida entera se alcanza *shock* establecido en el reglamento a los 61 años. A los 40 años el impacto aplicable debería ser considerablemente superior, 61 %, mientras que a edades superiores a los 60 dicho valor debería disminuir considerablemente.

En el caso de tratar con seguros con duraciones menores a 14 años en el caso de las mujeres y a 12 años en el caso de los hombres, a la hora de calcular las reservas no sería necesario aplicar un impacto tan elevado a la longevidad para ninguna edad. Para seguros de duraciones superiores a las citadas anteriormente, dependiendo de la edad, entonces deberemos aplicar un mayor o un menor impacto.

6. CONCLUSIONES

Partiendo de consideraciones teóricas, diversos autores han cuestionado la aplicación de un impacto constante a la hora de evaluar el riesgo de longevidad al que se enfrentan las aseguradoras de vida, tal como se propone en Solvencia II.

En base a los datos descargados de la *HMD* hemos construido diversas estructuras de datos en R con las que analizar las series históricas, comparar los países entre sí y relacionarlos en particular con España. Con esto hemos estudiado efectos como la rectangularización o la bajada de la tasa de mortalidad en neonatos. Los datos finalmente nos han permitido calcular la tasa de mortalidad de una Europa constituida por 21 países, que incluyen toda la zona Euro y algunos adicionales significativos, en el intervalo de tiempo en que todos y cada uno de estos países tienen datos en la *HMD*: a fecha de presentar este trabajo, de 1990 a 2017.

Con el fin de obtener una proyección de la mortalidad, hemos analizado teóricamente tres modelos: Lee Carter, CBD y P-splines. Para aplicarlos, hemos instalado en el entorno de *R-Studio* las librerías que nos han permitido realizar estas proyecciones en el tiempo en base a los datos de *HMD* de cada uno de los tres modelos. Con esto hemos comparado los resultados entre sí y frente a los datos reales. Hemos tomado como base del cálculo la probabilidad de fallecimiento por edad entre 1990 y 2011 y como datos de control la información entre 2012 y 2017. A partir de 2018 únicamente podemos comparar los resultados de los modelos entre ellos.

A partir de las proyecciones obtenidas, calculamos los factores de mejora de la mortalidad estimada y construimos un índice de longevidad para la población europea (ELI). Estos parámetros serían una alternativa a aplicar un porcentaje fijo del 20 % respecto del dato del último año del registro histórico. Y, por tanto, la tabla 5.3 completa este trabajo.

A la hora de aplicar el *shock* de longevidad, debemos tener en cuenta la edad del individuo y la duración del seguro a contratar, dado que, como vimos en el capítulo anterior, son los aspectos más significativos. Al realizar los cálculos, los resultados se ven influenciados principalmente por las edades de las personas y por la duración del periodo de tiempo considerado (lejanía respecto de los datos históricos, reales).

Como vimos en el capítulo anterior en el gráfico 5.19, un *shock* superior al 20 % en un seguro vida entera únicamente tendría sentido si el asegurado tiene menos de 65 años en el caso de las mujeres y 61 años en el caso de los hombres. La aplicación del 20 % propuesta en Solvencia II no se sostiene para ninguna de las edades consideradas si la duración del seguro es corta. La calibración propuesta de este valor fijo para la evaluación del riesgo biométrico de longevidad penaliza indebidamente la contratación de dicha contingencia.

Cuando se actualicen los datos post pandemia COVID-19, será recomendable realizar un nuevo estudio y así comprobar el efecto que ha tenido dicho evento en la esperanza de vida. Cabe esperar que el corte con el *shock* de Solvencia II se alcance a edades más tempranas que en el presente estudio, debido a que la esperanza de vida bajará en los próximos años debido a las secuelas que está causando este virus.

Podemos concluir que la aplicación del estrés en la población más joven que desea contratar seguros de larga duración debe ser más elevado que lo recomendado en la normativa de Solvencia II, mientras que en edades avanzadas el valor a aplicar debería ser mucho menor. Esto se debe a que las mejoras en la calidad de vida de los jóvenes, por unas condiciones de higiene generalizadamente buenas, nuevos avances en medicina por la aplicación de nuevas tecnologías, mejoras en el tratamiento y logística de conservación de los alimentos, y las normas y sistemas de seguridad y prevención de riesgos laborales, que tienen un impacto global más significativo sobre la longevidad si se disfrutan toda la vida que si sólo se ha tenido acceso a ellos a partir de cierto momento. Cada vez mayor porcentaje de la población alcanzará edades elevadas, sin que ello tenga que suponer que las vidas se alarguen indefinidamente: Recordemos que biológicamente existen pocos casos de personas supercentenarias de 110 años.

Dicho de otro modo, las personas que ya han llegado a edades avanzadas tienen un periodo vital más breve para disfrutar de todos los beneficios de las mejoras en la calidad de vida que se han ido logrando con el paso del tiempo, y eso se refleja en un menor impacto de las correcciones calculadas, mientras que en el futuro la mayor parte de las personas superarán los 84 años de vida, que es aproximadamente la esperanza de vida al nacer en países como España o Italia. Pero serán pocos los que vivan otros 15 años más.

Por tanto, un cálculo del SCR de vida (en particular respecto del riesgo de longevidad) que tenga en consideración un mayor nivel de detalle y, por tanto, unos cálculos efectuados con mayor precisión, tendrá efectos positivos para las empresas aseguradoras del ramo vida porque podrán ajustar mejor sus provisiones.

Por último, podemos sacar conclusiones sobre cómo afectaría la aplicación de los resultados de la tabla 5.3 como mecanismo de transferencia del riesgo vía reaseguro por parte de las entidades aseguradoras. Para las compañías reaseguradoras esto supone un vehículo adecuado de compensación de riesgos tomándolos en el espacio europeo. Si nos centramos en el caso de las aseguradoras españolas, éstas estarían cediendo el riesgo sobre una tabla más favorable que la utilizada a día de hoy en este país. Esta asimetría favorece al mercado español. Por su parte, el mercado reasegurador debe corregir en el precio de transferencia este efecto en caso de que sus expectativas de composición de la cartera aceptada difieran de la composición demográfica base.

BIBLIOGRAFÍA

- [1] EIOPA. (26 de ene. de 2013). «Mission and tasks,» [En línea]. Disponible en: https://www.eiopa.europa.eu/about/mission-and-tasks_en (Acceso: 05-02-2021).
- [2] A. Castañer y M. M. Claramunt, «Solvencia II,» *Universitat de Barcelona*, vol. 2, 2017.
- [3] F. A. Rodríguez, V. M. C. Juárez, M. L. D. Reguera y J. M. R.-P. del Castillo, «Proposal for the recalibration of mortality and longevity shocks under the Solvency II framework,» *Nota técnica*, Diciembre de 2017.
- [4] J. M. Rodríguez-Pardo del Castillo, I. Albarrán Lozano, F. Ariza Rodríguez, V. M. Cobreces Juárez y M. L. Durbán Reguera, «The Longevity Risk and its practical application to Solvency II: Advanced actuarial models for its management,» *Julio Castelo International Award*, 2015.
- [5] J. M. R.-P. del Castillo, I. A. Lozano, F. A. Rodríguez, V. M. C. Juárez y M. L. D. Reguera, *El riesgo de longevidad y su aplicación práctica a Solvencia II: Modelos actuariales avanzados para su gestión*, 1.^a ed. Fundación Mapfre, 2014.
- [6] I. N. de Estadística. (5 de jun. de 2020). «Esperanza de Vida,» [En línea]. Disponible en: [https://www.ine.es/ss/Satellite?L=es_ES&c=INESeccion_C&cid=1259926380048&p=1254735110672&pagename=ProductosYServicios/PYSLayout#:~:text=Es%20el%20n%C3%5C%BAmero%20medio%20de,a%20cada%20edad\)%20actualmente%20observado.](https://www.ine.es/ss/Satellite?L=es_ES&c=INESeccion_C&cid=1259926380048&p=1254735110672&pagename=ProductosYServicios/PYSLayout#:~:text=Es%20el%20n%C3%5C%BAmero%20medio%20de,a%20cada%20edad)%20actualmente%20observado.) (Acceso: 24-04-2021).
- [7] B. Mundial. (21 de feb. de 2020). «Esperanza de vida al nacer, total (años),» [En línea]. Disponible en: <https://datos.bancomundial.org/indicador/SP.DYN.LE00.IN?locations=ES> (Acceso: 25-05-2021).
- [8] A. V. Oppenheim, A. S. Willsky e I. T. Young, *Signals and Systems*. Prentice Hall Internatiiona Inc, 1984.
- [9] Wikipedia. (26 de abr. de 2021). «Supercentenario,» [En línea]. Disponible en: <https://es.wikipedia.org/wiki/Supercentenario> (Acceso: 27-04-2021).
- [10] —, (7 de abr. de 2021). «Senescencia,» [En línea]. Disponible en: <https://es.wikipedia.org/wiki/Senescencia> (Acceso: 18-06-2021).
- [11] D. R. A. H. Fernández, «Telómeros y telomerasas,» *Revista Cubana de Investigación Biomédica*, 1999.

- [12] B. O. del Estado. (17 de dic. de 2020). «Resolución de 17 de diciembre de 2020, de la Dirección General de Seguros y Fondos de Pensiones, relativa a las tablas de mortalidad y supervivencia a utilizar por las entidades aseguradoras y reaseguradoras, y por la que se aprueba la guía técnica relativa a los criterios de supervisión en relación con las tablas biométricas, y sobre determinadas recomendaciones para fomentar la elaboración de estadísticas biométricas sectoriales.» [En línea]. Disponible en: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-17154 (Acceso: 03-05-2021).
- [13] M. Durbán y D.-J. Lee, *Splines con Penalizaciones: Teoría y aplicaciones*. Universidad Carlos III de Madrid.
- [14] T. H. M. Database. (28 de mayo de 2021). «The Human Mortality Database,» [En línea]. Disponible en: <https://www.mortality.org/>.
- [15] U. Europea. (2020). «Senescencia,» [En línea]. Disponible en: https://europa.eu/european-union/about-eu/euro/which-countries-use-euro_es (Acceso: 18-06-2021).
- [16] M. Dimitrova, «¿El sistema inmune difiere entre los hombres y las mujeres?» *News Medical Life Sciences*, vol. 1, 2020.
- [17] S. Pulido, «La Gripe Española: la pandemia de 1918 que no comenzó en España,» *Gaceta Médica*, Enero de 2018.
- [18] P. España. (3 de mar. de 2019). «Falta de conciencia por el problema demográfico de España,» [En línea]. Disponible en: <https://www.pwc.es/es/sala-prensa/notas-prensa/2019/expertos-alertan-falta-conciencia-problema-demografico-espana.html> (Acceso: 30-05-2021).
- [19] I. N. de Estadística, «Movimiento Natural de la Población (MNP), Indicadores Demográficos Básicos (IDB),» *Notas de Prensa*, 2020.
- [20] A. M. Villegas, P. Millossovich y V. K. Kaishev, «StMoMo: An R Package for Stochastic Mortality Modeling,» *Insurance: Mathematics and Economics*, 2015.
- [21] C. G. Camarda, «MortalitySmooth: An R Package for Smoothing Poisson Counts with P-Splines,» *Journal of Statistical Software*, 2012.
- [22] A. M. Villegas, V. Kaishev y P. Millossovich, *Eleventh International Longevity Risk and Capital Markets*, 2015.
- [23] Y. Macias y M. Santolino, «Aplicación de los Modelos Lee-Carter y Renshawhaberman en los Seguros de Vida y Mixtos,» *Anales del Instituto de Actuarios Españoles*, vol. 1, 2018.
- [24] J. Gogola, «A comparison of Lee-Carter and Cairns-Blake-Dowd stochastic mortality model,» *Department of Mathematics and Quantitative Methods, Faculty of Economics and Administration, University of Pardubice*, vol. 1, 2014.

- [25] S. Nocito, «Stochastic Mortality Projections: A Comparison of the Lee-Carter and the Cairns-Blake-Dowd models using Italian Data,» *University of Studies of Turin*, vol. 1, 2015.
- [26] C. G. Camarda, «Smoothing methods for the analysis of mortality development,» *Universidad Carlos III de Madrid*, 2008.
- [27] Y. Liu y J. S.-H. Li, *The Locally Linear Cairns-Blake.Dowd Model: A Note On Delta-Nuga Hedging of Longevity Risk*. Cambridge University Press, 2016.
- [28] D. A. Arik, P. Y. Babad y D. Raphael, *Mortality by Causes of Death*. International Actuarial Association.

CÓDIGO

Obtención de datos

CÓDIGO 1: DatosGenerales.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library(stringr)
7
8 # Escogemos los países que queramos utilizar en el estudio
9 # En nuestro caso: países de la eurozona, UK y Dinamarca
10 coleccion_paises <- list(
11   Austria="AUT",
12   Belgium="BEL",
13   # Chipre no tiene datos en HMD
14   Denmark="DNK",
15   France="FRATNP",
16   Germany="DEUTNP",
17   Greece="GRC",
18   Estonia="EST",
19   Finland="FIN",
20   Italy="ITA",
21   Ireland="IRL",
22   Latvia="LVA",
23   Lithuania="LTU",
24   Luxembourg="LUX",
25   # Malta no tiene datos en HMD
26   Netherlands="NLD",
27   Portugal="PRT",
28   Slovakia="SVK",
29   Slovenia="SVN",
30   Spain="ESP",
31   Sweden="SWE",
32   Switzerland="CHE",
33   UK="GBR_NP"
34 )
35
36 # Cogemos los ficheros que necesitamos
37 coleccion_URLs <- list (
38   rate="Mx_1x1.txt",
39   exposure="Exposures_1x1.txt",
40   death="Deaths_1x1.txt",
41   pop="Population.txt"
```

```

42 # En nuestro caso no necesitamos los siguientes ficheros pero el
    procedimiento para obtenerlos es analogo
43 #Births="Births.txt",
44 #DeathsByLexisTriangles="Deaths_lexis.txt",
45 #Exposure2RiskByLexisTriangles="Exposures_lexis.txt",
46 #Females="fltper_1x1.txt",
47 #Males="mltper_1x1.txt",
48 #TotalBothSexes="bltper_1x1.txt",
49 #LifeExpectancyAtBirth="E0per.txt "
50 )
51
52 # Funcion para eliminar los espacios que se introducen al concatenar
    componentes de una URL
53 sin_espacios <- function (x) gsub("\\s", "", x)
54
55 # Usuario en la HMD para poder descargar los datos
56 usuario="Nombre de nuestro usuario"
57 # Para que nos salga una pestania en la que poner la contrasenia ,
    yasi tener mayor seguridad dado que no la dejamos en el codigo:
58 passw <- getPass::getPass(paste(usuario," escribe tu password de
    www.mortality.org: "))
59
60 # Funcion para comprobar que los datos estan bien cogidos, sin "NA"
61 valida <- function(p = NA){
62     if (is.na(p))
63         for (paisX in 1:length(coleccion_paises)) pinta(paisX)
64
65     else pinta(p)
66 }
67
68 # Funcion para comprobar los datos
69 pinta <- function(paisX){
70     print (c(paisX, miHMDdata[[paisX]]$country ))
71     print(c(miHMDdata[[paisX]]$country,"MIages",miHMDdata[[paisX]][["
        age"]][1],";",miHMDdata[[paisX]][["age"]][length(miHMDdata[[
        paisX]][["age"]])]) )
72     print(c(miHMDdata[[paisX]]$country,"MIyears",miHMDdata[[paisX]][["
        year"]][1],";", miHMDdata[[paisX]][["year"]][length(miHMDdata
        [[paisX]][["year"]])]) )
73     print(head(c(miHMDdata[[paisX]]$country,"MIpop Total", head(
        miHMDdata[[paisX]][["pop"]][["total"]], 1 ) )))
74     print(head(c(miHMDdata[[paisX]]$country,"MIrate Total",head(
        miHMDdata[[paisX]][["rate"]][["total"]], 1 ) )))
75
76 }
77
78 # Cargamos los datos de ejemplo de la HMD de su pagina web
79 lazyLoad(filebase= "Localizacion del fichero en nuestro ordenador",
80     enviro=parent.frame())
81
82 # Para ver que estructura tienen estos datos:

```



```

83 str(HMDdata)
84
85 # Una vez cargados los datos que queramos, los metemos en un
    ficheroHMD nuevo que debe cargarse despues de cargar la libreria
    . Asi mantenemos la misma estructura
86 miHMDdata <- HMDdata
87 # Partimos de la misma estructura de datos, pero vacia
88 miHMDdata <- NULL
89
90 # Extraemos los datos que queramos y los dejamos ordenados
91
92 # Recorremos los paises que queremos incluir en el estudio
93 for (paisX in 1:length(coleccion_paises)){
94
95     # Construimos la estructura tal como debe quedar, con datos que
        son claramente erroneos, para sustituirlos y poder identificar
        si no se ha sustituido bien
96     miHMDdata[[names(coleccion_paises)[[paisX]]]]<-
97         list(country=names(coleccion_paises)[[paisX]],
98             type="mortality",
99             year=matrix(c(1111,1111),nrow=1,ncol=2),
100             age=matrix(c(22,222,22222,2222),nrow=1,ncol=4),
101             pop=list(female=matrix(c(3333,333,33,33),nrow=1,ncol=4),
102                     male=matrix(c(55,66,77,88),nrow=1,ncol=4),
103                     total=matrix(c(11,112,113,114),nrow=1,ncol=4)),
104             death=list(female=matrix(c(444,444,44,44),nrow=1,ncol=4),
105                        male=matrix(c(10,11,22,33),nrow=1,ncol=4),
106                        total=matrix(c(10,11,22,33),nrow=1,ncol=4)),
107             exposure=list(female=matrix(c(55,55,55,55),nrow=1,ncol=4),
108                           male=matrix(c(10,11,22,33),nrow=1,ncol=4),
109                           total=matrix(c(10,11,22,33),nrow=1,ncol=4)),
110             rate=list(female=matrix(c(6666,666,66,66),nrow=1,ncol=4),
111                      male=matrix(c(100,11,22,33),nrow=1,ncol=4),
112                      total=matrix(c(100,11,22,33),nrow=1,ncol=4))
113     )
114
115     for (ficheroX in 1:length(coleccion_URLs)){
116         # Si queremos ver los datos que va utilizando:
117         #print(names(coleccion_URLs)[[ficheroX]]) #variable
118         #print( coleccion_URLs[[ficheroX]] )      #valor
119
120         # Descargamos todos los datos y los guardamos en un directorio
121         # Es mucho mas eficiente que estar descargandolos cada vez que
            los necesitemos
122         midirectorio<-"C:\\\\Windows\\Temp"
123         setwd(file.path(midirectorio))
124         tmp_dir = "MortalitySmooth.tmp"
125
126         # Camino para comprobar si tenemos ya el fichero o no
127         if (!file.exists(tmp_dir))
128             dir.create(file.path(midirectorio, tmp_dir))

```

```

129 setwd(file.path(midirectorio, tmp_dir))
130 # Vemos si tenemos el fichero descargado
131 fichero_tmp<- paste(names(coleccion_paises)[[paisX]],
132                     coleccion_URLs[ficheroX])
133 tengo_ya<-as.vector(list.files(recursive = F))
134
135 # Si no tenemos el fichero, lo descargamos
136 if (! fichero_tmp %in% tengo_ya ){
137     urlDescarga=sin_espacios(paste(
138         "https://",usuario,":",passw,"@www.mortality.org/hmd/",
139         coleccion_paises[[paisX]],"/STATS/",
140         coleccion_URLs[ficheroX]))
141     download.file(urlDescarga,fichero_tmp)
142
143     # Limpiamos los datos descargados quitando caracteres que nos
144     # llevan a errores
145     midata <- base::readLines(con=fichero_tmp, n=-1,
146                             ok=TRUE, warn=TRUE, encoding="UTF-8"
147                             )
148     sinlineasconmenos <- midata[grepl(pattern="[-]",
149                                     x=midata, invert=TRUE)]
150     sinmasnimenos <- gsub('\\+', ' ', sinlineasconmenos)
151     limpiado <- gsub('\\s\\.',' ', sinmasnimenos)
152     base::writeLines(limpiado,con=fichero_tmp, sep="\n")
153 }
154
155 #Guardamos el fichero descargado
156 descargado<- read.delim(fichero_tmp,
157                         skip= 2,
158                         header= TRUE,
159                         strip.white= TRUE,
160                         quote= "+",
161                         sep= ",",
162                         dec= ".",
163                         col.names=c("year","age","female","male"
164                                     ,"total")
165 )
166
167 # Cogemos los años distintos que se contemplen en los datos
168 listaaños<- sort(as.integer(unique(descargado$year)),
169                 decreasing = FALSE )
170
171 # Guardamos la longitud de la lista de años
172 años=length(listaaños)
173 # Nos aseguramos de que tienen el formato correcto
174 mode(listaaños)= "integer"
175 # Guardamos los años en los datos, donde corresponde
176 miHMDdata[[paisX]][["year"]] <- listaaños
177
178 # Realizamos el mismo proceso con las edades contempladas
179 listaedades<-sort(as.integer(unique(descargado$age)),
180                 decreasing = FALSE )

```

```

176     edades=length(listaedades)
177     mode(listaedades)= "integer"
178     miHMDdata[[paisX]][["age"]] = listaedades
179
180     # En la variable "population" hay un año mas que en el resto de
        variables
181     # Cogemos los datos dependiendo del sexo de la persona, y el
        conjunto de la población
182     ellas = matrix(descargado[["female"]],nrow=edades,
183                    ncol= length(descargado[["female"]])/edades,
184                    dimnames=list( miHMDdata[[paisX]][[4]],
185                                   miHMDdata[[paisX]][[3]] ) )
186     ellos = matrix(descargado[["male"]],nrow=edades,
187                    ncol= length(descargado[["male"]])/edades,
188                    dimnames=list( miHMDdata[[paisX]][[4]],
189                                   miHMDdata[[paisX]][[3]] ) )
190     todos = matrix(descargado[["total"]],nrow=edades,
191                    ncol= length(descargado[["total"]])/edades,
192                    dimnames=list( miHMDdata[[paisX]][[4]],
193                                   miHMDdata[[paisX]][[3]] ) )
194
195     # Nos aseguramos de que tienen el formato correcto
196     mode(ellas)= "double"
197     mode(ellos)= "double"
198     mode(todos)= "double"
199
200     # Lo guardamos en nuestra base de datos
201     miHMDdata[[paisX]][[ names(coleccion_URLs)[[ficheroX]] ]]<-
202         list(female = ellas, male = ellos, total= todos )
203
204 }
205 # Comprobamos la calidad de los datos descargados
206 valida(paisX)
207 }
208
209 # Por último salvamos la estructura de datos para poder cargarla en
        un futuro con mayor rapidez
210 saveRDS(miHMDdata, file="Ubicación donde guardar el fichero")

```

CÓDIGO 2: DatosEsperanzaVidaAlNacer.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library(stringr)
7
8 # Países que queremos estudiar
9 coleccion_países <- list(
10   Austria="AUT",
11   Belgium="BEL",
12   # Chipre no tiene datos en HMD
13   Denmark="DNK",
14   France="FRATNP",
15   Germany="DEUTNP",
16   Greece="GRC",
17   Estonia="EST",
18   Finland="FIN",
19   Italy="ITA",
20   Ireland="IRL",
21   Latvia="LVA",
22   Lithuania="LTU",
23   Luxembourg="LUX",
24   # Malta no tiene datos en HMD
25   Netherlands="NLD",
26   Portugal="PRT",
27   Slovakia="SVK",
28   Slovenia="SVN",
29   Spain="ESP",
30   Sweden="SWE",
31   Switzerland="CHE",
32   UK="GBR_NP"
33 )
34
35 # Función para eliminar los espacios que se introducen al
36   concatenar componentes de una URL
37 sin_espacios <- function (x) gsub("\\s", "", x)
38
39 # Usuario en la HMD para poder descargar los datos
40 usuario="Nombre de nuestro usuario"
41 # Para que nos salga una pestaña en la que poner la contraseña ,y
42   así tener mayor seguridad dado que no la dejamos en el código:
43 passw<-getPass::getPass(paste(usuario," escribe tu password de www.
44   mortality.org: "))
45
46 # Descargamos los datos de la esperanza de vida al nacer
47 EVN_coleccion_URLs <- list (
48   evn="E0per.txt"
49 )
```

```

47
48 # Estructura de datos vacia
49 evnHMDdata <- NULL
50 # Definimos el directorio en el que guardar los datos
51 midirectorio<-"Directorio que queremos"
52 setwd(file.path(midirectorio))
53 # Definimos la carpeta en la que guardarlo
54 tmp_dir = "NombreCarpeta.tmp"
55 # Si la carpeta ya existe, no la creamos
56 if (!file.exists(tmp_dir))
57   dir.create(file.path(midirectorio, tmp_dir))
58 setwd(file.path(midirectorio, tmp_dir))
59
60 # Mismo procedimiento que DatosGenerales.R
61 for(paisX in 1:length(coleccion_paises)){
62   # Llenamos la estructura con valores que identifiquemos con
        rapidez
63   evnHMDdata[[names(coleccion_paises)[[paisX]]] <-
64     list(country=names(coleccion_paises)[[paisX]],
65           year = matrix(c(1111,111111), nrow=1, ncol=2),
66           evn =list(female=matrix(c(3333,333,33,33), nrow=1, ncol
              =4),
67                     male = matrix(c(5,6,7,8), nrow=1, ncol=4),
68                     total = matrix(c(10,11,12,14), nrow=1, ncol=4))
69   )
70
71 # Sobreescribimos los valores anteriores con los datos
72 for (ficheroX in 1:length(EVN_coleccion_URLs)){
73   fichero_tmp<-paste( names(coleccion_paises)[[paisX]],
74                       EVN_coleccion_URLs[ficheroX] )
75   tengo_ya<-as.vector(list.files(recursive = FALSE))
76   if (! fichero_tmp %in% tengo_ya ){
77     urlDescarga = sin_espacios(paste("https://",usuario,":",
78                                     passw,"@www.mortality.org/hmd/",coleccion_paises[[
79                                     paisX]],"/STATS/",EVN_coleccion_URLs[ficheroX]))
80     download.file(urlDescarga,fichero_tmp)
81     midata <- base::readLines(con=fichero_tmp, n=-1, ok=TRUE,
82                             warn=TRUE, encoding="UTF-8")
83     sinlineasconmenos <- midata[grepl(pattern="[-]", x=midata,
84                                     invert=TRUE)]
85     sinmasnimenos <- gsub('\\+', ' ', sinlineasconmenos)
86     limpiado <- gsub('\\s\\.',' ', sinmasnimenos)
87     base::writeLines(limpiado,con=fichero_tmp, sep="\n")
88   }
89
90   descargado <-
91     read.delim(fichero_tmp,
92               skip = 2,
93               header = TRUE,
94               strip.white = TRUE,
95               quote="+",

```

```

92         sep = "",
93         dec = ".",
94         col.names = c("year", "female", "male", "total")
95     )
96
97     listaanyos <- sort(as.integer(unique(descargado$year)),
98                       decreasing = FALSE )
99     anyos=length(listaanyos)
100    mode(listaanyos)= "integer"
101    evnHMDdata[[paisX]][["year"]] <- listaanyos
102
103    evnHMDdata[[paisX]][[names(EVN_coleccion_URLs)[[ficheroX]]]]
104    <-
105    list(female = descargado$female,
106         male = descargado$male,
107         total= descargado$total )
108
109 }
110
111 # Por ultimo salvamos la estructura de datos para poder cargarla en
112    un futuro con mayor rapidez
113 saveRDS(evnHMDdata, file="Ubicacion donde guardar el fichero")

```

CÓDIGO 3: DatosLx.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library(stringr)
7
8 # Países que queremos estudiar
9 coleccion_países <- list(
10   Austria="AUT",
11   Belgium="BEL",
12   # Chipre no tiene datos en HMD
13   Denmark="DNK",
14   France="FRATNP",
15   Germany="DEUTNP",
16   Greece="GRC",
17   Estonia="EST",
18   Finland="FIN",
19   Italy="ITA",
20   Ireland="IRL",
21   Latvia="LVA",
22   Lithuania="LTU",
23   Luxembourg="LUX",
24   # Malta no tiene datos en HMD
25   Netherlands="NLD",
26   Portugal="PRT",
27   Slovakia="SVK",
28   Slovenia="SVN",
29   Spain="ESP",
30   Sweden="SWE",
31   Switzerland="CHE",
32   UK="GBR_NP"
33 )
34
35 # Función para eliminar los espacios que se introducen al
36   concatenar componentes de una URL
37 sin_espacios <- function (x) gsub("\\s", "", x)
38
39 # Usuario en la HMD para poder descargar los datos
40 usuario="Nombre de nuestro usuario"
41 # Para que nos salga una pestaña en la que poner la contraseña ,y
42   así tener mayor seguridad dado que no la dejamos en el código:
43 passw<-getPass::getPass(paste(usuario," escribe tu password de www.
44   mortality.org: "))
45
46 # Obtenemos los datos correspondientes a Lx de la base de datos
47 LX_coleccion_URLs <- list (
48   # Mujeres
49   lxF="fltper_1x1.txt",
```

```

47 # Hombres
48 lxM="mltper_1x1.txt",
49 # Total
50 lxT="bltper_1x1.txt"
51 )
52
53 # Partimos de una estructura vacia de datos
54 lxHMDdata <- NULL
55
56 # Definimos el directorio en el que guardar los datos
57 midirectorio<-"Directorio que queremos"
58 setwd(file.path(midirectorio))
59 # Definimos la carpeta en la que guardarlo
60 tmp_dir = "NombreCarpeta.tmp"
61 # Si la carpeta ya existe, no la creamos
62 if (!file.exists(tmp_dir))
63   dir.create(file.path(midirectorio, tmp_dir))
64 setwd(file.path(midirectorio, tmp_dir))
65
66 # Mismo procedimiento que en DatosGenerales.R
67 for (paisX in 1:length(coleccion_paises)) {
68   lxHMDdata[[names(coleccion_paises)[[paisX]]] <- list(country=
        names(coleccion_paises)[[paisX]], year = matrix(c
        (1111,111111), nrow=1, ncol=2), lx =list(female=matrix(c
        (333,333,33,33),nrow=1,ncol=4), male=matrix(c(5,6,7,8), nrow
        =1,ncol=4), total=matrix(c(10,11,12,14), nrow=1,ncol=4)))
69
70   for (ficheroX in 1:length(LX_coleccion_URLs)){
71     fichero_tmp<-paste(names(coleccion_paises)[[paisX]], LX_
        coleccion_URLs[ficheroX])
72     tengo_ya<-as.vector(list.files(recursive = FALSE))
73     if (! fichero_tmp %in% tengo_ya ){
74       urlDescarga=sin_espacios(paste("https://",usuario,":",
        passw,"@www.mortality.org/hmd/",coleccion_paises[[
        paisX]],"/STATS/",LX_coleccion_URLs[ficheroX]))
75       download.file(urlDescarga,fichero_tmp)
76       midata <- base::readLines(con=fichero_tmp, n=-1, ok=TRUE,
        warn=TRUE, encoding="UTF-8")
77       sinlineasconmenos <- midata[grepl(pattern="[-]", x=midata,
        invert=TRUE)]
78       sinmasnimenos <- gsub('\\+', ' ', sinlineasconmenos)
79       limpiado <- gsub('\\s\\.',' ', sinmasnimenos)
80       base::writeLines(limpiado,con=fichero_tmp, sep="\n")
81     }
82   }
83
84 # Se procesan por separado
85 descargado <- NULL
86 for (ficheroX in 1:length(LX_coleccion_URLs)){
87   fichero_tmp<-paste( names(coleccion_paises)[[paisX]], LX
        _coleccion_URLs[ficheroX])

```



```

88         descargado[[ficheroX]] <- read.delim(fichero_tmp, skip =
89             2, header = TRUE, strip.white = TRUE, quote="+",
            sep = "", dec = ".", col.names = c("Year", "Age", "mx"
            , "qx", "ax", "lx", "dx", "Lx", "Tx", "ex"),)
90     }
91
92     columnaelegida="lx"
93
94     # carga de datos y depuracion basica
95     listaanyos <- sort(as.integer(unique(descargado[[1]][["Year"
96         ]])), decreasing = FALSE)
97     anyos=length(listaanyos)
98     mode(listaanyos)= "integer"
99     lxHMDdata[[paisX]][["year"]] <- listaanyos
100
101     listaedades<-sort(as.integer(unique(descargado[[1]][["Age"]]))
102         , decreasing = FALSE )
103     edades=length(listaedades)
104     mode(listaedades)= "integer"
105     lxHMDdata[[paisX]][["age"]] = listaedades
106
107     ellas = matrix(descargado[[1]][[columnaelegida]],nrow=edades,
108         ncol= length(descargado[[1]][[columnaelegida]])/edades,
109         dimnames=list(listaedades, listaanyos))
110     ellos = matrix(descargado[[2]][[columnaelegida]],nrow=edades,
111         ncol= length(descargado[[2]][[columnaelegida]])/edades,
112         dimnames=list(listaedades, listaanyos))
113     todos = matrix(descargado[[3]][[columnaelegida]],nrow=edades,
114         ncol= length(descargado[[3]][[columnaelegida]])/edades,
115         dimnames=list(listaedades, listaanyos))
116
117     mode(ellas)= "double"
118     mode(ellos)= "double"
119     mode(todos)= "double"
120
121     lxHMDdata[[paisX]][[ columnaelegida]] <- list(female = ellas,
122         male = ellos, total= todos )
123 }
124
125 # Por ultimo salvamos la estructura de datos para poder cargarla en
126     un futuro con mayor rapidez
127 saveRDS(lxHMDdata, file="Ubicacion donde guardar el fichero")

```

CÓDIGO 4: EleccionColores.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6
7 # Relacionamos cada país con un color
8 # De este modo unificamos todos los gráficos
9 mis_colores <- list(
10     Austria      ="#8B2323",
11     Belgium      ="#00008B",
12     # Chipre no tiene datos en HMD
13     Denmark      ="#DA70DB",
14     France        ="#0000FF",
15     Germany       ="#8B864E",
16     Greece        ="#FFFFE0",
17     Estonia       ="#4D4D4D",
18     Finland       ="#FFE1FF",
19     Italy          ="#40E0D0",
20     Ireland       ="#006B00",
21     Latvia        ="#008E8B",
22     Lithuania     ="#E5E5E5",
23     Luxembourg    ="#EEEE00",
24     # Malta no tiene datos en HMD
25     Netherlands  ="#FF8C00",
26     Portugal      ="#FF6A6A",
27     Slovakia      ="#FFA460",
28     Slovenia      ="#C1FFC1",
29     Spain         ="#FF0000",
30     Sweden        ="#0FF000",
31     Switzerland  ="#B0C4DE",
32     UK            ="#FF1493",
33     Europe        ="#000000"
34 )
35
36 # De igual modo relacionamos cada edad con un color
37 colores = c(rainbow(30),rainbow(30),rainbow(30),rainbow(30))
38 tras<-length(mis_colores)
39 for (n in 1:111){
40     mis_colores[tras+n]=colores[n]
41     names(mis_colores)[tras+n] <- as.character(n-1)
42 }
43
44 #Lo salvamos para utilizarlo en los gráficos
45 saveRDS(mis_colores, file="Ubicacion donde guardar el fichero")
```

CÓDIGO 5: ConstrucciónEuropa.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6
7 # Cogemos los datos que dejamos guardados anteriormente
8 miHMDdata <- readRDS (file="Ubicacion donde guardarnos el fichero")
9 # Nombres de los países de la base de datos
10 coleccion_paises <- names(miHMDdata)
11
12 # Cogemos los datos de las lx
13 lxHMDdata <- readRDS (file="Ubicacion donde guardarnos el fichero")
14
15 # Estudiamos en que años tenemos datos en todos los países
16   seleccionados para crear una base de datos europea
17 # Primer año desde el que todos los países tienen datos
18 # Si quisieramos, podriamos coger un dato inicial distinto
19 desde=""
20 for(pais in names(miHMDdata))
21   for(tabla in c("pop","death","exposure","rate"))
22     for(sexo in c("female","male","total"))
23       if(desde < dimnames(miHMDdata)[[pais]][[tabla]][[sexo]]
24         [[2]][[1]])
25         desde <- dimnames(miHMDdata)[[pais]][[tabla]][[sexo
26           ]][[2]][[1]]
27
28 # Ultimo año en que todos los países tienen datos
29 # Si quisieramos, podriamos estudiar hasta otra fecha
30 hasta="9999"
31 for(pais in names(miHMDdata))
32   for(tabla in c("pop","death","exposure","rate"))
33     for(sexo in c("female","male","total"))
34       if(hasta > dimnames(miHMDdata)[[pais]][[tabla]][[sexo]]
35         [[2]][[length(dimnames(miHMDdata)[[pais]][[tabla]][[
36           sexo]])[2]]]))
37       hasta <- dimnames(miHMDdata)[[pais]][[tabla]][[sexo
38         ]][[2]][[length(dimnames(miHMDdata)[[pais]][[
39           tabla]][[sexo]])[2]]]]
40
41 # Lista de años coincidentes en todos los países
42 Lanyos <- c(desde:hasta)
43 # Longitud del intervalo a estudiar
44 Nanyos <- as.integer(hasta)-as.integer(desde)+1
45
46 # Edad máxima recogida común a todos los países
47 # En nuestro caso sabemos que esta edad es 110, pero podría ser que
48   variase en algún país
49 edadmaxima="999"
```

```

42 for(pais in names(miHMDdata))
43   for(tabla in c("pop","death","exposure","rate"))
44     for(sexo in c("female","male","total"))
45       if(edadmaxima > dimnames(miHMDdata[[pais]][[tabla]][[
46         sexo]])[[1]][length(dimnames(miHMDdata[[pais]][[
47           tabla]][[sexo]])[[1]])] )
48         edadmaxima <- dimnames(miHMDdata[[pais]][[tabla]][[
49           sexo]])[[1]][length(dimnames(miHMDdata[[pais
50             ]][[tabla]][[sexo]])[[1]])]
51
52 # Lista de edades coincidentes en todos los paises
53 Ledades <- c(0:edadmaxima)
54 # Longitud del rango a estudiar
55 Nedades<-as.integer(edadmaxima)+1
56
57 # Primero creamos una estructura vacia
58 euHMDdata <- NULL
59 # La rellenamos
60 euHMDdata[["Europe"]] <-
61   list(country = "Europe",
62         type = "mortality",
63         # Cogemos los anos comunes a todos los paises
64         year = matrix(Lanyos, nrow=1, ncol=Nanyos ),
65         # Cogemos las edades comunes a todos los paises
66         age = matrix(Ledades, nrow=1, ncol=Nedades ),
67         # El resto de datos los calculamos despues
68         pop      = NULL,
69         death    = NULL,
70         exposure = NULL,
71         rate     = NULL
72       )
73 # Igual para lx
74 eu_lxHMDdata <- NULL
75 eu_lxHMDdata[["Europe"]] <- list( country = "Europe", year = Lanyos,
76   lx = list( female = matrix(rep(0,length(Ledades)*length(Lanyos)
77     ), nrow=length(Ledades), ncol=length(Lanyos)), male = matrix(rep
78     (0,length(Ledades)*length(Lanyos)), nrow=length(Ledades), ncol=
79     length(Lanyos)), total = matrix(rep(0,length(Ledades)*length(
80     Lanyos)), nrow=length(Ledades), ncol=length(Lanyos))), age =
81     Ledades)
82
83 # Son parametros son acumulativos, se suman y se introducen en la
84   matriz
85 for (sexo in c("female","male","total"))
86   for(tabla in c("pop","death","exposure","rate")){
87     resultado <- matrix(0,nrow=Nedades, ncol=Nanyos,
88       dimnames=list(Ledades, Lanyos))
89     for (pais in names(miHMDdata)){
90
91       todo <- miHMDdata[[pais]][[tabla]][[sexo]]

```

```

82         x.ages <- which(!is.na(match(rownames(todo), Ledades)))
83         y.years <- which(!is.na(match(colnames(todo), Lanyos)))
84
85         parte <- todo[x.ages, y.years]
86
87         resultado <- resultado + parte
88     }
89     # Lo guardamos en su sitio
90     euHMDdata[["Europe"]][[tabla]][[sexo]] <- resultado
91 }
92
93 # La tasa (rate) la obtenemos dividiendo: death/exposure
94 for(sexo in c("female", "male", "total"))
95     euHMDdata[["Europe"]][["rate"]][[sexo]] <-
96     euHMDdata[["Europe"]][["death"]][[sexo]]/euHMDdata[["Europe"]][[
97         "exposure"]][[sexo]]
98
99 tabla = "pop"
100 # Calculamos la lx conjunta
101 for (sexo in c("female", "male", "total")){
102
103     a_pondero <- NULL
104     a_ponderar = eu_lxHMDdata [["Europe"]][["lx"]][[sexo]]
105
106     for ( pais in names(miHMDdata)){
107         todo <- miHMDdata[[pais]][[tabla]][[sexo]]
108         x.ages <- which(!is.na(match(rownames(todo), Ledades)))
109         y.years <- which(!is.na(match(colnames(todo), Lanyos)))
110         a_ponderar <- a_ponderar + ( lxHMDdata [[pais]][["lx"]][[sexo
111             ]][y.years] * miHMDdata[[pais]][[tabla]][[sexo]][,y.years
112             ])
113     }
114
115     tmp <- a_ponderar / euHMDdata[["Europe"]][[tabla]][[sexo]]
116     eu_lxHMDdata[["Europe"]][["lx"]][[sexo]] <- tmp
117 }
118
119 # Lo guardamos para utilizarlo en futuras ocasiones.
120 saveRDS(euHMDdata, file="Ubicacion donde guardar el fichero")
121 saveRDS(eu_lxHMDdata, file="Ubicacion donde guardar el fichero")

```

Gráficos

CÓDIGO 6: GraficoRectangularizacion.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library(shape)
7
8 # Para dar homogeneidad a todos los gráficos
9 mis_colores <- readRDS (file = "Localización del fichero")
10
11 # Cogemos los datos de Lx
12 lxHMDdata <- readRDS (file="Localización del fichero")
13 # Cogemos los nombres de nuestra base de datos
14 coleccion_paises <- names(lxHMDdata)
15
16 # Guardamos los gráficos que hagamos en el directorio deseado
17 midirectorio <- "Localización del directorio"
18 setwd(file.path(midirectorio))
19
20 # Creamos una carpeta dentro del directorio, si ya existe, entonces
  lo guardamos ahí
21 destino = "Rectangularización"
22 if (!file.exists(destino)) dir.create(file.path(midirectorio,
  destino))
23 setwd(file.path(midirectorio, destino))
24
25 # Función para dibujar gráficos de Lx para estudiar la
  rectangularización
26 dibuja_todo_LX <- function(pais, sexo, Dlineas, titulo, theta, phi,
  Dcolor){
27
28   # Cogemos los datos desde la fecha deseada y los ponemos sobre una
    línea completa pero nula, define la anchura con la que va a
    pintar, y cogemos en tramo de lxHMDdata que encaje
29
30   # Conjunto de datos
31   conjunto <- lxHMDdata[[pais]][["lx"]][[sexo]]
32
33   persp(x = lxHMDdata[[pais]][["age"]],
34         y = lxHMDdata[[pais]][["year"]],
35         z=conjunto,
36         xlab="Edades",
37         ylab="años",
38         zlab="lx",
39         main=paste("Rectangularización:",pais,"desde que hay datos,
    sexo:", sexo),
```

```

40     theta=theta,
41     phi=phi,
42     shade = NA,
43     ticktype = "detailed",
44     expand = 1.1,
45     col = drapecol(conjunto),
46     border="grey70",
47     box=TRUE)
48 }
49
50 # Funcion para guardar un grafico en el directorio establecido
51 dibuja_y_salva_LX <- function(pais, sexo, Dlineas, titulo, theta,
52     phi, Dcolor, destino ){
53     png(destino,width=6000,height=4000, res = 350)
54     imagen <-dibuja_todo_LX(pais, sexo, Dlineas, titulo, theta, phi,
55         Dcolor)
56     dev.off()
57 }
58
59 # Dibujamos Lx para obesrvar la rectangularizacion en Espana
60 dibuja_todo_LX(pais="Spain", sexo="total", Dlineas="lx", titulo =
61     paste("desde que hay datos"), theta=130, phi=20, Dcolor="black")
62
63 # Igual pero guardando el grafico
64 dibuja_y_salva_LX(pais="Spain", sexo="total", Dlineas="lx", titulo
65     = "desde que hay datos", theta=30, phi=30, Dcolor="black",
66     destino="Nombre del fichero.png")

```

CÓDIGO 7: GraficoEsperanzaVidaAlNacer.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library("png")
7
8 # Para dar uniformidad a todos los gráficos
9 mis_colores <- readRDS (file = "Localizacion del fichero")
10
11 # Cogemos los datos de EsperanzaVidaAlNacer
12 evnHMDdata <- readRDS (file="Localizacion del fichero")
13 # Cogemos los nombres de nuestra base de datos
14 coleccion_paises <- names(evnHMDdata)
15
16 # Guardamos los gráficos que hagamos en el directorio deseado
17 midirectorio <- "Localizacion del directorio"
18 setwd(file.path(midirectorio))
19
20 # Creamos una carpeta dentro del directorio, si ya existe, entonces
  lo guardamos ahí
21 destino = "EsperanzaVidaAlNacer"
22 if (!file.exists(destino)) dir.create(file.path(midirectorio,
  destino))
23 setwd(file.path(midirectorio, destino))
24
25 # Funcion para dibujar gráficos de la esperanza de vida
26 dibuja_ENV <- function(Dpaises, Dlineas, desde, hasta, titulo,
  leyenda, leyendaX, leyendaY){
27
28   # Cogemos los datos desde la fecha deseada y los ponemos sobre una
    línea completa pero nula, define la anchura con la que va a
    pintar, y cogemos en tramo de evnHMDdata que encaje
29
30   # Para realizar el gráfico tendremos en cuenta que no todos los
    países tienen datos desde la misma fecha, por lo que algunas
    líneas empiezan antes que otras
31
32   # Fijamos la escala a utilizar y los años a dibujar
33   escalaY = 100
34   years <- (desde:hasta)
35   ceros <- rep(x=NA, length.out= length(years))
36
37   # Escogemos los colores
38   color <- 1
39   colores <- c(unlist(mis_colores[1:22]), unlist(mis_colores[1:22]),
    unlist(mis_colores[1:22]))
40
41
```



```

42 resultado <- matplot(x=years, y=ceros, type = "l", lty = 3, col =
    'black', ylim = c(0, escalaY), cex.axis = 1, main = titulo,
    xlab = "Ao", ylab = "Edad esperable")
43 grid(col = 'black', lty = 1, lwd = 0.5)
44
45 for ( pais in Dpaises ) {
46     matiz <-1
47     for (linea in Dlineas) {
48         if ( hasta < evnHMDdata[[pais]][["year"]][[1]] | desde >
            evnHMDdata[[pais]][["year"]][[length( evnHMDdata[[pais]
                ][["year"]])]] ) break
49         # pero si la interseccin no es vaca, se pinta
50         if( desde < evnHMDdata[[pais]][["year"]][[1]]) {
51             primero <- 1
52             desplazamiento <- match(evnHMDdata[[pais]][["year"
                ]][[1]],years)
53         }
54         else {
55             primero <- match(desde, evnHMDdata[[pais]][["year"]])
56             desplazamiento = 1
57         }
58         if( hasta > evnHMDdata[[pais]][["year"]][[length(
            evnHMDdata[[pais]][["year"]])]]) {
59             ultimo <- length( evnHMDdata[[pais]][["year"]])
60         }
61         else {
62             ultimo <- match(hasta, evnHMDdata[[pais]][["year"]])
63         }
64         estaLinea <- ceros
65         for (i in primero:ultimo) estaLinea[desplazamiento+i-
            primero] <- evnHMDdata[[pais]][["evn"]][[linea]][i]
66         resultado <- lines(x=years, y=estaLinea, type = "l",
            lty = 5, col = colores[color*matiz], ylim = c(0,
            escalaY), cex.axis = 1, main = "", xlab = "", ylab
            = "")
67         matiz <- matiz + 1
68     }
69     color <- color + 1
70 }
71 resultado <- legend(leyendaX, leyendaY, legend = leyenda, cex =
    0.7, lty = 5, col = colores[1:(length(Dpaises)*length(
    Dlineas))])
72 }
73
74 # Funcion para guardar un grafico en el directorio establecido
75 dibuja_y_salva_ENV <- function(Dpaises, Dlineas, desde , hasta,
    titulo, leyenda, leyendaX, leyendaY, fichero){
76     png(fichero,width=2400,height=1600, res = 350)
77     imagen <-dibuja_ENV(Dpaises, Dlineas, desde, hasta, titulo,
        leyenda, leyendaX, leyendaY)
78     print (imagen)

```

```

79 dev.off()
80 }
81
82
83 # Dibujamos la esperanza de vida en Europa
84 # En caso de que queramos dibujar unicamente los totales
85 dibuja_ENV(Dpaises=names(evnHMDdata), Dlineas="total", desde ="1725"
86           , hasta ="2017", titulo = "ENV Europa", leyenda= c(names(
87             evnHMDdata) ),leyendaX=1725, leyendaY= 102)
88
89 # En caso de querer dibujar los totales y los sexos
90 dibuja_ENV(Dpaises=names(evnHMDdata),Dlineas = c("female", "male", "
91           total") , desde = "1908", hasta = "2018", titulo="ENV Espaa",
92           leyenda=c("female", "male", "total"),leyendaX=2000, leyendaY=40
93           )
94
95 # Igual pero guardando los graficos
96 dibuja_y_salva_ENV(Dpaises=names(evnHMDdata), Dlineas="total", desde
97                   ="1725", hasta ="2017", titulo = "ENV Europa", leyenda= c(names
98                     (evnHMDdata) ),leyendaX=1725, leyendaY= 110, fichero="Nombre
99                     fichero.png")
100
101 dibuja_y_salva_ENV(Dpaises="Spain",Dlineas = c("female", "male", "
102           total") , desde = "1908", hasta = "2018", titulo="ENV Espaa",
103           leyenda=c("female", "male", "total"),leyendaX=2000, leyendaY=40,
104           fichero="Nombre fichero.png" )

```

CÓDIGO 8: GraficoAnálisisHistoricoDecadas.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6
7 # Cogemos los datos que necesitamos
8 euHMDdata <- readRDS (file = "Localizcion del fichero")
9 miHMDdata <- readRDS (file = "Localizcion del fichero")
10 # Cogemos los nombres de nuestra base de datos
11 coleccion_paises <- names(miHMDdata)
12
13 # Introducimos los datos de Europa en nuestra base de datos
14 miHMDdata[["Europe"]] <- euHMDdata[["Europe"]]
15 coleccion_extendida_paises <- names(miHMDdata)
16
17 # Guardamos los graficos que hagamos en el directorio deseado
18 midirectorio <- "Localizacion del directorio"
19 setwd(file.path(midirectorio))
20
21 # Creamos una carpeta dentro del directorio, si ya existe, entonces
    lo guardamos ahi
22 destino = "Análisis historico"
23 if (!file.exists(destino)) dir.create(file.path(midirectorio,
    destino))
24 setwd(file.path(midirectorio, destino))
25
26 # Pais que queremos estudiar
27 miPais = "Spain"
28
29 # Edades en la base de datos
30 ages <- dimnames(miHMDdata[[miPais]][["death"]][["female"]])[[1]]
31
32 # Cogemos los datos necesarios para el analisis
33 DxMujeres <- miHMDdata[[miPais]][["death"]][["female"]]
34 DxHombres <- miHMDdata[[miPais]][["death"]][["male"]]
35
36 ExpMujeres <- miHMDdata[[miPais]][["exposure"]][["female"]]
37 ExpHombres <- miHMDdata[[miPais]][["exposure"]][["female"]]
38
39 RatesMujeres <- miHMDdata[[miPais]][["rate"]][["female"]]
40 RatesHombres <- miHMDdata[[miPais]][["rate"]][["female"]]
41
42 # Dibujamos el numero de fallecimientos
43
44 # Mujeres
45 png("Nombre fichero.png",width=2400,height=1600, res = 350)
46     matplot(ages, DxMujeres, type = "l", lty = 3, col = c(rep(1, 10),
        rep(2, 10),rep(3, 10), rep(4, 10),rep(5, 10),rep(6, 10),rep
```

```

(7, 10),rep(8, 10),rep(9, 8)), ylim = c(0, 17000), cex.axis =
  1, main = "N de Mujeres fallecidas", xlab = "Edad", ylab = "
  Nmero de defunciones")
47 legend(10, 17000, legend = c("1910-1919","1920-1929","1930-1939",
  "1940-1949", "1950-1959", "1960-1969", "1970-1979", "
  1980-1989", "1990-1999", "2000-2009", "2010-2018"), cex =
  0.7, lty = 3, col = c(1, 2, 3, 4, 5, 6, 7, 8, 9))
48 dev.off()
49
50 #Hombres
51 png("Nombre fichero.png",width=2400,height=1600, res = 350)
52 matplot(ages, DxHombres, type = "l", lty = 3, col = c(rep(1, 10),
  rep(2, 10),rep(3, 10), rep(4, 10),rep(5, 10),rep(6, 10),rep
  (7, 10),rep(8, 10),rep(9, 8)), ylim = c(0, 17000), cex.axis =
  1, main = "N de Hombres fallecidos", xlab = "Edad", ylab = "
  Nmero de defunciones")
53 legend(10, 17000, legend = c("1910-1919","1920-1929","1930-1939",
  "1940-1949", "1950-1959", "1960-1969", "1970-1979", "
  1980-1989", "1990-1999", "2000-2009", "2010-2018"), cex =
  0.7, lty = 3, col = c(1, 2, 3, 4, 5, 6, 7, 8, 9))
54 dev.off()
55
56
57 # Mismo grafico pero a partir de un cierto ano
58
59 # En este caso, desde la Espaa constitucional
60 desde = "1978"
61 # Cogemos los datos
62 anyos = dimnames(miHMDdata[[miPais]][["exposure"]][["female"]])[[2]]
63 primero_index <- match(desde,anyos)
64
65 DxMujeres <- miHMDdata[[miPais]][["death"]][["female"]][,primero_
  index:length(anyos)]
66 DxHombres <- miHMDdata[[miPais]][["death"]][["male"]][,primero_index
  :length(anyos)]
67
68 # Lo dibujamos
69
70 # Mujeres
71 png("Nombre fichero.png",width=2400,height=1600, res = 350)
72 matplot(x=ages, y=DxMujeres, type = "l", lty = 3, col = c(rep(1,
  10),rep(2, 10),rep(3, 10), rep(4, 10),rep(5, 10),rep(6, 10),
  rep(7, 10),rep(8, 10),rep(9, 8)), ylim = c(0, 14000), cex.
  axis = 1, main = "N de Mujeres fallecidas", xlab = "Edad",
  ylab = "Nmero de defunciones")
73 legend(85, 400000, legend = c("1978-1988","1988-1998", "1998-2008
  ", "2008-2018"), cex = 0.7, lty = 3, col = c(1, 2, 3, 4))
74 dev.off()
75
76 # Hombres
77 png("Nombre fichero.png",width=2400,height=1600, res = 350)

```

```

78     matplot(ages, DxHombres, type = "l", lty = 3, col = c(rep(1, 10),
      rep(2, 10),rep(3, 10), rep(4, 10),rep(5, 10),rep(6, 10),rep
      (7, 10),rep(8, 10),rep(9, 8)), ylim = c(0, 17000), cex.axis =
      1, main = "N de Hombres fallecidos", xlab = "Edad", ylab = "
      Nmero de defunciones")
79     legend(85, 400000, legend = c("1978-1988", "1988-1998", "1998-2008
      ", "2008-2018"), cex = 0.7, lty = 3, col = c(1, 2, 3, 4))
80 dev.off()
81
82
83 # Ahora dibujamos la exposicion
84
85 # Analogo al caso anterior
86 desde = "1978"
87 anyos = dimnames(miHMDdata[[miPais]][["exposure"]][["female"]])[2]
88 primero_index <- match(desde, anyos)
89
90 ExpMujeres <- miHMDdata[[miPais]][["exposure"]][["female"]][, primero
      _index:length(anyos)]
91 ExpHombres <- miHMDdata[[miPais]][["exposure"]][["male"]][, primero_
      index:length(anyos)]
92
93 png(paste("Nombre del fichero.png"), width=2400, height=1600, res =
      350)
94     matplot(x=ages, y=ExpMujeres, type = "l", lty = 3, col = c(rep(1,
      10),rep(2, 10),rep(3, 10), rep(4, 10),rep(5, 10),rep(6, 10),
      rep(7, 10),rep(8, 10),rep(9, 8)), ylim = c(0, 440000), cex.
      axis = 1, main = "Mujeres expuestas", xlab = "Edad", ylab = "
      Expuestas")
95     legend(85, 400000, legend = c("1978-1988", "1988-1998", "1998-2008
      ", "2008-2018"), cex = 0.7, lty = 3, col = c(1, 2, 3, 4))
96 dev.off()
97
98 png(paste("Nombre del fichero.png"), width=2400, height=1600, res =
      350)
99     matplot(x=ages, y=ExpHombres, type = "l", lty = 3, col = c(rep(1,
      10),rep(2, 10),rep(3, 10), rep(4, 10),rep(5, 10),rep(6, 10),
      rep(7, 10),rep(8, 10),rep(9, 8)), ylim = c(0, 440000), cex.
      axis = 1, main = "Hombres expuestos", xlab = "Edad", ylab = "
      Expuestos")
100    legend(85, 400000, legend = c("1978-1988", "1988-1998", "1998-2008
      ", "2008-2018"), cex = 0.7, lty = 3, col = c(1, 2, 3, 4))
101 dev.off()
102
103
104 # Por ultimo dibujamos la tasa de mortalidad
105
106 # Analogo a los casos anteriores
107 desde = "1978"
108
109 anyos = dimnames(miHMDdata[[miPais]][["exposure"]][["female"]])[2]

```

```

110 primero_index <- match(desde, anyos)
111
112 DxMujeres <- miHMDdata[[miPais]][["death"]][["female"]][, primero_
    index:length(anyos)]
113 DxHombres <- miHMDdata[[miPais]][["death"]][["male"]][, primero_index
    :length(anyos)]
114
115 ExpMujeres <- miHMDdata[[miPais]][["exposure"]][["female"]][, primero
    _index:length(anyos)]
116 ExpHombres <- miHMDdata[[miPais]][["exposure"]][["male"]][, primero
    _index:length(anyos)]
117
118 # Tasa de mortalidad, fallecidos/exposicion
119 MxtMujer <- DxMujeres/ExpMujeres
120 MxtHombre <- DxHombres/ExpHombres
121
122 png(paste("Nombre fichero.png"), width=2400, height=1600, res = 350)
123 matplot(ages, log(MxtMujer), type = "l", lty = 3, col = c(rep(1,
    10), rep(2, 10), rep(3, 10), rep(4, 10), rep(5, 10), rep(6, 10),
    rep(7, 10), rep(8, 10), rep(9, 8)), ylim = c(-10, 2), main = "
    log(m(x,t)) en Mujeres", xlab = "Edad", ylab = "ln(mxt)")
124 legend(85, 400000, legend = c("1978-1988", "1988-1998", "1998-2008",
    "2008-2018"), cex = 0.7, lty = 3, col = c(1, 2, 3, 4))
125 dev.off()
126
127 png(paste("log m hombres desde", desde, ".png"), width=2400, height
    =1600, res = 350)
128 matplot(ages, log(MxtHombre), type = "l", lty = 3, col = c(rep(1,
    10), rep(2, 10), rep(3, 10), rep(4, 10), rep(5, 10), rep(6, 10),
    rep(7, 10), rep(8, 10), rep(9, 8)), ylim = c(-10, 2), main = "
    log(m(x,t)) en Hombres", xlab = "Edad", ylab = "ln(mxt)")
129 legend(85, 400000, legend = c("1978-1988", "1988-1998", "1998-2008",
    "2008-2018"), cex = 0.7, lty = 3, col = c(1, 2, 3, 4))
130 dev.off()

```

CÓDIGO 9: GraficoPiramidePoblacion.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerias necesarias
6   # install.packages("readxl")
7   # install.packages("tidyverse")
8 library(curl)
9 library(readxl)
10 library(tidyverse)
11 library(ggplot2)
12
13 # Cogemos los datos
14 miHMDdata <- readRDS (file="Localizacion del fichero")
15 euHMDdata <- readRDS (file="Localizacion del fichero")
16
17 # Introducimos los datos de Europa en nuestra base de datos
18 miHMDdata[["Europe"]]<-euHMDdata[["Europe"]]
19
20 # Creamos una carpeta dentro del directorio, si ya existe, entonces
   lo guardamos ahi
21 midirectorio<-"Localizacion del directorio"
22 setwd(file.path(midirectorio))
23 destino = "PiramidePoblacional"
24 if (!file.exists(destino))   dir.create(file.path(midirectorio,
   destino))
25 setwd(file.path(midirectorio, destino))
26
27 # Funcion para guardar un grafico en el directorio establecido
28 piramide_vs_Europa_salva <- function(pais, fichero, anyo, rango,
   destino){
29   png(destino,width=2400,height=1600, res = 350)
30   graf <- piramide_vs_Europa(pais, fichero, anyo, rango)
31   print(graf+geom_point(colour = "white", size = 0.02) + theme(
     legend.position = "none") )
32   dev.off()
33   return(graf)
34 }
35
36 # Funcion para tratar valores que superan un limite
37 recorta<- function(datos, limite){
38   for (i in 1:111) {
39     if ( is.na(datos[[i]] ) ) datos[[i]] <- limite
40     # Si nos salimos de los lmites, lo pinta igualmente
41     if ( datos[[i]] > limite) datos[[i]] <- limite
42   }
43   return(datos)
44 }
45
```

```

46 # Funcion para dibujar graficos: piramides poblacionales
47 piramide_vs_Europa <- function(pais, fichero, anyo, rango){
48
49   # Buscamos el inicio de los datos de Europa
50   primero <- match("1990", dimnames(miHMDdata)[[pais]][[fichero]][["male"]])[[2]]
51
52   todoH <- miHMDdata[[pais]][[fichero]][["male"]]
53   todoM <- miHMDdata[[pais]][[fichero]][["female"]]
54   todoHE <- miHMDdata[["Europe"]][[fichero]][["male"]]
55   todoME <- miHMDdata[["Europe"]][[fichero]][["female"]]
56   mi.anyo <- which(!is.na(match(colnames(todoH), anyo)))
57   mi.anyoE <- which(!is.na(match(colnames(todoHE), anyo)))
58   mi.edades <- 0:110 # todas las edades
59
60   parteH <- recorta( todoH[ 1:111, mi.anyo], limite=rango)
61
62   parteM <- - recorta( todoM[ 1:111, mi.anyo], limite=rango )
63
64   parteHE <- recorta( todoHE[ 1:111, mi.anyoE], limite=rango)
65
66   parteME <- - recorta( todoME[ 1:111, mi.anyoE], limite=rango )
67
68   # Creamos una estructura lineal
69   ambos <- NULL
70   ambos <- data.frame(edad = c(mi.edades, mi.edades, mi.edades, mi.edades),
71     sexo=c(rep("male", length(parteH)), rep("female",
72       length(parteH)), rep("maleE", length(parteH)), rep("femaleE",
73       length(parteH))), valores=c(parteH, parteM, parteHE, parteME)
74   )
75
76   # Orden de magnitud
77   por = 10**floor((log10(abs(min(parteM)))))
78   de = as.integer(as.integer(min(parteM)/por)*por)*1.6
79   a = as.integer(as.integer(max(parteH)/por)*por)*1.6
80   trozos = floor((a-de)/por/2)
81
82   # Elaboracion de la grafica
83   grafico <- ggplot(ambos, aes(x = 'edad', y = 'valores', fill =
84     sexo)) +
85     geom_bar(data=subset(ambos, sexo == "maleE"), stat = "identity",
86       width = 0.5 , fill = "blue") +
87     geom_bar(data=subset(ambos, sexo == "femaleE"), stat = "identity",
88       width = 0.5 , fill = "pink") +
89     geom_bar(data=subset(ambos, sexo == "male"), stat = "identity",
90       width = 0.3 , fill = "cyan") +
91     geom_bar(data=subset(ambos, sexo == "female"), stat = "identity",
92       width = 0.3 , fill = "red") +
93     coord_flip() +
94     ggthemes::theme_tufte() +
95     theme(plot.title = element_text(family = "Arial", hjust =

```



```

    0.8, size = 18), axis.text.x = element_text(family = "
    Arial"), axis.text.y = element_text(family = "Arial")) +
87   labs(title = paste("Piramide de", fichero, "de", pais, "vs", "Europe
    el ao", anyo),
88   x = "",
89   y = "  Mujeres                      Hombres  ",
90   caption = "Fuente: datos de https://www.mortality.org/") +
91   scale_y_continuous(limits=c( -rango , rango))
92   print(grafico)
93
94   return(grafico)
95 }
96
97 # Funcion para guardar un grafico en el directorio establecido
98 piramide_vs_Europa_salva <- function(pais, fichero, anyo, rango,
    destino){
99   png(destino,width=2400,height=1600, res = 350)
100   graf <- piramide_vs_Europa(pais, fichero, anyo, rango)
101   print(graf+geom_point(colour = "white", size = 0.02) + theme(
    legend.position = "none") )
102   dev.off()
103   return(graf)
104 }
105
106 # Grafico piramide poblacional comparando Europa con Espania
107 g <- piramide_vs_Europa(pais="Spain", fichero="death", anyo="2015",
    rango=1000000)
108
109 # Igual para salvar el grafico
110 g <- piramide_vs_Europa_salva(pais="Spain", fichero="death", anyo="
    2015", rango=1000000, destino = paste("Nombre fichero.png"))
111
112 # Grafico cambiando distintos parametros
113 g <- piramide_vs_Europa( pais="Spain", fichero="pop", anyo="1990",
    rango=25000000)
114
115 # Grafico cambiando distintos parametros
116 g <- piramide_vs_Europa( pais="Spain", fichero="rate", anyo="1990",
    rango=0.5)
117
118 # Pintamos la tasa de mortalidad
119 for (anyo in 1990:2017 ) {
120   g <- piramide_vs_Europa_salva( pais="Spain", fichero="rate",
    anyo=anyo, rango=0.5, destino = paste("Nombre fichero",anyo,
    ".png"))
121   g
122 }
123
124 # Igual para Europa
125 for (anyo in 1990:2017){
126   g <- piramide_vs_Europa_salva( pais="Europe", fichero="rate",

```

```

    anyo=anyo, rango=0.5, destino = paste("Nombre fichero",anyo,".
    png"))
127   g
128 }
129
130 # Para estudiar distintos valores de manera comoda
131
132 fichero="pop"
133 rango=50000000
134
135 fichero="rate"
136 rango=0.5
137
138 fichero="death"
139 rango=50000
140
141 anyo=1991
142
143 pais="Spain"
144 for (anyo in 1990:2017){
145   g <- piramide_vs_Europa( pais=pais, fichero=fichero, anyo=anyo,
    rango=rango)
146   print(g+geom_point(colour = "white", size = 0.02) + theme(legend.
    position = "none") )
147   g <- piramide_vs_Europa_salva( pais=pais, fichero=fichero, anyo=
    anyo, rango=rango, destino = paste("Piramide Europa y Espaa",
    fichero,anyo,".png"))
148   print(c(pais,anyo))
149 }

```

CÓDIGO 10: GraficoHistoricoDeDatos.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library("png")
7 library("ggplot2")
8 library("ggthemes")
9 library("dplyr")
10 # En caso de querer animar graficos
11 library(gganimate)
12
13 # Para dar uniformidad a todos los grficos
14 mis_colores <- readRDS (file = "Localizcion del fichero")
15
16 # Cogemos los datos que necesitamos
17 euHMDdata <- readRDS (file = "Localizcion del fichero")
18 miHMDdata <- readRDS (file = "Localizcion del fichero")
19 # Cogemos los nombres de nuestra base de datos
20 coleccion_paises <- names(miHMDdata)
21
22 # Introducimos los datos de Europa en nuestra base de datos
23 miHMDdata[["Europe"]] <- euHMDdata[["Europe"]]
24 coleccion_extendida_paises <- names(miHMDdata)
25
26 # Creamos una carpeta dentro del directorio, si ya existe, entonces
   lo guardamos ahí
27 midirectorio<-"Localizacion del directorio"
28 setwd(file.path(midirectorio))
29 destino = "HistoricosDatos"
30 if (!file.exists(destino)) dir.create(file.path(midirectorio,
   destino))
31 setwd(file.path(midirectorio, destino))
32
33 # Funcion para seleccionar el pais que queramos, los datos que
   queramos (pop, death, expousure o rate), el sexo que queramos (
   male, female o total) y las edades y los anos deseados, con el
   angulo que queramos
34 selecciona_unaEdad_miHMDdata <- function(miPais, miFichero, miSexo,
   miEdad, miPeriodo){
35
36   # Periodo
37   miDesde <- miPeriodo[1]
38   miHasta <- miPeriodo[length(miPeriodo)]
39
40   # Edades
41   miEdades <- dimnames(miHMDdata[[miPais]][[miFichero]][[miSexo]])
     [[1]]
42
```

```

43 # Anos
44 miAnyos <- dimnames(miHMDdata[[miPais]][[miFichero]][[miSexo]])
    [[2]]
45
46 # Si el rango no es correcto, paramos:
47 if (miHasta < miAnyos[[1]] | miDesde > miAnyos[[length(miAnyos)]])
    break
48
49 periodo = miDesde:miHasta
50
51 # Si el rango es correcto, cogemos los datos
52 if (miDesde < miAnyos[[1]]) {
53     primero <- 1
54     desplazamiento <- match(miAnyos[[1]], periodo)
55 } else {
56     primero <- match(miDesde, miAnyos)
57     desplazamiento = -primero + 2
58 }
59 if (miHasta > miAnyos[[length(miAnyos)]]) {
60     ultimo <- length(miAnyos)
61 } else {
62     ultimo <- match(miHasta, miAnyos)
63 }
64
65 edad_index <- match(miEdad, miEdades)
66 estaLinea <- rep.int(x = NA, length(periodo))
67
68 for (i in primero:ultimo)
69     estaLinea[i + desplazamiento - 1] <- miHMDdata[[miPais]][[
        miFichero]][[miSexo]][edad_index, ][i]
70
71 return(estaLinea)
72 }
73
74 # Funcion para dibujar los datos
75 dibuja_miHMDdata <- function(datos, referente, titulo, subtitulo,
    leyenda) {
76
77     # Modificamos la estructura de datos para poder usar ggplot
78     graficable <- data.frame(anyo=NULL, serie= NULL, valores=NULL)
79     for (serie in dimnames(datos)[[2]][2:length(dimnames(datos)[[2]])
    ]){
80         graficaserie <- data.frame(anyo= datos$x[1]:datos$x[length(
            datos$x)], serie = rep(serie,length(datos$x)), valores =
            datos[serie])
81         colnames(graficaserie) <- c("anyo","serie","valores")
82         graficable <- bind_rows(grficable, graficaserie)
83     }
84
85     # Lo dibujamos
86     resultado <- NULL

```

```

87     resultado <- ggplot(graficable ,
88       aes(x = anyo , y = valores)) +
89       geom_line( aes(color=serie),size = 0.95,linetype = "twodash")
90       +
91       scale_color_manual(values= unlist(mis_colores))+
92       labs( x = "Aos", y = "Tasa", color = leyenda, title = titulo,
93         subtitle = subtítulo) +
94       theme_light()+
95       theme(legend.text = element_text(size = 9) ) +
96       theme(legend.key = element_rect(fill = "white", colour = "
97         white"))
98
99     # Para destacar un pais
100     if (!is.null(referente)){
101       destacada <- data.frame(x=datos$x, valores= datos[,referente
102         ,])
103       resultado <- resultado +
104         geom_line(data=destacada, aes(x=x, y= valores ), size =
105           3.85,linetype = "solid", alpha=0.3)
106     }
107     return(resultado)
108   }
109 }
110
111 # Dibuja la serie temporal por paises que tengamos en nuestra base
112 de datos
113 dibuja_series_temporales_por_paises_miHMDdata <- function(misPaises,
114   miSexo, miFichero, miEdad, miPeriodo, referente, titulo,
115   subtítulo) {
116
117   conjunto <- NULL
118   conjunto <- data.frame(x = miPeriodo)
119   # Recorremos todos los paises
120   for (miPais in misPaises) {
121     datosPais <-selecciona_unaEdad_miHMDdata(miPais = miPais, miSexo
122       = miSexo, miFichero = miFichero, miEdad = miEdad, miPeriodo
123       = miPeriodo)
124     elemento <- data.frame(pais = datosPais)
125     ifelse (is.null(conjunto), conjunto <- elemento, conjunto <-
126       cbind(conjunto, elemento))
127   }
128
129   colnames(conjunto) <- c("x", misPaises)
130   # Lo pintamos
131   imagen <- dibuja_miHMDdata(datos = conjunto, referente = referente
132     , titulo = titulo, subtítulo = subtítulo, leyenda="Paises")
133   return(imagen)
134 }
135
136 # Dibuja la serie temporal por edades
137 dibuja_series_temporales_por_edades_miHMDdata <- function(miPais,
138   miSexo, miFichero, misEdades, miPeriodo, referente, titulo,

```

```

    subtítulo) {
125
126 conjunto <- NULL
127 conjunto <- data.frame(x = miPeriodo)
128 # Recorremos todas las edades
129 for (miEdad in misEdades) {
130     datosPais <-selecciona_unaEdad_miHMDdata( miPais = miPais,
        miSexo = miSexo, miFichero = miFichero, miEdad = miEdad,
        miPeriodo = miPeriodo)
131     elemento <- data.frame(pais = datosPais)
132     ifelse (is.null(conjunto), conjunto <- elemento, conjunto <-
        cbind(conjunto, elemento))
133 }
134
135 colnames(conjunto) <- c("x", misEdades)
136 # Lo pintamos
137 imagen <- dibuja_miHMDdata(datos = conjunto, referente = referente
    , titulo = titulo, subtítulo = subtítulo , leyenda= "Edades")
138 return(imagen)
139 }
140
141 # Funcion para salvar los graficos
142 salva <- function(migrafico, fichero, directorio){
143     # Lugar para guardarlo
144     destino = directorio
145     if (!file.exists(destino)) dir.create(file.path(midirectorio,
        destino))
146     setwd(file.path(midirectorio, destino))
147     # Parametros del fichero que guardemos
148     png(fichero, width = 4800, height = 3200, res = 350)
149     imagen + geom_point()
150     print(imagen)
151     dev.off()
152     return(imagen)
153 }
154
155
156 # Ejemplos de uso
157
158 # CASO: dibujamos la tasa de fallecimiento en hombres de 25 años de
    todos los países entre 1910 y 2000, remarcando España
159 imagen <- NULL
160 imagen <- dibuja_series_temporales_por_paises_miHMDdata(misPaises =
    coleccion_paises, miSexo = "male", miFichero = "rate", miEdad =
    "25", miPeriodo = "1910":"2000", referente = "Spain", titulo =
    "Historico de las tasas de fallecimiento en Europa", subtítulo =
    "A los 25 años")
161 # Lo guardamos
162 salva(migrafico=imagen, fichero = paste("Titulo fichero.png"),
    directorio = paste(directorio,"Nombre de carpeta"))
163

```

```

164
165 # CASO: dibujamos del numero de fallecidos totales, a todas la
      edades, entre 1750 y 2020, remarcando Espana
166 for (edad in "0":"110"){
167     imagen <- NULL
168     imagen <- dibuja_series_temporales_por_paises_miHMDdata(misPaises
      = coleccion_paises, miSexo = "total", miFichero = "death",
      miEdad = edad, miPeriodo = "1750":"2020", referente = "Spain"
      , titulo = "Historico de fallecidos en Europa", subtítulo =
      paste("A los",edad,"aos"))
169     # Lo guardamos
170     salva(migrafo=imagen, fichero = paste("Fallecidos a los",edad,".
      png"), directorio = paste(directorio,"Nombre de carpeta") )
171 }
172
173
174 # CASO: dibujamos de los 15 a los 50 años la tasa de mortalidad
      femenina en España, entre 1910 y 2000, remarcando la edad de 30
      años
175 imagen<-NULL
176 imagen <- dibuja_series_temporales_por_edades_miHMDdata(miPais = "
      Spain", miSexo = "female", miFichero = "rate", misEdades = "15":
      "50", miPeriodo = "1910":"2000", referente = '30', titulo =
      paste("Historico de las tasas de fallecimiento en",pais),
      subtítulo = "Mujeres jvenes")
177 # Lo guardamos
178 salva( migrafo=imagen, fichero = paste(pais,"Titulo Fichero.png"),
      irectorio = paste(directorio,"Nombre carpeta"))
179
180
181 # CASO: dibujamos el numero de fallecidos totales entre 50 y 100
      años, en todos los países entre 1978 y 2020, remarcando los 65
      años
182 for (pais in coleccion_paises){
183     imagen<-NULL
184     imagen <- dibuja_series_temporales_por_edades_miHMDdata(miPais =
      pais, miSexo = "total", miFichero = "death", misEdades = "50":
      "100", miPeriodo = "1978":"2020", referente = '65', titulo =
      paste("Historico de fallecimientos en",pais), subtítulo = "
      total edades avanzadas")
185     # Lo guardamos
186     salva(migrafo=imagen, fichero = paste(pais,"Nombre fichero.png")
      , directorio = paste(directorio,"Nombre carpeta"))
187 }

```

CÓDIGO 11: Grafico3D.R

```
1 # Resumen de los datos disponible en https://www.mortality.org/
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library(grDevices)
7 library(shape)
8 library(plot3D)
9
10 # Para dar homogeneidad a todos los gráficos
11 mis_colores <- readRDS (file = "Localizacion del fichero")
12
13 # Cogemos los datos que necesitamos
14 euHMDdata <- readRDS (file = "Localizacion del fichero")
15 miHMDdata <- readRDS (file = "Localizacion del fichero")
16 # Cogemos los nombres de nuestra base de datos
17 coleccion_paises <- names(miHMDdata)
18
19 # Introducimos los datos de Europa en nuestra base de datos
20 miHMDdata[["Europe"]] <- euHMDdata[["Europe"]]
21 coleccion_extendida_paises <- names(miHMDdata)
22
23 # Guardamos los gráficos que hagamos en el directorio deseado
24 midirectorio <- "Localizacion del directorio"
25 setwd(file.path(midirectorio))
26
27 # Creamos una carpeta dentro del directorio, si ya existe, entonces
  lo guardamos ahí
28 destino = "Gráficos 3D"
29 if (!file.exists(destino)) dir.create(file.path(midirectorio,
  destino))
30 setwd(file.path(midirectorio, destino))
31
32 # Función para seleccionar el país que queramos, los datos que
  queramos (pop, death, exposure o rate), el sexo que queramos (
  male, female o total) y las edades y los años deseados, con el
  ángulo que queramos
33 selecciona_todasEdades_miHMDdata <- function(miPais, miFichero,
  miSexo, miEdad, miAnyos, miColor, theta, phi){
34
35   # Periodo
36   desde <- miAnyos[1]
37   hasta <- miAnyos[length(miAnyos)]
38
39   # Edades
40   miAges <- dimnames( miHMDdata[[miPais]][[miFichero]][[miSexo]] )
    [[1]]
41
42   #Años
```



```

43 miyears <- dimnames( miHMDdata[[miPais]][[miFichero]][[miSexo]] )
    [[2]]
44
45 # Si el rango no es correcto, paramos:
46 if (hasta < miyears[[1]] | desde > miyears[[length(miyears)]])
    break
47
48 # Si el rango es correcto, cogemos los datos
49 if (desde < miyears[[1]]) {
50     primero <- 1
51     desplazamiento <- match(miyears[[1]], ancho)
52 } else {
53     primero <- match(desde, miyears)
54     desplazamiento = -primero + 2
55 }
56 if (hasta > miyears[[length(miyears)]]) {
57     ultimo <- length(miyears)
58 } else {
59     ultimo <- match(hasta, miyears)
60 }
61
62 midata = miHMDdata[[miPais]][[miFichero]][[miSexo]][miAges[1]:
    miAges[length(miAges)],primero:ultimo]
63
64 return(midata)
65 }
66
67 # Funcion para dibujar
68 dibuja <- function(midata, mititulo, miColor, theta, phi){
69
70     # Colores
71     niveles<- length(mis_colores)
72     col.pal<-colorRampPalette(c("blue", "lightblue"))
73     colors<-col.pal(niveles)
74
75     z.facet.range<-cut(midata, niveles)
76
77     midibu <- persp(x = seq(0, 1, length.out = nrow(midata)), y = seq
        (0, 1, length.out = ncol(midata)), z=midata, xlab="Edad", ylab
        ="Aos", zlab="misDatos", main=mititulo, theta=theta, phi=phi,
        shade = 0.19, ticktype = "detailed", nticks=10, r=3, scale=
        TRUE, expand=0.9, col = drapecol(midata), border="grey70", box
        =TRUE)
78
79     return(midibu)
80 }
81
82 # Funcion para salvar los graficos
83 dibuja_y_salva <- function(midata, mititulo, miColor, theta, phi,
    destino){
84     png(destino,width=6000,height=4000, res = 350)

```

```

85     dibuja (midata=midata, mititulo=mititulo, miColor=miColor, theta=
        theta, phi=phi)
86     dev.off()
87 }
88
89
90 # Ejemplos de uso
91
92 # CASO: Poblacion Europa
93 midata <- selecciona_todasEdades_miHMDdata(miPais="Europe",
        miFichero="pop", miSexo="total", miEdad="0":"110", miAnyos="1990
        ":"2017")
94 dibuja(midata=midata, mititulo="Poblacion Europea ambos sexos",
        miColor="lightblue", theta=120, phi=20)
95 # Lo guardamos
96 dibuja_y_salva(midata=midata, mititulo="Poblacion Europea ambos
        sexos", miColor="lightblue", theta=150, phi=40, destino= "Nombre
        fichero.png")
97
98
99 # CASO: Animacion de fallecimientos en Espana
100 midata <- selecciona_todasEdades_miHMDdata(miPais="Spain", miFichero
        ="death", miSexo= "male", miEdad="0":"110", miAnyos="1990":"2017
        ")
101 # Giramos el grafico
102 for (i in 1:36){
103     dibuja(midata=midata, mititulo="Tasa de fallecimientos en Europa
        masculina", miColor="lightblue", theta=(100+10*i), phi= (20 +
        i ))
104     # Lo guardamos
105     dibuja_y_salva(midata= midata, mititulo="Fallecimientos en Europa
        masculina", miColor="lightblue", theta=(100+10*i), phi= (20 +
        i ), destino= paste("ANIMACION decesos de Europa",i,".png"))
106 }

```

Modelización

CÓDIGO 12: ModelosyELI.R

```
1 # Modelos LC, BCD y P-SPLINES a distintas edades
2 # (CC BY-NC-SA 4.0) vickymubar@gmail.com
3 # junio de 2021
4
5 # Librerías necesarias
6 library(xtable)
7 library("ilc")
8 library("forecast")
9 library(StMoMo)
10 library(MortalitySmooth)
11 library(svcM)
12 library("mgcv")
13 library(png)
14 library(shape)
15 library(plot3D)
16
17 #para dar homogeneidad a todos los graficos
18 mis_colores <- readRDS (file = "I:\\Dropbox\\TFM vicky\\R\\mis_
    colores.rds")
19
20 # Para dar uniformidad a todos los graficos
21 mis_colores <- readRDS (file = "Localizcion del fichero")
22
23 # Cogemos los datos que necesitamos
24 euHMDdata <- readRDS (file = "Localizcion del fichero")
25 miHMDdata <- readRDS (file = "Localizcion del fichero")
26 # Cogemos los nombres de nuestra base de datos
27 coleccion_paises <- names(miHMDdata)
28
29 # Introducimos los datos de Europa en nuestra base de datos
30 miHMDdata[["Europe"]] <- euHMDdata[["Europe"]]
31 coleccion_extendida_paises <- names(miHMDdata)
32
33 # Creamos una carpeta dentro del directorio, si ya existe, entonces
    lo guardamos ahí
34 midirectorio<-"Localizacion del directorio"
35 setwd(file.path(midirectorio))
36 destino = "Modelos"
37 if (!file.exists(destino)) dir.create(file.path(midirectorio,
    destino))
38 setwd(file.path(midirectorio, destino))
39
40 # Preparamos los datos, usaremos variables globales
41
42 miFichero <- "rate"
43 # Por ejemplo la esperanza de vida en Espana en 2019
```

```

44 miEdad <- "84"
45 miPais <- "Europe"
46 # Realmente esta variable se cambia en los bucles de despues
47 # Hay que guardar los resultados tras modelar cada vez, dado que lo
    hacen para un sexo concreto
48 miSexo <- "female"
49 desde <- "1990"
50 # Hasta donde consideramos datos historicos
51 hasta <- "2011"
52 # Ano final de la proyeccion
53 futuro <- "2025"
54 # Numero de anos de poyeccion
55 horizonte <- as.integer(futuro) - as.integer(hasta)
56 # Anos en base a los que hacemos los calculos
57 anyos_a_considerar <- desde:hasta
58 # Edades a estudiar
59 edades_a_considerar <- 40:100
60 # Numero de simulaciones
61 nsim <- 1000
62 # Edades para los graficos
63 rango=c("50","60","70","80","90")
64
65 anyos_datos_reales <- dimnames(miHMDdata[[miPais]][["exposure"]][[
    miSexo]])[[2]]
66 primero_index <- match(desde,anyos_datos_reales)
67 ultimo_index <- match(hasta,anyos_datos_reales)
68 if (ultimo_index > length(anyos_datos_reales)) print("NO TIENES
    DATOS")
69
70 todas_las_edades <- dimnames(miHMDdata[[miPais]][["death"]][[miSexo
    ]])[[1]]
71 Nanyos = as.integer(anyos_a_considerar)
72
73 todos_los_anyos <- desde:futuro
74
75 # Variables globales, se cambian con <<- en las funciones
76
77 # Para LC y CBD
78 LC <- NULL
79 LCfit <- NULL
80 LCres <- NULL
81 LCfor <- NULL
82 LCsim <- NULL
83
84 CBD <- NULL
85 CBDfit <- NULL
86 CBDres <- NULL
87 CBDfor <- NULL
88 CBDsim <- NULL
89
90 Pop <- NULL

```

```

91 Dx      <- NULL
92 Exp      <- NULL
93 Rate     <- NULL
94
95 miStMoMo <-NULL
96
97 # Para P-splines
98 x <- desde:hasta
99 y_death <- NULL
100 e_exp    <- NULL
101 z_deaths <- NULL
102 y.p spline <- NULL
103
104 miExp     <- NULL
105 fit1D     <- NULL
106 fit1Dfor  <- NULL
107 fit1Dfor1d <- NULL
108 fit1Dfor3d <- NULL
109 fit2D     <- NULL
110 fit2Dint  <- NULL
111 fit2Dpredicted <- NULL
112
113 # FUNCIONES de simulacion
114
115 # Para coger los datos necesarios en los modelos
116 seleccionar_datos_simulacion <- function(miPais, anyos_a_considerar,
117     miEdad, edades, sexo){
118
119     miSexo <-<- sexo
120
121     # Antes de guardar datos borramos lo anterior
122     Dx <-<- NULL
123     Dx <-<- miHMDdata[[miPais]][["death"]] [[miSexo]][,primero_index
124         :ultimo_index]
125     Exp <-<- NULL
126     Exp <-<- miHMDdata[[miPais]][["exposure"]][miSexo]][,primero_index
127         :ultimo_index]
128     Rate <-<- NULL
129     Rate <-<- miHMDdata[[miPais]][["rate"]] [[miSexo]][,primero_
130         index:ultimo_index]
131
132     y_death <-<- NULL
133     y_death <-<- Dx [match(miEdad, todas_las_edades ), ] #death
134         pais edad
135     e_exp <-<- NULL
136     e_exp <-<- Exp[match(miEdad, todas_las_edades ), ] #death pais
137         edad
138     z_deaths <-<- NULL
139     z_deaths <-<- Dx [match(edades_a_considerar,todas_las_edades),]
140     miExp <-<- NULL
141     miExp <-<- Exp[match(edades_a_considerar,todas_las_edades),]

```

```

136
137 y_extendido <- NULL
138 y_extendido <- miHMDdata[[miPais]][["death"]] [[miSexo]][ ,
    primero_index:length(anyos_datos_reales)][match(miEdad, edades
    ), ] # NO SON TODOS LOS DATOS REALES. SOLO HASTA 2011
139 e_extendido <- NULL
140 e_extendido <- miHMDdata[[miPais]][["exposure"]][miSexo][ ,
    primero_index:length(anyos_datos_reales)][match(miEdad, edades
    ), ]
141
142 y.pspline <-<- NULL
143 y.pspline <-<- rep(NA,length(todos_los_anyos))
144
145 for ( i in 1:length(y_extendido)) y.pspline[[i]] <-<- log(y_
    extendido/e_extendido)[[i]]
146
147 miyears <-<- dimnames(miHMDdata[[miPais]][["rate"]][miSexo])
    [[2]]
148 primero <-<- match (anyos_a_considerar[[1]], miyears )
149 ultimo <-<- match (anyos_a_considerar[[length(anyos_a_considerar)
    ]], miyears )
150
151 # Para LC y CBD
152 miDEMOGdata <-<- NULL
153 miDEMOGdata <-<- demogdata(data = Rate, pop = miHMDdata[[miPais]][["
    pop"]] [[miSexo]][,primero_index:ultimo_index], ages = edades,
    years = Nanyos, type = miHMDdata[[miPais]][["type"]], label =
    miHMDdata[[miPais]][["country"]], name = miSexo)
154
155 miStMoMo <-<- StMoMoData(miDEMOGdata)
156
157 y.pspline <-<- y.pspline
158 }
159
160
161 # Calculos del modelo LC
162 modelar_LC <-<- function(ages.fit=edades_a_considerar, years.fit=anyos
    _a_considerar){
163     # Para datos que siguen una Poisson
164     LC <-<- lc(link = "log")
165     LCfit <-<- fit(LC, data = miStMoMo, ages.fit = edades_a_considerar,
    years.fit=years.fit)
166     # Bondad del modelo = residuos
167     LCres <-<- residuals(LCfit)
168     LCfor <-<- forecast(LCfit, h = horizonte)
169     LCsim <-<- simulate(LCfit, nsim = nsim, h = horizonte)
170 }
171
172
173 # Calculos del modelo CBD

```

```

174 modelar_CBD <- function(ages.fit=edades_a_considerar, years.fit=
    anyos_a_considerar){
175     desviacion <- function(x,ages) x - mean(ages)
176     # Para datos que siguen una Poisson
177     CBD <- StMoMo(link="log",staticAgeFun=FALSE, periodAgeFun = c(
        "1",desviacion) )
178     CBDfit <- fit(CBD, data = miStMoMo, ages.fit = edades_a_
        considerar, years.fit=anyos_a_considerar)
179     CBDres <- residuals(CBDfit)
180     CBDfor <- forecast(CBDfit, h = horizonte)
181     CBDsim <- simulate(CBDfit, nsim = nsim, h = horizonte)
182 }
183
184
185 # Calculos del modelo P-SPLINES
186
187 # Para 1 dimension
188 modelar_PSPLINES_1D <- function(){
189
190     fit1D <- Mort1Dsmooth(x = x, y = y_death, offset = log(e_exp))
191
192     # Grados de libertad
193     df=6
194     df_label <- paste("Fix df=",df=df)
195     fit1Daic <- Mort1Dsmooth(x = x, y = y_death, offset = log(e_exp),
        method = 2)
196     fit1Dsub <- Mort1Dsmooth(x = x, y = y_death, offset = log(e_exp),
        method = 3, lambda = 10000)
197     fit1Dfix <- Mort1Dsmooth(x = x, y = y_death, offset = log(e_exp),
        method = 4, df = df)
198
199     k <- fit1D$ndx + fit1D$deg
200     mgam <- gam(y_death ~ offset(log(e_exp)) + s(x, bs = "ps", k = k
        ), family = "poisson")
201
202     fit1D1d <- Mort1Dsmooth(x = x, y = y_death, offset = log(e_exp
        ), pord = 1)
203     fit1D3d <- Mort1Dsmooth(x = x, y = y_death, offset = log(e_exp
        ), pord = 3)
204
205     fit1Dfor <- predict(fit1D, newdata = todos_los_anyos, se.fit =
        TRUE)
206     fit1Dfor1d <- predict(fit1D1d, newdata = todos_los_anyos)
207     fit1Dfor3d <- predict(fit1D3d, newdata = todos_los_anyos)
208 }
209
210 # Para 2 dimensiones
211 modelar_PSPLINES_2D <- function(){
212
213     fit2D <- Mort2Dsmooth(x = edades_a_considerar, y = Nanyos, Z = z_
        deaths, offset = log(miExp))

```

```

214
215 set.seed(1)
216 whi0 <- sample(x = 1: ((length(edades_a_considerar) * length(
    Nanyos))), size = nsim)
217 W <- matrix(1, length(edades_a_considerar), length(Nanyos))
218 W[whi0] <- 0
219
220 fit2Dint <-<- Mort2Dsmooth(x = edades_a_considerar, y = Nanyos, Z =
    z_deaths, offset = log(miExp), W = W)
221
222 fit2Dpredicted <-<- predict( object = fit2Dint, newdata = list(x=
    edades_a_considerar, y=todos_los_anyos), type = "link", se.fit
    = FALSE)
223 }
224
225
226 # Para comprobar que los resultados son correctos
227 comprobar_modelado_PSPLINES_1D <- function(fit1D, x, y, e){
228
229     summary(fit1D)
230     plot(fit1D)
231     lines(x, fit1Daic$logmortality, col = 3, lwd = 2)
232     lines(x, fit1Dsub$logmortality, col = 4, lwd = 2)
233     lines(x, fit1Dfix$logmortality, col = 5, lwd = 2)
234
235     legend("bottom", c("Actual", "BIC", "AIC", "Subjective",df_label),
        col = 1:5, pch = c(1, -1, -1, -1, -1), lwd = c(1, 2, 2, 2, 2)
        , lty=c(0, 1, 1, 1, 1), bty = "n")
236
237     lines(x, log(fitted(mgam)/e_exp), lwd = 2, lty = 2, col = "green4"
        )
238     legend("bottomleft", c("GCV/UBRE", "REML"), col = c("green4", "
        red4"), title = "library(\"mgcv\")", lwd = c(2, 2), lty = c(2,
        2), bty = "n")
239 }
240
241
242 # Para comprobar que los resultados son correctos
243 comprobar_modelado_PSPLINES_2D <- function(fit2D){
244     res2D <- residuals(fit2D, type = "pearson")
245     res.grid <- expand.grid(list(x = edades_a_considerar, y = Nanyos))
246     res.grid$res <- c(res2D)
247     levelplot(res ~ Nanyos * edades_a_considerar, data = res.grid, at
        = c(min(res2D), -2, -1, 1, 2, max(res2D)))
248     plot(fit2D, palette = "terrain.colors")
249 }
250
251
252 # FUNCIONES de representacion
253
254 # Grafico en 3D

```



```

255 dibuja <- function(midata, mititulo, miColor, theta, phi, zlim=NULL,
    destino=NULL){
256
257   if (! is.null(destino)) png(destino, width = 4000, height = 4000,
    res = 350)
258   if ( is.null(zlim)) zlim= range(midata)
259
260   midibu <- persp(x = seq(as.integer(unlist(dimnames(midata)[1]))
    [1], as.integer(unlist(dimnames(midata)[1]))[length(as.integer
    (unlist(dimnames(midata)[1])))]), length.out = nrow(midata)), y
    = seq(as.integer(unlist(dimnames(midata)[2]))[1], as.integer(
    unlist(dimnames(midata)[2]))[length(as.integer(unlist(dimnames
    (midata)[2])))]), length.out = ncol(midata)), z=midata, xlab="
    Edad", ylab="Anos", zlab="Datos obtenidos de las simulaciones"
    , zlim = zlim, main=mititulo, theta=theta, phi=phi, shade =
    0.19, ticktype = "detailed", nticks=10, r=3, scale=TRUE,
    expand=0.9, col = drapecol(midata), border="grey70", box=TRUE)
261
262   if (! is.null(destino)) dev.off()
263   midibu
264   return(midibu)
265 }
266
267
268 # Para dibujar P SPLINE
269 dibuja_pspline<-function(x, y, x.pspline, y.pspline, y.pspline.3d, y
    .pspline.1d, y.pspline.for, y.pspline.new){
270
271   plot(x, y, xlim = range(x.pspline), ylim = range(y.pspline,-4.0,
    -5.5), pch=16 )
272
273   compara_pspline(x, y, x.pspline, y.pspline, y.pspline.3d, y.
    pspline.1d, y.pspline.for, y.pspline.new)
274
275   legend("top", c(paste("Datos reales para el calculo hasta", hasta)
    , "d=2", "d=3", "d=1", paste("Datos reales para comparar hasta
    ",anyos_datos_reales[[length(anyos_datos_reales)]])), lty = c
    (0, 1, 1, 1, 0), col = 1:5, pch = c(1, -1, -1, -1, 1), lwd = c
    (1, 2, 2, 2, 1), bty = "n", cex = 0.7)
276 }
277
278
279 # Unicamente pinta sobre un grafico ya hecho
280 compara_pspline <- function(x, y, x.pspline, y.pspline, y.pspline.3d
    , y.pspline.1d, y.pspline.for, y.pspline.new){
281   lines(x, y , lwd = 1, col = 2)
282   lines(x.pspline, y.pspline, lwd = 5, col = 2)
283   lines(x.pspline, y.pspline + 2 * y.pspline.for, lwd=4, lty = 2,
    col = 2)
284   lines(x.pspline, y.pspline - 2 * y.pspline.for, lwd=4, lty = 2,
    col = 2)

```

```

285 lines(x.pspline, y.pspline.3d, lwd = 4, col = 3)
286 lines(x.pspline, y.pspline.1d, lwd = 4, col = 4)
287 lines(x.pspline, y.pspline.new, type="b", lwd = 2, col = 5)
288 }
289
290
291 # Para dibujar los modelos
292 dibuja_modelo <- function(x, xs, y, ys, ymin, ymax, ymin, ymax,
293                             ylabel, rango, titulo, cols, destino=NULL){
294
295     if (! is.null(destino)) png(destino, width = 2400, height = 1600,
296                               res = 350)
297
298     plot(x, y[rango[1],], xlim = range( x, xs ), ylim = range( ymin,
299                               ymax, ymin, ymax ), type = "l", xlab = "Ano", ylab = ylabel,
300                               main = titulo, lwd = 5)
301
302     compara_un_modelo ( x=x, xs=xs, y=y, ys=ys, rango=rango, cols=
303                               cols )
304
305     grid(nx = NULL, ny = NULL, col = "lightgray", lty = "dotted", lwd
306           = par("lwd"), equilogs = TRUE)
307
308     if (! is.null(destino)) dev.off()
309 }
310
311 # Para comparar un modelo
312 compara_un_modelo <- function(x, xs, y, ys, rango, cols){
313
314     for (Sedad in rango){
315         matlines(x, y[Sedad,], type = "l", lwd=5, lty = 1, col= cols[1])
316         matlines(xs, ys[Sedad, , 1:200], type = "l", lty = 5, col= cols)
317     }
318 }
319
320 # Para comparar varios modelos
321 compara_modelos <- function(x, xs, y, ys, ymin, ymax, ymin, ymax,
322                             ylabel, rango, titulo, cols, yHs, yHsmin, yHsmax, cols2, destino
323                             = NULL){
324
325     if (! is.null(destino)) png(destino, width = 2400, height = 1600,
326                               res = 350)
327
328     dibuja_modelo(x=x, xs=xs, y=y, ys=ys, ymin=ymin, ymax=ymax, ymin=
329                               ymin, ymax=ymax, ylabel=ylabel, rango=rango, titulo=titulo,
330                               cols=cols)
331
332     compara_un_modelo ( x, xs, y,yHs, rango, cols2)

```

```

325
326   if (! is.null(destino)) dev.off()
327 }
328
329
330 # Para comprobar que todo funciona
331 comprobar_ejecucion <- function(){
332
333   #LC
334   plot(LCfit)
335   plot(LCres, type="colourmap", reslim = c(-3.5,3.5))
336   plot(LCfor)
337
338   #CBD
339   plot(CBDfit)
340   plot(CBDres, type="colourmap", reslim = c(-3.5,3.5))
341   plot(CBDfor)
342
343   #PSPLINES
344   comprobar_modelado_PSPLINES_1D (fit1D = fit1D ,x=x, y=y_death, e=e
     _exp)
345
346   comprobar_modelado_PSPLINES_2D(fit2D = fit2D)
347
348
349   midata<- fit2Dpredicted
350
351   persp(x = seq(as.integer(unlist(dimnames(midata)[1]))[1], as.
     integer(unlist(dimnames(midata)[1]))[length(as.integer(unlist(
     dimnames(midata)[1]))]), length.out = nrow(midata)), y = seq(
     as.integer(unlist(dimnames(midata)[2]))[1], as.integer(unlist(
     dimnames(midata)[2]))[length(as.integer(unlist(dimnames(midata)
     [2]))]), length.out = ncol(midata)), z=midata, xlab="Edad",
     ylab="Anos", zlab="tasa de mortalidad", main=paste("PSPLINE
     Europa datos hasta 2011", miSexo), theta=220, phi=30, shade =
     0.19, ticktype = "detailed", nticks=10, r=3, scale=TRUE,
     expand=0.9, col = drapecol(midata), box=TRUE )
352 }
353
354 ### FIN DE LAS FUNCIONES
355
356
357 # Llamamos a los modelos
358
359 # para ambos sexos, comprobacion de que todo funciona
360 for (sexo in c( "female", "male" ) ) {
361
362   seleccionar_datos_simulacion(miPais=miPais, anyos=anyos_datos_
     reales, miEdad=miEdad, edades=todas_las_edades, sexo=sexo)
363

```

```

364     modelar_LC(ages.fit=edades_a_considerar, years.fit=anyos_a_
          considerar)
365
366     modelar_CBD(ages.fit=edades_a_considerar, years.fit=anyos_a_
          considerar)
367
368     modelar_PSPLINES_1D ()
369
370     modelar_PSPLINES_2D ()
371
372     comprobar_ejecucion()
373 }
374
375
376
377 ##### RESULTADO FINAL: comparacion de los 3 modelos entre si y
          contra datos reales
378
379 # Calculo final de la mejora de estos modelos
380 # Considera hasta=2011 un parametro
381
382 # Anos y edades a estudiar
383 periodo <- (as.integer(hasta)+1):futuro
384 Nedades <- 1:length(edades_a_considerar)
385 Nperiodo <- 1:length(periodo)
386
387 #Creamos la estructura en la que guarar los resultados
388 miRESULTADO <- NULL
389 miRESULTADO <- list(country ="Europe", type = "future_rate", year =
          futuro, ages = edades_a_considerar, datos_de_partida = list(
          female=matrix(c(1,11), nrow=1, ncol=2),male=matrix(c(5,6,7,8),
          nrow=1, ncol=4)), datos_de_control = list(female=matrix(c(1,11),
          nrow=1, ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)),
          RATEtradicional = list(female=matrix(c(1,11), nrow=1, ncol=2),
          male=matrix(c(5,6,7,8), nrow=1, ncol=4)), PromedioELI = list(
          female=matrix(c(2,22), nrow=1, ncol=2),male=matrix(c(0,1,2,1),
          nrow=1, ncol=4)), LCqxmedia = list(female=matrix(c(1,11), nrow
          =1, ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)), CBDqxmedia
          = list(female=matrix(c(1,11), nrow=1, ncol=2),male=matrix(c
          (5,6,7,8), nrow=1, ncol=4)), PSPLINES2Dpredicted = list(female=
          matrix(c(1,11), nrow=1, ncol=2),male=matrix(c(5,6,7,8), nrow=1,
          ncol=4)), LC_factor_mejora = list(female=matrix(c(1,11), nrow=1,
          ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)), CBD_factor_
          mejorav= list(female=matrix(c(4,44), nrow=1, ncol=2),male=matrix
          (c(0,1,2,1), nrow=1, ncol=4)), PSPLINESd2_factor_mejora = list(
          female=matrix(c(2,22), nrow=1, ncol=2),male=matrix(c(0,1,2,1),
          nrow=1, ncol=4)), shockELI = list(female=matrix(c(2,22), nrow=1,
          ncol=2),male=matrix(c(0,1,2,1), nrow=1, ncol=4)), ELI_factor_
          mejora = list(female=matrix(c(2,22), nrow=1, ncol=2),male=matrix
          (c(0,1,2,1), nrow=1, ncol=4)))
390

```

```

391 # Lo rellenos
392 for (miSexo in c("female","male")){
393
394   miRESULTADO[["LCqxmedia"]][[miSexo]] <- array(c(rep(0.00000001,
     length(edades_a_considerar)*length(Nperiodo))), dim =c(length(
     edades_a_considerar),length(Nperiodo)), dimnames=list (edades_
     a_considerar, periodo))
395   miRESULTADO[["CBDqxmedia"]][[miSexo]] <- miRESULTADO[["LCqxmedia"
     ]][[miSexo]]
396   miRESULTADO[["datos_de_partida"]][[miSexo]] <- miHMDdata[[miPais
     ]][["rate"]][[miSexo]][(edades_a_considerar[1]+1):(edades_a_
     considerar[length(edades_a_considerar)]+1),1:length(desde:
     hasta) ]
397   miRESULTADO[["datos_de_control"]][[miSexo]] <- miHMDdata[[miPais
     ]][["rate"]][[miSexo]][(edades_a_considerar[1]+1):(edades_a_
     considerar[length(edades_a_considerar)]+1),(length(desde:hasta
     )+1):length(colnames(miHMDdata[[miPais]][["rate"]][[miSexo]])
     )]
398
399   for (anyo_a_estimar in futuro ) {
400
401     referencia <- miHMDdata[[miPais]][["rate"]][[miSexo]][edades_
     a_considerar+1,hasta]
402     diferimiento <- 1 / (as.integer(anyo_a_estimar) - as.integer(
     hasta) )
403
404     seleccionar_datos_simulacion(miPais = miPais, anyos=anyos_datos_
     reales, miEdad = miEdad, edades=todas_las_edades, sexo=
     miSexo)
405     modelar_LC(ages.fit=edades_a_considerar, years.fit=anyos_a_
     considerar)
406     modelar_CBD(ages.fit=edades_a_considerar, years.fit=anyos_a_
     considerar)
407     modelar_PSPLINES_1D ()
408     modelar_PSPLINES_2D ()
409
410     miRESULTADO[["RATEtradicional"]][[miSexo]] <- miRESULTADO[["
     datos_de_partida"]][[miSexo]]
411     ultima_real <- length(desde:hasta)
412
413     for (extra in (ultima_real+1):length(desde:futuro))
414       miRESULTADO[["RATEtradicional"]][[miSexo]] = cbind(
         miRESULTADO[["RATEtradicional"]][[miSexo]],(miRESULTADO
         [["RATEtradicional"]][[miSexo]][,ultima_real]*0.8))
415
416     colnames( miRESULTADO[["RATEtradicional"]][[miSexo]] ) <- desde:
     futuro
417
418   # LC y CBD:
419   for ( edad in Nedades )
420     for ( anyo in Nperiodo ){

```

```

421     miRESULTADO[["LCqxmedia"]][[miSexo]] [edad, anyo] <- mean(
      LCsim[["rates"]][edad, anyo,])
422     miRESULTADO[["CBDqxmedia"]][[miSexo]] [edad, anyo] <- mean(
      CBDsim[["rates"]][edad, anyo,])
423   }
424   # P-splines
425   miRESULTADO[["PSPLINES2Dpredicted"]][[miSexo]] <- exp(
     fit2Dpredicted)
426
427   miRESULTADO[["LC_factor_mejora"]][[miSexo]] <- (miRESULTADO[["
     LCqxmedia"]][[miSexo]][, anyo_a_estimar] / referencia) ^
     diferimiento - 1
428   miRESULTADO[["CBD_factor_mejora"]][[miSexo]] <- (miRESULTADO[["
     CBDqxmedia"]][[miSexo]][, anyo_a_estimar] / referencia) ^
     diferimiento - 1
429   miRESULTADO[["PSPLINESd2_factor_mejora"]][[miSexo]] <- (exp(
     fit2Dpredicted[, anyo_a_estimar]) / referencia) ^
     diferimiento - 1
430 }
431 }
432
433 # Promedio de los 3 resultados
434 # calculo principal
435 for (miSexo in c("female", "male")){
436   miRESULTADO[["PromedioELI"]][[miSexo]] <-array (c(rep
     (0.000000000000000001, length(edades_a_considerar)*length(
     Nperiodo))), dim=c(length(edades_a_considerar), length(
     Nperiodo)), dimnames=list (edades_a_considerar, periodo))
437   for (anyo in 1:length(periodo)){
438     tmp_df <- structure(list(miRESULTADO[["LCqxmedia"]][[miSexo
     ]][,1], miRESULTADO[["CBDqxmedia"]][[miSexo]][, anyo],
     miRESULTADO[["PSPLINES2Dpredicted"]][[miSexo]][, (21+anyo)]),
     class = "data.frame", row.names = c(NA, -61L))
439     tmp_matriz <- as.matrix(tmp_df)
440     miRESULTADO[["PromedioELI"]][[miSexo]][, anyo] <- apply(tmp_
     matriz[, c(1,2,3)], 1, mean, na.rm = TRUE)
441   }
442
443   tmp_df <- structure(list(miRESULTADO[["LC_factor_mejora"]][[miSexo
     ]], miRESULTADO[["CBD_factor_mejora"]][[miSexo]], miRESULTADO
     [["PSPLINESd2_factor_mejora"]][[miSexo]]), class = "data.frame
     ", row.names = c(NA, -61L))
444   tmp_matriz <- as.matrix(tmp_df)
445   miRESULTADO[["ELI_factor_mejora"]][[miSexo]] <- apply(tmp_
     matriz[, c(1,2,3)], 1, mean, na.rm = TRUE)
446 }
447 }
448
449
450 # Representacion de factores de mejora
451 fmLCH=miRESULTADO[["LC_factor_mejora"]][["male"]]

```

```

452 fmLCM=miRESULTADO[["LC_factor_mejora"]][["female"]]
453 fmCBDH=miRESULTADO[["CBD_factor_mejora"]][["male"]]
454 fmCBDM=miRESULTADO[["CBD_factor_mejora"]][["female"]]
455 fmPSH=miRESULTADO[["PSPLINESd2_factor_mejora"]][["male"]]
456 fmPSM=miRESULTADO[["PSPLINESd2_factor_mejora"]][["female"]]
457 fmELIH=miRESULTADO[["ELI_factor_mejora"]][["male"]]
458 fmELIM=miRESULTADO[["ELI_factor_mejora"]][["female"]]
459 plot(x= edades_a_considerar, y=fmELIH, xlim = range(40,100), ylim =
      range(-0.08,0.010) , pch=16 , main="Factores de mejora male y
      female")
460 lines(x=edades_a_considerar, y=fmELIM, lwd = 2, lty=1, col = "orange
      ")
461 lines(x=edades_a_considerar, y=fmCBDM, lwd = 2, lty=2, col = "cyan")
462 lines(x=edades_a_considerar, y=fmCBDM, lwd = 2, lty=2, col = "blue")
463 lines(x=edades_a_considerar, y=fmPSM, lwd = 2, lty=5, col = "green")
464 lines(x=edades_a_considerar, y=fmPSM, lwd = 2, lty=11, col = "yellow
      ")
465 lines(x=edades_a_considerar, y=fmLCM, lwd = 4, lty=8, col = "magenta
      ")
466 lines(x=edades_a_considerar, y=fmLCH, lwd = 4, lty=9, col = "pink")
467
468
469 # Calculo de antenamo los productos (1-lambda) que luego se
      necesitaran
470 ELILambda <- NULL
471 ELILambda <- matrix(rep(NA,2*61),61,2)
472 ELI_Lambdas_triangularizada <- NULL
473 ELI_Lambdas_triangularizada <- list(female= matrix(rep(NA, length(
      edades_a_considerar)*length(edades_a_considerar)), nrow = length
      (edades_a_considerar), ncol = length(edades_a_considerar)), male
      = matrix(rep(NA, length(edades_a_considerar)*length(edades_a_
      considerar)), nrow = length(edades_a_considerar), ncol = length(
      edades_a_considerar)))
474 colnames(ELILambda)<-c("female","male")
475 rownames(ELILambda)<-edades_a_considerar
476
477 for ( miSexo in c("female","male")){
478   for ( anyo_a_estimar in futuro ) {
479
480     referencia <- miHMDdata[[miPais]][["rate"]][[miSexo]][edades_a_
      considerar+1,hasta]
481     diferimiento <- 1 / (as.integer(anyo_a_estimar) - as.integer(
      hasta) )
482
483     ELILambda[,miSexo] <- (( miRESULTADO[["PromedioELI"]][[miSexo
      ]][,anyo_a_estimar] / referencia) ^ (diferimiento)) - 1
484
485     ELI_Lambdas_triangularizada[[miSexo]][,1] <- (1- (ELILambda[,
      miSexo] * (-1)))
486     medias.desplazada <- ( 1 -( ELILambda[,miSexo] * (-1)))
487

```

```

488     for ( i in 2:length(edades_a_considerar)){
489         medias.desplazada <- c(medias.desplazada[(2):length(edades_
490             a_considerar)],NA)
491         ELI_Lambdas_triangularizada[[miSexo]][,i] <- ELI_Lambdas_
492             triangularizada[[miSexo]][,(i-1)] * medias.desplazada
493     }
494     rownames(ELI_Lambdas_triangularizada[[miSexo]]) <- edades_a_
495         considerar
496     colnames(ELI_Lambdas_triangularizada[[miSexo]]) <- 1:length(
497         edades_a_considerar)
498 }
499 }
500
501 # Exportar a LaTeX
502 print(xtable((ELILambda*(-100)), type = "latex"), file = "Nombre
503     fichero.tex")
504
505 # Y finalmente el calculo del SHOCK del modelo ELI en base a la
506     media de LC, CBD y P-splines para Europa
507 qx <- NULL
508 lx <- NULL
509 lx0 <- NULL
510 f_lsexo <- NULL
511 f_edad_base_al_contratar <- NULL
512 tramo_asegurable <- edades_a_considerar[1:(length(edades_a_
513     considerar)-1)]
514 ELI_shock <- NULL
515 ELI_shock <- list( female=matrix(0,nrow=(length(edades_a_considerar)
516     -1), ncol=(length(edades_a_considerar)-1), dimnames=list( edades
517     _a_considerar[1:(length(edades_a_considerar)-1)],1:(length(
518     edades_a_considerar)-1))), male=matrix(0,nrow=(length(edades_a_
519     considerar)-1), ncol=(length(edades_a_considerar)-1), dimnames=
520     list( edades_a_considerar[1:(length(edades_a_considerar)-1)],
521     1:(length(edades_a_considerar)-1))))
522
523 lx_desde_qx <- NULL
524 lx_desde_qx <- list( female=rep(0.0,length(todas_las_edades)), male=
525     rep(0.0,length(todas_las_edades)))
526
527 for (miSexo in c("female","male")){
528     lx_desde_qx[[miSexo]][1] <- 100000
529     for( ed_lx in 2:length(todas_las_edades) )
530         lx_desde_qx[[miSexo]][ed_lx] <- lx_desde_qx[[miSexo]][ed_lx-1]*
531             (1-euHMDdata[["Europe"]][["rate"]][miSexo][ed_lx-1],
532             hasta] )
533 }
534
535 plot(x=1:111, y=eu_lxHMDdata[["Europe"]][["lx"]][miSexo][,hasta],
536     main=paste("Diferencia de lx"), lwd = 1.5 , pch=16, xlab = "
537     duracion" )

```



```

521 lines(x=1:111, y=lx_desde_qx[[miSexo]], lwd = 1.5, lty=1, col = "
      blue")
522
523
524 # Calculo recursivo de los productos (1 - lambda prima)
525 f_lprima <- function(anyo_contrato){
526   if ( anyo_contrato == 0 ) return ( lx0 )
527   else return( f_lprima (anyo_contrato -1 ) * ( 1 - qx[anyo_
      contrato+f_edad_base_al_contratar] * ELI_Lambdas_
      triangularizada[[miSexo]][f_edad_base_al_contratar-39,anyo_
      contrato]))
528 }
529
530 # Calculo de shock
531 for (miSexo in c("female","male") ) {
532   qx_base <- NULL
533   qx_esperada <- NULL
534   qx_base <- matrix(0,nrow=(length(edades_a_considerar)-1), ncol
      =(length(edades_a_considerar)-1), dimnames=list( "1":"60","40"
      : "99"))
535   qx_esperada <- matrix(0,nrow=(length(edades_a_considerar)-1), ncol
      =(length(edades_a_considerar)-1), dimnames=list( "1":"60","40"
      : "99"))
536   f_lsexo <- miSexo
537
538   for ( edad_base_al_contratar in tramo_asegurable){
539     f_edad_base_al_contratar <- edad_base_al_contratar
540     qx <- euHMDdata[["Europe"]][["rate"]][[miSexo]][,hasta]
541     lx <- lx_desde_qx[[miSexo]]
542     lx0 <- lx[(edad_base_al_contratar+1)]
543
544     for ( anyo_contrato in 1:(length(tramo_asegurable)- edad_base_al
      _contratar+tramo_asegurable[1])){
545       qx_base[anyo_contrato,edad_base_al_contratar-39] <- ( 1
      - ( lx[edad_base_al_contratar+anyo_contrato+1] / lx0))
546       qx_esperada[anyo_contrato,edad_base_al_contratar-39] <- ( 1
      - ( f_lprima( anyo_contrato ) / lx0 ) )
547       ELI_shock[[miSexo]][edad_base_al_contratar-39,anyo_contrato]
      <- (( qx_base[anyo_contrato,(edad_base_al_contratar-39)]
      / qx_esperada[anyo_contrato,(edad_base_al_contratar-39)])
      - 1)
548     }
549   }
550 }
551
552 # Caso Vida Entera
553 ELI_shock_vida_entera <- NULL
554 ELI_vida_entera <- data.frame(female = matrix(rep(NA,length(edades_a
      _considerar)), length(edades_a_considerar)), male = matrix(rep(
      NA,length(edades_a_considerar)), length(edades_a_considerar)))
555

```

```

556 for ( miSexo in c("female","male"))
557     for ( i in 1:(length(edades_a_considerar)-1))
558         ELI_shock_vida_entera[[miSexo]][i] <- max(ELI_shock[[
                    miSexo]][i,])
559
560 # Preparamos tabla para Latex
561 rownames(ELI_shock_vida_entera) <- edades_a_considerar
562 factor=100
563 ELI_shock_tabla_resultado = cbind(ELI_shock[["female"]][,1]*factor,
                    ELI_shock[["male"]][,1]*factor, ELI_shock[["female"]][,5]*factor
                    , ELI_shock[["male"]][,5]*factor, ELI_shock[["female"]][,10]*
                    factor, ELI_shock[["male"]][,10]*factor, ELI_shock_vida_entera[[
                    "female"]]*factor, ELI_shock_vida_entera[["male"]]*factor)
564 colnames(ELI_shock_tabla_resultado) <- c("M1", "H1", "M5", "H5", "M10",
                    "H10", "MVE", "HVE")
565 print(xtable(ELI_shock_tabla_resultado, type = "latex"), file = "
                    Nombre fichero.tex")
566
567
568 # Ejemplos de algunos graficos del documento
569
570 # ELI_shock 2D
571 plot(x=1:60, y=ELI_shock[[miSexo]][1,], main=paste("ELI_shock para
                    varias edades"), lwd = 1.5 , pch=16, xlab = "duracion" )
572 lines(x=1:60, y=ELI_shock[[miSexo]][5,], lwd = 1.5, lty=1, col = "
                    blue")
573 lines(x=1:60, y=ELI_shock[[miSexo]][20,], lwd = 1.5, lty=1, col = "
                    red")
574
575 # Diferencia entre lx y lx_desde_qx
576 plot(x=1:111, y=eu_lxHMDdata[["Europe"]][["lx"]][miSexo][,hasta],
                    main=paste("diferencia entre lx (puntos) y lx_desde_qx (azul)",
                    edad_base_al_contratar), lwd = 1.5 , pch=16 )
577 lines(x=1:111, y=lx_desde_qx[miSexo], lwd = 1.5, lty=1, col = "
                    blue")
578 lines(x=1:111, y=(eu_lxHMDdata[["Europe"]][["lx"]][miSexo][,hasta]
                    - lx_desde_qx[miSexo])), lwd = 1.5, lty=1, col = "red")
579
580 # Comparacion de las qx en 3D
581 dibuja(qx_base, "qx_base", miColor="red", theta=240, phi=30, zlim=
                    NULL, destino=NULL)
582 dibuja(qx_esperada, "qx_esperada", miColor="blue", theta=240, phi
                    =30, zlim=NULL, destino=NULL)
583
584 # comparacion de qx en 2D
585 plot(x=1:60, y=rep(0,60), main="qx esperada (rojo) vs qx (azul)",
                    ylim = range(0,1), lwd = 1.5 , pch=16 )
586 for ( duracion_seguro in c(1,5,10,15,20,25,30,35,40)){
587     lines(x=1:60, y=qx_esperada[duracion_seguro,], lwd = 1.5, lty=i,
                    col = "red")

```

```

588   lines(x=1:60, y=qx_base[duracion_seguro,], lwd = 0.9, lty=i, col =
        "blue")
589 }
590
591 # Shocks en 2D y 3D
592 for (miSexo in c("female","male")) {
593
594   # ELI SHOCK en 3d
595   for (angulo in 30)
596     dibuja(midata= ELI_shock[[miSexo]], mititulo = paste("ELI shock
        para ", ifelse(miSexo == "female", "mujeres", "hombres")),
        miColor = "lightblue", theta = angulo, phi=20, zlim=range
        (-0.001,0.8), destino=paste("ELI_shock 3D",miSexo,".png"))
597
598   # ELI SHOCK en 2d
599   png(paste("ELI_shock 2D",ifelse(miSexo=="female", "mujeres", "
        hombres") ,".png"), width = 3000, height = 3000, res = 350)
600   tradicion=rep(20,length(1:58))
601   plot(x="40":"97", y=tradicion, xlim = range(40,100), ylim = range
        (-5,75), xlab = "edades", ylab = "ELI_shock", pch=16, main=
        paste("Indice Europeo de Longevidad",ifelse(miSexo=="female",
        "mujeres", "hombres")) ) )
602   grid(col = 'grey80', lty = 1, lwd = 0.5)
603   for (i in 1:60) lines(x="40":"97", y=100*ELI_shock[[miSexo]][1:58,
        i], lwd = 0.4, lty=4, col = i)
604   lines(x="40":"97", y=100*ELI_shock[[miSexo]][1:58,1], lwd=3, lty
        =1, col="coral")
605   lines(x="40":"97", y=100*ELI_shock[[miSexo]][1:58,10], lwd=3, lty
        =1, col="darkgoldenrod2")
606   lines(x="40":"97", y=100*ELI_shock[[miSexo]][1:58,20], lwd=3, lty
        =1, col="cyan3")
607   lines(x="40":"97", y=100*ELI_shock[[miSexo]][1:58,30], lwd=3, lty
        =1, col="darkorchid1")
608   lines(x="40":"97", y=100*ELI_shock_vida_entera[[miSexo]][1:58],
        lwd=6, lty=1, col="red")
609   legend("topright", c("20%", "1 ano", "10 anos", "20 anos", "30
        anos", "Vida Entera"), col = c("black","coral1","
        darkgoldenrod2","cyan3","darkorchid1","red"), pch = c(1, -1,
        -1, -1, -1, -1), lwd = c(6, 2, 2, 2, 2, 2), lty=c(0, 1, 1, 1,
        1, 1), bty = "n")
610   dev.off()
611 }
612
613
614 # Ejemplos de estudios
615
616 # LC en escala logaritmica
617 dibuja_modelo(x=LCfit$years, xs=LCsim$years, y=log(Rate), ys=log(
        LCsim$rates), ymin= log(Rate[rango[1],]), ymax= log(Rate[rango[
        length(rango)],]), ymin=log(LCsim$rates[rango[1], , 1:200]),
        ymax=log(LCsim$rates[rango[length(rango)], , 1:200]), ylabel="

```

```

    Logaritmo de la tasa de fallecimientos", rango=rango, titulo=
    paste(miPais,"LC tasas de mortalidad a los", paste( unlist(rango
    ), collapse=', '),ifelse(miSexo=="female", "mujeres", "hombres")
    ), cols = hcl.colors(10, "Green-Orange"))
618 # De forma manual ponemos datos en el grafico
619 # Hay que ajustar dependiendo de lo representado
620 text(2011.5, -1.7, 50)
621 text(2011.5, -2.8, 60)
622 text(2011.5, -3.8, 70)
623 text(2011.5, -4.7, 80)
624 text(2011.5, -5.6, 90)
625
626
627 # Analogo para CBD
628 dibuja_modelo(x=CBDfit$years, xs=CBDsim$years, y=log(Rate), ys=log(
    CBDsim$rates), ymin= log(Rate[rango[1],]), ymax= log(Rate[rango[
    length(rango)],]), ymin=log(CBDsim$rates[rango[1], , 1:200]),
    ymax=log(CBDsim$rates[rango[length(rango)], , 1:200]), ylabel="
    Logaritmo de la tasa de fallecimientos", rango=rango, titulo=
    paste(miPais,"CBD tasas de mortalidad de ",ifelse(miSexo=="
    female", "mujeres", "hombres")," a los", paste( unlist(rango),
    collapse=', ')), cols = hcl.colors(10, "Purple-Blue"))
629 text(2011.5, -1.9, 50)
630 text(2011.5, -3.25, 60)
631 text(2011.5, -4.5, 70)
632 text(2011.5, -5.3, 80)
633 text(2011.5, -6.3, 90)
634
635
636 # Comparacion entre LC y CBD
637 compara_modelos(x=CBDfit$years, xs=CBDsim$years, y=log(Rate), ys=log
    (CBDsim$rates), ymin= log(Rate[rango[1],]), ymax= log(Rate[rango
    [length(rango)],]), ymin=log(CBDsim$rates[rango[1], , 1:200]),
    ymax=log(CBDsim$rates[rango[length(rango)], , 1:200]), ylabel="
    Logaritmo de la tasa de fallecimientos", rango=rango, titulo=
    paste(miPais,"LC vs CBD tasas de mortalidad de ",ifelse(miSexo==
    "female", "mujeres", "hombres")," a los", paste( unlist(rango),
    collapse=', ')), cols = hcl.colors(10, "Terrain"), yHs=log(LCsim
    $rates), yHsmin=log(LCsim$rates[rango[1], , 1:200]), yHsmax=log(
    LCsim$rates[rango[length(rango)], , 1:200]), cols2 = hcl.colors
    (10, "Pastel 1"))
638 text(2011.5, -1.65, 90)
639 text(2011.5, -2.75, 80)
640 text(2011.5, -3.85, 70)
641 text(2011.5, -4.7, 60)
642 text(2011.5, -5.5, 50)
643
644
645 # modelo P SPLINE
646
647 png("PSPLINE.png", width = 2400, height = 1600, res = 350)

```

```

648
649 dibuja_pspline(x = x, y = log(y_death/e_exp), x.pspline = todos_los_
    anos, y.pspline = fit1Dfor$fit, y.pspline.3d = fit1Dfor3d, y.
    pspline.1d = fit1Dfor1d, y.pspline.for = fit1Dfor$se.fit, y.
    pspline.new = y.pspline)
650
651 dev.off()
652
653
654 # Los 3 modelos juntos
655
656 png("LC BCD PSPLINES comparando 3 modelos.png", width = 3600, height
    = 2400, res = 350)
657
658 rango=c(miEdad)
659
660 compara_modelos(x=CBDfit$years, xs=CBDsim$years, y=log(Rate), ys=log
    (CBDsim$rates), ymin= log(Rate[rango[1],]), ymax= log(Rate[rango
    [length(rango)],]), ymin=log(CBDsim$rates[rango[1], , 1:200]),
    ymax=log(CBDsim$rates[rango[length(rango)], , 1:200]), ylabel="
    Logaritmo de la tasa de fallecimientos", rango=rango, titulo=
    paste("Europa", "P splines vs LC vs CBD: log(tasa de mortalidad)
    a los", paste(unlist(rango), collapse=', '), "anos"), cols = hcl
    .colors(10, "Terrain"), yHs=log(LCsim$rates), yHsmin=log(LCsim$
    rates[rango[1], , 1:200]), yHsmax=log(LCsim$rates[rango[length(
    rango)], , 1:200]), cols2 = hcl.colors(10, "Pastel 1"))
661
662 # Y anadimos p-splines:
663 compara_pspline(x = x, y = log(y_death/e_exp) , x.pspline = todos_
    los_anos, y.pspline = fit1Dfor$fit, y.pspline.3d = fit1Dfor3d,
    y.pspline.1d = fit1Dfor1d, y.pspline.for = fit1Dfor$se.fit, y.
    pspline.new = y.pspline)
664
665 # La leyenda
666 legend("bottomleft", c(paste("Datos reales para el calculo hasta",
    hasta), "P SPLINES d=2", "P SPLINES d=3", "P SPLINES d=1", paste
    ("Datos reales para comparar hasta", anyos_datos_reales[[length(
    anyos_datos_reales)]]), "Lee Carter", "Cairns, Blake and Dowd" ),
    lty = c(0, 1, 1, 1, 0,0,0), col = 1:5, pch = c(1, -1, -1, -1,
    1,0,0), lwd = c(1, 2, 2, 2, 1,0,0), bty = "n", cex = 0.7)
667
668 dev.off()
669
670 # Si queremos estudiar los datos para cada edad, basta con hacer un
    bucle e introducir las funciones anteriores
671
672
673
674 # ELI COMPAREMOS LOS RESULTADOS, porque es el valor del trabajo
675 # Hasta 2017 tenemos datos ciertos.
676 # desde 2012 tenemos datos estimados

```

```

677 # vamos a comparar de 2012 a 2017 la bondad de cada resultado
678
679 for (miSexo in c("female","male")){
680
681   dibuja(midata=miRESULTADO[["RATEtradicional"]][[miSexo]], mititulo
     =paste("Tasa de mortalidad para", ifelse(miSexo == "female", "
     mujeres", "hombres"),"hasta 2025 metodo tradicional"), miColor
     ="lightblue", theta=185, phi=20, zlim=range(0,0.55), destino=
     paste("PROYECCION a 2025 Tradicional",miSexo,".png"))
682
683   dibuja(midata=cbind(miRESULTADO[["datos_de_partida"]][[miSexo]],
     miRESULTADO[["LCqxmedia"]][[miSexo]]), mititulo = paste("Tasa
     de mortalidad para", ifelse(miSexo == "female", "mujeres", "
     hombres"),"hasta 2025 por LC"), miColor = "lightblue", theta =
     185, phi = 20, zlim = range(0,0.55), destino = paste("
     PROYECCION a 2025 LC",miSexo,".png"))
684
685   dibuja(midata=cbind(miRESULTADO[["datos_de_partida"]][[miSexo]],
     miRESULTADO[["CBDqxmedia"]][[miSexo]]), mititulo = paste("Tasa
     de mortalidad para", ifelse(miSexo == "female", "mujeres", "
     hombres"),"hasta 2025 por CBD"), miColor = "lightblue", theta
     = 185, phi = 20, zlim = range(0,0.55), destino = paste("
     PROYECCION a 2025 CBD",miSexo,".png"))
686
687   dibuja(midata=miRESULTADO[["PSPLINES2Dpredicted"]][["male"]],
     mititulo = paste("Tasa de mortalidad para", ifelse(miSexo == "
     female", "mujeres", "hombres"),"hasta 2025 por P-SPLINES"),
     miColor = "lightblue", theta = 185, phi = 20, zlim = range
     (0,0.55), destino = paste("PROYECCION a 2025 PSPLINES",miSexo,
     ".png"))
688 }
689
690 # Diferencia respecto a los datos reales
691 # Estructura
692 miERRORES <- NULL
693 miERRORES <- list(RealRate = list(female=matrix(c(1,11), nrow=1,
     ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)), TradRate =
     list(female=matrix(c(1,11), nrow=1, ncol=2),male=matrix(c
     (5,6,7,8), nrow=1, ncol=4)), LCRate = list(female=matrix(c(1,11)
     , nrow=1, ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)),
     CBDRate = list(female=matrix(c(1,11), nrow=1, ncol=2),male=
     matrix(c(5,6,7,8), nrow=1, ncol=4)), PSPLINESRate = list(female=
     matrix(c(1,11), nrow=1, ncol=2),male=matrix(c(5,6,7,8), nrow=1,
     ncol=4)), Traderror = list(female=matrix(c(1,11), nrow=1, ncol
     =2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)), LCerror = list(
     female=matrix(c(1,11), nrow=1, ncol=2),male=matrix(c(5,6,7,8),
     nrow=1, ncol=4)), CBDerror = list(female=matrix(c(4,44), nrow=1,
     ncol=2),male=matrix(c(0,1,2,1), nrow=1, ncol=4)), PSPLINESerror
     = list(female=matrix(c(2,22), nrow=1, ncol=2),male=matrix(c
     (0,1,2,1), nrow=1, ncol=4)), Traderror_relativo = list(female=
     matrix(c(1,11,111,111), nrow=1, ncol=4),male=matrix(c(5,6,7,8),

```

```

nrow=1, ncol=4)), LCerror_relativo = list(female=matrix(c
(1,11,111,111), nrow=1, ncol=4),male=matrix(c(5,6,7,8), nrow=1,
ncol=4)), CBDerror_relativo = list(female=matrix(c(4,44,444,444)
, nrow=1, ncol=4),male=matrix(c(0,1,2,1), nrow=1, ncol=4)),
PSPLINESerror_relativo = list(female=matrix(c(2,22,222,222),
nrow=1, ncol=4),male=matrix(c(0,1,2,1), nrow=1, ncol=4)))

694
695 # Calculo de errores
696 for (miSexo in c("female","male")){
697
698   miERRORES[["RealRate"]][[miSexo]] <- miHMDdata[[miPais]][["rate"
    ]][[miSexo]][41:101,23:28]
699   miERRORES[["TradRate"]][[miSexo]] <- miRESULTADO[["RATEtradicional
    "]][[miSexo]][,23:28]
700   miERRORES[["LCRate"]][[miSexo]] <- miRESULTADO[["LCqxmedia"]][[
    miSexo]][,1:6]
701   miERRORES[["CBDRate"]][[miSexo]] <- miRESULTADO[["CBDqxmedia"]][[
    miSexo]][,1:6]
702   miERRORES[["PSPLINESRate"]][[miSexo]] <- miRESULTADO[["
    PSPLINES2Dpredicted"]][[miSexo]][,23:28]
703
704   #Absolutos
705   miERRORES[["Traderror"]][[miSexo]] <- (miERRORES[["TradRate"]][[
    miSexo]] - miERRORES[["RealRate"]][[miSexo]])
706   miERRORES[["LCerror"]][[miSexo]] <- (miERRORES[["LCRate"]][[miSexo
    ]]] - miERRORES[["RealRate"]][[miSexo]])
707   miERRORES[["CBDerror"]][[miSexo]] <- (miERRORES[["CBDRate"]][[
    miSexo]] - miERRORES[["RealRate"]][[miSexo]])
708   miERRORES[["PSPLINESerror"]][[miSexo]]<- (miERRORES[["PSPLINESRate
    "]][[miSexo]] - miERRORES[["RealRate"]][[miSexo]])
709
710   #Relativos
711   miERRORES[["Traderror_relativo"]][[miSexo]] <- ( miERRORES[["
    Traderror"]][[miSexo]] / miERRORES[["RealRate"]][[miSexo]])
712   miERRORES[["LCerror_relativo"]][[miSexo]] <- ( miERRORES[["LCerror
    "]][[miSexo]] / miERRORES[["RealRate"]][[miSexo]])
713   miERRORES[["CBDerror_relativo"]][[miSexo]] <- ( miERRORES[["
    CBDerror"]][[miSexo]] / miERRORES[["RealRate"]][[miSexo]])
714   miERRORES[["PSPLINESerror_relativo"]][[miSexo]]<- ( miERRORES[["
    PSPLINESerror"]][[miSexo]] / miERRORES[["RealRate"]][[miSexo
    ]]])
715
716 }
717
718 # Los graficos de los errores se realizan de manera analoga a la
    comprobacion de la bondad de los modelos, con un bucle por sexo
719
720
721 # Comparacion entre modelos
722
723 miCOMPARACION <- NULL

```



```

724 miCOMPARACION <- list(TradRate = list(female=matrix(c(1,11), nrow=1,
      ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)), LCRate = list
      (female=matrix(c(1,11), nrow=1, ncol=2),male=matrix(c(5,6,7,8),
      nrow=1, ncol=4)), CBDRate = list(female=matrix(c(1,11), nrow=1,
      ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4)), PSPLINESRate =
      list(female=matrix(c(1,11), nrow=1, ncol=2),male=matrix(c
      (5,6,7,8), nrow=1, ncol=4)), LCvsTrad = list(female=matrix(c
      (1,11), nrow=1, ncol=2),male=matrix(c(5,6,7,8), nrow=1, ncol=4))
      , CBDvsTrad = list(female=matrix(c(1,11), nrow=1, ncol=2),male=
      matrix(c(5,6,7,8), nrow=1, ncol=4)), PSPLINEvsTrad = list(female
      =matrix(c(4,44), nrow=1, ncol=2),male=matrix(c(0,1,2,1), nrow=1,
      ncol=4)), CBDvsLC = list(female=matrix(c(2,22), nrow=1, ncol=2)
      ,male=matrix(c(0,1,2,1), nrow=1, ncol=4)), PSPLINESvsLC = list(
      female=matrix(c(2,22), nrow=1, ncol=2),male=matrix(c(0,1,2,1),
      nrow=1, ncol=4)), PSPLINESvsCBD = list(female=matrix(c(2,22),
      nrow=1, ncol=2),male=matrix(c(0,1,2,1), nrow=1, ncol=4)),
      LCvsTrad_relativo = list(female=matrix(c(1,11), nrow=1, ncol=2),
      male=matrix(c(5,6,7,8), nrow=1, ncol=4)), CBDvsTrad_relativo =
      list(female=matrix(c(1,11), nrow=1, ncol=2),male=matrix(c
      (5,6,7,8), nrow=1, ncol=4)), PSPLINEvsTrad_relativo = list(
      female=matrix(c(4,44), nrow=1, ncol=2),male=matrix(c(0,1,2,1),
      nrow=1, ncol=4)), CBDvsLC_relativo = list(female=matrix(c(2,22),
      nrow=1, ncol=2),male=matrix(c(0,1,2,1), nrow=1, ncol=4)),
      PSPLINESvsLC_relativo = list(female=matrix(c(2,22), nrow=1, ncol
      =2),male=matrix(c(0,1,2,1), nrow=1, ncol=4)), PSPLINESvsCBD_
      relativo = list(female=matrix(c(2,22), nrow=1, ncol=2),male=
      matrix(c(0,1,2,1), nrow=1, ncol=4)))

725
726 # Bucle de calculo
727 for (miSexo in c("female","male")){
728
729   miCOMPARACION[["TradRate"]][[miSexo]] <- miRESULTADO[["
      RATEtradicional"]][[miSexo]][,23:36]
730   miCOMPARACION[["LCRate"]][[miSexo]] <- miRESULTADO[["LCqxmedia"
      ]][[miSexo]]
731   miCOMPARACION[["CBDRate"]][[miSexo]] <- miRESULTADO[["CBDqxmedia"
      ]][[miSexo]]
732   miCOMPARACION[["PSPLINESRate"]][[miSexo]] <- miRESULTADO[["
      PSPLINES2Dpredicted"]][[miSexo]][,23:36]
733
734 # Absolutos
735   miCOMPARACION[["LCvsTrad"]][[miSexo]] <-(miCOMPARACION[["LCRate"
      ]][[miSexo]] - miCOMPARACION[["TradRate"]][[miSexo]])
736   miCOMPARACION[["CBDvsTrad"]][[miSexo]] <-(miCOMPARACION[["CBDRate"
      ]][[miSexo]] - miCOMPARACION[["TradRate"]][[miSexo]])
737   miCOMPARACION[["PSPLINESvsTrad"]][[miSexo]] <-(miCOMPARACION[["
      PSPLINESRate"]][[miSexo]] - miCOMPARACION[["TradRate"]][[
      miSexo]])
738
739   miCOMPARACION[["CBDvsLC"]][[miSexo]] <-(miCOMPARACION[["CBDRate"
      ]][[miSexo]] - miCOMPARACION[["LCRate"]][[miSexo]])

```



```

740 miCOMPARACION[["PSPLINESvsLC"]][[miSexo]] <-(miCOMPARACION[["
    PSPLINESRate"]][[miSexo]] - miCOMPARACION[["LCRate"]][[miSexo
    ]])
741 miCOMPARACION[["PSPLINESvsCBD"]][[miSexo]] <-(miCOMPARACION[["
    PSPLINESRate"]][[miSexo]] - miCOMPARACION[["CBDRate"]][[miSexo
    ]])
742
743 # Relativos
744 miCOMPARACION[["LCvsTrad_relativo"]][[miSexo]] <-(miCOMPARACION[["
    LCvsTrad"]][[miSexo]] / miCOMPARACION[["TradRate"]][[miSexo]])
745 miCOMPARACION[["CBDvsTrad_relativo"]][[miSexo]] <-(miCOMPARACION[["
    CBDvsTrad"]][[miSexo]] / miCOMPARACION[["TradRate"]][[miSexo
    ]])
746 miCOMPARACION[["PSPLINESvsTrad_relativo"]][[miSexo]] <-(
    miCOMPARACION[["PSPLINESvsTrad"]][[miSexo]] / miCOMPARACION[["
    TradRate"]][[miSexo]])
747
748 miCOMPARACION[["CBDvsLC_relativo"]][[miSexo]] <-(miCOMPARACION[["
    CBDvsLC"]][[miSexo]] / miCOMPARACION[["LCRate"]][[miSexo]])
749 miCOMPARACION[["PSPLINESvsLC_relativo"]][[miSexo]] <-(
    miCOMPARACION[["PSPLINESvsLC"]][[miSexo]] / miCOMPARACION[["
    LCRate"]][[miSexo]])
750 miCOMPARACION[["PSPLINESvsCBD_relativo"]][[miSexo]] <-(
    miCOMPARACION[["PSPLINESvsCBD"]][[miSexo]] / miCOMPARACION[["
    CBDRate"]][[miSexo]])
751
752 }
753
754 # Para representarlo se haria un bucle analogo al caso anterior
755
756 # Ultima comparacion: el modelo ELI vs el 20%
757 for (miSexo in c("male","female")) {
758
759     dibuja(midata= cbind(miRESULTADO[["datos_de_partida"]][[miSexo
        ]],miRESULTADO[["PromedioELI"]][[miSexo]]), mititulo=paste("
        PromedioELI de modelos", ifelse(miSexo == "female", "mujeres
        ", "hombres")), miColor = "lightblue", theta = 180, phi =
        20, zlim = range(0,0.55), destino = paste("Rate ELI",miSexo,
        ".png"))
760
761 # Comparando en 2D para que se aprecie
762 png(paste("DOM macro ELI vs Trad",ifelse(miSexo=="female", "
    mujeres", "hombres") ,".png"), width = 3000, height = 3000,
    res = 350)
763
764 estos_colores=c("blue","green","yellow","orange","red")
765 anyos <- as.integer(seq(length=4, from=2010, to=2025))
766
767 plot(x=-edades_a_considerar, y=miRESULTADO[["RATEtradicional"
    ]][[miSexo]][,"2020"], main="Comparacion del modelo
    tradicional con ELI", xlim=c(-100,-40), ylim=c(0,0.3), ylab=

```

```

    "tasa de fallecimientos")
768 grid(col = 'grey80', lty = 1, lwd = 0.5)
769
770 for ( anyo in 1:length(anyos) )
771   lines(x = -edades_a_considerar, y = miRESULTADO[["PromedioELI"
    ]][[miSexo]][,anyo], lwd = 1.5+anyo/2, lty = 2+anyo/2, col
    = estos_colores[anyo])
772
773 lines(x = -edades_a_considerar, y = miRESULTADO[["
    RATEtradicional"]][[miSexo]][,"2020"] , lwd = 3, lty = 1,
    col = "grey30")
774
775 legend("topright", legend=c(as.character(anyos),"Tradicional"),
    lty = 6, col = estos_colores, pch = c(1, -1, -1, -1, 1), lwd
    = c(3, 2, 2.5, 3, 3.5), bty = "n", cex = 0.7)
776
777 dev.off()
778
779 # A mayor detalle
780 png(paste("DOM micro ELI vs Trad",ifelse(miSexo=="female", "
    mujeres", "hombres") ,".png"), width = 3000, height = 3000,
    res = 350)
781
782 estos_colores=c("blue","green","yellow","orange","red")
783 anyos <- as.integer(seq(length=4, from=2010, to=2025))
784 plot(x = -edades_a_considerar, y=miRESULTADO[["RATEtradicional"
    ]][[miSexo]][,"2020"],
785      main="Comparacion del modelo tradicional con ELI",
786      xlim=c(-55,-40),                      ylim=c(0.001,0.004)
    ,
787      ylab="tasa de fallecimientos")
788 grid(col = 'grey80', lty = 1, lwd = 0.5)
789
790 for ( anyo in 1:length(anyos) )
791   lines(x = -edades_a_considerar,y = miRESULTADO[["PromedioELI"
    ]][[miSexo]][,anyo], lwd = 1.5+anyo/2, lty = 2+anyo/2, col
    = estos_colores[anyo])
792
793 lines(x = -edades_a_considerar,y = miRESULTADO[["RATEtradicional
    "]][[miSexo]][,"2020"] , lwd = 3, lty = 1, col = "grey30")
794
795 legend("topright", legend = c(as.character(anyos),"Tradicional")
    , lty = 6, col = estos_colores, pch = c(1, -1, -1, -1, 1),
    lwd = c(3, 2, 2.5, 3, 3.5), bty = "n", cex = 0.7)
796
797 dev.off()
798 }

```