

Máster Universitario Ciencias Actuariales y Financieras
2022-2024

Trabajo Fin de Máster

“Riesgo climático: Impacto de las altas temperaturas en la mortalidad en función del género y la edad mediante la modelización de series temporales”

Silvia Pedrero Rico

Tutora

Raquel Pérez Calderón

Madrid, mayo de 2024.

DETECCIÓN DEL PLAGIO

La Universidad utiliza el programa **Turnitin Feedback Studio** para comparar la originalidad del trabajo entregado por cada estudiante con millones de recursos electrónicos y detecta aquellas partes del texto copiadas y pegadas. Copiar o plagiar en un TFM es considerado una **Falta Grave**, y puede conllevar la expulsión definitiva de la Universidad.



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

RESUMEN

El impacto de las altas temperaturas en la mortalidad española toma cada vez más importancia en un contexto de cambio climático. Este trabajo busca predecir la mortalidad de las personas en edades afectadas por el verano del 2022 para materializar el riesgo, desde un punto de vista actuarial, a causa de un aumento por encima de lo esperado en la mortalidad.

En primer lugar, se lleva a cabo un análisis mediante modelos GLM para identificar para qué edades, diferenciando por el género, se experimenta un impacto en la mortalidad a causa de altas temperaturas. A continuación, mediante la modelización de dos series temporales (mortalidad y temperatura) usando metodología VAR, la cual permite relacionar estas dos series, se estudia la causalidad y función impulso-respuesta entre la mortalidad diaria para un rango de edad vulnerable y la temperatura máxima diaria para ese mismo periodo de tiempo. Finalmente, se predice la mortalidad futura para estas edades teniendo en cuenta los valores pasados de la temperatura y de la mortalidad.

Los resultados muestran como las altas temperaturas causan a la mortalidad tomando los datos del 2022 y cómo usando los valores pasados diarios de ese año de la variable climática, se estima un aumento en la mortalidad diaria futura. Además se demuestra como existe un mayor efecto en la mortalidad de las mujeres, sobre todo para edades avanzadas, más que para los hombres.

Palabras Clave: Mortalidad, altas temperaturas, modelo VAR, función impulso-respuesta.

*A mis padres,
no hay vidas para tanto agradecimiento,
todo lo que soy es por vosotros.*

ÍNDICE GENERAL

1. INTRODUCCIÓN.	1
2. REVISIÓN DE LA LITERATURA.	4
3. DESCRIPCIÓN DE LA MUESTRA E HIPÓTESIS DE PARTIDA.	7
3.1. DESCRIPCIÓN DE LA MUESTRA.	7
3.2. HIPÓTESIS DE PARTIDA	10
4. METODOLOGÍA.	19
4.1 PRIMERA PARTE METODOLÓGICA.	19
4.2 SEGUNDA PARTE METODOLÓGICA.	25
4.2.1 Estacionariedad.	27
4.2.2 Modelo VAR.	28
4.2.3 Causalidad de Granger.	29
4.2.4 Función Impulso-Respuesta.	30
4.2.5. Descomposición de la varianza.	30
4.2.6. Residuos del modelo.	31
4.2.7. Predicción de la mortalidad.	33
5. RESULTADOS.	34
5.1 RESULTADOS PARA LA PRIMERA PARTE METODOLÓGICA.	34
5.2 RESULTADOS PARA LA SEGUNDA PARTE METODOLÓGICA.	48
5.2.1 Estacionariedad.	49
5.2.2 Modelo VAR.	52
5.2.3 Causalidad de Granger	56
5.2.4 Función Impulso-Respuesta.	59
5.2.5 Descomposición de la varianza.	61
5.2.6 Residuos del modelo.	64
5.2.7 Predicción de mortalidad.	67
5.3. RESULTADOS PARA MODELOS VAR CON FRECUENCIA ANUAL.	75
6. CONCLUSIÓN.	77
BIBLIOGRAFÍA.	80
ANEXO.	85
A1. Gráficos para el año 2021 en los meses fríos.	85
A2. Código de R.	85

ÍNDICE DE FIGURAS

Figura 3.1 Exceso de mortalidad por mes en España	10
Figura 3.2 Relación entre mortalidad y temperatura en España 2021.....	12
Figura 3.3 Relación entre mortalidad y temperatura en España 2022	12
Figura 3.4 Relación entre mortalidad y temperatura en España 2022 desde octubre a diciembre.	15
Figura 3.5 Relación entre mortalidad y temperatura en España 2022 desde enero a marzo.....	15
Figura 3.6 Gráfico de dispersión entre mortalidad y temperatura para el año 2022	16
Figura 3.7 Gráfico de dispersión entre mortalidad y temperatura para el año 2022	17
Figura 3.8 Gráfico de dispersión entre mortalidad y temperatura para los meses de abril a septiembre.....	17
Figura 5.1 Histograma de la mortalidad para conocer su distribución.....	34
Figura 5.2 Resultados del modelo GLM para la edad 25.....	35
Figura 5.3 Resultados del modelo GLM para la edad 40.....	36
Figura 5.4 Resultados del modelo GLM para la edad 74.....	36
Figura 5.5 Resultados del modelo GLM para la edad 75.....	37
Figura 5.6 Resultados del modelo GLM para la edad 79.....	38
Figura 5.7 Resultados del modelo GLM para la edad 80.....	38
Figura 5.8 Resultados del modelo GLM para la edad 83.....	38
Figura 5.9 Resultados del modelo GLM para la edad 90.....	39
Figura 5.10 Resultados del modelo GLM para la edad 95.....	39
Figura 5.11 Resultados del modelo GLM para la edad 101.....	39
Figura 5.12 Resultados del modelo GLM para la edad 23.....	40
Figura 5.13 Resultados del modelo GLM para la edad 40.....	41
Figura 5.14 Resultados del modelo GLM para la edad 39.....	41
Figura 5.15 Resultados del modelo GLM para la edad 48.....	42
Figura 5.16 Resultados del modelo GLM para la edad 58.....	42
Figura 5.17 Resultados del modelo GLM para la edad 61.....	43
Figura 5.18 Resultados del modelo GLM para la edad 72.....	43
Figura 5.19 Resultados del modelo GLM para la edad 74.....	44
Figura 5.20 Resultados del modelo GLM para la edad 79.....	45
Figura 5.21 Resultados del modelo GLM para la edad 80.....	45
Figura 5.22 Resultados del modelo GLM para la edad 83.....	46
Figura 5.23 Resultados del modelo GLM para la edad 89.....	46
Figura 5.24 Resultados del modelo GLM para la edad 101.....	47
Figura 5.25 Resultados del test Dickey & Fuller para cada rango de mortalidad	49
Figura 5.26 Resultados del test Phillips & Perron para cada rango de mortalidad	49
Figura 5.27 Resultados del test KPSS para cada rango de mortalidad	50
Figura 5.28 Resultado de las diferencias óptimas para la mortalidad M85	51
Figura 5.29 Resultados de los test para la mortalidad M85 con una diferencia.....	51
Figura 5.30 Resultados de los test para la serie de temperatura máxima.....	52
Figura 5.31 Resultados criterios información M85	53
Figura 5.32 Resultados criterio información M7584.....	53

Figura 5.33 Resultados criterios información H4564	53
Figura 5.34 Resultados criterios información H6574	53
Figura 5.35 Resultados del modelo VAR H4564 con dos retardos	54
Figura 5.36 Resultados del modelo VAR H4564 con un retardo	55
Figura 5.37 Resultados del modelo VAR M85 con un retardo.....	55
Figura 5.38 Resultados del modelo VAR M7584 con un retardo.....	56
Figura 5.39 Resultados del modelo VAR H6574 con un retardo	56
Figura 5.40 Resultados de causalidad de Granger para M85.....	57
Figura 5.41 Resultados de causalidad de Granger para M7584.....	57
Figura 5.42 Resultados de causalidad de Granger para H4564	58
Figura 5.43 Resultados de causalidad de Granger para H6574	58
Figura 5.44 Resultados de función impulso respuesta M85	59
Figura 5.45 Resultados de función impulso respuesta M7584	59
Figura 5.46 Resultados de función impulso respuesta H4564	60
Figura 5.47 Resultados de función impulso respuesta H6574	60
Figura 5.48 Resultados de la descomposición de la varianza ante un shock en la temperatura para M85.....	62
Figura 5.49 Resultados de la descomposición de la varianza ante un shock en la temperatura para M7584.....	62
Figura 5.50 Resultados de la descomposición de la varianza ante un shock en la temperatura para H4564	63
Figura 5.51 Resultados de la descomposición de la varianza ante un shock en la temperatura para H6574	63
Figura 5.52 Resultados de test para comprobar la normalidad en los residuos.....	64
Figura 5.53 Resultados de test para comprobar la independencia en los residuos.....	65
Figura 5.54 Resultados de test para comprobar la homocedasticidad en los residuos tomando los residuos de la mortalidad.	66
Figura 5.55 Resultados de test para comprobar la homocedasticidad en los residuos tomando los residuos de la temperatura máxima.	66
Figura 5.56 Resultados de test para conocer la media de los errores.....	67
Figura 5.57 Número de defunciones diarias las mujeres a partir de 85 años en 2022	68
Figura 5.58 Variación de defunciones diarias las mujeres a partir de 85 años en 2022.....	68
Figura 5.59 Valor esperado de las defunciones diarias para mujeres a partir de 85 años	69
Figura 5.60 Número de defunciones diarias las mujeres de 75 a 84 años en 2022.....	70
Figura 5.61 Valor esperado de las defunciones diarias para mujeres de 75 a 84 años.....	71
Figura 5.62 Número de defunciones diarias los hombres de 45 a 64 años en 2022.....	72
Figura 5.63 Valor esperado de las defunciones diarias para hombres de 45 a 64 años.....	72
Figura 5.64 Número de defunciones diarias los hombres de 65 a 74 años en 2022.....	73
Figura 5.65 Valor esperado de las defunciones diarias los hombres de 65 a 74 años.....	73
Figura 5.66 Causalidad de Granger entre mortalidad y altas temperaturas para mujeres de 85 años.....	76
Figura 5.67 Modelo VAR (1) para mujeres de 85 años.	76

1. INTRODUCCIÓN.

El cambio climático y sus efectos directos en la salud humana son cada vez más estudiados debido a las grandes repercusiones que pueden ocurrir en pocos años si no se aplican soluciones eficientes. Europa se posiciona como un área climática clave para aplicar políticas de prevención y mitigación, ya que esta también sufrió y sufre en la actualidad de manera directa los efectos climatológicos.

En el verano del 2022 se experimentaron temperaturas extremadamente altas en la mayor parte de Europa, especialmente al sur del continente en los países colindantes con el mar. En España el número de fallecidos en el mes de julio superó al mes de diciembre, donde la mortalidad española tendía siempre a tener uno de los picos más altos.

Estos datos reflejan una correlación positiva entre defunciones y altas temperaturas y cómo comprender estas puede ser muy útil para entender los posibles valores futuros que puede tomar la mortalidad, siendo esta la motivación del estudio. Estimar la mortalidad en función de variables significativas como las climáticas permite, desde un punto de vista actuarial, gestionar mejor el riesgo, anticiparse a situaciones de estrés y asegurar así que el capital es suficiente para cubrir siniestros inesperados en seguros vida riesgo o que las primas son suficientes para cubrir la demanda en seguros de salud.

El objetivo de este trabajo es conocer, diferenciando entre hombres y mujeres, para qué edades las olas de calor tienen un impacto significativo en la mortalidad. Determinadas las edades vulnerables, será para estas para las que se deba predecir su mortalidad futura ya que cuentan con mayor incertidumbre al estar influidas por la temperatura máxima, la cual antes no tenía tanto impacto y que en el futuro se espera que continúen creciendo por informes como el de la Organización Mundial de la Salud (OMS) publicado en el 2023.

Para realizar el estudio, en primer lugar, se van a utilizar modelos lineales generalizados (Generalized Lineal Models, GLM) para conocer las edades afectadas por las temperaturas. El uso de estos modelos surge a raíz de la falta de linealidad entre las variables mortalidad y temperatura y la estimación de la distribución de mortalidad mediante Kernel, donde los resultados se ajustan a una Poisson.

Con los datos de la mortalidad mensual del 2022 de las mujeres para los meses de abril a septiembre de cada edad desde los 0 años hasta los 101 años, se realizan 101 modelos (uno para cada edad) siendo la variable dependiente la mortalidad (para esa edad) y la variable explicativa la temperatura. En cada modelo se analiza si la variable independiente tiene impacto significativo en la mortalidad. Este proceso se repite con la mortalidad mensual de los hombres para cada edad desde los 0 años hasta los 101 años.

En cuanto a esta primera parte metodológica, los resultados indican como hay diferencia en las edades afectadas según el género. Para las mujeres, a partir de los 80 años en adelante, el calor tiene un impacto muy significativo en la mortalidad. Para los hombres el tramo más afectado por esta variable climática se encuentra entre los 50 a 75 años.

Además, es notable que las mujeres tienen más edades vulnerables que los hombres, siendo en estos menor el impacto y menos significativo.

Estas conclusiones son acordes a las estudiadas en otros trabajos, como en Ballester et al. (2023) donde en general para todos los países europeos existe la diferencia de género y edad cuando se estudia el impacto de las olas de calor.

Las edades afectadas por altas temperaturas serán de interés para el estudio de predicción, y por ello, en segundo lugar, se va a llevar a cabo la modelización entre dos series temporales. Estas van a ser la temperatura máxima y la mortalidad diaria, también en el año 2022, para esas edades vulnerables afectadas por las temperaturas.

Los modelos VAR, siendo una extensión de los ARIMA, permiten relacionar series temporales, pudiendo así crear la función respuesta al impulso y poder predecir una de las series (mortalidad) con la información de la otra (temperatura máxima).

Entonces, una vez se conocen a qué edades el calor del 2022 afecta a la mortalidad se separan por rangos de edad y se obtiene su mortalidad diaria de mayo a septiembre de la base de datos del Sistema de Monitorización de la Mortalidad Diaria, comenzando así el segundo estudio. Para las mujeres se toma la mortalidad diaria de las edades de 75 a 84 años y de 85 en adelante. Para los hombres se toma la mortalidad diaria de las edades de 45 a 64 años y de 65 a 74 años. Estas son las edades en las que en la primera parte metodológica ha salido la variable temperatura como explicativa de la mortalidad.

Con los datos diarios de mortalidad de los cuatro grupos de edad (por tanto, cuatro series temporales) y los datos diarios de temperatura máxima para el mismo periodo, mediante la metodología VAR, se estudia la relación entre la serie temporal de mortalidad y la serie temporal de temperatura, teniendo así cuatro modelos VAR, uno por cada grupo de mortalidad diaria. En los cuatro se estudia por separado la Causalidad de Granger y en todos se concluye que las altas temperaturas causan efecto en la mortalidad y que, por lo tanto, la variable climática y sus valores pasados deben ser utilizados para predecir el número de fallecidos. Las funciones impulso respuesta de cada modelo también concluyen lo mismo. Además, un shock inesperado en la temperatura máxima diaria con datos del 2022 provoca un aumento en la mortalidad que durante varios periodos.

Ambos resultados afirman que si se quiere predecir la mortalidad para esas edades es importante considerar los datos observados (presentes y pasados) de la temperatura máxima, y es así como se realiza la estimación de los fallecidos futuros en este estudio.

En comparación con el estudio de mortalidad y temperatura diaria para el año 2022, se realiza la misma metodología pero datos anuales, desde el 1970 al 2022. Se toma la probabilidad de fallecimiento para una edad vulnerable por cada año y la temperatura máxima anual. Los datos reflejan, como si se toman series históricas anuales, el impacto de las altas temperaturas no es significativo en la mortalidad.

Este trabajo confirma que, dada la situación de calor extremo a consecuencia del cambio climático en el año 2022 en España esto está afectando a ciertas edades, según su género. El efecto es que la mortalidad puede aumentar más de lo esperado en dichas edades y por eso es relevante la predicción de mortalidad para las mismas. Conocer los posibles valores futuros de mortalidad para edades afectadas por altas temperaturas es también importante a nivel actuarial, ya que se puede predecir el riesgo y por tanto hacer una gestión de este más eficiente.

Además queda comprobado que si en vez de coger los datos diarios del 2022 se toman datos anuales con un histórico desde 1970, aunque haya estudios anteriores al siglo XXI que midan el impacto de las olas de calor en el número de fallecidos, se demuestra como los resultados son algo particular del año 2022. Aun así, en un contexto de cambio climático, puede que sea a partir de este año donde el efecto de la temperatura sea significativo en los años siguientes.

2. REVISIÓN DE LA LITERATURA.

El impacto de la temperatura en la mortalidad española es algo muy estudiado desde científicos a economistas y el cómo paliar los efectos ha sido tema de debate en muchos artículos, en los que se han planteado soluciones y las posibles causas en caso de que estas no se apliquen de una manera óptima.

En estudios previos al siglo XXI, Rubio et al. (1997) comentaban que conocer los patrones de temperatura era muy importante en términos de mortalidad para aquellas edades y género afectados por esta variable.

Hay que mencionar que las temperaturas extremas bajas han sido las culpables de que el número de defunciones sea mayor en los meses fríos que en los meses cálidos. Si se toman datos de fallecidos mensuales para distintos años, se observa como en los meses de diciembre y enero el pico es mucho más alto que para el resto de los meses, a excepción del año 2022, donde el mayor número de fallecidos es en julio y agosto (además de enero). Así que los expertos antes del 2022 defendían que una política eficaz sería la enfocada en bajas temperaturas (Carmona et al., 2016) y prevenir los efectos de esta en la salud.

Se podría pensar que si el cambio climático, causado por un nivel muy alto de gases de efecto invernadero, está provocando que las temperaturas suban (Achebak et al., 2020), los planes contra las bajas temperaturas no serían del todo necesarios, pero la realidad es distinta, ya que no hay evidencias suficientes en España de que esto esté sucediendo y que por tanto el número de fallecidos en meses de invierno haya disminuido.

Por tanto, sin olvidar que los meses fríos son causantes de un aumento en la mortalidad, este trabajo se centra en el hecho, en principio anómalo, del año 2022, donde el exceso de mortalidad por temperatura en los meses de verano es comparable con los meses de invierno y donde se demuestra que las olas de calor han causado un impacto en la salud.

Aunque las temperaturas extremas altas sean un tema de actualidad también hay estudios antiguos. En el año 2003, los meses de verano estuvieron azotados por olas de calor persistentes y con temperaturas muy elevadas en toda Europa, lo cual afectó directamente al número de personas fallecidas.

Para el caso de España, empezaron a estudiarse las distintas características que puede tener una persona para ser vulnerable a altas temperaturas en función de su género, edad, lugar de residencia... Este último pudo ser muy determinante para ese año ya que la mayoría de la sociedad empezó a vivir en las ciudades y estas conservan mucho más el calor (Martínez Navarro et al., 2004) potenciando así el gran número de fallecidos. Esta ola de calor en un periodo con temperaturas altas, y sus efectos en la mortalidad, impulsaron planes de prevención y de anticipación partiendo de la hipótesis de que las posibles temperaturas extremas futuras serían similares (Díaz et al., 2018) y, por ende, sus soluciones serían homogéneas. Años más tarde se demostraría que esa hipótesis de partida no era del todo acertada ya que, por ejemplo, la temperatura la cual indica a partir de qué grado (°C) tiene un impacto significativo en la mortalidad, varía con el tiempo (Pliego et al., 2020).

Tras este acontecimiento, y la posibilidad de haberlo anticipado, diecinueve años más tarde sucedió un hecho que ha sido el único comparable con los efectos del calor del 2003 (Ballester et al., 2023) y cuyo resultado ha sido aún más grave, la ola extrema de calor que afectó a casi toda Europa, destacando países como España, en el verano del 2022. Esta afectó a personas vulnerables en función del género, edad, enfermedad previa... Este artículo empieza a marcar unas pautas claras de como en función de ser hombre o mujer, la edad a la que les afecta el calor extremo varía.

El “*Sistema de Monitorización de la Mortalidad Diaria*” fue capaz de demostrar como en el verano del 2022 la mortalidad de toda España aumentó por encima de lo esperado de manera homogénea en la mayor parte de la península, algo que años anteriores no se había podido verificar por falta de impacto relevante (León-Gómez et al., 2023).

Las mujeres son las gran afectadas por las olas de calor debido a sus características físicas como señaló la “*Agencia de Seguridad de la Salud del Reino Unido*” (UKHSA).

El estudio de por qué a las mujeres de edades avanzadas, embarazadas o con alguna enfermedad les afecta más las olas de calor fue y sigue siendo algo muy analizado, por ejemplo, por Sánchez-Guevara et al. (2021) donde demostraron las características de las mujeres hacen que estas sean las más afectadas a partir de una cierta edad por estas temperaturas elevadas.

Como en este trabajo, se realizaron estudios mediante series temporales (Achebak et al., 2020) al ser un buen método cuando se evalúan estas variables, donde se media la relación entre la mortalidad de las mujeres y las altas temperaturas en España. Se tenían evidencias para afirmar que esta relación era mucho más fuerte que para los hombres y como la edad era otra variable de interés.

Este cambio climático está causado por la gran cantidad de gases de efecto invernadero los cuales provocan un impacto directo en las altas temperatura y, por ende, en el aumento de la mortalidad, ha supuesto también efectos indirectos, por ejemplo, en el ámbito de salud laboral en los sectores de trabajo más vulnerables (Carretero, 2023) donde se ha detectado un aumento en el número de fallecidos cuando los trabajadores tenían que estar expuestos a temperaturas muy elevadas durante largos periodos de tiempo.

Otros efectos que tienen las olas de calor es agravar las enfermedades cardiovasculares o respiratorias de hombres y mujeres. La Organización Mundial de la Salud (OMS, 2021) estima que los efectos climatológicos adversos van a provocar cambios en las enfermedades generando así un aumento de alrededor de medio millón de muertes más de lo esperado al año (Linares et al., 2017).

Frente a esta situación climatológica conocida y por informes como el de la OMS (2024), que comentan como esto no va a ser algo temporal, a menos que se actúe con la disciplina suficiente, se les da a los países europeos el requisito de hacer planes de prevención adecuados teniendo en cuenta que son estos mismos, los que en cierta parte, provocan esta situación (Gómez y Rodríguez, 2021) muy dañina para el clima.

Por tanto, el impacto de las temperaturas extremas, tanto por ser bajas como por ser elevadas, tienen un impacto significativo en la mortalidad. Aunque el efecto de las temperaturas cálidas no hubiera sido tan significativo en las defunciones en un pasado, es en el verano del 2022 donde se evidencian las consecuencias del cambio climático y como en estos meses se observa un mayor número de fallecidos que en invierno. Por ello, en términos actuariales, existe una importancia de las aseguradoras en predecir la mortalidad futura de las personas vulnerables (conociendo su edad y género). Hay que tener en cuenta que el cambio climático continúa y la cantidad elevada de gases efecto invernadero también y, con ello, el aumento de la temperatura máxima.

3. DESCRIPCIÓN DE LA MUESTRA E HIPÓTESIS DE PARTIDA.

3.1. DESCRIPCIÓN DE LA MUESTRA.

Este trabajo gira en torno a la mortalidad y la temperatura en España, en particular para el verano del 2022. Antes de desarrollar los modelos de predicción, se van a realizar estudios previos. Para todo ello se toman distintas bases de datos, las cuales se van a enumerar en orden de aplicación.

3.1.1. Descripción de la muestra para la hipótesis de partida.

Como parte introductoria para las hipótesis de partida, se va a graficar cómo el verano del 2022 ha supuesto unas cifras de mortalidad por encima de lo esperado y, para conseguirlo, se toman los datos de Eurostat (EUROSTAT, 2024), donde con frecuencia mensual, medido en porcentaje, representa las defunciones desde enero de 2020 hasta septiembre del 2023.

Continuando con el estudio en las hipótesis de partida, se analiza el comportamiento mensual de la mortalidad en base a la temperatura de manera general (sin separar por género ni edad) en los años 2021 y 2022 entre los meses de abril y septiembre¹.

Para ello, se toman datos mensuales de temperatura en España de los datos abiertos de la Agencia Estatal de Meteorología, del archivo llamado “TM_MES_2021” y “TM_MES_2022” (AEMET, s.f.). y para la parte de mortalidad, se usa la base de datos de defunciones mensuales que proporciona el INE (Instituto Nacional de Estadística, s.f.). Esto se va a realizar para corroborar que existe una relación entre ambas variables, además de estudiar si esta relación es o no lineal e intentar conocer la función de distribución de la mortalidad. Todo esto se utilizará como análisis previo del apartado 4 del trabajo.

¹ También se comenta brevemente para meses fríos usando la misma base de datos.

Una vez realizado un contexto y estudios previos, la metodología y parte central de este estudio va a estar dividido en dos bloques con bases de datos y frecuencia distintas.

La primera parte metodológica del trabajo se centra en conocer cómo la mortalidad se comporta por las altas temperaturas en función del género y edad. La segunda parte metodológica está basada en el estudio de causalidad y predicción de la mortalidad usando modelos VAR unidireccionales (miden el impacto unidireccional de la serie temporal alta temperatura en la serie temporal mortalidad)

3.1.2. Descripción de la muestra para la primera parte metodológica.

Indagando en el primer bloque metodológico, se realizan modelos de regresión (GLM) con el fin de conocer el impacto de la temperatura en la mortalidad en el año 2022 según características individuales y para conseguirlo, se ha necesitado tomar la mortalidad mensual para cada edad, desde los 0 años hasta los 101, diferenciando los datos por hombres y mujeres.

Se han generado 101 modelos de mujeres, uno para cada edad, y 101 modelos de hombres, uno para cada edad, donde se estudia la relación entre la variable dependiente, la mortalidad (por edad), y la variable independiente, la temperatura, para los meses de abril a septiembre y así conocer a qué edades les afecta el aumento de la temperatura.

Estos datos se obtienen del Instituto Nacional de Estadística (INE, s.f) donde se toma para cada edad su mortalidad en los meses cálidos (abril-septiembre) primero para mujeres y después para hombres en 2022. En la parte de la temperatura mensual se siguen tomando los mismos datos que antes de la AEMET en el archivo “TM_MES_2022”.

3.1.3. Descripción de la muestra para la segunda parte metodológica.

Continuando con el estudio, y empezando con el segundo bloque metodológico, se elaboran modelos VAR, los cuales se van a usar para estudiar la causalidad y predecir la mortalidad futura. Las series temporales contarán con datos diarios de los meses de mayo a septiembre del 2022, para seguir con la línea del trabajo.

La parte de mortalidad diaria será estudiada gracias a la base de datos de MoMo (MoMo, 2024), donde se tomará la mortalidad diaria por rangos de edad, en las mujeres para el grupo de 75-84 y +85 años y para hombres para 45-64 y 65-74 años (consiguiendo así cuatro series temporales de mortalidad diaria).

Los rangos seleccionados son a raíz de los resultados de la primera parte metodológica. En esta se estudió para qué edades, en función del género, la temperatura influía en la mortalidad y es ese el motivo de que se tomen solo esos grupos de edades por género, ya que son los afectados por la situación climática y, por consiguiente, en los que interesa estimar el número de fallecidos futuro.

Para obtener las temperaturas máximas diarias se han tomado las de Córdoba como base del estudio, al ser una de las provincias ejemplo donde el calor es más elevado y persistente, y en un enfoque de cambio climático, donde las temperaturas tienden a estar aumentando, pueden ser estas algo representativo tanto del presente como del futuro.

La temperatura diaria se encuentra en los datos abiertos de la AEMET sobre datos meteorológicos diarios (AEMET, s.f).

Al final de los resultados del segundo bloque metodológico, se comentará brevemente un segundo modelo VAR estudiado con datos anuales, donde se toma una serie histórica desde el 1970 al 2022 tanto de mortalidad como de temperatura máxima anual.

Se toman las series temporales de mortalidad para algunas de las edades que los modelos GLM hayan concluido que existe un impacto significativo de la temperatura, igual que cuando se tenían datos diarios.

Los datos de mortalidad en este caso se obtienen de Human Mortality Database (Human Mortality Database, 2021), donde se toman para cada género la q_x correspondiente a cada edad desde el 1970 al 2021. Por falta todavía de los datos del 2022, se hace una estimación.

Los datos de temperatura anual histórica se obtienen a través de una descarga de pago de la base de datos de la AEMET, en la que se ha realizado un estudio previo en cada año para conocer la temperatura máxima anual en cada comunidad autónoma, cogiendo una provincia como referencia, y haciendo una media entre todas estas. Este es el procedimiento aproximado debido a la gran cantidad de datos (AEMET, s.f).

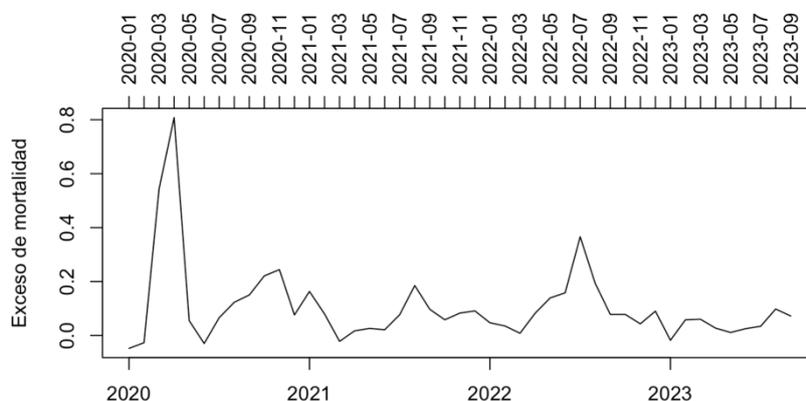
3.2. HIPÓTESIS DE PARTIDA

El verano del 2022 se ha considerado uno de los más calurosos de los últimos años, pudiéndose comparar con los años 2003 y 2016, aunque el porcentaje de muertos es más elevado incluso en este reciente año. Entre los meses de verano, se ha estimado que más de un tercio de la población fallecida fue a causa de las altas temperaturas observándose en el mes de julio un incremento por encima de lo esperado.

Hay estudios a nivel nacional que han conseguido demostrar este suceso (León et al., 2022) en personas de elevada edad o con enfermedades delicadas conociendo así, por edad y género, cuánto ha sido el exceso de fallecidos comparado con el esperado.

Para corroborar que este evento, algo inesperado si lo comparamos con los años 2020 y 2021, ha supuesto un aumento en las defunciones por encima de lo que se esperaba y es acorde a los distintos trabajos realizados en el país, se toman datos de Eurostat (2024), el cual ha recopilado datos de manera semanal sobre las defunciones. La figura 3.1 generada sobre los meses de España arroja los siguientes resultados hasta final del verano del 2023.

Figura 3.1
Exceso de mortalidad por mes en España



Fuente: Elaboración propia

El eje de abscisas representa los años de estudio y el eje de ordenadas el exceso de mortalidad, donde multiplicado por 100% se obtiene el porcentaje de exceso.

Se observa como existen cuatro picos representativos de exceso de mortalidad por encima de lo esperado. Los dos primeros pertenecen a la primera y segunda ola del COVID en el año 2020. El tercero pertenece al verano del 2021, donde ya se empezaban a contabilizar

fallecimientos a causa de un aumento de las temperaturas. Sin embargo, este pasó a un segundo plano si lo comparamos con el cuarto y último pico representativo, julio del 2022, donde el exceso de mortalidad representa casi un 50%, siendo el segundo más alto por debajo del inicio de la pandemia (García López, 2023)

Aunque este gráfico únicamente se centre en España, países europeos como Portugal, Italia y Grecia, experimentaron porcentajes similares en los meses de más calor (Ballester et al., 2023).

España es un país con gran variedad de temperaturas y, aunque el impacto de las altas temperaturas esté empezando a ser ahora significativo en la mortalidad, durante muchos años se ha estudiado el impacto de temperaturas extremas, tanto frías como cálidas, en el país y la necesidad de implementar planes para prevenir estos excesos.

Tras un estudio de mortalidad mensual en España durante distintos años pasados, es evidente que las cifras más altas siempre se han producido en los meses de invierno, en específico, diciembre y enero. Aun así es evidente como el 2022 ha recordado como es igual de importante las políticas frente ambos sucesos.

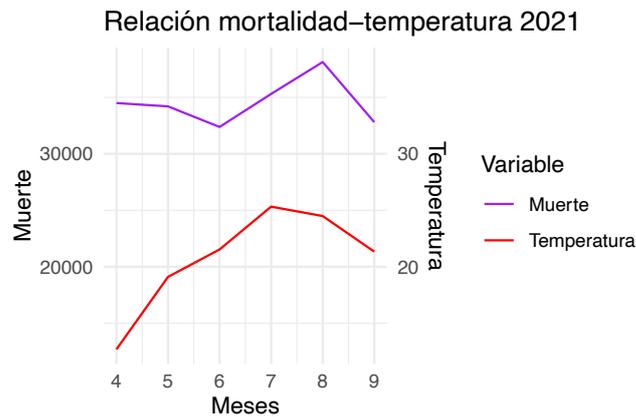
Por otro lado, si se observa la mortalidad mensual en los años 2021 y 2022 dada por el Instituto Nacional de Estadística, se empieza a ver una tendencia donde las tasas en los meses de verano empiezan a equipararse a la de los meses de invierno, siendo incluso en el año 2022 en algunos días superiores. Por ejemplo, el día 18 de julio en España se observaron 1.761 muertes siendo el día con más fallecidos, por encima de lo observado muchos días en enero (el mes con tasas más elevadas).

El Informe Mensual Climatológico de Julio de 2022 de la Agencia Estatal de Meteorología representa cómo en toda España experimentaba la categoría más extrema de temperatura siendo superior a la temperatura que había al final de siglo veinte e inicio del siglo veintiuno (AEMET, 2022).

Estos resultados son distintos a los que se muestran en el informe del 2021 (AEMET, 2021) donde pocas zonas cuentan con calor extremo y es por lo general, más moderado.

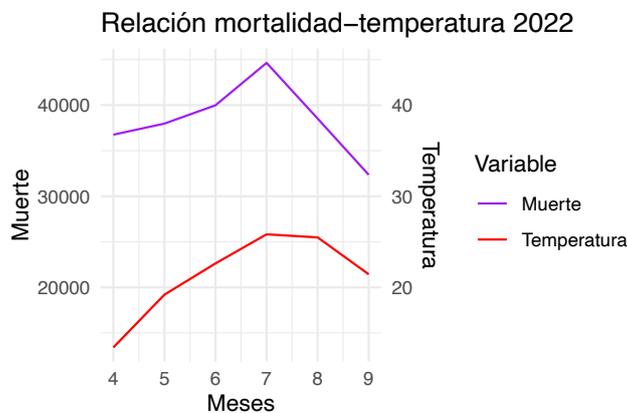
Para compararlo analíticamente, se ha estudiado la relación entre la temperatura y la mortalidad para los meses cálidos, de abril a septiembre, para los años 2021 y 2022 y así analizar sus diferencias y similitudes.

Figura 3.2
Relación entre mortalidad y temperatura en España 2021



Fuente: Elaboración propia

Figura 3.3
Relación entre mortalidad y temperatura en España 2022



Fuente: Elaboración propia

Para el año 2022, se puede ver en la figura 3.3 una fuerte relación de cómo a medida que va aumentando la temperatura, se experimenta un aumento en la mortalidad, siendo el punto máximo el número 7 que corresponde al mes de julio.

Para el año 2021, en la figura 3.2, la relación no es tan directa pero sí que se empieza a ver una clara tendencia de que el aumento del calor, a medida que nos acercamos a julio y agosto, está influyendo en el número de fallecidos.

Estos gráficos van en consonancia con los informes de la AEMET donde destacaban el año 2022 como característico de esto y el 2021 como un aviso de lo que va a suceder. Para el futuro estudio, y asegurarse de que existe una relación alta entre ambas variables en los meses cálidos, se ha realizado, con datos mensuales, el habitual coeficiente de correlación de Pearson, usado para medir la relación entre la mortalidad y temperatura al ser dos variables aleatorias numéricas.

Está definido como

$$\rho_{mortalidad,temp} = \frac{\sigma_{mortalidad \times temp}}{\sigma_{mortalidad} \times \sigma_{temp}} \quad (1)$$

siendo el resultado² para el 2022 de 0.4729103 y para el 2021 de 0.3533139.

En primer lugar, se puede considerar que la correlación no es del todo elevada si tenemos en cuenta la relación que se observa en los gráficos de las figuras 3.2 y 3.3 comentados anteriormente, sobre todo para el año 2022.

Por ello, se realiza otra prueba de correlación. En este caso, el denominado Coeficiente de correlación de Spearman definido por

$$\rho = 1 - \frac{6 \sum \text{diferencia entre los grupos}^2}{\text{número de grupos}(\text{número de grupos}^2 - 1)} \quad (2)$$

Estos grupos se asignarán en función de características comunes entre ambas variables aplicándose la misma valoración, cuando se cumpla la misma condición y, por ende, estando en el mismo grupo.

El objetivo no es que la relación entre ambas variables sea lineal como en la correlación de Pearson, lo que se busca es estudiar si estas se relacionan de la misma manera al mismo tiempo y agregarlas por grupos de movimientos similares.

² Resultados obtenidos a través de la función *cor* en R.

En este caso, resulta³ el coeficiente de Spearman para los meses cálidos del año 2022 una correlación de 0.7714286 y para el 2021 de 0.4857143.

La principal distinción entre un coeficiente de correlación y el otro es que mientras que el de Pearson solo evalúa la relación lineal entre variables, el de Spearman lo hace para cualquier relación. Además, el primero está más afectado por los valores atípicos pudiendo, por tanto, infraestimar el resultado de correlación.

Otra conclusión que se puede sacar de por qué el de Spearman se acerca más a lo visto en el gráfico, es a raíz de la distribución de los datos, ya que si estos no siguen una *Normal* (más adelante se demostrará esto) o la relación no lineal, este es el coeficiente (2) que mejor va a medir cómo la mortalidad y temperatura están relacionadas.

Antes de continuar con el estudio, es importante remarcar que esos coeficientes surgen de la toma del rango de meses entre abril y septiembre, pues si el estudio se realiza para los meses fríos, de enero-marzo y octubre-diciembre, las correlaciones son negativas y muy elevadas (para el 2022 entre los meses de enero y marzo es de -0.9869361 y para los meses de octubre y diciembre de -0.7591465). Se observa cómo a medida que baja la temperatura, la mortalidad aumenta.

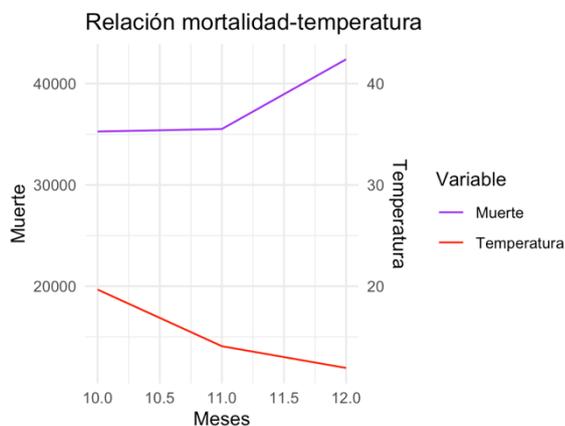
Gracias a los datos del Instituto Nacional de Estadística se puede corroborar como hasta el 2022, el mayor número de muertes en España era en los meses de enero y diciembre, siendo, por ejemplo, en el 2021 el número de defunciones era un 28% mayor en enero que en agosto (el mes más cálido). Como se sabe, esto en el año siguiente, 2022, va a cambiar, ya que será en verano el mayor porcentaje.

³ Resultados obtenidos a través de la función *cor* en R aplicando `method = "spearman"`.

Las figuras 3.4 y 3.5 de los meses fríos del 2022 representan lo siguiente.⁴

Figura 3.4

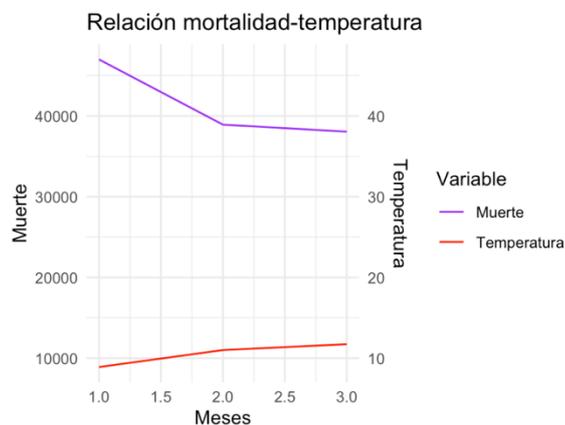
Relación entre mortalidad y temperatura en España 2022 desde octubre a diciembre.



Fuente: Elaboración propia

Figura 3.5

Relación entre mortalidad y temperatura en España 2022 desde enero a marzo.



Fuente: Elaboración propia

En la figura 3.4 los puntos 10, 11 y 12 corresponden a los meses octubre, noviembre y diciembre respectivamente.

En la figura 3.5 los puntos 1, 2 y 3 corresponden a los meses enero, febrero y marzo respectivamente.

⁴ En el Anexo también se encuentran los gráficos de meses fríos para el año 2021.

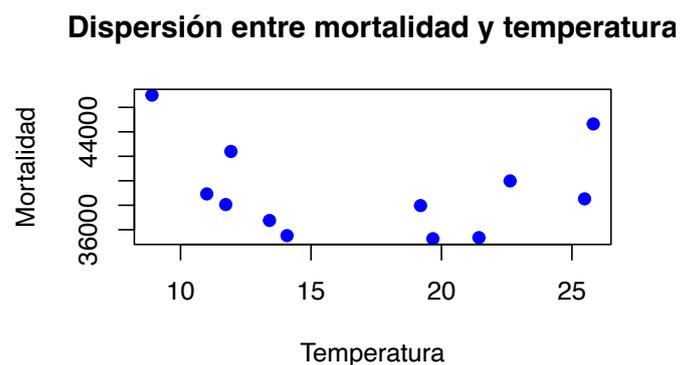
En el gráfico superior (octubre-diciembre) se observa que a partir de noviembre hay una bajada de temperatura y un aumento de la mortalidad. Este se mantiene elevado hasta febrero (gráfico inferior, punto 2 del eje de abscisas) cuando ya empieza a subir algo la temperatura y baja la mortalidad.

Es por esto que este trabajo, aun considerando que la mortalidad en España en el 2022 también es a causa de temperaturas muy bajas, se centra en los meses cálidos y el impacto de las altas temperaturas, al ser un fenómeno más reciente a raíz del cambio climático y, por tanto, algo más inesperado desde un punto de vista actuarial.

Retomando el estudio del por qué el coeficiente de Spearman se ajusta más a los datos, se puede deducir que los datos no sigan una distribución *Normal* o que la relación carece de linealidad. Para confirmarlo, se va a estudiar si la relación es lineal usando gráficos de dispersión e intentando ajustar una línea de regresión lineal. Además, en el siguiente apartado, mediante Kernel, se va a tener una idea de cómo es la distribución de mortalidad estimada.

En primer lugar, si tomamos todos los meses del año 2022, el gráfico de dispersión arroja lo siguientes resultados en la figura 3.6.

Figura 3.6
Gráfico de dispersión entre mortalidad y temperatura para el año 2022



Fuente: Elaboración propia

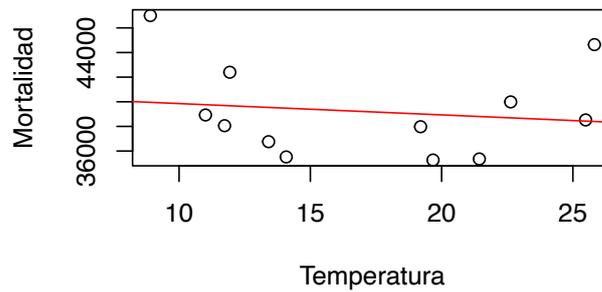
La figura 3.6 muestra como no hay una relación lineal entre las variables al haber un número elevado en mortalidad tanto para temperaturas muy bajas como altas, y como no se puede crear un patrón o línea. Además, el coeficiente de correlación de Pearson nos da

un resultado de -0.1457972 si se toman los doce meses, por lo que hay poca correlación lineal. Por consiguiente, no muestran una relación lineal significativa.

La figura 3.7 intenta ajustar todo el año 2022 a un modelo de regresión lineal.

Figura 3.7
Gráfico de dispersión entre mortalidad y temperatura para el año 2022

Gráfico de dispersión de Temperatura vs Morta

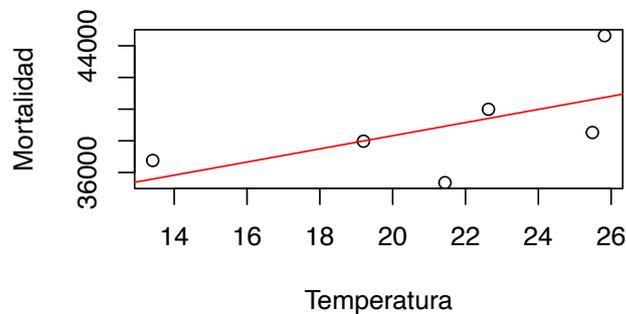


Fuente: Elaboración propia

El gráfico nos confirma al trazar la línea de regresión que la relación no es lineal. Es importante comentar que si el estudio se centra solo en meses cálidos, de abril a septiembre, en el siguiente gráfico de dispersión, la figura 3.8, se puede ajustar en mayor medida una línea de regresión. Aun así, se puede ver como todavía hay datos muy por encima o muy por debajo de la diagonal.

Figura 3.8
Gráfico de dispersión entre mortalidad y temperatura para los meses de abril a septiembre

Gráfico de dispersión de Temperatura vs Morta



Fuente: Elaboración propia

El coeficiente de correlación de Pearson (1) puede no ser el más apropiado para medir la relación entre las variables de interés cuando estas no siguen linealidad y existen valores atípicos a raíz de los meses de enero, julio y diciembre.

El coeficiente de correlación de Spearman (2), el cual no mide si la relación es lineal y es más robusto a valores atípicos, es más correcto para los datos de este trabajo. Este proporciona para los meses cálidos del 2022 un coeficiente de 0.7714286 entre la relación de la mortalidad y la temperatura, reflejando así una clara relación entre ambas.

Este apartado, al comprobar la relación estrecha que siguen ambas variables en este año, resalta la utilidad de conocer el impacto de las temperaturas y comprobar si estas pueden ser de utilidad para predecir la mortalidad futura y, desde un punto de vista actuarial, estimar el aumento de mortalidad por encima de lo esperado.

4. METODOLOGÍA.

Las olas de calor y el impacto en la mortalidad española ha sido caso de estudio desde principios del siglo. Como se comentaba anteriormente, los años 2003 y 2016 ya experimentaron aumentos significativos en la mortalidad a raíz del calor, pero es ahora en el 2022 donde se empieza a realmente poner el foco en esos planes de prevención poco efectivos y estudiar cómo las personas vulnerables están muy afectadas cada vez a temperaturas más elevadas y persistentes.

Previamente, se demostró cómo en el año 2022 la mortalidad en los meses de verano estaba superando la mortalidad en los meses de invierno, algo inusual comparado con los años anteriores. Esto indica como el cambio climático está pudiendo cambiar los patrones de salud y longevidad de la población. Que en ese año las altas temperaturas fueran generalizadas en toda España, alcanzándose límites que se pensaban inalcanzables, ya no solo por la intensidad si no por la durabilidad, está haciendo, entre otras, que las personas vulnerables se enfrenten a un agravamiento de su condición a raíz de esta situación anómala. Esto tiene un efecto directo en cuanto riesgo de suscripción en seguros de salud, puesto que puede generar un aumento en la demanda de servicios por encima de lo esperado y que la prima para cubrir esto sea insuficiente.

Por otro lado, en seguros de vida que cubran la mortalidad, si la persona asegurada se encuentra en aquellas edades afectadas más por el calor alto y persistente, puede que la probabilidad de fallecimiento aumente y, con ello, la cantidad de reservas que la aseguradora debe tener, con la consecuente disminución de fondos propios.

4.1 PRIMERA PARTE METODOLÓGICA.

La primera parte del estudio tiene como objetivo conocer para que edades el impacto de altas temperaturas es perjudicial en la mortalidad, dando lugar a un aumento de esta. Para ello, se van a tomar los datos mensuales de temperatura de los meses cálidos de España (abril a septiembre) en el año 2022. Para el análisis se han omitido los meses fríos, al estar este trabajo enfocado en el impacto de las temperaturas extremas de calor en la mortalidad, ya que si se incluyeran esos meses, habría que hacer dos estudios por separado para no sesgar las muertes en los meses de frío con las muertes en los meses de calor, y viceversa.

La idea metodológica es la siguiente. Se van a generar 101 modelos para mujeres y 101 modelos para hombres, tomando como variable dependiente la mortalidad en función de la edad y como variable explicativa la temperatura.

Cada modelo estará formado por la mortalidad mensual a la edad i para el género j de abril a septiembre y la temperatura mensual de abril a septiembre.

Donde $i \in [0, 101]$ y $j \in [mujer, hombre]$

$$mortalidad_{i,mujer} = cte + coeficiente * temperatura$$

$$mortalidad_{i,hombre} = cte + coeficiente * temperatura$$

El subíndice i representa la edad, y se tiene la mortalidad desde la edad 0 hasta la 101 para cada sexo, por lo que hay 101 modelos para mujer (uno por mortalidad a la edad i) y 101 modelos para hombre (uno por mortalidad a la edad i).

Se toman tantas edades como posibles hay en la base de datos del Instituto Nacional de Estadística, para así hacer un estudio completo. Además, el motivo de estas dos variables (edad y género) es a raíz de diversos estudios que reflejan una diferencia en las edades a las que les afecta el calor en función del género.

El objetivo no es solo conocer si la temperatura causa un efecto en la mortalidad sino conocer también qué edad causa ese efecto y a qué sexo pertenece. Esto da un análisis más útil para poder usarlo actuarialmente, donde la edad es siempre una variable muy importante y conocer si la mortalidad de las personas a una determinada edad va a estar influenciada por el calor puede aplicarse en muchos estudios.

Una vez contada la idea metodológica, se va a llevar a la práctica. El primer paso es determinar qué modelo de regresión va a ser el utilizado para el estudio. Una vez obtenido, se realiza un bucle en R⁵, que para cada edad, dé los resultados del modelo y así poder ver si la variable explicativa (temperatura) es significativa para la variable dependiente (mortalidad) y en qué grado de significatividad.

Para seleccionar el tipo de modelo de regresión, uno de los más comunes son los modelos lineales,

⁵ En el anexo se encuentra el código.

(3)

$$y = \beta_0 + \beta_1 x + \varepsilon$$

donde $\varepsilon \sim N(0, \sigma^2)$

se estudia la relación entre la variable endógena (mortalidad) y la exógena (temperatura). Sin embargo, en este tipo de modelos se puede estar incumpliendo dos condiciones, entre otras, fundamental de ellos.

La primera es que la relación debe ser lineal y en el apartado anterior se ha podido ver que si se tomaba la muestra de los doce meses, mediante el gráfico de dispersión y correlación de Pearson, indicaban una falta de linealidad. Para el caso de solo los seis meses cálidos, sucedía algo similar, aunque se podía trazar una línea algo más clara, la mayoría de los valores se encontraban o por encima o por debajo de esta.

Otra condición que debe suponerse es que los residuos sigan una distribución normal. En primer lugar, la mortalidad está influida por diversos factores y su variabilidad puede generar que los residuos no cumplan con esa condición. Además, ya se ha visto como esta puede estar afectada por efectos extremos, generando así una mayor cantidad en los extremos de la distribución. Por ende, no se puede asegurar que la distribución del error siga una Normal.

La falta del cumplimiento de condiciones necesarias para usar modelos de regresión lineal clásicos, lleva a aplicar Modelos Lineales Generalizados, siendo una buena opción al no cumplir con algunos supuestos del modelo anterior.

Los modelos líneas habituales cuentan con un avance, los modelos GLM(*Generalized Linear Models*), en los cuales no hace falta que los errores sigan una $\sim Normal$, y así la variable respuesta, en este caso mortalidad, no requiere que esa sea su distribución.

Estos modelos tienen como característico la función de enlace la cual está determinada por la variable explicativa y tomará la que mejor se ajuste a sus características.

En cuanto al modelo lineal generalizado, se va a seguir cumpliendo esta idea,

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (4)$$

pero la diferencia viene por la parte del componente sistemático η donde se conoce que este en LM (Lineal Models) era igual a la esperanza de la variable dependiente y ahora en los generalizados no se va a cumplir exactamente esa igualdad, como se indica en estudios de datos con este tipo de modelos. Por la parte del modelo lineal habitual (López-González y Ruiz-Soler, 2011) existe una *relación directa* entre ambos valores comentados, mientras que por GLM se encuentra entre estos la conocida *función de enlace* (dependiendo de la variable respuesta), muy característica en estos modelos.

En el caso de los lineales

$$\eta_i = \text{esperanza matemática de la variable dependiente}_i$$

En el caso de los lineales generalizados, con la función enlace será

$$\eta_i = g(\text{esperanza matemática de la variable dependiente}_i)$$

La función de enlace $g()$ es conocida por medir la relación entre la variable dependiente (en este caso la mortalidad) y la explicativa (en este caso la temperatura) sin la necesidad de que esta sea lineal.

Para aplicar una función de enlace adecuada, se va a estudiar la distribución de mortalidad, donde se van a tomar los datos para todas las edades mensuales de España en el año 2022 con todos los meses, al ser lo más genérico.

Se sabe que los valores que se tienen son valores aleatorios, y que aunque el número de muertes, por ejemplo, en mayo han sido de 36.759, las cantidades cercanas a este tienen alta probabilidad de que hubieran salido o que salgan si la situación es parecida.

El método de estimación Kernel mide la probabilidad de obtener un valor según la cercanía de los valores próximos. Esto permite estimar la función de densidad de los datos asignando a cada uno de ellos una variable aleatoria continua dándole así distribución continua y una probabilidad, haciendo que los datos cercanos al obtenido cuenten con alta probabilidad de suceder.

Por ende, se está suponiendo que la probabilidad de que el número de muertes en mayo hubiera sido de 36.760 es muy alta y no 0, como sí que en otros métodos se obtendría (al no ser exactamente esa cantidad).

Mediante la programación en R, se realiza primero un Kernel Uniforme con el fin de no sesgar los resultados y después de uno Gaussiano. Para cada uno se va a determinar cuántos valores se deben incluir a la hora de medir puntos cercanos mediante h .

En el primer caso, se considerará para el Kernel Uniforme,

$$h = 1.06 \frac{\widehat{\text{desviación típica muestra}}}{\sqrt[5]{\text{tamaño muestra}}} \quad (5)$$

En el caso de Kernel Gaussiano,

$$h = 0.9 \min(\widehat{\text{desviación típica muestra}}; q_{75} - q_{25} / 1.34) \frac{1}{\sqrt[5]{\text{tamaño muestra}}} \quad (6)$$

Donde $q_{75} - q_{25}$ es la resta de cuantiles.

Es importante esto ya que si h es muy alta va a suavizar demasiado la función de densidad y se van a dar probabilidades altas a números muy lejanos, y si, por el contrario, es muy baja, va a considerar muy pocos datos cercanos y va a tender a una distribución empírica.

Una vez realizados ambos, se generarán escenarios partiendo de la secuencia generada de posibles valores y la función de distribución obtenida. Se crean números aleatorios de una Uniforme (0,1) donde a cada valor de la función de distribución se le asigna uno de estos números aleatorios para cada uno de los escenarios, donde finalmente el resultado es el de la figura 5.1. En esta se observa como la distribución de número de fallecidos puede acercarse a una distribución de Poisson, por lo que será esta la seleccionada, tomando como función de enlace el logaritmo, algo habitual con esta distribución.

Por tanto, ya conocido el tipo de modelo y la función de enlace, en primer lugar, se realizan los modelos para las edades de 0 a 101 en las mujeres en España para el año 2022. Se usan los datos de abril a septiembre proporcionados por el Instituto Nacional de Estadística (INE) en su página web acerca de las defunciones y las temperaturas mensuales dados por la Agencia Estatal de Meteorología (AEMET).

Mediante la realización de un bucle, se aplica para cada edad de las mujeres un modelo generalizado usando la función *glm* siendo la variable dependiente la mortalidad a la edad i y la explicativa la temperatura, añadiendo la parte de la distribución que será Poisson, por el gráfico de la figura 5.1 y la función de enlace *log* en cada modelo.

Una vez realizado para todas las edades, se obtiene a partir de qué momento las mujeres están más afectadas por las altas temperaturas y a qué grado de significatividad. Estudios reciente como en Ballester et al., 2023 explican como los rangos de edad comprendidos entre hombres y mujeres a los que les afecta la temperatura son distintos. Ellos concluyen con que a las mujeres en Europa les afecta más el calor y las edades más perjudiciales para ellas son a partir de los 80 años. En este trabajo y con los modelos GLM por edad y sexo, se va a ver en los resultados si coincide con sus conclusiones para el territorio español.

En segundo lugar, se realiza los modelos para las edades de 0 a 101 en los hombres en España para el año 2022, usando los datos de abril a septiembre proporcionados por el Instituto Nacional de Estadística (INE) en su página web acerca de las defunciones y las temperaturas mensuales dados por la Agencia Estatal de Meteorología (AEMET), como se ha utilizado para las mujeres.

Mediante la realización del mismo bucle, se va a aplicar para cada edad de los hombres un modelo generalizado usando la función *glm* siendo la variable dependiente la mortalidad a la edad i (ahora del otro género) y la explicativa será la temperatura, añadiendo la parte de la distribución que será Poisson y la función de enlace *log* en cada modelo.

4.2 SEGUNDA PARTE METODOLÓGICA.

Hay una evidencia de que la temperatura extrema causa un aumento en la mortalidad y ese aumento es algo inesperado y heterogéneo, no va a afectar a todo el mundo igual. Conocer a qué edades, en función del género, les va a provocar un aumento en su mortalidad es el primer paso para reducir la incertidumbre.

Una vez ya conocidas las edades vulnerables, es solo para estas para las que se debe estudiar el valor futuro de mortalidad por tres motivos:

1. Para estas edades sí existe una respuesta en la mortalidad ante un shock en la temperatura.
2. El resto de las edades no van a experimentar ningún cambio en su mortalidad.
3. Las altas temperaturas siguen creciendo, lo que da a intuir, que la mortalidad de las edades afectadas también.

El primer estudio ya se cubre en la primera parte metodológica, permitiendo conocer las edades, en función del género, afectadas por las olas de calor.

Para poder llevar a cabo la segunda parte, la parte predictiva, se va a intentar estimar el proceso estocástico (una sucesión de variables aleatorias ordenadas en el tiempo), usando como base de estudio series temporales.

Como se comentó en el punto tres del proyecto, se toman series temporales diarias de mortalidad y temperatura para demostrar cómo esta última puede ayudar a predecir la mortalidad en los meses calurosos.

En términos de predicción de series temporales, los modelos ARIMA son los más usados debido a su efectividad, donde estudia la dependencia entre una variable y sus retardos y la dependencia entre una variable y el retardo de los errores de la predicción.

$$y_t = cte + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \gamma_p y_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \phi_q \varepsilon_{t-q} \quad (7)$$

donde $\varepsilon \sim i. i. d. (0, \sigma^2)$

Siendo p la parte autorregresiva, d la parte media móvil y teniendo en cuenta que se pueden tomar diferencias para que se cumpla la condición de estacionariedad.

La limitación de estos modelos es que solo se centran en una serie temporal y, como el estudio gira en torno a la mortalidad y a las altas temperaturas se opta por los modelos VAR, que permite ver la relación entre series temporales, pudiendo así medir la causalidad y estudiar el impacto del cambio climático en la mortalidad mediante funciones impulso respuesta.

El Modelo Vector Autorregresivo (VAR) es un modelo multivariante que analiza series temporales donde la ecuación del modelo tiene tanto los retardos de la variable dependiente como los retardos de las variables explicativas.

En este caso, al contemplar solo el efecto de las temperaturas sobre la mortalidad y no al revés, se opta por un VAR unidireccional.

En este ejemplo se puede ver como se incluyen también los retardos de temperatura para explicar la mortalidad, y existirán tantos como el modelo determine.

$$\begin{aligned}
 Y_{mortalidad_t} = & Y_{mortalidad_0} + \gamma_{mortalidad,mortalidad} Y_{mortalidad_{t-1}} \\
 & + \gamma_{mortalidad,altatemperatura} Y_{altatemperatura_{t-1}} + \dots + \varepsilon_{mortalidad_t}
 \end{aligned}
 \tag{8}$$

Conocidos los rangos de edad para los que es interesante predecir la mortalidad futura, se crean cuatro modelos VAR combinando cada una de las series temporales de mortalidad diaria con la serie temporal de temperatura máxima.

Modelo VAR 1 (mortalidad diaria del rango 1 y temperatura máxima)

Modelo VAR 2 (mortalidad diaria del rango 2 y temperatura máxima)

Modelo VAR 3 (mortalidad diaria del rango 3 y temperatura máxima)

Modelo VAR 4 (mortalidad diaria del rango 4 y temperatura máxima)

Cada rango i estará formado un grupo de edades, dos para mujeres y dos para hombres. Antes de la predicción, se siguen una serie de pasos previos, estudios de estacionariedad, de causalidad, de los residuos, obtención de las funciones respuesta impulso y descomposición de la varianza para así conseguir el objetivo esperado.

Usar modelos VAR con datos diarios es algo muy utilizado en otros trabajos. Por ejemplo, Linares-Gil et al. (2017) comentaban que usar mortalidad diaria es mejor indicador si se habla de altas temperaturas.

Es por ello que, para realizar predicciones eficientes con cada modelo se siguen una serie de pasos previos como son: el estudio de estacionariedad, de causalidad, de residuos, obtención de las funciones respuesta impulso y descomposición de la varianza; para así conseguir el objetivo esperado, predecir.

4.2.1 Estacionariedad.

En primer lugar, se debe estudiar si las series temporales (la mortalidad por distintas edades y género y la temperatura) son estacionarias. Esto surge a raíz de la necesidad de los modelos VAR para estimar y predecir de manera eficiente, ya que suponen que tanto en media como varianza las series que se toman van a ser estacionarias, al igual que los modelos ARIMA antes comentados.

Para ello se toman todas las series de mortalidad en función de la edad y género que se tienen, además de la temperatura y se realizan distintos test para comprobarlo, ya que gráficamente no se puede concluir nada debido a que la serie que se toma es una sola muestra de un proceso estocástico y este resultado es aleatorio.

Los test más comunes son el test de Dickey-Fuller, el test de Phillips-Perron y el procedimiento KPSS, siendo distinto a los dos primeros por no ser un test de raíz unitaria. En cuanto al test de Dickey-Fuller y Phillips-Perron, parten de la base de que la hipótesis nula determina si la serie tiene raíz unitaria mediante un modelo autorregresivo de orden 1, AR(1).

Dickey & Fuller (1979) determinan que si el coeficiente que acompaña al primer retardo en valor absoluto es igual 1, la serie temporal no es estacionaria, pero que esta tiende a ser estacionaria si ese coeficiente es menor, en valor absoluto, a 1.

Por tanto, las hipótesis que marca son,

$$H_0 \quad |\rho| = 1$$

$$H_1 \quad |\rho| < 1$$

Siendo $|\rho|$ el primer retardo del modelo AR(1) en absoluto.

Phillips & Perron (1988) realizaron el mismo estudio con ayuda de los diversos artículos que escribieron antes Dickey & Fuller, y partiendo de la misma idea que el de estos, mejoraron algunas limitaciones como, por ejemplo, los problemas con la autocorrelación. Mientras el test de DF establece que no hay relación entre los valores observados con sus pasados, el test de PP sí permite esta relación y esto ayuda a identificar si la serie es estacionaria, aun con presencia de correlación.

Por último, el test KPSS (Kwiatkowski et al., 1992) sigue mejorando el test de Dickey & Fuller, y el estudio aquí es directamente la estacionariedad, ver cómo se ha generado la serie y si ha sido bajo un proceso estacionario, y no tanto por $|\rho|$.

Por ello, se realiza cada test en cada serie temporal mediante un bucle, con el fin de comprobar la estacionariedad de las series temporales. En caso de que no cumplan con esta idea, se tomarán tantas diferencias como sea necesario,⁶ y se volverán a realizar los tres test de nuevo, para comprobar, que una vez aplicadas las diferencias, si cumplen con la condición.

El comando *ndiffs* va a determinar, de cada serie que necesite aplicar diferencias, la cantidad que se debe diferenciar para ser estacionarias.

4.2.2 Modelo VAR.

Una vez todas las series temporales son estacionarias, se estudia de qué orden es el modelo VAR (cuántos retardos de cada variable de deben incluir) tomando la mortalidad como variable dependiente, y los retardos de esta y de la temperatura como variables explicativas.

En el siguiente apartado, se comprobará mediante la causalidad de Granger si el efecto de las variables explicativas es significativo y si por tanto la alta temperatura ayuda a predecir la mortalidad para cada rango de edad en función del género.

Para elegir el mejor modelo en función de los retardos que se pueden añadir, se van a usar el Criterio de información de Akaike y el Criterio de información bayesiano, ya que el mejor modelo será el que minimice estos dos criterios.

⁶ En el anexo se encuentra el código para ello.

$$AIC(\text{modelo}) = -2 \log ML(\text{modelo}) + 2 * \text{número parámetro} \quad (9)$$

$$BIC(\text{modelo}) = -2 \log ML(\text{modelo}) + \log(\text{observaciones}) * \text{núm parámetros} \quad (10)$$

Donde $ML(\text{modelo})$ representa la función del parámetro por Máxima Verosimilitud (Maximum Likelihood).

Cada uno de los criterios cuenta con limitaciones. Por ejemplo, AIC, cuando la serie es pequeña, aplica más parámetros al modelo de los que realmente necesita y BIC más que añadir parámetros de más, lo que hace es penalizar cuando son demasiados.

Combinando cada serie de mortalidad con la temperatura, se he estudiado, teniendo en cuenta las ventajas y limitaciones de cada criterio, y determinando cuántos retardos se deben añadir para explicar mejor el modelo, y en función de eso seleccionar el modelo VAR óptimo.

4.2.3 Causalidad de Granger.

Una vez todas las series temporales son estacionarias, y se conoce el orden del modelo, se debe estudiar la causalidad de Granger. En distintos trabajos con distintas variables, por ejemplo en Ashenfelter, et al. (1982) estudian la causalidad con modelos VAR.

Granger, en el año 1969, establece que si se tienen evidencias de que si una variable causa a otra, los valores pasados de la causante ayudan a predecir los valores de la causada. En este trabajo, se comprueba que si la alta temperatura causa la mortalidad, se van a usar los retardos de temperatura para ayudar a predecir la mortalidad.

Para ello, se va a aplicar *grangertest* entre cada serie de mortalidad y la temperatura máxima, para así estudiar si las altas temperaturas pueden ayudar a predecir la mortalidad en los meses cálidos.

4.2.4 Función Impulso-Respuesta.

Una vez que se sabe que las altas temperaturas sí que causan un impacto en la mortalidad, para analizar cómo se comportan y se relacionan distintas variables a lo largo del tiempo, la función de respuesta impulso (IRF, por sus siglas en inglés) es la técnica más utilizada. Cuando una variable exógena experimenta un cambio repentino, esta función recrea cómo la variable endógena se ve afectada por este efecto anómalo.

En este trabajo, la variable a estudiar es la mortalidad tras un impacto inesperado de la temperatura. El tiempo que tarda la mortalidad en volver a su estado habitual antes de que sucediera el shock es muy relevante para conocer la magnitud del impacto y si este dura mucho en el tiempo.

Se ha aplicado en cada rango de edad de hombres y mujeres, la función $irf()$ en R de cada modelo VAR, siendo la variable de respuesta la mortalidad y estudiándose 8 periodos en adelante para calcular la respuesta al impulso. Además de calcular intervalos de confianza mediante el método Bootstrap para las estimaciones de la función respuesta impulso y medir así la incertidumbre, donde a mayor amplitud mayor será esta última.

4.2.5. Descomposición de la varianza.

Una vez estudiadas las IRF, es interesante el impacto de las altas temperaturas en la variabilidad de la mortalidad mediante la técnica de descomposición de la varianza .

Durante 50 periodos se va a identificar qué parte de las fluctuaciones de la mortalidad son atribuibles a cambios en las temperaturas extremas y cuáles a las propias variaciones de la mortalidad por otros factores.

Para ello se aplica la función *fevd* en cada modelo de mortalidad y temperatura especificando en qué variable se quiere hacer la descomposición de la varianza, en este caso, ante un impacto inesperado en la temperatura.

4.2.6. Residuos del modelo.

Finalmente, antes de la predicción, para que los modelos sean válidos, se debe comprobar si el término de error cumple ciertas condiciones importantes⁷.

4.2.6.1. Normalidad de los residuos.

En primer lugar, se va a conocer la distribución de los errores los cuales deben seguir normalidad para asegurar precisión.

Para ello se realizan dos test, Cramer-von Mises (CVM) y Anderson-Darling (AD), ambos muy usados para empezar a conocer la distribución, en este caso, del residuo.

Ambos procedimientos parten de la idea de estudiar la diferencia entre la función de distribución estimada y la real. En este caso, se va a tomar como distribución estimada una Normal y se va a comparan con la distribución real desconocida, en el caso de que el pvalor sea mayor al nivel de significatividad, no hay evidencias para rechazar la hipótesis de normalidad y, por tanto, los residuos seguirán una distribución normal.

Para ello, se aplicará a cada residuo de cada modelo *cvm.test* y *ad.test*, usando la Normal como distribución para la hipótesis nula con media 0 y de desviación típica el error.

La diferencia entre CVM y AD, aunque estudian lo mismo, es que el primero da más valor a los datos de la parte central donde suele haber más datos y es el segundo el que da más peso a las colas de la distribución, más común en ciencias actuariales.

4.2.6.2. Residuos incorrelados.

En segundo lugar, se estudia si no hay autocorrelación en los residuos mediante el test de Ljung-box, que parte de la hipótesis nula que establece ausencia de autocorrelación.

Para cada variable del modelo, tomando *Box.test* con diez retardos, se va a aplicar este procedimiento y en caso de que no se pueda rechazar la hipótesis nula de no autocorrelación se puede asegurar que los modelos están capturando bien los movimientos de cada serie al ser los errores independientes.

⁷ Que cada término de error del modelo cumpla con las cuatro condiciones implica que el modelo tiene una alta capacidad predictiva y no está infraestimando o subestimando.

4.2.6.3. Residuos homocedásticos.

En tercer lugar, los residuos tienen que distribuirse a lo largo del tiempo sin mostrar patrones y, por lo tanto, cumplir la condición de homogeneidad.

Para cada rango de edad se genera,

- Un modelo de regresión lineal donde la variable dependiente es el residuo al cuadrado de la mortalidad y las variables explicativas los retardos de temperatura máxima y la misma mortalidad.

$$\varepsilon^2_{mortalidad} = \gamma_{t-1}^{mortalidad} + \gamma_{t-1}^{temperatura\ max} \quad (11)$$

- Un modelo de regresión donde la variable dependiente es el residuo al cuadrado de la temperatura máxima y las variables explicativas los retardos de la misma temperatura máxima y la mortalidad.

$$\varepsilon^2_{temperatura\ max} = \gamma_{t-1}^{mortalidad} + \gamma_{t-1}^{temperatura\ max} \quad (12)$$

Con esto se calcula el test Breush-Pagan, siendo la diferencia entre el tamaño de los residuos por la parte autorregresiva del modelo, todo ello multiplicado por el valor del R^2 del modelo de los errores generado previamente.

4.2.6.4. Media cero.

Finalmente, se comprueba si la media del término de error es nula. En caso de que no sea así, la estimación no será correcta y sobreestimaré o infraestimaré.

Obtenida la media de los residuos, se comprueba si se encuentra dentro de la región crítica, calculado a partir de los cuantiles 2,5% y 97,5%. En caso de que la media esté dentro de este, la hipótesis de esperanza nula se cumple.

Una vez comprobadas estas cuatro condiciones, se puede asegurar que los residuos de los modelos están incorrelados, son homocedásticos, con media cero y que siguen una distribución normal.

4.2.7. Predicción de la mortalidad.

Una vez comprobados que los modelos VAR cumplen con la condición de causalidad, donde los valores pasados de las altas temperaturas son significativos para predecir la mortalidad, y que los errores siguen todas las condiciones necesarias de normalidad, homocedasticidad, media cero e incorrelación, se realiza la predicción para obtener los valores futuros de la mortalidad teniendo en cuenta las altas temperaturas.

El uso de modelos VAR permite predecir una variable de interés con ayuda de otra, la cual afecte significativamente a la primera. A nivel actuarial esto es muy relevante a la hora de poder estimar la mortalidad futura teniendo la alta temperatura.

Poder medir el riesgo y encontrar unos valores por los que puede fluctuar la mortalidad futura en función de la edad a raíz del cambio climático, tiene un impacto muy importante en las aseguradoras ya que conocerá si su capital, en caso de seguros donde se asegura la mortalidad, es suficiente para cubrir los sucesos posibles o si debe provisionar una cantidad mayor. Para seguros de salud, como se sabe que las temperaturas afectan a las personas con enfermedades cardiovasculares o respiratorias, el estudio es el mismo. Hay que conocer si las primas son suficientes para los eventos que pueden suceder, a raíz de las olas de calor extremas.

Es importante considerar los resultados de Descomposición de la varianza, ya que muestran cómo el aumento de las temperaturas afecta a la mortalidad, pero este impacto es uno de los muchos factores que influyen en esta.

Para ello, se van a utilizar Point forecast, valor esperado de la variable aleatoria para un periodo futuro determinado y Interval forecast, intervalo de confianza por el que el proceso estocástico va a fluctuar y en el que se va a encontrar el verdadero valor de la estimación.

Tanto Point forecast como Interval Forecast se harán para los cuatro modelos VAR(1)⁸ del estudio para tener así una visión más completa del estudio, tanto como por el valor esperado como por la incertidumbre que este puede tener.

⁸ Un dato relevante es que cada serie temporal cuenta con 153 observaciones, lo que va a influir en la capacidad de predicción del modelo.

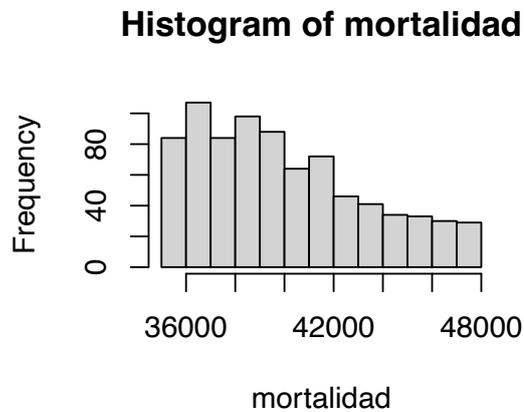
5. RESULTADOS.

5.1 RESULTADOS PARA LA PRIMERA PARTE METODOLÓGICA.

La figura 5.1 representa la estimación de la distribución de la mortalidad mediante el método Kernel, con el objetivo de establecer una función de enlace óptima de los modelos GLM en función de la variable dependiente, mortalidad.

Figura 5.1

Histograma de la mortalidad para conocer su distribución



Fuente: Elaboración propia

Kernel es una estimación, pero visualmente se aprecia como la mortalidad puede acercarse a otro tipo de distribuciones que no sean la Normal, siendo una de las posibles Poisson.

Diversos estudios que han analizado la mortalidad usando modelos GLM y a la hora de modelar la mortalidad como variable explicativa asumen que sigue esta distribución Poisson (Puig, X., Ginebra, J., & Gispert, R., 2005). Esta se suele utilizar cuando la variable explicativa es categórica y, en este caso, representa el número de fallecidos en los meses cálidos para una edad específica, cuya frecuencia puede ser ajustada por una distribución de Poisson. Es por lo que en el trabajo se va a utilizar, ya conocida la familia, una función de enlace logarítmica. Esta es la más usada cuando la variable dependiente sigue este tipo de distribución.

Una vez conocida la familia del modelo, se obtienen los resultados del primer estudio. Con los modelos GLM para las mujeres se va a ver para qué edades la temperatura es significativa en la mortalidad.

Debido a la gran cantidad de modelos, se va a ir observando por rango de edad, empezando de los 0 a los 25. Para la edad 25 se observa como la temperatura no es significativa en la figura 5.2, aunque para ninguna de todas las edades de este rango lo es.

Figura 5.2

Resultados del modelo GLM para la edad 25.

```
Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadM)

Deviance Residuals:
    151     152     153     154     155     156
 0.20807 -0.33790 -0.08716  0.05574  0.09924  0.04926

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.33935     1.43594  -0.236   0.813
Temperatura  0.06575     0.06332   1.038   0.299
```

Fuente: Elaboración propia

Aunque se observa que el coeficiente de temperatura es positivo, es decir, tiene una relación positiva con la mortalidad, el p-valor es muy elevado y, por tanto, no tenemos evidencias suficientes para rechazar la hipótesis nula, siendo esta:

$$H_0: \beta_i = 0 \tag{13}$$

donde β_i corresponde al coeficiente de la temperatura, en este caso.

Para el rango de edad entre 26 y 51, representando aquí el modelo para la edad 40, se ve que la variable temperatura tampoco es significativa. La única edad donde la temperatura provoca un efecto es para la edad de 49 años, algo anormal ya que se ve como en los años cercanos no hay ninguna edad más donde la variable independiente sea relevante.

Por tanto, para edades tempranas de las mujeres, las olas de calor no tienen un efecto claro y directo⁹.

⁹ Los resultados para los hombres demuestran como las altas temperaturas sí que van a causar efecto a edades más tempranas.

Figura 5.3
Resultados del modelo GLM para la edad 40

```
Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadM)

Deviance Residuals:
    241     242     243     244     245     246
-0.15269 -0.06457 -0.09870 -0.39212  0.15338  0.53184

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.12107    0.59709   3.552 0.000382 ***
Temperatura  0.02405    0.02700   0.891 0.373117
```

Fuente: Elaboración propia

Para este rango de edad y, en particular, para la edad 40, se sigue viendo un coeficiente positivo, pero no significativo en la variable explicativa por su elevado p-valor.

Si se considera el siguiente, entre los 50 y 75 años, continúa sin ser relevante la variable temperatura, aunque se observa otra edad algo incorrelada con el resto de las edades cercanas que es la edad¹⁰ 67, algo no muy característico al ser algo extraordinario.

Sin embargo, sí es a partir de las últimas edades de este grupo, en particular para la edad 74 y 75 donde se empieza a observar como la significatividad de la variable explicativa, temperatura, empieza a ser algo usual, como se observa en la figura 5.4 y figura 5.5.

Figura 5.4
Resultados del modelo GLM para la edad 74

```
Resumen del modelo para Edad 74 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadM)

Deviance Residuals:
    445    446    447    448    449    450
 0.80795 -1.37238  1.98362 -0.07695  0.25383 -1.69184

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.212391    0.135700  38.411 < 2e-16 ***
Temperatura  0.016399    0.006169   2.658  0.00785 **
```

Fuente: Elaboración propia

¹⁰ Es la única de las pocas edades antes de los 70 significativa para la mujer, resultado distinto si lo comparamos con el de los hombres.

En este caso, se observa como la temperatura tiene un efecto de aumento en la mortalidad por su coeficiente positivo (0.016399) y a un nivel de significatividad del 1%, se tienen evidencias para poder rechazar la hipótesis nula (13). Por lo tanto, al rechazarla se concluye que el coeficiente es significativo.

Figura 5.5

Resultados del modelo GLM para la edad 75

```

Resumen del modelo para Edad 75 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
    data = datos_edadM)

Deviance Residuals:
    451     452     453     454     455     456
-0.02443  1.54170 -0.56098  1.72683 -1.01699 -1.76402

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  4.886114   0.141948   34.42 < 2e-16 ***
Temperatura  0.030494   0.006392    4.77 1.84e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Fuente: Elaboración propia

Similar a los resultados en la figura 5.4, en este caso con un mayor nivel de confianza al 0,1% de significatividad, se afirma en la figura 5.5 que para la edad 75, las altas temperaturas afectan a la mortalidad para el género femenino.

Se va a observar como a medida que aumenta la edad, también aumenta el grado de confianza para rechazar la hipótesis nula.

A partir de los 76 años para las mujeres, se sigue el patrón del rango anterior, donde la variable explicativa ya empieza a influir en la dependiente de manera muy habitual. A partir de la edad 80 en adelante, la temperatura va a ser significativa a un nivel de confianza, en la mayoría, muy alto hasta la 101, la última edad del estudio.

Aquí se va a representar algunas edades, donde se confirma una relación entre ambas variables cuando las mujeres cuentan con edades avanzadas.

Figura 5.6

Resultados del modelo GLM para la edad 79

Resumen del modelo para Edad 79 :

Call:

```
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadM)
```

Deviance Residuals:

475	476	477	478	479	480
1.34477	-0.54208	-2.17215	2.46945	0.01261	-1.22784

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	5.572946	0.111604	49.935	< 2e-16 ***
Temperatura	0.018218	0.005067	3.596	0.000324 ***

Fuente: Elaboración propia

Figura 5.7

Resultados del modelo GLM para la edad 80

Resumen del modelo para Edad 80 :

Call:

```
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadM)
```

Deviance Residuals:

481	482	483	484	485	486
0.9216	0.3051	-1.9156	2.0005	0.1854	-1.5921

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	5.572334	0.115190	48.38	< 2e-16 ***
Temperatura	0.014472	0.005243	2.76	0.00578 **

Fuente: Elaboración propia

Figura 5.8

Resultados del modelo GLM para la edad 83

Resumen del modelo para Edad 83 :

Call:

```
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadM)
```

Deviance Residuals:

499	500	501	502	503	504
0.5963	2.2172	-1.2416	3.8276	-1.7057	-3.9980

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	5.739338	0.102350	56.075	< 2e-16 ***
Temperatura	0.018620	0.004646	4.008	6.12e-05 ***

Fuente: Elaboración propia

Figura 5.9

Resultados del modelo GLM para la edad 90

Resumen del modelo para Edad 90 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
data = datos_edadM)

Deviance Residuals:
541 542 543 544 545 546
1.476 1.965 -1.692 4.968 -1.232 -5.849

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) 6.653086 0.068981 96.448 < 2e-16 ***
Temperatura 0.011164 0.003147 3.547 0.000389 ***

Fuente: Elaboración propia

Figura 5.10

Resultados del modelo GLM para la edad 95

Resumen del modelo para Edad 95 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
data = datos_edadM)

Deviance Residuals:
571 572 573 574 575 576
1.1784 1.5098 -0.5712 3.9159 -1.2144 -5.1245

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) 6.184981 0.085792 72.093 < 2e-16 ***
Temperatura 0.013075 0.003909 3.345 0.000823 ***

Fuente: Elaboración propia

Figura 5.11

Resultados del modelo GLM para la edad 101

Resumen del modelo para Edad 101 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
data = datos_edadM)

Deviance Residuals:
607 608 609 610 611 612
3.1190 -1.2687 -0.4113 5.2172 -0.9509 -6.2922

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) 5.62725 0.10287 54.701 < 2e-16 ***
Temperatura 0.02470 0.00465 5.312 1.08e-07 ***

Fuente: Elaboración propia

Para las figuras 5.6, 5.7, 5.8, 5.9, 5.10 y 5.11 los resultados representan lo mismo. Se identifica como para mujeres a partir de los 80 años, el coeficiente de temperatura es positivo y significativo en la mayoría a un nivel de confianza del 1%, lo que indica que se tiene mucha evidencia para rechazar la hipótesis que dice que el coeficiente de la temperatura no es significativo.

De acuerdo con los resultados que presentaban sobre el impacto del calor del verano del 2022 en la mortalidad de las mujeres en Ballester et al. (2023) y diversos estudios, se puede afirmar que es a partir de los 80 años, cuando las mujeres empiezan a experimentar los efectos del cambio climático y, con ello, el aumento de las temperaturas en los meses cálidos de nuestro país.

Las siguientes figuras van a representar el impacto de la temperatura en los hombres, siguiendo la misma metodología y como para el estudio de las mujeres, debido a la gran cantidad de modelos, se va a ir analizando por rango de edad.

De manera similar al análisis anterior, en la figura 5.12 no se observa ninguna relación significativa entre la temperatura y la mortalidad, aunque se podría resaltar que en los hombres de 22 años y 23 años se encuentra algo de relevancia de la variable independiente. Esto puede ser motivo del golpes de calor al realizar actividades al aire libre sin una buena hidratación y con una falsa sensación de invulnerabilidad.

Figura 5.12
Resultados del modelo GLM para la edad 23

```
Resumen del modelo para Edad 23 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)

Deviance Residuals:
    139     140     141     142     143     144
 0.8424 -0.8757  0.3094 -0.7163  1.2731 -1.0439

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.15259    0.87245   0.175  0.8612
Temperatura  0.09503    0.03790   2.507  0.0122 *
```

Fuente: Elaboración propia

Para el rango de edad entre 26 y 51 años, representando aquí el modelo como ejemplo para la edad 40, figura 5.13, (en la media del tramo), se percibe que tampoco la temperatura es significativa. Por el contrario, sí se observa como para la edad 39 y edad 48, aun con un nivel del 5%, la variable alta temperatura afecta al número de fallecidos a esas edades. Ya se empiezan a encontrar las primeras diferencias con las mujeres, ya que para edades más tempranas la temperatura empieza a provocar un impacto perjudicial.

Figura 5.13
Resultados del modelo GLM para la edad 40

```

Resumen del modelo para Edad 40 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)

Deviance Residuals:
    241    242    243    244    245    246
-1.4514  3.4821 -1.0383 -0.4561  0.4338 -1.6638

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.49220    0.49873   4.997 5.82e-07 ***
Temperatura  0.02338    0.02256   1.036  0.3

```

Fuente: Elaboración propia

Aunque sigue habiendo edades a las que los hombres no se ven afectados por la temperatura al tener un p-valor muy elevado (como se ve en esta figura 5.13), se observa en la siguiente figura 5.14 como para edades cercanas, sí podemos empezar a rechazar la hipótesis de no significatividad, a diferencia con los modelos de las mujeres, donde para estas edades, se podía seguir concluyendo que la temperatura no tenía efecto en la mortalidad.

Figura 5.14
Resultados del modelo GLM para la edad 39

```

Resumen del modelo para Edad 39 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log")
     data = datos_edadHOMBRES)

Deviance Residuals:
    235    236    237    238    239    240
-0.29514  0.45321 -1.40823  0.12517  0.01741  0.98796

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.07774    0.50011   4.155 3.26e-05 ***
Temperatura  0.04766    0.02228   2.139  0.0324 *

```

Fuente: Elaboración propia

Algo que resulta poco común en las mujeres, pero que empieza a observarse en los hombres de este rango de edad, como se ve en la figura 5.14, es el inicio de la afectación por las temperaturas a edades más tempranas.

Esta idea se sigue mostrando en la figura 5.15, donde la variable explicativa tiene también un coeficiente positivo y significativo en la variable mortalidad.

Figura 5.15
Resultados del modelo GLM para la edad 48

```

Resumen del modelo para Edad 48 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)

Deviance Residuals:
    289    290    291    292    293    294
-0.08444  0.34609 -0.24938 -0.04682  0.15628 -0.12719

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  3.38612    0.30448  11.121  <2e-16 ***
Temperatura  0.02893    0.01373   2.108  0.0351 *

```

Fuente: Elaboración propia

Comentando las edades entre los 50 y 75 años, se empiezan a ver las primeras diferencias con los modelos de mujeres. Para estas edades, en las mujeres hasta los 74 años no era normal ver un impacto significativo. Sin embargo, en los hombres para edades más tempranas, como es el rango que empieza en 50 años, ya existen varios modelos que muestran como un impacto en la temperatura afecta a la mortalidad.

Figura 5.16
Resultados del modelo GLM para la edad 58

```

Resumen del modelo para Edad 58 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)

Deviance Residuals:
    349    350    351    352    353    354
 0.14995  0.12152 -0.76351  1.18668 -0.77762  0.04898

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  4.507876    0.184657  24.412  < 2e-16 ***
Temperatura  0.021682    0.008364   2.592  0.00953 **

```

Fuente: Elaboración propia

A la edad 58 de los hombres, en la figura 5.16, a un nivel de confianza del 1%, se puede concluir que la temperatura afecta significativamente en su mortalidad, conclusión que no se podía tomar en los modelos de mujeres.

En las siguientes figuras se siguen mostrando modelos de hombres más jóvenes que las mujeres a los que les influye la temperatura.

Figura 5.17

Resultados del modelo GLM para la edad 61

Resumen del modelo para Edad 61 :

Call:
 glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
 data = datos_edadHOMBRES)

Deviance Residuals:
 367 368 369 370 371 372
 0.5606 0.2178 -1.1982 1.4747 -0.1469 -0.9661

Coefficients:
 Estimate Std. Error z value Pr(>|z|)
 (Intercept) 4.904420 0.160861 30.489 <2e-16 ***
 Temperatura 0.014472 0.007322 1.976 0.0481 *

Fuente: Elaboración propia

Figura 5.18

Resultados del modelo GLM para la edad 72

Resumen del modelo para Edad 72 :

Call:
 glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
 data = datos_edadHOMBRES)

Deviance Residuals:
 433 434 435 436 437 438
 0.91723 -0.01191 -0.83829 1.37621 0.44270 -1.95262

Coefficients:
 Estimate Std. Error z value Pr(>|z|)
 (Intercept) 5.583995 0.118280 47.210 <2e-16 ***
 Temperatura 0.010603 0.005399 1.964 0.0495 *

Fuente: Elaboración propia

Figura 5.19

Resultados del modelo GLM para la edad 74

Resumen del modelo para Edad 74 :

```
Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)
```

```
Deviance Residuals:
    445     446     447     448     449     450
 0.6969  0.6375 -1.1200  1.9625 -0.2214 -2.0448
```

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.561919   0.117050  47.517  <2e-16 ***
Temperatura  0.013177   0.005333   2.471   0.0135 *
```

Fuente: Elaboración propia

En las figuras 5.17, 5.18 y 5.19, se siguen demostrando las discrepancias entre hombres y mujeres, donde es para el género masculino algo relevante la significatividad de la temperatura, entre unos niveles del 5% y 1% de confianza, en la mortalidad, mientras que para las mujeres no existían evidencias para poder rechazar la hipótesis de no significatividad.

Finalmente, a partir de los 76 años para los hombres, se observan tanto diferencias como similitudes, en comparación con las edades de las mujeres.

La primera diferencia radica en que a partir de esta edad, es poco común encontrar edades en las que los hombres se vean afectados por la temperatura, a diferencia de las mujeres, que es a partir de esta, donde el impacto del calor empieza a ser influyente.

Aun así, hay similitudes; por ejemplo, para ciertas edades muy avanzadas, los hombres también les afecta las olas de calor extremas causado por el agravamiento de enfermedades cardíacas y de otras patologías.

En consecuencia, aquí se representan edades, en la figura 5.20 y figura 5.21 donde en las mujeres la temperatura afectaba en su mortalidad y para hombres no es significativa y las figuras 5.22 y 5.23 para las edades 89 y 101, al ser unas de las pocas significativas para este rango, y comunes con el otro género.

Figura 5.20

Resultados del modelo GLM para la edad 79

Resumen del modelo para Edad 79 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
data = datos_edadHOMBRES)

Deviance Residuals:
475 476 477 478 479 480
1.2233 0.2031 -0.7594 1.5254 1.0563 -3.3793

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) 6.100244 0.097332 62.674 <2e-16 ***
Temperatura 0.003019 0.004467 0.676 0.499

Fuente: Elaboración propia

Comparándolo con la figura 5.6, en esta el p-valor era menor al 0,01% haciendo así que se pudiera rechazar la hipótesis de no significatividad y concluir que para la edad 79 en las mujeres las altas temperaturas afectaban a la mortalidad. Por el contrario, en esta figura 5.20, el resultado es totalmente distinto, se percibe como el p-valor es muy elevado indicando que no hay evidencia suficiente para negar la hipótesis nula de falta de significatividad en los hombres para esa misma edad.

Figura 5.21

Resultados del modelo GLM para la edad 80

Resumen del modelo para Edad 80 :

Call:
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
data = datos_edadHOMBRES)

Deviance Residuals:
481 482 483 484 485 486
1.3096 -0.8255 0.1063 2.3984 -0.5583 -2.5399

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) 6.167256 0.093789 65.757 <2e-16 ***
Temperatura 0.003449 0.004303 0.802 0.423

Fuente: Elaboración propia

En esta figura 5.21, que representa la mortalidad para la edad 80 de los hombres, se puede estudiar en comparación con la figura 5.7 para las mujeres. Esta representaba como la mortalidad de la mujer para la edad 80 años estaba influenciada por las temperaturas, por el contrario, en la figura 5.21 no se observa el mismo efecto, y evidencia la falta de relevancia de la variable explicativa en la mortalidad del hombre al tener un pvalor tan elevado.

Figura 5.22*Resultados del modelo GLM para la edad 83*

Resumen del modelo para Edad 83 :

Call:

```
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)
```

Deviance Residuals:

499	500	501	502	503	504
0.7585	0.6153	-0.1550	3.9215	-2.5878	-2.7870

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	6.019261	0.098848	60.894	<2e-16 ***
Temperatura	0.006026	0.004526	1.331	0.183

Fuente: Elaboración propia

El efecto que se sigue evidenciando es el mismo. Cuando se cambia el género, aunque la edad sea la misma, el efecto de la temperatura cambia, y como se observa en esta figura 5.22, no hay evidencia para decir que el coeficiente es distinto de cero y, por ende, no son significativas las altas temperaturas en los hombres de 83 años.

Estas siguientes figuras representan dos edades de similitud entre hombres y mujeres, resultado razonable si se considera que en el 2022 las olas de calor fueron muy prolongadas y con temperaturas muy elevadas y estas afectan directamente a la salud, pudiendo así afectar también a hombres en situaciones vulnerables.

Figura 5.23*Resultados del modelo GLM para la edad 89*

Resumen del modelo para Edad 89 :

Call:

```
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)
```

Deviance Residuals:

535	536	537	538	539	540
0.9914	0.7853	-1.0622	2.9668	-0.7934	-3.0160

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	6.159985	0.083382	73.877	< 2e-16 ***
Temperatura	0.017980	0.003786	4.749	2.05e-06 ***

Fuente: Elaboración propia

Figura 5.24

Resultados del modelo GLM para la edad 101

Resumen del modelo para Edad 101 :

Call:

```
glm(formula = TasaMortalidad ~ Temperatura, family = poisson(link = "log"),
     data = datos_edadHOMBRES)
```

Deviance Residuals:

607	608	609	610	611	612
0.4248	1.7381	-0.4043	2.0546	-0.5541	-3.5811

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	4.177751	0.210810	19.818	< 2e-16 ***
Temperatura	0.025573	0.009524	2.685	0.00725 **

Fuente: Elaboración propia

En las figuras 5.23 y 5.24, sí hay evidencias suficientes, a un nivel de confianza del 0,1% y 1% respectivamente, para concluir que las altas temperatura sí afectan a los hombres de estas edades muy avanzadas.

Finalmente, todos estos resultados demuestran como a los hombres les influye menos el calor extremo, ya que hay menos edades en las que la temperatura es significativa y se observa como esta se concentra en edades más tempranas, si lo comparamos con las edades de las mujeres perjudicadas.

Aun así, es importante recordar como a ciertas edades muy avanzadas, los hombres también están afectados por estos efectos y como se concluyó con las mujeres, tanto para los planes de prevención como desde el punto de vista actuarial, es relevante el impacto de las olas de calor persistentes que está experimentando el país y la diferencia entre el género y la edad.

5.2 RESULTADOS PARA LA SEGUNDA PARTE METODOLÓGICA.

Gracias al primer estudio realizado se concluye que para las mujeres, las edades más afectadas por altas temperaturas son:

- Entre los 75 y 84 años.
- Igual o mayor de 85 años.

En el caso de los hombres, se observa el efecto de la temperatura en edades:

- Entre 45 y 64 años,
- Entre los 65 y 74 años.

Estos cuatro grupos de edades son para los que se va a tener los datos de la mortalidad diaria, teniendo así cuatro series temporales:

- Igual o mayor de 85 años, siendo la primera serie temporal de mortalidad (M85).
- Entre los 75 y 84 años, siendo la segunda serie temporal de mortalidad (M7584).
- Entre 45 y 64 años, siendo la tercera serie temporal de mortalidad (H4564).
- Entre los 65 y 74 años, siendo la cuarta serie temporal de mortalidad (H6574).

Estas series temporales de mortalidad diaria se toman para los meses de mayo a septiembre. Cada una de estas se combinará con la serie temporal temperatura máxima diaria para los mismos meses.

Finalmente, se obtendrán cuatro modelos VAR formado por cada serie temporal de mortalidad y temperatura diaria.

Modelo VAR 1 (mortalidad diaria M85 y temperatura máxima)

Modelo VAR 2 (mortalidad diaria M7584 y temperatura máxima)

Modelo VAR 3 (mortalidad diaria H4564 y temperatura máxima)

Modelo VAR 4 (mortalidad diaria M6574 y temperatura máxima)

5.2.1 Estacionariedad.

Siguiendo la metodología, la primera parte es comprobar la estacionariedad de las series temporales de mortalidad y temperatura máxima mediante los tres test comentados.

Figura 5.25

Resultados del test Dickey & Fuller para cada rango de mortalidad

	Adf. test
M+85	0.54800875
M7584	0.47526985
H4564	0.27199366
H6574	0.06065776

Fuente: Elaboración propia

Observando los resultados de la figura 5.25 y partiendo de que la hipótesis nula es la no estacionariedad, a excepción de la última, a un nivel de 10% de confianza, el resto de las series temporales, en principio, son no estacionarias, ya que no hay suficiente evidencia para rechazar la hipótesis nula de no estacionariedad.

Figura 5.26

Resultados del test Phillips & Perron para cada rango de mortalidad

	pp. test
M+85	0.1543322
M7584	0.0100000
H4564	0.0100000
H6574	0.0100000

Fuente: Elaboración propia

Usando el procedimiento PP, es notable una diferencia en las tres últimas series de mortalidad, donde sí se tienen evidencias para rechazar la hipótesis nula y, por ende, asumir que son estacionarias.

Como se comentó en la parte metodológica, el test de Phillips & Perron es una mejora del test de Dickey & Fuller en términos de autocorrelación y heterocedasticidad, y es por ello que se toman como definitivos los resultados del test PP que determina que las últimas tres series son estacionarias.

Finalmente, se comprueba con el último test, el único que no es de raíz unitaria, para comprobar que los resultados son acordes con la decisión tomada.

Figura 5.27

Resultados del test KPSS para cada rango de mortalidad

	kpss.test
M+85	0.01803623
M7584	0.02379010
H4564	0.04403755
H6574	0.04946781

Fuente: Elaboración propia

Partiendo de la hipótesis nula: la serie es estacionaria, se confirma que no hay suficientes evidencias para rechazar esa hipótesis en ninguna de las cuatro series, y por lo tanto, se dice que las series son estacionarias.

Las hipótesis nulas de cada test para tener en cuenta son,

$$\begin{array}{l} \text{Phillips \& Perron} \\ H_0 \text{ No estacionariedad} \end{array} \quad (14)$$

$$\begin{array}{l} \text{KPSS} \\ H_0 \text{ Estacionariedad} \end{array} \quad (15)$$

Para las series temporales de mortalidad en mujeres de 75 a 84 años, hombres de 45 a 64 años y hombres de 65 a 74 años, ambos test arrojan el mismo resultado.

Por la parte del test PP, hay evidencias para rechazar la hipótesis nula por lo que las series son estacionarias.

Por la parte del test KPSS, no hay evidencias para rechazar la hipótesis nula por lo que las series son estacionarias.

Las hipótesis nula de cada test son opuestas, por lo que si se rechaza la hipótesis nula en una prueba y en la otra no, se demuestra que ambas pruebas coinciden en su conclusión.

Con la mortalidad de mujeres iguales o mayores de 85 años, PP determina que no se puede rechazar la hipótesis nula y KPSS también indica que no hay evidencias para rechazarla. Ambos coinciden en no poder rechazar su hipótesis nula. Como se ha comentado antes, si las hipótesis nulas son opuestas, y en ninguna de las dos pruebas se rechaza la hipótesis nula, los resultados obtenidos son distintos para cada prueba, ya que

el primero indica que la serie temporal de mortalidad M85 es no estacionaria y el segundo test indica que sí es estacionaria.

Finalmente, se van a tomar los resultados del test de Phillips & Perron, ya que este es robusto a la heterocedasticidad y autocorrelación además de no ser tan sensible como el test KPSS a la elección del futuro modelo.

Se concluiría así que la mortalidad de mujeres igual o mayores de 85 años es no estacionaria, y, por consiguiente, se debe tomar diferencias solo en esta serie, con el fin, de convertirla en estacionaria.

Por tanto, la única serie temporal en la que se deben tomar diferencias es en la mortalidad de mujeres igual o mayores de 85 años, donde aplicado *ndiffs*, este marca el número de diferencias necesarias para que la serie sea estacionaria. En este caso 1 diferencia como marca la figura 5.28.

Figura 5.28
Resultado de las diferencias óptimas para la mortalidad M85

	Serie	Dif_Optimas
1	Mortalidad M85	1

Fuente: Elaboración propia

Tras tomar una diferencia en esta serie¹¹, se repiten los tres test de nuevo para verificar que se cumple el objetivo y se estudia su pvalor en la figura 5.29.

Figura 5.29
Resultados de los test para la mortalidad M85 con una diferencia

	Prueba	P_Value
1	ADF	0.01
2	PP	0.01
3	KPSS	0.10

Fuente: Elaboración propia

¹¹ Código de la toma de una diferencia en el Anexo.

Los resultados de la figura 5.29 muestran como una vez se toma una diferencia en esa serie temporal, los tres test establecen estacionariedad en la misma.

Tanto los test ADF y PP al ser la hipótesis nula la no estacionariedad, y tener evidencias para rechazarla, por tanto indican que la serie es estacionaria. Como para el test KPSS al ser la hipótesis nula la estacionariedad, y no tener evidencias para rechazarla, por tanto indica que la serie es estacionaria. Todas concluyen lo mismo.

Siguiendo el mismo método para la serie temporal de temperatura máxima diaria, los resultados de los tres test utilizados se muestran en la figura 5.30.

Figura 5.30

Resultados de los test para la serie de temperatura máxima

	Prueba	P_Value
1	KPSS	0.03932086
2	ADF	0.40746836
3	PP	0.05247631

Fuente: Elaboración propia

Aunque de Dickey – Fuller (ADF) se obtenga que es no estacionaria, los test de Phillips & Perron y KPSS asumen estacionariedad en la serie temporal, y como ambos test corrigen las limitaciones del primero, se considera esta idea.

Por tanto, de las cinco series temporales que se tienen (cuatro de mortalidad y una de temperatura), se ha sido necesario tomar diferencias solo en la mortalidad de mujeres mayores a 85 años y, una vez tomadas, se obtiene que todas las series son estacionarias.

5.2.2 Modelo VAR.

Una vez se tienen las series de mortalidad y temperatura máxima con la condición de estacionariedad, se comprueba qué orden para cada modelo VAR (M85-TempMáxima, M7485-TempMáxima, H4564-TempMáxima, H6574-TempMáxima) es el que mejor se ajusta a los datos.

Se ha realizado para cada combinación, mortalidad_M85_diff con TempMaxima, mortalidadM_7584 con TempMaxima, mortalidadH_4564 con TempMaxima y mortalidadH_6574 con TempMaxima, un procedimiento para seleccionar con cuántos retardos el modelo VAR es mejor, siguiendo el criterio de información AIC y BIC.

Figura 5.31
Resultados criterios información M85

	Lag	AIC	BIC
1	1	2163.967	2182.070
2	2	2147.519	2177.625
3	3	2147.519	2177.625
4	4	2163.967	2182.070
5	5	2163.967	2182.070

Fuente: Elaboración propia

Figura 5.32
Resultados criterio información M7584

	Lag	AIC	BIC
1	1	1951.365	1969.509
2	2	1938.607	1968.780
3	3	1919.894	1962.043
4	4	1919.894	1962.043
5	5	1919.894	1962.043

Fuente: Elaboración propia

Figura 5.33
Resultados criterios información H4564

	Lag	AIC	BIC
1	1	1885.420	1903.563
2	2	1870.022	1900.195
3	3	1870.022	1900.195
4	4	1870.022	1900.195
5	5	1870.022	1900.195

Fuente: Elaboración propia

Figura 5.34
Resultados criterios información H6574

	Lag	AIC	BIC
1	1	1897.526	1915.67
2	2	1897.526	1915.67
3	3	1897.526	1915.67
4	4	1897.526	1915.67
5	5	1897.526	1915.67

Fuente: Elaboración propia

En la mayoría de las figuras, los criterios de información establecen el coger dos retardos (lags) de las variables, pero cuando luego se toma el modelo con esos dos retardos, solo es el primer retardo de la temperatura el significativo. Por ello, aunque en algunos modelos los criterios mejoren (se minimice el resultado) con el segundo retardo, se opta por solo añadir retardos significativos y que realmente influyan en la variable de interés, mortalidad, y, por tanto, se toma solo un retardo en cada modelo VAR.

Un ejemplo de lo ahora comentado; para los hombres entre 45 y 64 años tanto el AIC como BIC establecen que el mejor modelo VAR es el que contiene dos retardos pero la figura 5.35 representa el siguiente resultado.

Figura 5.35
Resultados del modelo VAR H4564 con dos retardos

```

Estimation results for equation mortalidadH_4564:
=====
mortalidadH_4564 = mortalidadH_4564.l1 + TempMaxima.l1 + mortalidadH_4564.l2 + TempMaxima.l2
+ const

      Estimate Std. Error t value Pr(>|t|)
mortalidadH_4564.l1  0.16092   0.08031   2.004 0.046948 *
TempMaxima.l1       0.33648   0.38687   0.870 0.385866
mortalidadH_4564.l2 0.24689   0.07999   3.087 0.002422 **
TempMaxima.l2       0.17003   0.38850   0.438 0.662279
const               38.20281   9.77031   3.910 0.000141 ***

```

Fuente: Elaboración propia

Se observa como al meter dos retardos de ambas variables (TempMaxima.l1, TempMaxima.l2, mortalidadH_4564.l1, mortalidadH_4564.l2) ni el primer retardo de temperatura máxima ni el segundo son significativos al tener un p-valor superior a cualquier nivel de confianza, y, por consiguiente, ninguno causa efecto en la mortalidad.

Al contrario, como se observa la figura 5.36, si se añade solo un retardo de ambas variables (TempMaxima.l1, mortalidadH_4564.l1) ambos son significativos para predecir la mortalidad (el pvalor < al nivel de significancia).

Figura 5.36

Resultados del modelo VAR H4564 con un retardo

```
Estimation results for equation mortalidadH_4564:
=====
mortalidadH_4564 = mortalidadH_4564.l1 + TempMaxima.l1 + const

      Estimate Std. Error t value Pr(>|t|)
mortalidadH_4564.l1  0.22462   0.07978   2.815  0.00553 **
TempMaxima.l1       0.57234   0.20164   2.838  0.00517 **
const               53.27530   8.79033   6.061  1.06e-08 ***
```

Fuente: Elaboración propia

Como el objetivo del trabajo es demostrar cómo la temperatura máxima afecta a la mortalidad, se opta solo por usar retardos significativos.

Finalmente se obtiene que los cuatro modelos siguen un VAR de orden 1, VAR (1), al ser solo el primer retardo significativo.

En las siguientes figuras se va a ilustrar¹² como en todos los modelos VAR de orden 1, el retardo de la temperatura es significativo para el valor de la mortalidad y por tanto, se sigue afirmando la hipótesis de que para estas edades conocer el valor pasado de la temperatura máxima ayudará a predecir la mortalidad futura.

Figura 5.37

Resultados del modelo VAR M85 con un retardo

```
Estimation results for equation mortalidad_M85_diff:
=====
mortalidad_M85_diff = mortalidad_M85_diff.l1 + TempMaxima.l1 + const

      Estimate Std. Error t value Pr(>|t|)
mortalidad_M85_diff.l1 -0.35957   0.07687  -4.677  6.5e-06 ***
TempMaxima.l1         1.05732   0.51819   2.040  0.0431 *
const                 -38.02619  18.55150  -2.050  0.0422 *
```

Fuente: Elaboración propia

¹² El modelo VAR para H4564 es el ilustrado en la figura 5.36 de la página anterior.

Figura 5.38

Resultados del modelo VAR M7584 con un retardo

Estimation results for equation mortalidadM_7584:
=====

$$\text{mortalidadM_7584} = \text{mortalidadM_7584.l1} + \text{TempMaxima.l1} + \text{const}$$

	Estimate	Std. Error	t value	Pr(> t)	
mortalidadM_7584.l1	0.51202	0.06829	7.498	5.38e-12	***
TempMaxima.l1	0.92483	0.25933	3.566	0.000487	***
const	31.85802	10.16407	3.134	0.002075	**

Fuente: Elaboración propia

Figura 5.39

Resultados del modelo VAR H6574 con un retardo

Estimation results for equation mortalidadH_6574:
=====

$$\text{mortalidadH_6574} = \text{mortalidadH_6574.l1} + \text{TempMaxima.l1} + \text{const}$$

	Estimate	Std. Error	t value	Pr(> t)	
mortalidadH_6574.l1	0.29239	0.07824	3.737	0.000265	***
TempMaxima.l1	0.50199	0.20769	2.417	0.016859	*
const	61.17315	9.88924	6.186	5.67e-09	***

Fuente: Elaboración propia

5.2.3 Causalidad de Granger

La presentación de las figuras 5.36, 5.37, 5.38 y 5.39 ya da a intuir como sí existe una relación positiva y significativa para la predicción de la mortalidad en función de la temperatura máxima, pero para comprobarlo, se estudia la causalidad de Granger.

Como se comentaba en la parte metodológica, si se demuestra que las temperaturas extremas causan la mortalidad, se puede decir que las altas temperaturas ayudan a predecir los valores futuros de la mortalidad para ciertas edades en función del género.

Para las mujeres de edades iguales o superiores a 85 años, se observaba en la figura 5.37 como el retardo de temperatura máxima era significativo en la mortalidad. El resultado del test de causalidad muestra lo siguiente en la figura 5.40.

Figura 5.40*Resultados de causalidad de Granger para M85*

Granger causality test

Model 1: mortalidad_M85_diff ~ Lags(mortalidad_M85_diff, 1:1) + Lags(TempMaxima, 1:1)

Model 2: mortalidad_M85_diff ~ Lags(mortalidad_M85_diff, 1:1)

	Res.Df	Df	F	Pr(>F)
1	148			
2	149	-1	4.1633	0.04309 *
3	150	-1	19.7166	1.739e-05 ***

Fuente: Elaboración propia

En esta figura se confirma como a un nivel de confianza del 5% se tienen evidencias para decir que las altas temperaturas causan la mortalidad.

Esto viene a raíz de que la hipótesis nula del test es la no causalidad y, por tanto, si el pvalor es menor que el nivel de confianza, se tienen indicios para rechazar esta hipótesis y, en consecuencia, decir que sí que existe causalidad.

Para las mujeres de edades de 75 a 84 años, se esperan que las conclusiones de la figura 5.41 sean similares a los que se obtuvieron en la figura 5.38, la cual evidenciaba la existencia de una relación positiva y significativa entre la temperatura pasada y la mortalidad.

Figura 5.41*Resultados de causalidad de Granger para M7584*

Granger causality test

Model 1: mortalidadM_7584 ~ Lags(mortalidadM_7584, 1:1) + Lags(TempMaxima, 1:1)

Model 2: mortalidadM_7584 ~ Lags(mortalidadM_7584, 1:1)

	Res.Df	Df	F	Pr(>F)
1	149			
2	150	-1	12.718	0.0004873 ***
3	151	-1	85.993	< 2.2e-16 ***

Fuente: Elaboración propia

A un nivel de confianza del 0,1%, se tiene evidencia suficiente para afirmar que las altas temperaturas causan un aumento en la mortalidad para este rango de edad y, por tanto, que pueden ser utilizadas para estimar el número futuro de mujeres fallecidas para edades de 75 a 84 años.

Las conclusiones para los hombres son bastante similares a las de las mujeres para ambos rangos de edad, como se observa en las figuras 5.42 y 5.43.

La figura 5.36 afirmaba que el valor pasado de las temperaturas máximas era significativo en la mortalidad de los hombres de edades entre 45 y 64 años.

Siguiendo el test de causalidad, se puede afirmar con la figura 5.42, que para este rango de edad a un nivel de significatividad del 1%, el aumento de las temperaturas máximas en España afecta a la mortalidad. Por tanto, su información es útil para predecir.

Figura 5.42

Resultados de causalidad de Granger para H4564

Granger causality test

Model 1: mortalidadH_4564 ~ Lags(mortalidadH_4564, 1:1) + Lags(TempMaxima, 1:1)

Model 2: mortalidadH_4564 ~ Lags(mortalidadH_4564, 1:1)

	Res.Df	Df	F	Pr(>F)
1	149			
2	150	-1	8.0566	0.0051667 **
3	151	-1	13.9245	0.0002695 ***

Fuente: Elaboración propia

Finalmente, la figura 5.38 revelaba el impacto significativo de la temperatura máxima en la mortalidad de los hombres de 64 a 75 años, ya que el retardo de alta temperatura servía para predecir la mortalidad. La figura 5.43 va en consonancia con estos resultados.

Figura 5.43

Resultados de causalidad de Granger para H6574

Granger causality test

Model 1: mortalidadH_6574 ~ Lags(mortalidadH_6574, 1:1) + Lags(TempMaxima, 1:1)

Model 2: mortalidadH_6574 ~ Lags(mortalidadH_6574, 1:1)

	Res.Df	Df	F	Pr(>F)
1	149			
2	150	-1	5.8418	0.01686 *
3	151	-1	19.8158	1.655e-05 ***

Fuente: Elaboración propia

A un nivel de confianza del 5%, se afirma que la variable de temperatura máxima causa la mortalidad para los hombres en este rango de edad y, por lo tanto, es útil tomar la variable climática para conocer el futuro del número de fallecidos.

Una vez estudiados todos los modelos y confirmar que se cumple el test de causalidad de Granger entre cada serie temporal de mortalidad y temperatura máxima, hay evidencia suficiente para afirmar que la temperatura sí es una variable relevante para ayudar a predecir la mortalidad futura en el año 2022.

5.2.4 Función Impulso-Respuesta.

Una vez comprobada la causalidad de Granger, afirmando que las altas temperaturas sí que causan un impacto en la mortalidad, se realiza la función impulso respuesta de la mortalidad ante un shock en la temperatura.

Para ello se aplica *irf* en cada modelo, dando como *response* a la mortalidad, para estudiar en ocho retardos el impacto en la variable y el tiempo que tarda en volver a su valor habitual.

Las figuras 5.44, 5.45, 5.46 y 5.47 reflejan cómo la mortalidad responde a un shock inesperado en la temperatura máxima.

El eje de abscisas representa el momento del choque (el periodo) y en el eje de ordenadas el impacto en la mortalidad.

Figura 5.44

Resultados de función impulso respuesta M85

\$TempMaxima	
mortalidad_M85_diff	
[1,]	0.0000000
[2,]	2.5575679
[3,]	1.2830002
[4,]	1.4126804
[5,]	1.0944695
[6,]	0.9738346
[7,]	0.8176316
[8,]	0.7029975
[9,]	0.5985210

Fuente: Elaboración propia

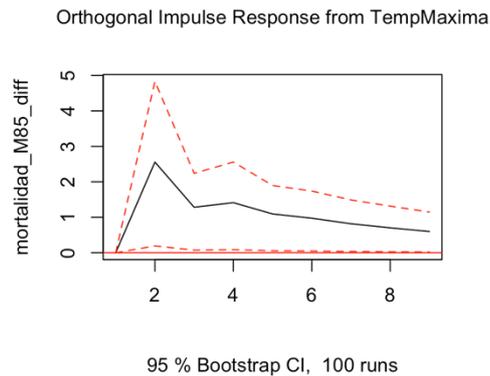


Figura 5.45

Resultados de función impulso respuesta M7584

\$TempMaxima	
mortalidadM_7584	
[1,]	0.0000000
[2,]	1.0189533
[3,]	1.3921057
[4,]	1.3768918
[5,]	1.2649733
[6,]	1.1103550
[7,]	0.9271257
[8,]	0.7424215
[9,]	0.5936535

Fuente: Elaboración propia

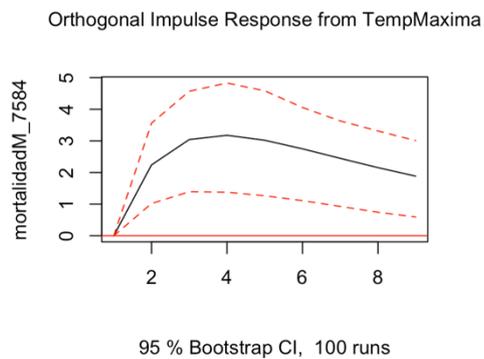
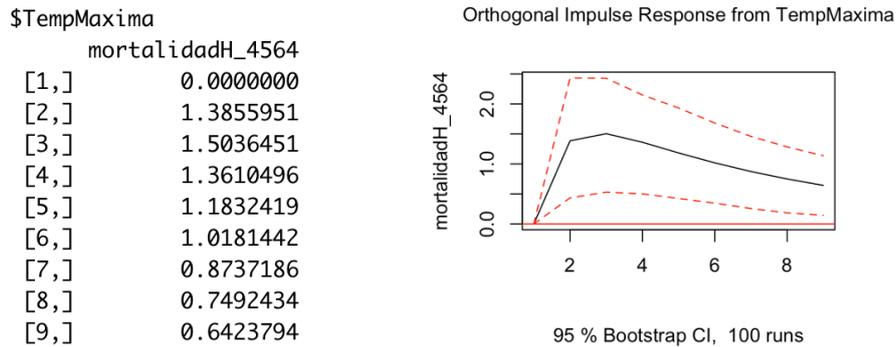
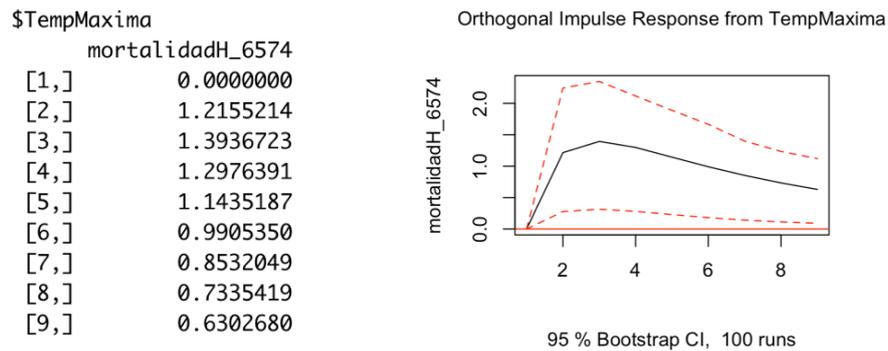


Figura 5.46
Resultados de función impulso respuesta H4564



Fuente: Elaboración propia

Figura 5.47
Resultados de función impulso respuesta H6574



Fuente: Elaboración propia

En primer lugar, observando los resultados en estas figuras se puede evidenciar una respuesta positiva y directa de la alta temperatura sobre la mortalidad, entendiéndose una correlación positiva entre las mismas.

Por ejemplo, en el caso de mujeres con edades superiores a 85 años, la figura 5.44, en el periodo 0 [1,] no se produce ningún cambio en la mortalidad ya que es el momento del choque, por lo que su efecto todavía no se ve reflejado. Será a partir del periodo 1 [2,] donde se evidencie la innovación en la temperatura sobre la mortalidad. Se observa como a medida que pasan los periodos, el impacto va disminuyendo progresivamente, como era de esperar, hasta tender al valor que tenía antes de este suceso.

Gráficamente se demuestra que en el periodo uno, un impacto de la temperatura máxima ha supuesto un aumento en tres unidades aproximadamente en la mortalidad de las mujeres de igual o mayor edad a 85 años, como respuesta a ese choque inesperado.

En cuanto al intervalo de confianza al 95%, se ve como los primeros periodos son amplios debido a una mayor incertidumbre en la estimación y cómo a medida que el valor tiende a la situación habitual, la diferencia es menor y más ajustada.

Esto proporciona indicios de cómo el shock de la temperatura máxima afecta de manera significativa a la mortalidad, ya que hay una persistencia de respuesta durante varios periodos aunque este aumento va decreciendo.

Para el resto de los rangos se experimenta una situación similar, un shock inesperado en la temperatura máxima diaria, ha supuesto un aumento por encima de lo esperado en la mortalidad durante unos periodos, hasta que finalmente tiende a su valor habitual, por tanto se evidencia una respuesta positiva y significativa.

Aplicando esto a seguros de riesgo donde se cubra la mortalidad los resultados arrojan conclusiones interesantes. Existen evidencias de que las personas puedan fallecer por encima de lo esperado con anterioridad al estar en un momento de cambio climático, donde las olas de calor y altas temperaturas cada vez son más usuales y, por ende, se deberá provisionar más capital con el fin de poder cumplir con esta situación “de estrés”. Un aumento de la mejor estimación del pasivo supondría a su vez una disminución de los fondos propios por ese aumento en las reservas (mayor SCR) para cubrir ese incremento en la mortalidad, por ahora, inesperado.

5.2.5 Descomposición de la varianza.

Conocido el impacto que tienen las temperaturas máximas en la mortalidad de distintas personas en función de su edad y género, también es relevante conocer qué cantidad de cambios en la mortalidad vienen por este efecto climático y saber si depende de más variables. Los planes de actuación frente a las olas de calor para obtener mejoras en la salud están muy interesados en esto, ya que si se conoce que uno de los motivos de los cambios en la mortalidad viene por las altas temperaturas, se puede solucionar el problema de una manera más eficiente.

Figura 5.48*Resultados de la descomposición de la varianza ante un shock en la temperatura para M85*

	mortalidad_M85_diff		TempMaxima
[1,]	1.0000000	[1,]	0.000000000
[2,]	0.9939055	[2,]	0.006094546
[3,]	0.9924770	[3,]	0.007522972
[4,]	0.9906799	[4,]	0.009320130
[5,]	0.9895960	[5,]	0.010403976
[6,]	0.9887396	[6,]	0.011260354
[7,]	0.9881363	[7,]	0.011863664
[8,]	0.9876909	[8,]	0.012309078
[9,]	0.9873683	[9,]	0.012631726
[10,]	0.9871326	[10,]	0.012867355
[11,]	0.9869611	[11,]	0.013038898
[12,]	0.9868360	[12,]	0.013163968
[13,]	0.9867449	[13,]	0.013255112
[14,]	0.9866784	[14,]	0.013321552
[15,]	0.9866300	[15,]	0.013369982

Fuente: Elaboración propia

Se muestran solo 15 de los 50 periodos al ser algo representativo y, aunque se conozca que las altas temperaturas afectan a la mortalidad significativamente, para las mujeres de edad igual o superior a 85 años, los resultados muestran como la temperatura máxima afecta en la mortalidad y este movimiento va aumentando. Sin embargo, es notable que el impacto viene en mayor parte por otros shocks en la propia mortalidad y como esta se ve afectada por estos a lo largo del tiempo.

En adelante se van a ver el resto de los resultados de 15 periodos en adelante para los otros tres modelos, donde el efecto es muy similar al de la figura 5.48.

Figura 5.49*Resultados de la descomposición de la varianza ante un shock en la temperatura para M7584*

	mortalidadM_7584	TempMaxima
[1,]	1.0000000	0.00000000
[2,]	0.9815072	0.01849276
[3,]	0.9518446	0.04815544
[4,]	0.9217386	0.07826139
[5,]	0.8960659	0.10393409
[6,]	0.8757973	0.12420271
[7,]	0.8603388	0.13966116
[8,]	0.8487236	0.15127644
[9,]	0.8400458	0.15995419
[10,]	0.8335723	0.16642772
[11,]	0.8287413	0.17125874
[12,]	0.8251320	0.17486795
[13,]	0.8224322	0.17756778
[14,]	0.8204102	0.17958977
[15,]	0.8188943	0.18110568

Fuente: Elaboración propia

A medida que pasan los periodos en la figura 5.49 el impacto de las altas temperaturas en la mortalidad aumenta, hasta que explica en el periodo 15 un 20% de los cambios que se experimentan en el número de defunciones de las mujeres entre 75 y 84 años.

Figura 5.50

Resultados de la descomposición de la varianza ante un shock en la temperatura para H4564

	mortalidadH_4564	TempMaxima
[1,]	1.0000000	0.0000000
[2,]	0.9868998	0.01310020
[3,]	0.9719938	0.02800616
[4,]	0.9600915	0.03990850
[5,]	0.9512843	0.04871567
[6,]	0.9448665	0.05513352
[7,]	0.9401954	0.05980462
[8,]	0.9367898	0.06321017
[9,]	0.9343022	0.06569781
[10,]	0.9324822	0.06751782
[11,]	0.9311490	0.06885099
[12,]	0.9301716	0.06982841
[13,]	0.9294545	0.07054549
[14,]	0.9289282	0.07107182
[15,]	0.9285417	0.07145829

Fuente: Elaboración propia

El impacto de la temperatura en los cambios de la mortalidad para hombres de edades de 45 a 64 en la figura 5.50 es mucho menor que el que se explica para mujeres en las figuras 5.48 y 5.49 para las distintas edades.

Esta diferencia de resultados muestra que es mucho mayor el porcentaje explicado por altas temperaturas en las mujeres que en los hombres, en consonancia tanto con lo observado en este trabajo como en otros estudios publicados.

Figura 5.51

Resultados de la descomposición de la varianza ante un shock en la temperatura para H6574

	mortalidadH_6574	TempMaxima
[1,]	1.0000000	0.0000000
[2,]	0.9909491	0.00905088
[3,]	0.9794688	0.02053123
[4,]	0.9697019	0.03029808
[5,]	0.9622431	0.03775686
[6,]	0.9567203	0.04327973
[7,]	0.9526635	0.04733655
[8,]	0.9496870	0.05031301
[9,]	0.9475017	0.05249833
[10,]	0.9458956	0.05410444
[11,]	0.9447141	0.05528594
[12,]	0.9438443	0.05615573
[13,]	0.9432036	0.05679641
[14,]	0.9427315	0.05726853
[15,]	0.9423835	0.05761654

Fuente: Elaboración propia

Comparando la figura 5.51 con la figura 5.50, ambas para hombres, hay un mayor cambio justificado por las temperaturas extremas en la mortalidad, al ser edades más avanzadas y, por tanto, más susceptibles al calor extremo. No obstante, sigue siendo mucho menor que el que se proporcionaba en la mortalidad de las mujeres.

En consecuencia, el estudio de descomponer la varianza, una vez se ha comprobado que existe causalidad de las altas temperaturas diarias en el número de fallecidos diarios, muestra que esta variable climática impacta más y explica un mayor porcentaje en la mortalidad de las mujeres que en los hombres, como otros artículos ya mencionados concluían.

Además, se necesitan tener en cuenta otros componentes que también tienen un impacto significativo en la mortalidad como lugar de residencia, salud, poder adquisitivo... aunque el fin de este trabajo es solo demostrar que las altas temperaturas son una variable también relevante.

5.2.6 Residuos del modelo.

Para cada uno de los cuatro modelos VAR del trabajo se van a obtener los residuos del modelo, comprobando en los siguientes apartados si cumplen con los requisitos suficientes comentados en la parte metodológica para aceptar los modelos y, finalmente, realizar la predicción.

5.2.6.1. Normalidad de los residuos.

Con los test Cramer-von Mises y Anderson-Darling, que miden la distancia entre la distribución estimada y la teórica, se obtienen los siguientes resultados para los cuatro modelos, M85, M7584, H4564 y H6574.

Figura 5.52

Resultados de test para comprobar la normalidad en los residuos

	Grupo	CVM_Statistic	CVM_p_value	AD_Statistic	AD_p_value
1	85	0.6480898	0.2398075	3.438230	0.2511192
2	7584	0.4822151	0.5265529	2.609291	0.5353357
3	4564	0.4595673	0.5766598	3.425623	0.2542674
4	6574	0.4881818	0.5135551	3.183336	0.3219658

Fuente: Elaboración propia

Partiendo de la hipótesis nula que dicta que la distribución es una Normal, se tienen evidencias suficientes en ambos test para no rechazar la hipótesis de normalidad al tener un pvalor mayor que cualquier nivel de significancia en todos los grupos. Por tanto, se puede afirmar que los residuos de los cuatro modelos cumplen con la condición de normalidad.

5.2.6.2. Residuos incorrelados.

Tomando como hipótesis nula la independencia de los residuos, la siguiente figura mostrará si hay evidencias para aceptar esta condición o si, por el contrario, se muestra que el proceso no es ruido blanco y hay dependencia.

Figura 5.53
Resultados de test para comprobar la independencia en los residuos

	Grupo	Variable	Statistic	p_value
1	85	Primera	7.156035	0.710636750
2	85	Segunda	7.156035	0.710636750
3	7584	Primera	26.436288	0.003195661
4	7584	Segunda	26.436288	0.003195661
5	4564	Primera	19.320979	0.036370153
6	4564	Segunda	19.320979	0.036370153
7	6574	Primera	14.019474	0.172105170
8	6574	Segunda	14.019474	0.172105170

Fuente: Elaboración propia

A la vista de los resultados de la figura 5.53, para todos los modelos, y para cada retardo (de mortalidad y temperatura, primera y segunda variable respectivamente) no hay suficiente evidencia para rechazar la hipótesis nula (independencia), al tener todos un pvalor por encima de los niveles de significatividad. Por consiguiente, se acepta que los residuos no tienen autocorrelación y son independientes.

5.2.6.3. Homocedasticidad.

Para grupo de edad y género se generan dos modelos de regresión lineal, donde la diferencia entre uno y otra será la variable dependiente. En primer lugar, serán los residuos de la mortalidad al cuadrado y para el otro, los residuos de la temperatura

máxima al cuadrado, los dos con las mismas variables explicativas (los valores pasados de ambas).

Una vez se tienen los modelos de regresión de los residuos, con el test BP se verifica la condición, teniendo este como hipótesis nula que los residuos son homocedásticos.

Figura 5.54

Resultados de test para comprobar la homocedasticidad en los residuos tomando los residuos de la mortalidad.

	Modelo	p_value
1	M85	0.1777944
2	M7584	0.4236561
3	H4564	0.6777691
4	H6574	0.9702936

Fuente: Elaboración propia

Partiendo del modelo $\varepsilon^2_{mortalidad} = \gamma_{t-1}^{mortalidad} + \gamma_{t-1}^{temperatura}$, la figura 5.54 representa como no hay evidencias suficientes para rechazar la hipótesis nula (homocedasticidad), ya que el p-valor de todos los modelos es superior a todos los niveles de significatividad y, por tanto, se asume que los errores de la mortalidad son homocedásticos.

Figura 5.55

Resultados de test para comprobar la homocedasticidad en los residuos tomando los residuos de la temperatura máxima.

	Modelo	p_value
1	M85	0.197108287
2	M7584	0.423656053
3	H4564	0.001248519
4	H6574	0.661746007

Fuente: Elaboración propia

En este caso, el modelo de regresión del error es $\varepsilon^2_{temperatura} = \gamma_{t-1}^{mortalidad} + \gamma_{t-1}^{temperatura}$, la figura 5.55 representa que no hay evidencias suficientes para rechazar la hipótesis nula (homocedasticidad), ya que el p-valor de nuevo es superior a todos los niveles de significatividad. Se asume que los errores de la temperatura para todos los modelos son homocedásticos.

5.2.6.4. Media centrada en cero.

Para verificar que las predicciones van a ajustarse a los valores observados se comprueba que la media del error de los cuatro modelos VAR es nula.

Figura 5.56

Resultados de test para conocer la media de los errores

	Modelo	media
1	M85	-2.211772e-16
2	M7584	-1.270628e-16
3	H4564	4.108056e-17
4	H6574	-2.449805e-16

Fuente: Elaboración propia

La media del error en los cuatro modelos de combinación de series temporales es cero como muestra la figura 5.56, asegurando así predicciones sin sesgo y confirmando que los residuos se comportan como ruido blanco.

5.2.7 Predicción de mortalidad.

Una vez comprobado que las series temporales son estacionarias, que se ha tomado un modelo VAR con un número óptimo de retardos en función de las características de la muestra, que los residuos de este cumplen con cuatro condiciones fundamentales y que una de las series temporales (máxima temperatura) causa unidireccionalmente a la otra serie temporal (mortalidad), se puede realizar la predicción de la variable de interés con la información que proporciona la otra variable.

El modelo predictivo para cada modelo VAR permite obtener la mortalidad diaria futura para días cercanos de mujeres y hombres en función de los rangos de edad que utiliza MoMo (“*Sistema de Monitorización de la Mortalidad Diaria*”).

Los resultados de las figuras usan la serie temporal de 153 observaciones (datos de mayo a septiembre) de mortalidad para predecir la mortalidad futura, pero teniendo en cuenta las altas temperaturas diarias. Se va a poder estudiar la tendencia que está prediciendo el modelo y cómo la mortalidad va a crecer en comparación con los últimos datos conocidos, a medida que se tienen esos registros de aumento en la temperatura.

La figura 5.57 muestra los datos de mortalidad conocidos de las mujeres de edad igual o superior a 85 años sin tomar diferencias y la figura 5.58 muestra los datos de mortalidad conocidos de las mujeres de edad igual o superior a 85 años tomando una diferencia.

Figura 5.57

Número de defunciones diarias las mujeres a partir de 85 años en 2022

```
Time Series:
Start = 1
End = 153
Frequency = 1
[1] 341 377 353 360 391 360 343 363 393 394 396 397 391 401 409 374
[17] 399 424 396 423 360 370 363 324 330 347 355 326 314 335 352 377
[33] 339 358 304 293 360 388 358 358 347 333 370 407 385 442 457 470
[49] 414 380 376 418 363 340 346 332 387 318 390 380 342 364 344 363
[65] 383 390 390 375 375 383 404 478 469 510 553 543 548 589 579 502
[81] 533 516 494 454 502 429 451 465 463 431 405 440 453 422 440 451
[97] 396 387 390 354 446 397 419 431 423 344 344 371 314 348 332 320
[113] 385 347 380 337 361 328 300 360 301 307 347 327 322 310 288 269
[129] 335 293 306 300 306 333 335 350 327 320 288 262 298 323 299 316
[145] 300 306 271 281 296 311 317 273 322
```

Fuente: Elaboración propia

Figura 5.58

Variación de defunciones diarias las mujeres a partir de 85 años en 2022

```
Time Series:
Start = 2
End = 153
Frequency = 1
[1] 36 -24 7 31 -31 -17 20 30 1 2 1 -6 10 8 -35 25 25
[18] -28 27 -63 10 -7 -39 6 17 8 -29 -12 21 17 25 -38 19 -54
[35] -11 67 28 -30 0 -11 -14 37 37 -22 57 15 13 -56 -34 -4 42
[52] -55 -23 6 -14 55 -69 72 -10 -38 22 -20 19 20 7 0 -15 0
[69] 8 21 74 -9 41 43 -10 5 41 -10 -77 31 -17 -22 -40 48 -73
[86] 22 14 -2 -32 -26 35 13 -31 18 11 -55 -9 3 -36 92 -49 22
[103] 12 -8 -79 0 27 -57 34 -16 -12 65 -38 33 -43 24 -33 -28 60
[120] -59 6 40 -20 -5 -12 -22 -19 66 -42 13 -6 6 27 2 15 -23
[137] -7 -32 -26 36 25 -24 17 -16 6 -35 10 15 15 6 -44 49
```

Fuente: Elaboración propia

Los resultados de la figura 5.58 muestran la variación de mortalidad de mujeres a partir de los 85 años, ya que en el punto 5.2.1, “Estacionariedad”, los test concluían que en esta serie temporal había que tomar diferencias para eliminar la tendencia.

Al tomar una diferencia, la serie se genera a partir de:

$$\begin{aligned}
 t = 1 & \quad mortalidad_t - mortalidad_{t-1} \\
 t = 2 & \quad mortalidad_t - mortalidad_{t-1} \\
 t = 3 & \quad mortalidad_t - mortalidad_{t-1} \\
 & \quad \dots
 \end{aligned}$$

La variabilidad diaria de la mortalidad implica que pueda haber diferencias negativas ya que un día puede haber dos fallecidos menos un día, al día siguiente cinco fallecidos más y al próximo ocho menos. Contar con datos diarios hace que estas variaciones sean más sensibles que si se tomaran mensuales o anuales.

Es importante remarcar que todos los estudios previos de causalidad y función impulso respuesta como el modelo VAR cuentan con la variable mortalidad en diferencias, *mortalidad_M85_diff*.

Los valores futuros de variación que se obtienen a través de la predicción (con ayuda de los valores de la temperatura) se encuentran en la figura 5.59 hasta el periodo 34. Los siguientes periodos que no se muestran en esta figura siguen presentando un aumento notable en la mortalidad.

Una aclaración importante es que a partir de la figura para mujeres entre 75-84 años, no se ha tomado diferencias y por tanto los resultados que salen, no son la varianza de un día a otro, si no realmente la predicción de mortalidad diaria.

Figura 5.59

Valor esperado de las defunciones diarias para mujeres a partir de 85 años

	fcst	lower	upper
[1,]	-25.1941691	-85.38496	34.99662
[2,]	2.0488182	-62.16147	66.25910
[3,]	-6.5950952	-71.25309	58.06290
[4,]	-2.7392872	-67.52408	62.04550
[5,]	-3.4042668	-68.22804	61.41951
[6,]	-2.5784045	-67.43262	62.27581
[7,]	-2.3640724	-67.23813	62.50999
[8,]	-2.0082148	-66.89722	62.88079
[9,]	-1.7652493	-66.66498	63.13448
[10,]	-1.5363590	-66.44395	63.37124
[11,]	-1.3484764	-66.26179	63.56484
[12,]	-1.1853976	-66.10288	63.73209
[13,]	-1.0470927	-65.96762	63.87343
[14,]	-0.9286742	-65.85142	63.99407
[15,]	-0.8276812	-65.75204	64.09668
[16,]	-0.7414095	-65.66695	64.18413
[17,]	-0.6677627	-65.59416	64.25863
[18,]	-0.6048759	-65.53190	64.32215
[19,]	-0.5511830	-65.47866	64.37630
[20,]	-0.5053379	-65.43315	64.42247
[21,]	-0.4661942	-65.39425	64.46186
[22,]	-0.4327722	-65.36100	64.49546
[23,]	-0.4042354	-65.33260	64.52413
[24,]	-0.3798700	-65.30832	64.54858

[25,]	-0.3590660	-65.28759	64.56946
[26,]	-0.3413030	-65.26988	64.58727
[27,]	-0.3261364	-65.25475	64.60247
[28,]	-0.3131868	-65.24182	64.61545
[29,]	-0.3021300	-65.23079	64.62653
[30,]	-0.2926894	-65.22136	64.63598
[31,]	-0.2846287	-65.21331	64.64405
[32,]	-0.2777463	-65.20643	64.65094
[33,]	-0.2718699	-65.20056	64.65682
[34,]	-0.2668524	-65.19555	64.66184

Fuente: Elaboración propia

Los valores negativos de la figura 5.59 son la representación de cómo cambia la variabilidad de la mortalidad de un periodo a otro. En este caso, lo que se obtiene es la variación en la tasa de crecimiento o decrecimiento según el momento futuro. A medida que pasa el tiempo esta predicción tiende a 0, representando que existe un cambio en la mortalidad diaria hasta que esta variación en los días se estabiliza.

El signo negativo surge a raíz de los últimos días de la muestra donde la diferencia entre un día y otro es negativa (ya que son los días finales de septiembre) y al contar con datos diarios, aunque la mortalidad siga siendo alta, la variabilidad de un día a otro es tan sutil que influye a la hora de las diferencias y por eso se ve tanta variación.

La figura 5.60 muestra los datos diarios de mortalidad conocidos de las mujeres de edad entre 75 y 85 años. A primera vista se observa como el número de fallecidos diarios es menor que a partir de los 85 años por causas naturales.

Figura 5.60

Número de defunciones diarias las mujeres de 75 a 84 años en 2022

```

Time Series:
Start = 1
End = 153
Frequency = 1
[1] 132 132 137 125 131 140 121 126 159 120 133 141 133 118 128 122 145
[18] 134 154 121 145 123 132 133 134 136 136 137 108 144 152 126 114 119
[35] 115 112 124 132 136 127 132 119 98 114 128 142 132 149 163 148 116
[52] 127 126 145 114 113 121 156 123 111 143 144 131 140 132 162 133 124
[69] 116 132 128 153 158 158 164 165 175 197 189 152 169 146 172 164 168
[86] 178 153 155 178 142 152 135 153 150 140 132 148 122 142 136 145 142
[103] 130 149 153 114 120 126 113 120 125 118 135 135 123 148 134 116 108
[120] 131 134 134 105 113 115 116 99 136 134 114 116 113 100 121 117 109
[137] 113 117 105 128 99 112 122 108 119 110 105 124 117 133 136 107 111

```

Fuente: Elaboración propia

Figura 5.61*Valor esperado de las defunciones diarias para mujeres de 75 a 84 años*

	fcst	lower	upper
[1,]	115.3273	86.85198	143.8026
[2,]	118.3463	86.03543	150.6572
[3,]	120.6053	86.84727	154.3633
[4,]	122.3887	87.79217	156.9852
[5,]	123.8494	88.65767	159.0412
[6,]	125.0744	89.43509	160.7136
[7,]	126.1163	90.13764	162.0950
[8,]	127.0101	90.77454	163.2456
[9,]	127.7805	91.35112	164.2099
[10,]	128.4464	91.87107	165.0218
[11,]	129.0229	92.33778	165.7081
[12,]	129.5225	92.75476	166.2903
[13,]	129.9556	93.12578	166.7855
[14,]	130.3312	93.45469	167.2078
[15,]	130.6570	93.74536	167.5687
[16,]	130.9396	94.00155	167.8777
[17,]	131.1848	94.22683	168.1427
[18,]	131.3974	94.42455	168.3703
[19,]	131.5819	94.59779	168.5660
[20,]	131.7420	94.74939	168.7345

Fuente: Elaboración propia

Figura 5.61*Continuación: Valor esperado de las defunciones diarias para mujeres de 75 a 84 años*

[21,]	131.8808	94.88187	168.8798
[22,]	132.0013	94.99754	169.0050
[23,]	132.1058	95.09843	169.1131
[24,]	132.1964	95.18638	169.2065
[25,]	132.2751	95.26299	169.2872
[26,]	132.3433	95.32969	169.3570
[27,]	132.4025	95.38773	169.4173
[28,]	132.4539	95.43821	169.4695
[29,]	132.4984	95.48211	169.5147
[30,]	132.5371	95.52026	169.5539
[31,]	132.5706	95.55342	169.5878
[32,]	132.5997	95.58223	169.6172
[33,]	132.6249	95.60726	169.6426
[34,]	132.6468	95.62899	169.6647

Fuente: Elaboración propia

La figura 5.60 representa la mortalidad diaria conocida para las mujeres de edades de 75 a 84 años, y como a partir de estos valores observados, se predice la mortalidad futura diaria con los datos de altas temperaturas en la figura 5.61.

Se observa una gran diferencia con la figura 5.59, ya que en esta serie no están tomadas diferencias (no se necesitaban) y por eso esta predicción sí mide exactamente el aumento de la mortalidad diaria, y no la variabilidad de los días.

La figura 5.62 muestra los datos diarios de mortalidad conocidos de los hombres entre los 45 y 64 años, siendo estos los de menor cuantía en comparación con todos los datos diarios de todas las series de mortalidad estudiadas.

Figura 5.62

Número de defunciones diarias los hombres de 45 a 64 años en 2022

Time Series:
 Start = 1
 End = 153
 Frequency = 1

[1]	88	92	93	93	86	97	92	79	80	96	93	95	89	87	112	107	81
[18]	92	115	91	100	93	110	78	99	90	85	98	81	81	119	96	98	90
[35]	85	78	100	86	91	90	93	70	84	85	100	100	100	116	103	109	92
[52]	106	93	85	98	84	83	102	114	84	96	97	107	93	98	90	102	103
[69]	94	92	114	109	116	101	131	108	103	110	125	107	126	102	105	116	104
[86]	118	92	99	108	96	110	119	84	113	87	105	80	93	87	122	106	99
[103]	111	95	108	105	100	89	83	77	82	82	102	93	83	92	82	73	96
[120]	79	68	111	79	79	67	87	97	119	96	93	74	84	91	72	95	92
[137]	87	82	85	80	89	83	91	88	84	92	83	95	92	91	88	84	96

Fuente: Elaboración propia

Figura 5.63

Valor esperado de las defunciones diarias para hombres de 45 a 64 años

	fcst	lower	upper
[1,]	91.32257	68.37071	114.2744
[2,]	90.81471	67.08756	114.5419
[3,]	91.17742	67.23036	115.1245
[4,]	91.67024	67.57163	115.7689
[5,]	92.13420	67.92359	116.3448
[6,]	92.54139	68.24841	116.8344
[7,]	92.89264	68.53921	117.2461
[8,]	93.19426	68.79648	117.5920
[9,]	93.45295	69.02262	117.8833
[10,]	93.67476	69.22053	118.1290
[11,]	93.86492	69.39315	118.3367
[12,]	94.02795	69.54329	118.5126
[13,]	94.16772	69.67359	118.6619
[14,]	94.28755	69.78646	118.7886
[15,]	94.39027	69.88407	118.8965
[16,]	94.47834	69.96838	118.9883
[17,]	94.55385	70.04113	119.0666
[18,]	94.61857	70.10383	119.1333
[19,]	94.67407	70.15783	119.1903
[20,]	94.72164	70.20431	119.2390
[21,]	94.76243	70.24429	119.2806
[22,]	94.79740	70.27866	119.3161
[23,]	94.82737	70.30821	119.3465
[24,]	94.85307	70.33359	119.3726
[25,]	94.87511	70.35539	119.3948
[26,]	94.89400	70.37410	119.4139
[27,]	94.91019	70.39017	119.4302
[28,]	94.92407	70.40396	119.4442
[29,]	94.93598	70.41579	119.4562
[30,]	94.94618	70.42595	119.4664
[31,]	94.95493	70.43466	119.4752
[32,]	94.96243	70.44213	119.4827
[33,]	94.96886	70.44854	119.4892
[34,]	94.97437	70.45404	119.4947

Fuente: Elaboración propia

Para los hombres el efecto es el mismo que para las mujeres, y esto se observa en la figura 5.63, cuya tendencia es un incremento en la mortalidad, en función de los datos conocidos de temperatura y mortalidad diaria.

Finalmente, la figura 5.64 muestra los datos diarios de mortalidad conocidos para los hombres de edades comprendidas entre los 65 y 74 años.

Figura 5.64

Número de defunciones diarias los hombres de 65 a 74 años en 2022

```
Time Series:
Start = 1
End = 153
Frequency = 1
[1] 113 118 108 125 124 112 111 94 127 120 109 137 116 93 119 106 131
[18] 118 98 108 113 132 109 98 105 103 105 88 93 103 103 123 115 98
[35] 84 107 112 120 105 113 112 91 100 118 108 105 148 111 98 98 114
[52] 117 136 122 97 91 100 119 131 107 117 127 99 118 124 118 124 123
[69] 122 117 122 129 104 140 130 113 117 158 144 131 140 122 103 104 105
[86] 127 117 118 123 105 113 111 129 125 128 121 98 95 101 104 96 122
[103] 118 139 106 105 104 102 107 115 112 112 126 120 114 108 108 105 87
[120] 104 113 118 101 111 106 80 121 108 101 98 106 106 106 110 108 93
[137] 104 87 92 107 95 107 114 101 112 97 99 116 105 111 98 116 109
```

Fuente: Elaboración propia

Figura 5.65

Valor esperado de las defunciones diarias los hombres de 65 a 74 años

	fcst	lower	upper
[1,]	107.5004	83.62088	131.3799
[2,]	107.5544	82.51262	132.5962
[3,]	107.9868	82.69141	133.2822
[4,]	108.4693	83.03131	133.9072
[5,]	108.9156	83.37543	134.4558
[6,]	109.3082	83.69277	134.9235
[7,]	109.6480	83.97717	135.3188
[8,]	109.9406	84.22899	135.6523
[9,]	110.1922	84.45048	135.9339
[10,]	110.4084	84.64446	136.1723
[11,]	110.5941	84.81380	136.3743
[12,]	110.7536	84.96124	136.5459
[13,]	110.8906	85.08935	136.6918
[14,]	111.0082	85.20047	136.8160
[15,]	111.1093	85.29671	136.9219
[16,]	111.1961	85.37996	137.0123
[17,]	111.2707	85.45191	137.0895
[18,]	111.3348	85.51402	137.1555
[19,]	111.3898	85.56761	137.2120
[20,]	111.4370	85.61381	137.2603

Fuente: Elaboración propia

Figura 5.65

Continuación: Valor esperado de las defunciones diarias los hombres de 65 a 74 años

[21,]	111.4776	85.65362	137.3017
[22,]	111.5125	85.68791	137.3371
[23,]	111.5425	85.71743	137.3675
[24,]	111.5682	85.74285	137.3935
[25,]	111.5903	85.76471	137.4158
[26,]	111.6093	85.78352	137.4350
[27,]	111.6256	85.79969	137.4514
[28,]	111.6396	85.81360	137.4655
[29,]	111.6516	85.82556	137.4776
[30,]	111.6619	85.83584	137.4880
[31,]	111.6708	85.84468	137.4969
[32,]	111.6784	85.85227	137.5045
[33,]	111.6850	85.85880	137.5111
[34,]	111.6906	85.86441	137.5167

Fuente: Elaboración propia

La figura 5.65 representa una pequeña diferencia con el resto de los resultados donde el crecimiento de la mortalidad es positivo, pero mucho más lento que si lo comparamos con el resto, pudiendo ser por diversos motivos. Uno de ellos es que se ha evidenciado un menor impacto del calor en la mortalidad en los hombres que en las mujeres. Es, por ello, que hay menos edades a las que los hombres se vean afectados por estas temperaturas y puede ser la causa de que esta variable climática no sea de tanta utilidad para predecir, como sí lo es para las mujeres. Esta idea va en consonancia con la figura 5.51, la cual refleja que, aunque el calor explique un porcentaje de la mortalidad, este es menor para hombres que para mujeres.

5.3. RESULTADOS PARA MODELOS VAR CON FRECUENCIA ANUAL.

El estudio entre la mortalidad y la temperatura suele estudiarse con frecuencia diaria. Por ejemplo, Achebak et al. (2020) comentaban como es mucho más apropiado medir esta relación con datos diarios debido a las características de ambas variables. Aunque se afirme que es preferible la utilización de estos y en este trabajo se realice de esta manera, se ha querido probar con datos anuales para analizar los posibles efectos.

Una vez realizados los modelos GLM que sirven para estudiar a qué edad, en función del género, las altas temperaturas tienen un impacto en la mortalidad, se escogen tres edades para cada sexo y se toma la q_x correspondiente a cada edad desde el 1970 al 2022. Cada serie temporal de la mortalidad según la edad se va a combinar con la temperatura anual para el mismo tramo de años.

El proceso metodológico es exactamente el mismo que el de la segunda parte metodológica, aunque existe un código específico para este estudio en el Anexo. Primero se comprueba que las series son estacionarias, qué modelo VAR es el apropiado..., pero la principal diferencia viene por el estudio de la Causalidad de Granger y los resultados obtenidos del mismo.

La causalidad de Granger estudia si una variable ayuda a predecir otra, al tener un efecto causal en la misma. En este caso, si las altas temperaturas ayudan a predecir la mortalidad por un impacto causal del clima sobre la probabilidad de fallecer.

Se estudia la mortalidad para las edades 59, 72 y 74 para hombres y 80, 85, 90 y 94 para las mujeres. En todos los modelos VAR se va a obtener el mismo resultado. Por ello, se va a representar en las siguientes dos figuras el modelo VAR (1) para las series temporales altas temperaturas y probabilidad de fallecer para mujeres de 85 años, siendo esta edad y género representativo de la muestra.

Figura 5.66

Causalidad de Granger entre mortalidad y altas temperaturas para mujeres de 85 años.

```
Granger causality test

Model 1: mortalidad_85_diff ~ Lags(mortalidad_85_diff, 1:1) + Lags(TempMaxima,
1:1)
Model 2: mortalidad_85_diff ~ Lags(mortalidad_85_diff, 1:1)
Res.Df Df      F      Pr(>F)
1      47
2      48 -1  1.093   0.3012
3      49 -1 29.034 2.123e-06 ***
```

Fuente: Elaboración propia

Si se compara el resultado de este test con cualquiera de las series temporales diarias se observa que en esta figura no hay evidencias suficientes como para rechazar la hipótesis de que la temperatura no sirve para predecir la mortalidad, al tener un pvalor del 0.3012. Se concluye que las altas temperaturas no tienen un efecto causal en la mortalidad anual y, por tanto, no sirven de ayuda para predecir la probabilidad de fallecer dentro de un año.

Figura 5.67

Modelo VAR (1) para mujeres de 85 años.

```
Estimation results for equation mortalidad_85_diff:
=====
mortalidad_85_diff = mortalidad_85_diff.l1 + TempMaxima.l1 + const

              Estimate Std. Error t value Pr(>|t|)
mortalidad_85_diff.l1 -5.478e-01  1.057e-01  -5.184  4.3e-06 ***
TempMaxima.l1         9.219e-05  3.649e-04   0.253  0.802
const                 -5.021e-03  1.061e-02  -0.473  0.638
```

Fuente: Elaboración propia

La figura 5.67 representa un modelo VAR (1) ya que los criterios de información AIC y BIC determinaron que era el óptimo para la edad 85. Se observa como el valor pasado de alta temperatura no es significativo para la variable explicativa mortalidad y por tanto no es interesante usar el retardo climático para predecir la mortalidad ya que no le provoca ningún efecto relevante.

En consecuencia, no se tienen evidencias que con series históricas anuales la temperatura sea relevante a la hora de explicar o predecir la mortalidad para ninguna de las siete edades de hombres y mujeres estudiadas en el trabajo. Por el contrario, ya se sabe que si la frecuencia es con datos diarios del 2022 el efecto de causalidad es significativo y, con ello, la utilidad de las altas temperaturas para predecir la mortalidad diaria.

6. CONCLUSIÓN.

Conocer cómo se comporta la mortalidad a causa de las temperaturas extremas ha sido un tema estudiado desde antes del siglo XXI por autores como Díez (1996) donde analizaba el impacto de las temperaturas frías como cálidas en función del género y edad. El año 2003 supuso un punto de inflexión ya que se alcanzaron temperaturas extremadamente altas teniendo un efecto directo y significativo en la mortalidad española. Esto hizo que el gobierno del momento implantara en el año siguiente un plan preventivo frente a los efectos de las altas temperaturas en la salud.

Con el paso de los años, unas hipótesis imprecisas y el aumento de la temperatura a causa del cambio climático se empezó a dar una mayor prioridad a cómo mitigar el impacto de las olas de calor en las personas vulnerables.

Un año antes del verano con mayor número de fallecidos, López-Bueno et al. (2021) destacaban la importancia de conocer la temperatura la cual afectaba a las personas en términos de salud, trabajo y mortalidad.

Este trabajo ha supuesto un estudio en España tanto para comprobar qué edades, en función del género, son las más perjudicadas por el verano del 2022 y demostrar cómo un shock en las altas temperaturas supone un impacto en el número de fallecidos, pudiendo así utilizarse esta variable climatológica para predecir la mortalidad futura.

Ballester et al. (2023) diferenciaron cómo la ola de calor que asoló Europa afectó de distinta manera a hombres y mujeres teniendo mucha mayor repercusión en las segundas.

Para comprobarlo y ver el impacto en España en particular, se empieza por un análisis de correlación entre las temperaturas y mortalidad de los meses de abril a septiembre de los años 2021 y 2022 donde se concluye que aunque el primero fue un posible aviso, es al siguiente año, en el 2022, donde se observa una correlación positiva elevada entre ambas variables, siendo uno de los pocos años que cuenta con un mayor número de fallecidos en julio que en el mes de diciembre y con valores muy cercanos a los de enero.

Es por ello por lo que el trabajo se centra en los meses cálidos de este año para conocer las edades que están más afectadas por esta situación de temperaturas extremas y observar si existen diferencias entre las edades de mujeres y hombres.

Una vez realizado un modelo lineal generalizado (GLM) para cada edad en función de su mortalidad y las temperaturas mensuales, se confirma una diferencia en las edades en las que hay un impacto significativo de la variable independiente, temperatura, en la variable dependiente, mortalidad, en función del sexo.

Para las mujeres, es a partir de la edad 76 hasta la última edad de la muestra 101, donde la temperatura tiene una relación positiva significativa con la mortalidad, y hay suficiente evidencia para rechazar la hipótesis de no significatividad para este rango de edades.

Para los hombres, las edades afectadas son más tempranas, comprendidas entre los 50 años y 75 aproximadamente y es a partir de esta edad donde la temperatura deja de ser relevante. Este resultado contrasta con el obtenido para las mujeres, para quienes la temperatura empieza a ser significativa justo a partir de la edad 76.

Esto refleja la primera gran diferencia entre hombres y mujeres, además son estas las que durante más edades se ven afectadas por las olas de calor debido a sus condiciones hormonales y particulares, causado por una mayor retención del calor que los hombres.

Esto demuestra que no todas las edades cuentan con un impacto del calor en la mortalidad, por eso se realiza ese estudio previo de identificar las edades que son vulnerables a los efectos extremos. Una vez conocidas, el análisis se enfoca en conocer la mortalidad futura para esos rangos de edad ya que en un contexto de cambio climático y aumento de las emisiones de gases de efecto invernadero la tendencia es que la temperatura máxima continúe aumentando y por ende, el riesgo de un aumento en la mortalidad por encima de lo esperado para esas edades afectadas.

A raíz del uso de modelos VAR se puede analizar la conexión entre las dos series temporales con frecuencia diaria. Al combinar cada una de las series temporales de mortalidad para cada edad con la temperatura máxima, se obtienen cuatro modelos VAR (uno por cada rango de edad estudiado) y se obtiene gracias a las funciones impulso-respuesta como un shock en la variable impulso, temperatura máxima, genera una respuesta de aumento en la mortalidad durante unos periodos hasta que tiempo después vuelve a su valor habitual. Además por la causalidad de Granger se puede afirmar como los valores pasados de la temperatura máxima sirven para predecir la mortalidad futura en todos los rangos de edad vulnerables.

Es importante recordar que los datos diarios son muy utilizados en los modelos de series temporales cuando se quieren estudiar variables climatológicas y en este trabajo se demuestra además como si en vez de tomar frecuencia diaria se toma frecuencia anual para edades sensibles al calor, con datos históricos desde el 1970, no existen evidencias de que la temperatura máxima ayude a predecir la mortalidad futura para el año que viene.

Por tanto, el impacto del cambio climático y con ello el aumento de las temperaturas tienen un efecto importante en este país. En el año 2022 se ve un nuevo patrón de mortalidad donde las altas temperaturas están causando un aumento en las cifras de la primera. Se comprueba como no todas las edades están afectadas y como una de las diferencias viene por el género. Las mujeres mayores a 76 años y los hombres de edades de 50 a 75 son los más perjudicados, aunque son las primeras las que realmente sufren los efectos climáticos. Es para estas edades para las que se desea conocer la mortalidad futura, o al menos una aproximación, ya que van a estar impactadas por una variable específica que se espera que siga estando en un futuro, las altas temperaturas máximas.

Con el trabajo se ha conseguido que de un hecho genérico como es el impacto de las altas temperaturas en la mortalidad, se conozcan las edades afectadas diferenciado por hombres y mujeres. Con esta información, se ha podido estimar la mortalidad futura para estas edades, y con ello estimar el posible riesgo de un aumento en el número de fallecidos.

Este aumento es solo una estimación ya que se ha demostrado que el impacto del calor ha sido algo particular del año 2022 ya que si se toman datos anuales desde un histórico de más de 30 años para predecir no se consiguen evidencias de que la variable temperatura haya supuesto un efecto significativo en la mortalidad.

Esta idea se refuerza con los datos ya conocidos del 2023, donde se vuelve a picos de mortalidad mucho más altos en diciembre y enero que en verano, dando a entender que aunque la temperatura siga aumentando a consecuencia del cambio climático (y seguramente el número de fallecidos por encima de lo esperado en los meses con olas de calor), el impacto que genera en la mortalidad no va a ser homogéneo a lo largo de los años.

BIBLIOGRAFÍA.

- Achebak, H., Devolder, D., Ingole, V., & Ballester, J. (2020). Reversal of the seasonality of temperature-attributable mortality from respiratory diseases in Spain. *Nature communications*, 11(1), 2457.
- AEMET (s.f). *Estadística de las variables meteorofenológicas. Estadística Anual 2022* [Dataset].https://www.aemet.es/es/datos_abiertos/estadisticas/estadistica_meteorofenologicas
- AEMET (s.f). *Open Data / Datos diarios* [Dataset]. <https://datosclima.es/Aemethistorico/Meteosingleday.php>
- Agencia Estatal de Meteorología (2021). *Resumen anual climatológico*. https://www.aemet.es/documentos/es/serviciosclimaticos/vigilancia_clima/resumenes_climat/anuales/res_anual_clim_2021.pdf
- Agencia Estatal de Meteorología (2022). *Informe mensual climatológico*. https://www.aemet.es/documentos/es/serviciosclimaticos/vigilancia_clima/resumenes_climat/mensuales/2022/res_mens_clim_2022_07.pdf
- Ashenfelter, O., & Card, D. (1982). Time series representations of economic variables and alternative models of the labour market. *The Review of Economic Studies*, 49(5), 761-782.
- Balacco, H. (1986). Algunas consideraciones sobre la definición de causalidad de Granger en el análisis econométrico. *Económica*, 32.
- Ballester, J., Quijal-Zamorano, M., Méndez Turrubiates, R. F., Pegenaute, F., Herrmann, F. R., Robine, J. M., ... & Achebak, H. (2023). Heat-related mortality in Europe during the summer of 2022. *Nature medicine*, 29(7), 1857-1866.
- Carmona, R., Díaz, J., Mirón, I. J., Ortiz, C., Luna, M. Y., & Linares, C. (2016). Mortality attributable to extreme temperatures in Spain: A comparative analysis by city. *Environment International*, 91, 22-28.
- Carretero, E. S. (2023). Desafíos e impactos para la generación de una cultura de seguridad y salud laboral en el sector de la construcción española. *Gestión de la seguridad y la Salud en el Trabajo*, 5(1), 53-59.
- Cayuela, L. (2009). Modelos lineales generalizados (GLM). *Materiales de un curso del R del IREC*.

- Coeficiente de correlación de Pearson. (18 de octubre, 2023). En Wikipedia. https://es.wikipedia.org/w/index.php?title=Coeficiente_de_correlaci3n_de_Pearson&action=history
- Coeficiente de correlaci3n de Spearman. (18 de octubre, 2023). En Wikipedia. https://es.wikipedia.org/wiki/Coeficiente_de_correlaci3n_de_Spearman
- Díaz, J., Carmona, R., Mir3n, I. J., Luna, M. Y., & Linares, C. J. E. I. (2018). Time trend in the impact of heat waves on daily mortality in Spain for a period of over thirty years (1983–2013). *Environment international*, 116, 10-17.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427-431
- Díez, F. B. (1996). Meteorología y salud. La relaci3n entre la temperatura ambiental y la mortalidad. *Rev Esp Salud P3blica*, 70(3), 251-259.
- Díaz, J., Sáez, M., Carmona, R., Mir3n, I. J., Barcel3, M. A., Luna, M. Y., & Linares, C. (2019). Mortality attributable to high temperatures over the 2021–2050 and 2051–2100 time horizons in Spain: Adaptation and economic estimate. *Environmental research*, 172, 475-485. <https://doi.org/10.1016/j.envres.2019.02.041>
- Efron, B. (2013). A 250-year argument: Belief, behavior, and the bootstrap. *Bulletin of the American Mathematical Society*, 50(1), 129-146.
- Eurostat (2024). *Excess mortality by month*. [Dataset] https://ec.europa.eu/eurostat/databrowser/view/DEMO_MEXRT_custom_3098_01/bookmark/table?lang=en&bookmarkId=26981184-4241-4855-b18e-8647fc8c0dd2
- García López, L. (27 de junio, 2023). La mortalidad en Espa3a alcanza el peor dato desde que hay registro, solo por detr3s de la pandemia. *ABC*. <https://www.abc.es/sociedad/mortalidad-espana-alcanza-peor-dato-registro-solo-20230627135516-nt.html#vca=compartirrrss&vso=abc&vmc=rrss&vli=fixed-link>.
- G3mez, I. J., & Rodr3guez, S. M. S. (2021). Cobertura en la prensa europea de la adaptaci3n de las ciudades a las olas de calor y al cambio clim3tico. *Revista Mediterr3nea de Comunicaci3n: Mediterranean Journal of Communication*, 12(1), 45-63.

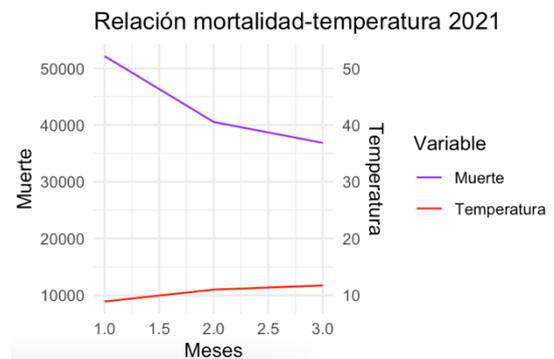
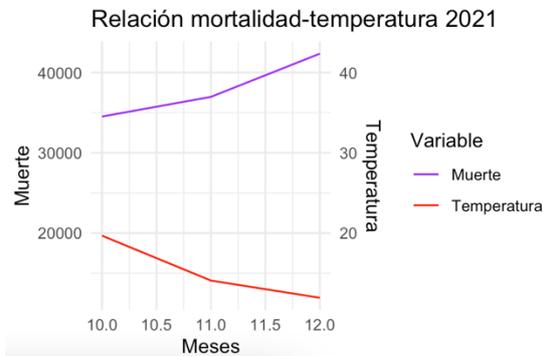
- Gómez, I. J., & Rodríguez, S. M. S. (2021). Cobertura en la prensa europea de la adaptación de las ciudades a las olas de calor y al cambio climático. *Revista Mediterránea de Comunicación: Mediterranean Journal of Communication*, 12(1), 45-63.
- Human Mortality DataBase (2021). *Life tables* [Dataset]. <https://www.mortality.org/Country/Country?cntr=ESP>
- Instituto Nacional de Estadística. (s.f.). *Defunciones mensuales, acumuladas y diferencia absoluta del acumulado por sexo y edad. Total nacional y provincias. 2009-2023*. <https://www.ine.es/jaxiT3/Tabla.htm?t=62281&L=0>.
- Instituto Nacional de Estadística. (s.f.). *Defunciones por edad, sexo y mes de defunción*. <https://www.ine.es/jaxiT3/Tabla.htm?t=31912&L=0>
- Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2012). *Loss models: from data to decisions* (Vol. 715). John Wiley & Sons.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of econometrics*, 54(1-3), 159-178.
- Leon, I., Frías, L., Delgado, C., & Larrauri, A. (30 de septiembre, 2022). Informe MoMo. *Excesos de mortalidad por todas las causas y atribuibles a excesos de temperatura en España. 1 de junio a 30 de septiembre de 2022*. Centro Nacional de Epidemiología. CIBERESP. ISCIII.
- Leon-Gomez, I., Sanz, C. D., & Larrauri, A. (2022). Implementación de un nuevo modelo MoMo, para la estimación de excesos de mortalidad por todas las causas y atribuibles a temperatura.
- Linares-Gil, C., Carmona-Alferez, R., Ortiz Burgos, C., & Diaz-Jimenez, J. (2017). *Temperaturas extremas y salud. Cómo nos afectan las olas de calor y de frío*. Instituto de Salud Carlos III (ISCIII). Escuela Nacional de Sanidad (ENS).
- Linares-Gil, C., Osa, J. D. L., & Diaz-Jimenez, J. (2023). Investigar los efectos de las altas temperaturas en la salud va más allá de calcular muertes.
- Lopez-Bueno, J. A., Diaz, J., Follos, F., Vellón, J. M., Navas, M. A., Culqui, D., ... & Linares, C. (2021). Evolution of the threshold temperature definition of a heat wave vs. evolution of the minimum mortality temperature: a case study in Spain during the 1983–2018 period. *Environmental Sciences Europe*, 33, 1-10.

- López-González, E., & Ruiz-Soler, M. (2011). Análisis de datos con el Modelo Lineal Generalizado. Una aplicación con R. *Revista española de pedagogía*, 59-80.
- Martínez Navarro, F., Simón-Soria, F., & López-Abente, G. (2004). Valoración del impacto de la ola de calor del verano de 2003 sobre la mortalidad. *Gaceta Sanitaria*, 18(suppl 1), 250-258.
- Mayoral, M. A. M., & Socuélamos, J. M. (2001). *Modelos lineales generalizados*. Universidad Miguel Hernández.
- McCullagh, P. (2019). *Generalized linear models*. Routledge.
- MoMo (2024). *Mortalidad notificada, observada, esperada y atribuible a temperatura* [Dataset]. https://momo.isciii.es/panel_momo/
- Novales, A. (2011). Modelos vectoriales autoregresivos (VAR). *Universidad Complutense*, 1-26.
- Organización Meteorológica Mundial (19 de marzo, 2024). Los indicadores del cambio climático alcanzaron niveles sin precedentes en 2023: OMM. <https://wmo.int/es/news/media-centre/los-indicadores-del-cambio-climatico-alcanzaron-niveles-sin-precedentes-en-2023-omm>
- Organización Mundial de la Salud. (30 de octubre, 2021). *Cambio climático*. <https://www.who.int/es/news-room/fact-sheets/detail/climate-change-and-health>.
- Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *biometrika*, 75(2), 335-346.
- Pliego, F. F., Gil, C. L., Graña, J. M. V., Bueno, J. A. L., Rico, M. Y. L., Martínez, G. S., & Jiménez, J. D. (2020). Evolución de la temperatura de mínima mortalidad en Madrid y Sevilla en el periodo 1983-2018. *Revista de Salud Ambiental*, 20(1).
- Puig, X., Ginebra, J., & Gispert, R. (2005). Análisis de la evolución temporal de la mortalidad mediante modelos lineales generalizados. *Gaceta Sanitaria*, 19(6), 481-485.
- Rubio, J. M., Pérez, I. M., Jiménez, J. D., & Odriozola, J. A. (1997). Influencia de variables atmosféricas sobre la mortalidad por enfermedades respiratorias y cardiovasculares en los mayores de 65 años de la Comunidad de Madrid. *Gaceta Sanitaria*, 11(4), 164-170.
- Robine, J. M., Cheung, S. L. K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J. P., & Herrmann, F. R. (2008). Death toll exceeded 70,000 in Europe during the summer of 2003. *Comptes Rendus. Biologies*, 331(2), 171-178.

- Sánchez-Guevara, C., Bueno, J. A. L., Peiró, M. N., Gil, C. L., & Fernández, A. S. (2021).
Salud en los barrios: Impacto de las temperaturas extremas. *Revista de Salud Ambiental*, 21(1), 65-73.
- Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.
- Vicente Maldonado, J(2023). *Apuntes Series Temporales*. Master Ciencias Actuariales.
UC3M.

ANEXO.

A1. Gráficos para el año 2021 en los meses fríos.



Ambos gráficos reflejan lo mismo que las figuras 3.4 y 3.5 para los meses fríos del 2022. Se observa como a medida que baja la temperatura, la mortalidad aumenta. Es por ello que en este trabajo se estudian solo los meses cálidos, ya que en caso de que se quisieran medir también estos, habría que hacerlo por separado, para no sesgar los resultados.

A2. Código de R.

```
#EXCESO DE MORTALIDAD
```

```
#Graficar el exceso de mortalidad para ver cómo en verano del 2022 hubo un aumento.
```

```
Excesodemortalidad<-read.csv2(file.choose(),header = FALSE,nrows = 45)
```

```
Fecha<-ts(Excesodemortalidad$V1)
```

```
exceso<-ts(Excesodemortalidad$V2)
```

```
Exceso<-exceso/100
```

```
data=ts((Excesodemortalidad$V2/100),freq=12,start=c(2020,1))
```

```
plot(data,main="", xlab="", ylab="Exceso de mortalidad")
```

```
axis(3, at = time(data), labels = Excesodemortalidad$V1, las = 2)
```

```
#RELACION TEMP Y MORTALIDAD
```

```
#2022#
```

```
Mortalidad<-read.csv2(file.choose(),header = FALSE)
```

```
Temperaturas<-read.csv2(file.choose(),header = FALSE)
```

```

Mortalidadts<-ts(Mortalidad$V2)[4:9]
Temperaturasts<-ts(Temperaturas$V2)[4:9]
serie_temporal <- ts(data = cbind(Mortalidadts, Temperaturasts), start = 4)
library(ggplot2)
df <- data.frame(Fecha = time(serie_temporal), Muerte = Mortalidadts, Temperatura =
Temperaturasts)

ggplot(df, aes(x = Fecha)) +
  geom_line(aes(y = Muerte, color = "Muerte")) +
  geom_line(aes(y = Temperatura * 1000, color = "Temperatura")) +
  scale_y_continuous(sec.axis = sec_axis(~./1000, name = "Temperatura")) +
  labs(title = "Relación mortalidad-temperatura 2022", x = "Meses", y = "Muerte", color
= "Variable") +
  scale_color_manual(values = c(Muerte = "purple", Temperatura = "red")) +
  theme_minimal() +
  xlim(min(df$Fecha), max(df$Fecha))
cor(Mortalidadts, Temperaturasts)
#se basa en los rangos de los datos y es menos sensible a los valores atípicos
correlation_spearman <- cor(Mortalidadts, Temperaturasts, method = "spearman")
print(correlation_spearman)

#2021#
Mortalidad21<-read.csv2(file.choose(),header = FALSE)
Temperaturast21<-read.csv2(file.choose(),header = FALSE)
Mortalidadts21<-ts(Mortalidad21$V2)[4:9]
Temperaturasts21<-ts(Temperaturas21$V2)[4:9]
serie_temporal21 <- ts(data = cbind(Mortalidadts21, Temperaturasts21), start = 4)
library(ggplot2)
install.packages("ggplot2")
library(ggplot2)
df21 <- data.frame(Fecha21 = time(serie_temporal21), Muerte = Mortalidadts21,
Temperatura = Temperaturasts21)
ggplot(df21, aes(x = Fecha21)) +
  geom_line(aes(y = Muerte, color = "Muerte")) +

```

```

geom_line(aes(y = Temperatura * 1000, color = "Temperatura")) +
scale_y_continuous(sec.axis = sec_axis(~./1000, name = "Temperatura")) +
labs(title = "Relación mortalidad-temperatura 2021", x = "Meses", y = "Muerte", color
= "Variable") +
scale_color_manual(values = c(Muerte = "purple", Temperatura = "red")) +
theme_minimal() +
xlim(min(df21$Fecha21), max(df21$Fecha21))
cor(Mortalidadts21, Temperaturasts21)
correlation_spearman21 <- cor(Mortalidadts21, Temperaturasts21, method =
"spearman")
print(correlation_spearman21)

```

```
#Invierno#
```

```
#2022 Enero-Marzo
```

```
MortalidadtsInv<-ts(Mortalidad$V2)[c(1:3)]
```

```
TemperaturastsInv<-ts(Temperaturas$V2)[c(1:3)]
```

```
serie_temporalInv <- ts(data = cbind(MortalidadtsInv, TemperaturastsInv))
```

```
dfInv <- data.frame(Fecha = time(serie_temporalInv), Muerte = MortalidadtsInv,
Temperatura = TemperaturastsInv)
```

```
ggplot(dfInv, aes(x = Fecha)) +
```

```
geom_line(aes(y = Muerte, color = "Muerte")) +
```

```
geom_line(aes(y = Temperatura * 1000, color = "Temperatura")) +
```

```
scale_y_continuous(sec.axis = sec_axis(~./1000, name = "Temperatura")) +
```

```
labs(title = "Relación mortalidad-temperatura", x = "Meses", y = "Muerte", color =
"Variable") +
```

```
scale_color_manual(values = c(Muerte = "purple", Temperatura = "red")) +
```

```
theme_minimal() +
```

```
xlim(min(dfInv$Fecha), max(dfInv$Fecha))
```

```
cor(MortalidadtsInv, TemperaturastsInv)
```

```
#2022 Octubre-Diciembre
```

```
MortalidadtsInv<-ts(Mortalidad$V2)[c(10:12)]
```

```
TemperaturastsInv<-ts(Temperaturas$V2)[c(10:12)]
```

```

serie_temporalInv <- ts(data = cbind(MortalidadtsInv, TemperurastsInv))
dfInv <- data.frame(Fecha = time(serie_temporalInv), Muerte = MortalidadtsInv,
Temperatura = TemperurastsInv)

ggplot(dfInv, aes(x = Fecha)) +
  geom_line(aes(y = Muerte, color = "Muerte")) +
  geom_line(aes(y = Temperatura * 1000, color = "Temperatura")) +
  scale_y_continuous(sec.axis = sec_axis(~./1000, name = "Temperatura")) +
  labs(title = "Relación mortalidad-temperatura", x = "Meses", y = "Muerte", color =
"Variable") +
  scale_color_manual(values = c(Muerte = "purple", Temperatura = "red")) +
  theme_minimal() +
  xlim(min(dfInv$Fecha), max(dfInv$Fecha))
cor(MortalidadtsInv, TemperurastsInv)

#2021
MortalidadtsInv21<-ts(Mortalidad$V2)[c(1:3, 10:12)]
TemperurastsInv21<-ts(Temperuras$V2)[c(1:3, 10:12)]
serie_temporalInv21 <- ts(data = cbind(MortalidadtsInv21, TemperurastsInv21))

dfInv21 <- data.frame(Fecha21 = time(serie_temporalInv21), Muerte =
MortalidadtsInv21, Temperatura= TemperurastsInv21)

ggplot(dfInv21, aes(x = Fecha21)) +
  geom_line(aes(y = Muerte, color = "Muerte")) +
  geom_line(aes(y = Temperatura * 1000, color = "Temperatura")) +
  scale_y_continuous(sec.axis = sec_axis(~./1000, name = "Temperatura")) +
  labs(title = "Relación mortalidad-temperatura", x = "Meses", y = "Muerte", color =
"Variable") +
  scale_color_manual(values = c(Muerte = "purple", Temperatura = "red")) +
  theme_minimal() +
  xlim(min(dfInv21$Fecha21), max(dfInv21$Fecha21))

cor(MortalidadtsInv21, TemperurastsInv21)

```

```

#VER SI ES RELACION LINEAL
#Grafico dispersión
RELACION<-read.csv2(file.choose(),header = TRUE)
plot(RELACION$TEMP[c(4:9)], RELACION$MORT[c(4:9)],
     main = "Dispersión entre mortalidad y temperatura",
     xlab = "Temperatura", ylab = "Mortalidad",
     pch = 19, col = "blue")

dtRELACION <- data.frame(Edad = RELACION$MES, TasaMortalidad =
RELACION$MORT, Temperatura = RELACION$TEMP)

#Correlacion
cor(RELACION$TEMP[c(4:9)],RELACION$MORT[c(4:9)])
correlation_spearman<- cor(RELACION$TEMP[c(4,9)], RELACION$MORT[c(4,9)],
method = "spearman")
print(correlation_spearman21)

#Relacion lineal
modelo <- lm(RELACION$MORT[c(4:9)] ~ RELACION$TEMP[c(4:9)], data =
RELACION)
residuos<-residuals(modelo)
shapiro.test(residuos)

plot(RELACION$TEMP[c(4:9)], RELACION$MORT[c(4:9)],
     xlab = "Temperatura", ylab = "Mortalidad",
     main = "Gráfico de dispersión de Temperatura vs Mortalidad")
abline(modelo, col = "red")
legend("downright", legend = "Línea de Regresión", col = "red", lty = 1)

#DISTRIBUCION DE LA MORTALIDAD#
datos<-RELACION$MORT
n<-length(datos)
max(datos)
hist(datos)

```

```

#Uniform Kernel
h2=0.9*min(sd(datos),(quantile(datos,0.75)-quantile(datos,0.25))/1.34)*(1/n^(1/5))
f<-c()
punto=seq(35000,48000,1)
for (j in 1:length(punto)){
  f[j]=(1/(h2*n))*sum(dunif((punto[j]-datos)/h2,0,2))
}
df<-data.frame(cbind(punto,f))
plot(df)
plot(cumsum(f),xaxt='n')

#Gaussian Kernel
h1=1.06*sd(datos)/(n^(1/5))
f<-c()
punto=seq(35000,48000,1)
for (j in 1:length(punto)){
  f[j]=(1/(h1*n))*sum(dnorm((punto[j]-datos)/h1,0,1))
}
df<-data.frame(cbind(punto,f))
plot(df)
plot(cumsum(f),xaxt='n')

aux<-cbind(punto,cumsum(f))
escenarios=1000
mortalidad<-c()
for (i in 1:escenarios){
  mortalidad[i]<-aux[which(round(aux[,2],3) == round(runif(1,0,1),3)),1][1]
}
mortalidad<-va[complete.cases(va)]
hist(mortalidad)

```

```

#MODELOS REGRESION MUJERES Y HOMBRES#

#MUJERES
bbddMUJERES<-read.csv2(file.choose(),header = TRUE)

dtModeloMUJERES <- data.frame(Edad = bbddMUJERES$EDAD, TasaMortalidad =
bbddMUJERES$MORTALIDAD, Temperatura = bbddMUJERES$TEMP)

dtModeloMUJERES <- na.omit(dtModeloMUJERES)

modelosM <- list()

for (edad in unique(dtModeloMUJERES$Edad)) {
  datos_edadM <- subset(dtModeloMUJERES, Edad == edad)

  modeloM <- glm(TasaMortalidad ~ Temperatura, data = datos_edadM, family =
poisson(link="log"))
  modelosM[[paste("Edad", edad)]] <- modeloM
}

for (i in 1:length(modelosM)) {
  cat("Resumen del modelo para", names(modelosM)[i], ":\n")
  print(summary(modelosM[[i]]))
}

for (i in 1:50) {
  cat("Resumen del modelo para", names(modelosM)[i], ":\n")
  print(summary(modelosM[[i]]))
}

vermodelo<-modelosM[[59]]
print(summary(vermodelo))
cat("Resumen del modelo para 25:\n")
print(summary(modelosM[[70]]))

```

```

#HOMBRES
bbddHOMBRES<-read.csv2(file.choose(),header = TRUE)

dtModeloHOMBRES <- data.frame(Edad = bbddHOMBRES$EDAD, TasaMortalidad =
bbddHOMBRES$MORTALIDAD, Temperatura = bbddHOMBRES$TEMP)

dtModeloHOMBRES <- na.omit(dtModeloHOMBRES)

modelosHOMBRES <- list()

for (edad in unique(dtModeloHOMBRES$Edad)) {
  datos_edadHOMBRES <- subset(dtModeloHOMBRES, Edad == edad)

  modeloHOMBRES <- glm(TasaMortalidad ~ Temperatura, data =
datos_edadHOMBRES,family = poisson(link="log"))

  modelosHOMBRES[[paste("Edad", edad)]] <- modeloHOMBRES
}

for (i in 1:50) {
  cat("Resumen del modelo para", names(modelosHOMBRES)[i], ":\n")
  print(summary(modelosHOMBRES[[i]]))
}

for (i in 1:length(modelosHOMBRES)) {
  cat("Resumen del modelo para", names(modelosHOMBRES)[i], ":\n")
  print(summary(modelosHOMBRES[[i]]))
}

vermodelo<-modelosHOMBRES[[59]]
print(summary(vermodelo))

```

```
# SERIES TEMPORALES DATOS ANUALES
```

```
#PAQUETES
```

```
install.packages("tidyverse")
```

```
install.packages("lubridate")
```

```
install.packages("car")
```

```
install.packages("urca")
```

```
install.packages("tseries")
```

```
install.packages("astsa")
```

```
install.packages("forecast")
```

```
install.packages("foreign")
```

```
install.packages("timsac")
```

```
install.packages("vars")
```

```
install.packages("lmtest")
```

```
install.packages("mFilter")
```

```
install.packages("dynlm")
```

```
install.packages("nlme")
```

```
install.packages("quantmod")
```

```
install.packages("xts")
```

```
library(fpp2)
```

```
library(matrixStats)
```

```
library(tseries)
```

```
library(gridExtra)
```

```
library(knitr)
```

```
library(readxl)
```

```
install.packages("grid.arrange")
```

```
library(lmtest)
```

```
library(tidyverse)
```

```
library(lubridate)
```

```
library(car)
```

```
library(urca)
```

```
library(tseries)
```

```
library(astsa)
```

```
library(forecast)
```

```

library(foreign)
library(timsac)
library(vars)
library(lmtest)
library(mFilter)
library(dynlm)
library(nlme)
library(quantmod)
library(xts)

#Lectura de datos
mortalidadST <- read.csv2(file.choose(), header = TRUE, skip = 0,)
Tabla_mortalidad=ts(mortalidadST)
temperaturaMaxST <- read.csv2(file.choose(), header = TRUE, skip = 0,)
TempMaxima=ts(temperaturaMaxST[c(1:52),2])
length(TempMaxima)

mortalidad_59<-ts(Tabla_mortalidad[c(22:74),6])
mortalidad_72<-ts(Tabla_mortalidad[c(22:74),7])
mortalidad_74<-ts(Tabla_mortalidad[c(22:74),8])
mortalidad_80<-ts(Tabla_mortalidad[c(22:74),2])
mortalidad_85<-ts(Tabla_mortalidad[c(22:74),3])
mortalidad_90<-ts(Tabla_mortalidad[c(22:74),4])
mortalidad_94<-ts(Tabla_mortalidad[c(22:74),5])

install.packages("graphics", dependencies = TRUE)

#ESTACIONARIO EN MEDIA
plotedad59=autoplot(mortalidad_59) +
  ggtitle("mortalidad_59") +
  ylab("q_59")+
  theme(plot.title = element_text(hjust = 0.5))

```

```
plottedad72=autoplot(mortalidad_72) +  
  ggtitle("mortalidad_72") +  
  ylab("q_72")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plottedad74=autoplot(mortalidad_74) +  
  ggtitle("mortalidad_74") +  
  ylab("q_74")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plottedad80=autoplot(mortalidad_80) +  
  ggtitle("mortalidad_80") +  
  ylab("q_80")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plottedad85=autoplot(mortalidad_85) +  
  ggtitle("mortalidad_85") +  
  ylab("q_85")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plottedad90=autoplot(mortalidad_90) +  
  ggtitle("mortalidad_90") +  
  ylab("q_90")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plottedad94=autoplot(mortalidad_94) +  
  ggtitle("mortalidad_94") +  
  ylab("q_94")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
quartz()  
grid.arrange(plottedad59, plottedad72, plottedad74,plottedad80, plottedad85,  
plottedad90,plottedad94, ncol = 3)
```

```
#Comprobar si son estacionarios
```

```
#Mortalidad
```

```
adf=c() #H0: No Estacionariedad
```

```
for (i in 2:8){
```

```
    adf[i]=adf.test((Tabla_mortalidad)[c(22:74),i])$p.value
```

```
}
```

```
Resultados_adf=data.frame(Adf.test=c(adf[2],adf[3],adf[4],adf[5],adf[6],adf[7],adf[8]))
```

```
rownames(Resultados_adf)<-
```

```
c("Edad_80","Edad_85","Edad_90","Edad_94","Edad_59","Edad_72","Edad_74")
```

```
Resultados_adf
```

```
#NR(H0) por lo tanto da indicio de que no son estacionarios.
```

```
pp=c() #H0->No Estacionariedad
```

```
for (i in 2:8){
```

```
    pp[i]=pp.test((Tabla_mortalidad)[c(22:74),i])$p.value
```

```
}
```

```
Resultados_pp=data.frame(pp.test=c(pp[2],pp[3],pp[4],pp[5],pp[6],pp[7],pp[8]))
```

```
rownames(Resultados_pp)<-
```

```
c("Edad_80","Edad_85","Edad_90","Edad_94","Edad_59","Edad_72","Edad_74")
```

```
Resultados_pp
```

```
#NR(H0) por lo tanto da indicio de que no son estacionarios.
```

```

kpss=c() #H0-> Estacionariedad

for (i in 2:8){

  kpss[i]=kpss.test((Tabla_mortalidad)[c(22:74),i])$p.value

}

Resultados_kpss=data.frame(kpss.test=c(kpss[2],kpss[3],kpss[4],kpss[5],kpss[6],kpss[7],kpss[8]))
rownames(Resultados_kpss)<-
c("Edad_80","Edad_85","Edad_90","Edad_94","Edad_59","Edad_72","Edad_74")
Resultados_kpss
#R(H0) por lo tanto da indicio de que no son estacionarios.
#aqui h0 es estacionario.

#Temperatura

TestKPSStemp<-kpss.test(TempMaxima)$p.value
#NR(H0) por lo tanto da indicio de que es estacionaria.
TestADFtemp<-adf.test(TempMaxima)$p.value
TestPPtemp<-pp.test(TempMaxima)$p.value

#Diferencias para que sean estacionarios
diferencias=c()

for (i in 2:8){

  diferencias[i]=ndiffs((Tabla_mortalidad)[c(22:74),i])

}

Resultados_logdiferencias=data.frame(Numero_de_diferencias=c(diferencias[2],diferencias[3],diferencias[4],diferencias[5],diferencias[6],diferencias[7],diferencias[8]))

```

```
rownames(Resultados_logdiferencias)<-  
c("Edad_80","Edad_85","Edad_90","Edad_94","Edad_59","Edad_72","Edad_74")  
Resultados_logdiferencias
```

```
Diferenciastemp<-ndiffs(TempMaxima)
```

```
#comprobar si tras diferenciar las series, son estacionarias
```

```
lgmortalidad_59<-log(mortalidad_59)  
lgmortalidad_72<-log(mortalidad_72)  
lgmortalidad_74<-log(mortalidad_74)  
ndiffs(lgmortalidad_59)
```

```
mortalidad_59_diff<-diff(lgmortalidad_59,lag=2)  
mortalidad_72_diff<-diff(lgmortalidad_72,lag=1)  
mortalidad_74_diff<-diff(lgmortalidad_74,lag=1)  
mortalidad_80_diff<-diff(mortalidad_80,lag=2)  
mortalidad_85_diff<-diff(mortalidad_85,lag=1)  
mortalidad_90_diff<-diff(mortalidad_90,lag=1)  
mortalidad_94_diff<-diff(mortalidad_94,lag=1)
```

```
mortalidad_59_diff<-ts(mortalidad_59_diff)  
mortalidad_72_diff<-ts(mortalidad_72_diff)  
mortalidad_74_diff<-ts(mortalidad_74_diff)  
mortalidad_80_diff<-ts(mortalidad_80_diff)  
mortalidad_85_diff<-ts(mortalidad_85_diff)  
mortalidad_90_diff<-ts(mortalidad_90_diff)  
mortalidad_94_diff<-ts(mortalidad_94_diff)
```

```
#Gráficos autoplot de los modelos
```

```
plot1_diff=autoplot(mortalidad_59_diff) +  
  ggtitle("mortalidad_59_diff") +  
  ylab("q_59")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plot2_diff=autoplot(mortalidad_72_diff) +  
  ggtitle("mortalidad_72_diff") +  
  ylab("q_72")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plot3_diff=autoplot(mortalidad_74_diff) +  
  ggtitle("mortalidad_74_diff") +  
  ylab("q_74")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plot4_diff=autoplot(mortalidad_80_diff) +  
  ggtitle("mortalidad_80_diff") +  
  ylab("q_80")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plot5_diff=autoplot(mortalidad_85_diff) +  
  ggtitle("mortalidad_85_diff") +  
  ylab("q_85")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plot6_diff=autoplot(mortalidad_90_diff) +  
  ggtitle("mortalidad_90_diff") +  
  ylab("q_90")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
plot7_diff=autoplot(mortalidad_94_diff) +  
  ggtitle("mortalidad_94_diff") +  
  ylab("q_94")+  
  theme(plot.title = element_text(hjust = 0.5))
```

```
quartz()  
grid.arrange(plot1_diff, plot2_diff, plot3_diff, plot4_diff, plot5_diff, plot6_diff,  
plot7_diff, ncol = 3)
```

```

adf_diff_59<-adf.test(mortalidad_59_diff)$p.value
adf_diff_72<-adf.test(mortalidad_72_diff)$p.value
adf_diff_74<-adf.test(mortalidad_74_diff)$p.value
adf_diff_80<-adf.test(mortalidad_80_diff)$p.value
adf_diff_85<-adf.test(mortalidad_85_diff)$p.value
adf_diff_90<-adf.test(mortalidad_90_diff)$p.value
adf_diff_94<-adf.test(mortalidad_94_diff)$p.value

```

```

Resultados_adf_diff <- data.frame(
  nombre = c("adf_diff_59", "adf_diff_72", "adf_diff_74", "adf_diff_80", "adf_diff_85",
"adf_diff_90", "adf_diff_94"),
  pvalor = c(adf_diff_59, adf_diff_72, adf_diff_74, adf_diff_80, adf_diff_85,
adf_diff_90, adf_diff_94)
)

```

```

kable(Resultados_adf_diff)

```

```

pp_diff_59<-pp.test(mortalidad_59_diff)$p.value
pp_diff_72<-pp.test(mortalidad_72_diff)$p.value
pp_diff_74<-pp.test(mortalidad_74_diff)$p.value
pp_diff_80<-pp.test(mortalidad_80_diff)$p.value
pp_diff_85<-pp.test(mortalidad_85_diff)$p.value
pp_diff_90<-pp.test(mortalidad_90_diff)$p.value
pp_diff_94<-pp.test(mortalidad_94_diff)$p.value

```

```

Resultados_pp_diff <- data.frame(
  nombre = c("pp_diff_59", "pp_diff_72", "pp_diff_74", "pp_diff_80", "pp_diff_85",
"pp_diff_90", "pp_diff_94"),
  pvalor = c(pp_diff_59, pp_diff_72, pp_diff_74, pp_diff_80, pp_diff_85, pp_diff_90,
pp_diff_94)
)

```

```

kable(Resultados_pp_diff)

```

```

kpss_diff_59<-kpss.test(mortalidad_59_diff)$p.value
kpss_diff_72<-kpss.test(mortalidad_72_diff)$p.value

```

```

kpss_diff_74<-kpss.test(mortalidad_74_diff)$p.value
kpss_diff_80<-kpss.test(mortalidad_80_diff)$p.value
kpss_diff_85<-kpss.test(mortalidad_85_diff)$p.value
kpss_diff_90<-kpss.test(mortalidad_90_diff)$p.value
kpss_diff_94<-kpss.test(mortalidad_94_diff)$p.value

```

```

Resultados_kpss_diff <- data.frame(
  nombre = c("kpss_diff_59", "kpss_diff_72", "kpss_diff_74", "kpss_diff_80",
"kpss_diff_85", "kpss_diff_90", "kpss_diff_94"),
  pvalor = c(kpss_diff_59, kpss_diff_72, kpss_diff_74, kpss_diff_80, kpss_diff_85,
kpss_diff_90, kpss_diff_94)
)

```

```

kable(Resultados_kpss_diff)

```

```

#Causalidad de Granger

```

```

TempMaxima=ts(temperaturaMaxST[c(3:54),2])

```

```

length(TempMaxima)

```

```

grangertest(mortalidad_85_diff~TempMaxima,orden=1)

```

```

#cambiando mortalidad por edad

```

```

library(lmtest)

```

```

install.packages("lmtest")

```

```

##### Proceso VaR #####

```

```

VaR59<-cbind(mortalidad_59_diff,TempMaxima)

```

```

VaR72<-cbind(mortalidad_72_diff,TempMaxima)

```

```

VaR74<-cbind(mortalidad_74_diff,TempMaxima)

```

```

VaR80<-cbind(mortalidad_80_diff,TempMaxima)

```

```

VaR85<-cbind(mortalidad_85_diff,TempMaxima)

```

```
VaR90<-cbind(mortalidad_90_diff,TempMaxima)
VaR94<-cbind(mortalidad_94_diff,TempMaxima)
library(vars)
```

```
modelo_var85 <- VAR(VaR85, lag.max = 1)
summary(modelo_var85)
BIC59<- BIC(modelo_var59)
AIC59<- AIC(modelo_var59)
```

```
summary(modelo_var94)
modelo_var72 <- VAR(VaR72, lag.max = 10, ic = "AIC")
modelo_var74<- VAR(VaR74, lag.max = 10, ic = "AIC")
modelo_var80 <- VAR(VaR80, lag.max = 10, ic = "AIC")
modelo_var85 <- VAR(VaR85, lag.max = 10, ic = "AIC")
modelo_var90 <- VAR(VaR90, lag.max = 10, ic = "AIC")
modelo_var94 <- VAR(VaR94, lag.max = 10, ic = "AIC")
```

```
modelo_var90 <- VAR(VaR90, lag.max = 2, ic = "AIC")
```

#Para seleccionar con cuantos retardos cada modelo es mejor CRITERIOS DE INFO:

AIC Y BIC

```
aic_values <- numeric(10)
for (lag in 1:10) {
  modelo_var <- VAR(VaR90, lag.max = lag, ic = "AIC")

  aic_values[lag] <- AIC(modelo_var)
}
print(aic_values)
```

#SERIES TEMPORALES DATOS DIARIOS 2022

#Lectura de datos

```
mortalidadST <- read.csv2(file.choose(), header = TRUE, skip = 0,)
```

```
Tabla_mortalidad=ts(mortalidadST)
```

```
temperaturaMaxST <- read.csv2(file.choose(), header = TRUE, skip = 0,)[,2]
```

```
TempMaxima=ts(temperaturaMaxST)[c(32:184)]
```

```
length(TempMaxima)
```

```
cor(TempMaxima,mortalidadM_85)
```

```
mortalidadM_85<-ts(Tabla_mortalidad[c(32:184),3])
```

```
mortalidadM_7584<-ts(Tabla_mortalidad[c(32:184),4])
```

```
mortalidadH_4564<-ts(Tabla_mortalidad[c(32:184),5])
```

```
mortalidadH_6574<-ts(Tabla_mortalidad[c(32:184),6])
```

```
install.packages("graphics", dependencies = TRUE)
```

#ESTACIONARIO EN MEDIA

```
plotM85=autoplot(mortalidadM_85) +
```

```
  ggtitle("mortalidad_Mujeres+85") +
```

```
  ylab("Mortalidad")+
```

```
  theme(plot.title = element_text(hjust = 0.5))
```

```
plotM7584=autoplot(mortalidadM_7584) +
```

```
  ggtitle("mortalidad_Mujeres75-84") +
```

```
  ylab("Mortalidad")+
```

```
  theme(plot.title = element_text(hjust = 0.5))
```

```
plot4564=autoplot(mortalidadH_4564) +
```

```
  ggtitle("mortalidad_Hombres45-64") +
```

```
  ylab("Mortalidad")+
```

```
  theme(plot.title = element_text(hjust = 0.5))
```

```

plot6574=autoplot(mortalidadH_6574) +
  ggtitle("mortalidad_65-74") +
  ylab("Mortalidad")+
  theme(plot.title = element_text(hjust = 0.5))

quartz()
grid.arrange(plotM85, plotM7584, plot4564,plot6574, ncol = 3)

#Comprobar si son estacionarios

## Mortalidad ##
adf=c() #HO: No Estacionariedad
for (i in 3:6){

  adf[i]=adf.test((Tabla_mortalidad)[c(32:184),i])$p.value

}

Resultados_adf=data.frame(Adf.test=c(adf[3],adf[4],adf[5],adf[6]))
rownames(Resultados_adf)<-c("M+85","M7584","H4564","H6574")
Resultados_adf

pp=c() #HO: No Estacionariedad
for (i in 3:6){

  pp[i]=pp.test((Tabla_mortalidad)[c(32:184),i])$p.value

}

Resultados_pp=data.frame(pp.test=c(pp[3],pp[4],pp[5],pp[6]))
rownames(Resultados_pp)<-c("M+85","M7584","H4564","H6574")
Resultados_pp

```

```

kpss=c()
for (i in 3:6){

  kpss[i]=kpss.test((Tabla_mortalidad)[c(32:184),i])$p.value

}

Resultados_kpss=data.frame(kpss.test=c(kpss[3],kpss[4],kpss[5],kpss[6]))
rownames(Resultados_kpss)<-c("M+85","M7584","H4564","H6574")
Resultados_kpss

#NR(H0) al 0,01 por lo tanto da indicio de que son estacionarios.
#aqui h0 es estacionario.

#Temperatura

TestKPSStemp<-kpss.test(TempMaxima)$p.value
#NR(H0) por lo tanto da indicio de que es estacionaria.
TestADFtemp<-adf.test(TempMaxima)$p.value
TestPPtemp<-pp.test(TempMaxima)$p.value

resultados_estacionariedad <- data.frame(
  Prueba = c("KPSS", "ADF", "PP"),
  P_Value = c(TestKPSStemp, TestADFtemp, TestPPtemp),
  stringsAsFactors = FALSE
)

# Muestra el data frame
print(resultados_estacionariedad)

```

```

#Diferencias para que sean estacionarios
Diferenciasmort<-ndiffs(mortalidadM_85)

datos_tabla <- data.frame(
  Serie = "Mortalidad M85",
  Dif_Optimas = Diferenciasmort,
  stringsAsFactors = FALSE
)

print(datos_tabla)

#comprobar si tras diferenciar las series, son estacionarias
mortalidad_M85_diff<-diff(mortalidadM_85,lag=1)
mortalidad_M85_diff<-ts(mortalidad_59_diff)

adf_diff_85<-adf.test(mortalidad_M85_diff)$p.value

pp_diff_85<-pp.test(mortalidad_M85_diff)$p.value

kpss_diff_85<-kpss.test(mortalidad_M85_diff)$p.value

resultados_pruebas <- data.frame(
  Prueba = c("ADF", "PP", "KPSS"),
  P_Value = c(adf_diff_85, pp_diff_85, kpss_diff_85),
  stringsAsFactors = FALSE
)

print(resultados_pruebas)

```

```
#Causalidad de Granger
```

```
#MUJER +85
```

```
TempMaxima=ts(temperaturaMaxST[c(33:184)])
```

```
length(mortalidad_M85_diff)
```

```
grangertest(mortalidad_M85_diff~TempMaxima,orden=1)
```

```
install.packages("vars")
```

```
library(vars)
```

```
#EL MODLEO VAR SE VE LUEGO PERO PARA COMPROBAR
```

```
var<-cbind(mortalidadM_7584,TempMaxima)
```

```
var_model <- VAR(var, lag.max = 1)
```

```
summary(var_model)
```

```
#MUJER 75-84
```

```
TempMaxima=ts(temperaturaMaxST[c(32:184)])
```

```
length(mortalidadM_7584)
```

```
grangertest(mortalidadM_7584~TempMaxima,orden=1)
```

```
#HOMBRE 45-64
```

```
TempMaxima=ts(temperaturaMaxST[c(32:184)])
```

```
length(mortalidadH_4564)
```

```
grangertest(mortalidadH_4564~TempMaxima,orden=1)
```

```

#HOMBRE 65-74
TempMaxima=ts(temperaturaMaxST[c(32:184)])

length(mortalidadH_6574)

grangertest(mortalidadH_6574~TempMaxima,orden=1)

##### Proceso VaR #####
#Variables a cambiar
VaRM85<-cbind(mortalidad_M85_diff,TempMaxima)
VaRM7584<-cbind(mortalidadM_7584,TempMaxima)
VaRH4564<-cbind(mortalidadH_4564,TempMaxima)
VaRH6574<-cbind(mortalidadH_6574,TempMaxima)
length(TempMaxima)

#Para seleccionar con cuantos retardos cada modelo es mejor CRITERIOS DE INFO:
AIC
aic_values <- numeric(5)
for (lag in 1:5) {
  modelo_var <- VAR(VaRH4564, lag.max = lag, ic = "AIC")

  aic_values[lag] <- AIC(modelo_var)
}
print(aic_values)

#BIC
bic_values <- numeric(5)
for (lag in 1:5) {
  modelo_var <- VAR(VaRH4564, lag.max = lag, ic = "AIC")

  bic_values[lag] <- BIC(modelo_var)
}
print(bic_values)

```

```

resultados <- data.frame(
  Lag = 1:5,
  AIC = aic_values,
  BIC = bic_values
)

print(resultados)

#Modelos con sus retardos
modeloM85 <- VAR(VaRM85, lag.max = 1)
summary(modeloM85)
modeloM7584 <- VAR(VaRM7584, lag.max = 1)
summary(modeloM7584)
modeloH4564 <- VAR(VaRH4564, lag.max = 1)
summary(modeloH4564)
modeloH6574 <- VAR(VaRH6574, lag.max = 1)
summary(modeloH6574)

##### Residuos #####
res_85 <- residuals(modeloM85)
res_7584 <- residuals(modeloM7584)
res_4564 <- residuals(modeloH4564)
res_6574 <- residuals(modeloH6574)

library(goftest)

##### 1-NORMALIDAD
#H0 es Normalidad
cvm_85 <- cvm.test(res_85, "pnorm", 0, sd(res_85), estimated = TRUE)
ad_85 <- ad.test(res_85, "pnorm", 0, sd(res_85), estimated = TRUE)
cvm_7584 <- cvm.test(res_7584, "pnorm", 0, sd(res_7584), estimated = TRUE)
ad_7584 <- ad.test(res_7584, "pnorm", 0, sd(res_7584), estimated = TRUE)
cvm_4564 <- cvm.test(res_4564, "pnorm", 0, sd(res_4564), estimated = TRUE)

```

```

ad_4564 <- ad.test(res_4564, "pnorm", 0, sd(res_4564), estimated = TRUE)
cvm_6574 <- cvm.test(res_6574, "pnorm", 0, sd(res_6574), estimated = TRUE)
ad_6574 <- ad.test(res_6574, "pnorm", 0, sd(res_6574), estimated = TRUE)
#NR(HO), cumplen con normalidad.

resultados <- data.frame(
  Grupo = c("85", "7584", "4564", "6574"),
  CVM_Statistic = c(cvm_85$statistic, cvm_7584$statistic, cvm_4564$statistic,
cvm_6574$statistic),
  CVM_p_value = c(cvm_85$p.value, cvm_7584$p.value, cvm_4564$p.value,
cvm_6574$p.value),
  AD_Statistic = c(ad_85$statistic, ad_7584$statistic, ad_4564$statistic,
ad_6574$statistic),
  AD_p_value = c(ad_85$p.value, ad_7584$p.value, ad_4564$p.value, ad_6574$p.value)
)

print(resultados)

##### 2-AUTOCORRELACION
library(stats)
library(tseries)

res_85_segunda_variable <- res_85[, 2]
length(res_85_segunda_variable)
res_85_primera_variable <- res_85[, 1]
tabla1_85 <- Box.test(res_85_primera_variable, lag = 10, type = "Ljung-Box")
tabla2_85 <- Box.test(res_85_segunda_variable, lag = 10, type = "Ljung-Box")

autocorrelacion_85=tabla_85$p.value

res_7584_segunda_variable <- res_7584[, 2]
length(res_7584_segunda_variable)
res_7584_primera_variable <- res_7584[, 1]
tabla1_7584 <- Box.test(res_7584_primera_variable, lag = 10, type = "Ljung-Box")

```

```

tabla2_7584 <- Box.test(res_7584_primera_variable, lag = 10, type = "Ljung-Box")

autocorrelacion_7584=tabla_7584$p.value

res_4564_segunda_variable <- res_4564[, 2]
length(res_4564_segunda_variable)
res_4564_primera_variable <- res_4564[, 1]
tabla1_4564 <- Box.test(res_4564_primera_variable, lag = 10, type = "Ljung-Box")
tabla2_4564 <- Box.test(res_4564_primera_variable, lag = 10, type = "Ljung-Box")

autocorrelacion_4564=tabla_4564$p.value

res_6574_segunda_variable <- res_6574[, 2]
length(res_6574_segunda_variable)
res_6574_primera_variable <- res_6574[, 1]
tabla1_6574 <- Box.test(res_6574_primera_variable, lag = 10, type = "Ljung-Box")
tabla2_6574 <- Box.test(res_6574_primera_variable, lag = 10, type = "Ljung-Box")

autocorrelacion_6574=tabla_6574$p.value
#NR(HO), no estan correlados

resultados <- data.frame(
  Grupo = c("85", "85", "7584", "7584", "4564", "4564", "6574", "6574"),
  Variable = c("Primera", "Segunda", "Primera", "Segunda", "Primera", "Segunda",
"Primera", "Segunda"),
  Statistic = c(tabla1_85$statistic, tabla2_85$statistic,
    tabla1_7584$statistic, tabla2_7584$statistic,
    tabla1_4564$statistic, tabla2_4564$statistic,
    tabla1_6574$statistic, tabla2_6574$statistic),
  p_value = c(tabla1_85$p.value, tabla2_85$p.value,
    tabla1_7584$p.value, tabla2_7584$p.value,
    tabla1_4564$p.value, tabla2_4564$p.value,
    tabla1_6574$p.value, tabla2_6574$p.value)
)

```

```

print(resultados)

install.packages("tseries")
library(tseries)

VaRM85<-cbind(mortalidad_M85_diff,TempMaxima)
VaRM7584<-cbind(mortalidadM_7584,TempMaxima)
VaRH4564<-cbind(mortalidadH_4564,TempMaxima)
VaRH6574<-cbind(mortalidadH_6574,TempMaxima)

##### 3-Heterocedasticidad BP. H0: Homocedasticidad
ec85<-data.frame(cbind(res_85_segunda_variable ^2,lag(mortalidad_M85_diff[1:151],-
1),lag(TempMaxima[1:151],-1)))
fitaux85<-lm(ec85$X1~ec85$X2+ec85$X3)
df<-2
bp.statistic85<-(length(res_85)-df)*summary(fitaux85)$r.squared
bp.pvalue85<-1-pchisq(bp.statistic85,df)
bp.pvalue_85=bp.pvalue85

ec7584<-data.frame(cbind(res_7584_segunda_variable
^2,lag(mortalidadM_7584[1:151],-1),lag(TempMaxima[1:151],-1)))
fitaux7584<-lm(ec7584$X1~ec7584$X2+ec7584$X3)
df<-2
bp.statistic7584<-(length(res_7584)-df)*summary(fitaux7584)$r.squared
bp.pvalue7584<-1-pchisq(bp.statistic7584,df)
bp.pvalue_7584=bp.pvalue7584

ec4564<-data.frame(cbind(res_4564_segunda_variable
^2,lag(mortalidadH_4564[1:151],-1),lag(TempMaxima[1:151],-1)))
fitaux4564<-lm(ec4564$X1~ec4564$X2+ec4564$X3)
df<-2
bp.statistic4564<-(length(res_4564)-df)*summary(fitaux4564)$r.squared
bp.pvalue4564<-1-pchisq(bp.statistic4564,df)

```

```
bp.pvalue_4564=bp.pvalue4564
```

```
ec6574<-data.frame(cbind(res_6574_segunda_variable  
^2,lag(mortalidadH_6574[1:151],-1),lag(TempMaxima[1:151],-1)))  
fitaux6574<-lm(ec6574$X1~ec6574$X2+ec6574$X3)  
df<-2  
bp.statistic6574<-(length(res_6574)-df)*summary(fitaux6574)$r.squared  
bp.pvalue6574<-1-pchisq(bp.statistic6574,df)  
bp.pvalue_6574=bp.pvalue6574
```

```
resultados <- c(bp.pvalue_85, bp.pvalue_7584, bp.pvalue_4564, bp.pvalue_6574)  
nombres <- c("M85", "M7584", "H4564", "H6574")  
tabla_resultados <- data.frame(Modelo = nombres, p_value = resultados)  
print(tabla_resultados)
```

```
#####
```

```
ec85<-data.frame(cbind(res_85_primera_variable ^2,lag(mortalidad_M85_diff[1:151],-  
1),lag(TempMaxima[1:151],-1)))  
fitaux85<-lm(ec85$X1~ec85$X2+ec85$X3)  
df<-2  
bp.statistic85<-(length(res_85)-df)*summary(fitaux85)$r.squared  
bp.pvalue85<-1-pchisq(bp.statistic85,df)  
bp.pvalue_85=bp.pvalue85
```

```
ec7584<-data.frame(cbind(res_7584_primera_variable  
^2,lag(mortalidadM_7584[1:151],-1),lag(TempMaxima[1:151],-1)))  
fitaux7584<-lm(ec7584$X1~ec7584$X2+ec7584$X3)  
df<-2  
bp.statstic7584<-(length(res_7584)-df)*summary(fitaux7584)$r.squared  
bp.pvalue7584<-1-pchisq(bp.statistic7584,df)  
bp.pvalue_7584=bp.pvalue7584
```

```
ec4564<-data.frame(cbind(res_4564_primera_variable  
^2,lag(mortalidadH_4564[1:151],-1),lag(TempMaxima[1:151],-1)))
```

```

fitaux4564<-lm(ec4564$X1~ec4564$X2+ec4564$X3)
df<-2
bp.statistic4564<-(length(res_4564)-df)*summary(fitaux4564)$r.squared
bp.pvalue4564<-1-pchisq(bp.statistic4564,df)
bp.pvalue_4564=bp.pvalue4564

ec6574<-data.frame(cbind(res_6574_primera_variable
^2,lag(mortalidadH_6574[1:151],-1),lag(TempMaxima[1:151],-1)))
fitaux6574<-lm(ec6574$X1~ec6574$X2+ec6574$X3)
df<-2
bp.statistic6574<-(length(res_6574)-df)*summary(fitaux6574)$r.squared
bp.pvalue6574<-1-pchisq(bp.statistic6574,df)
bp.pvalue_6574=bp.pvalue6574

resultados <- c(bp.pvalue_85, bp.pvalue_7584, bp.pvalue_4564, bp.pvalue_6574)
nombres <- c("M85", "M7584", "H4564", "H6574")
tabla_resultados <- data.frame(Modelo = nombres, p_value = resultados)
print(tabla_resultados)

##### 4-MEDIA
media_85=c()
qnorm(c(0.025,0.975),0,sd(res_85)/sqrt(length(res_85))) #Regi?n Cr?tica
intervalo_85=qnorm(c(0.025,0.975),0,sd(res_85)/sqrt(length(res_85)))
media_85=mean(res_85)

media_7584=c()
qnorm(c(0.025,0.975),0,sd(res_7584)/sqrt(length(res_7584))) #Regi?n Cr?tica
intervalo_7584=qnorm(c(0.025,0.975),0,sd(res_7584)/sqrt(length(res_7584)))
media_7584=mean(res_7584)

media_6574=c()
qnorm(c(0.025,0.975),0,sd(res_6574)/sqrt(length(res_6574))) #Regi?n Cr?tica
intervalo_6574=qnorm(c(0.025,0.975),0,sd(res_6574)/sqrt(length(res_6574)))
media_6574=mean(res_6574)

```

```

media_4564=c()
qnorm(c(0.025,0.975),0,sd(res_4564)/sqrt(length(res_4564))) #Regi?n Cr?tica
intervalo_4564=qnorm(c(0.025,0.975),0,sd(res_4564)/sqrt(length(res_4564)))
media_4564=mean(res_4564)

```

```

resultados <- c(media_85, media_7584, media_6574, media_4564)
nombres <- c("M85", "M7584", "H4564", "H6574")
tabla_resultados <- data.frame(Modelo = nombres, p_value = resultados)
print(tabla_resultados)
install.packages("knitr")
library(knitr)

```

```

Resultados_Mujer85=data.frame(Resultados_85=c(autocorrelacion_85,cvm_85$p.value
,ad_85$p.value,bp.pvalue_85,intervalo_85[1],media_85,intervalo_85[2]))
rownames(Resultados_Mujer85)<-
c("autocorrelacion","cvm","ad","bp.pvalue","limite_inferior","media","limite_superior"
)
kable(Resultados_Mujer85)

```

```

Resultados_Mujer7584=data.frame(Resultados_7584=c(autocorrelacion_7584,cvm_758
4$p.value,ad_7584$p.value,bp.pvalue_7584,intervalo_7584[1],media_7584,intervalo_7
584[2]))
rownames(Resultados_Mujer7584)<-
c("autocorrelacion","cvm","ad","bp.pvalue","limite_inferior","media","limite_superior"
)
kable(Resultados_Mujer7584)

```

```

Resultados_Hombre6574=data.frame(Resultados_6574=c(autocorrelacion_6574,cvm_6
574$p.value,ad_6574$p.value,bp.pvalue_6574,intervalo_6574[1],media_6574,intervalo
_6574[2]))

```

```
rownames(Resultados_Hombre6574)<-
c("autocorrelacion","cvm","ad","bp.pvalue","limite_inferior","media","limite_superior"
)
kable(Resultados_Hombre6574)
```

```
Resultados_Hombre4564=data.frame(Resultados_4564=c(autocorrelacion_4564,cvm_4
564$p.value,ad_4564$p.value,bp.pvalue_4564,intervalo_4564[1],media_4564,intervalo
_4564[2]))
```

```
rownames(Resultados_Hombre4564)<-
c("autocorrelacion","cvm","ad","bp.pvalue","limite_inferior","media","limite_superior"
)
kable(Resultados_Hombre4564)
```

```
##### Funcion respuesta impulso #####
```

```
#Como la temperatura sí que causa un impacto en la mortalidad (granger)
```

```
#Impulso respuesta de la mortalidad ante una innovacion en la temperatura
```

```
varM85_IRF <- irf(modeloM85, response = "mortalidad_M85_diff", n.ahead = 8, boot =
TRUE)
```

```
plot(varM85_IRF)
```

```
varM7584_IRF <- irf(modeloM7584, response = "mortalidadM_7584", n.ahead = 8, boot
= TRUE)
```

```
plot(varM7584_IRF)
```

```
varh4564_IRF <- irf(modeloH4564, response = "mortalidadH_4564", n.ahead = 8, boot
= TRUE)
```

```
plot(varh4564_IRF)
```

```
varH6574_IRF <- irf(modeloH6574, response = "mortalidadH_6574", n.ahead = 8, boot
= TRUE)
```

```
plot(varH6574_IRF)
```

```

#Descomposición de la varianza ante una innovacion en la temp
var85_var<-fevd(modeloM85,n.ahead=50)$mortalidad_M85_diff

var7584_var<-fevd(modeloM7584,n.ahead=50)$mortalidadM_7584

var4564_var<-fevd(modeloH4564,n.ahead=50)$mortalidadH_4564

var6574_var<-fevd(modeloH6574,n.ahead=50)$mortalidadH_6574

#Se observa que aunque la tempertura afecte lo que mas le afecta a su mortalidad es ella
misma

##### Prediccion #####

#M85
fcast_VAR <- predict(modeloM85, n.ahead = 153)
prediccion_dia_siguiete <- fcast_VAR$fcst$mortalidad[, "fcst"][153]
prediccion_discreta <- round(prediccion_dia_siguiete)
plot(fcast_VAR)
mean_mortalidad <- fcast_VAR$fcst$mortalidad[, "fcst"]
l_mortalidad <- fcast_VAR$fcst$mortalidad[, "lower"][153]
u_mortalidad <- fcast_VAR$fcst$mortalidad[, "upper"][153]

# Value at Risk al 99%
VaR <- qnorm(0.99, mean = mean_mortalidad, sd = sd_mortalidad)

#M7584
fcast_VAR <- predict(modeloM7584, n.ahead = 153)
prediccion_dia_siguiete <- fcast_VAR$fcst$mortalidad[, "fcst"][153]

plot(fcast_VAR)
mean_mortalidad <- fcast_VAR$fcst$mortalidad[, "fcst"]
l_mortalidad <- fcast_VAR$fcst$mortalidad[, "lower"]
u_mortalidad <- fcast_VAR$fcst$mortalidad[, "upper"]

```

```
# Value at Risk al 99%  
VaR <- qnorm(0.99, mean = mean_mortalidad, sd = sd_mortalidad)
```

```
#H4564  
fcast_VAR <- predict(modeloH4564, n.ahead = 153)  
plot(fcast_VAR)  
mean_mortalidad <- fcast_VAR$fcst$mortalidad[, "fcst"]  
l_mortalidad <- fcast_VAR$fcst$mortalidad[, "lower"]  
u_mortalidad <- fcast_VAR$fcst$mortalidad[, "upper"]
```

```
# Value at Risk al 99%  
VaR <- qnorm(0.99, mean = mean_mortalidad, sd = sd_mortalidad)
```

```
#H6574  
fcast_VAR <- predict(modeloH6574, n.ahead = 360)  
plot(fcast_VAR)  
mean_mortalidad <- fcast_VAR$fcst$mortalidad[, "fcst"]  
l_mortalidad <- fcast_VAR$fcst$mortalidad[, "lower"]  
u_mortalidad <- fcast_VAR$fcst$mortalidad[, "upper"]
```

```
# Value at Risk al 99%  
VaR <- qnorm(0.99, mean = mean_mortalidad, sd = sd_mortalidad)
```