

Estudio del impacto de las altas temperaturas en la mortalidad diaria en España

Máster Universitario en Modelización e
Investigación Matemática, Estadística y
Computación



Universidad
Zaragoza

Alonso Pueyo Gómez
Trabajo de fin de máster
Universidad de Zaragoza
Tutora: Jaione Etxeberria

26 de junio de 2025

Resumen

Índice general

Resumen	III
1. Introducción	1
2. Marco Teórico	3
3. Modelo y aplicación	7
Bibliografía	17

Capítulo 1

Introducción

Capítulo 2

Marco Teórico

La relación que existe entre la exposición a una temperatura extremadamente alta y la mortalidad diaria no es lineal. El efecto que conlleva el exceso de temperatura no se manifiesta de una manera inmediata. Por el contrario, existe un cierto retardo, y los efectos secundarios de dicha exposición pueden tardar un tiempo en ocurrir. Cuando se quiere estudiar la relación entre un suceso y su consiguiente efecto, conociendo que puede existir un retardo, se busca definir un modelo no lineal de retardo distribuido, o Distributed Lag Non-Linear Model en inglés (DLNM). Este tipo de modelos representan la relación no lineal entre la exposición al factor que se está estudiando y la respuesta que se produce, así como el retardo correspondiente. Fueron introducidos por primera vez por A. Gasparrini, B. Armstrong y M.G. Kenward en un trabajo del año 2010 donde se incluyen todos los detalles y hasta un ejemplo que estudia el efecto de la temperatura en la ciudad de Nueva York. Más adelante, Antonio Gasparrini ha seguido haciendo estudios con este tipo de modelos, creando también la librería *dlnm* del programa RStudio, que emplearemos a lo largo de este trabajo. Los modelos DLNM han sido útiles en multitud de proyectos a nivel internacional.

Se define una dimensión temporal adicional a través de lo que se conoce como base cruzada, un espacio bidimensional de funciones que describe tanto la forma de la relación como la dimensión de retraso, y que determina la estructura temporal del efecto estudiado.

Un método más sencillo que el que vamos a emplear nosotros es el empleo de modelos de retardo distribuido (DLM). La principal diferencia es que estos modelos no estudian la no linealidad que existe entre el efecto y las consecuencias posteriores. Sin embargo, resultan muy útiles cuando se está estudiando la estructura de retardo de efectos lineales. Cuando no existe tal linealidad, tienen limitaciones. Una solución a esto consiste en relajar los supuestos y extender la metodología a los DLNM. Queremos emplear este tipo de métodos cuando la relación es no lineal. Los modelos no lineales de retardo distribuidos a los que queremos llegar unifican los modelos anteriores más sencillos y sus variantes en un único tipo de modelización.

Ahora vamos a plantear el modelo básico general que se emplea. En nuestro caso estamos estudiando las muertes diarias producidas por un exceso de calor. Se trata de un modelo jerárquico. Primero, los datos son conteos que siguen una distribución de Poisson, y después, vamos a modelizar la media. Por consiguiente, emplearemos la regresión de Poisson. Así, el modelo general para cada dato Y_t , con $t = 1, \dots, n$ viene dado por la siguiente expresión:

$$g(\mu_t) = \alpha + \sum_{j=1}^J s_j(x_{tj}, \beta_j) + \sum_{k=1}^K \gamma_k u_{tk} \quad (2.1)$$

Se tiene que cada μ_t es la esperanza de las variables Y_t y la función g es una función de enlace monótona. Los parámetros β_j son los predictores lineales, y las funciones s_j demuestran la relación suavizada entre la variable x_{tj} y su correspondiente parámetro β_j . En otras palabras, esa suma estudia la linealidad entre los datos que se están estudiando y las variables que los explican. El parámetro α es el intercepto, una constante que nos sirve para ajustar. Por último, las variables u_{tk} incluyen otros predictores con efectos lineales que vienen especificados por los coeficientes γ_k .

Como ya hemos mencionado con anterioridad, los datos Y_t son conteos diarios que siguen una distribución de Poisson. Por tanto, se tiene que $E(Y_t) = \text{Var}(Y_t) = \mu_t$. Además, la función g monótona que hemos mencionado en (2.1) resulta ser la función logaritmo. El primer sumatorio, que incluía las relaciones lineales entre las variables x_{tj} y sus parámetros, se puede englobar una función $s(x_t)$ que explica la relación entre x_t y $g(\mu_t)$. Se trata de una suma de términos lineales. Para ello, se tiene que elegir un espacio de funciones donde s sea uno de sus elementos. Este espacio se conocerá como base. Las funciones de esta base transforman las variables originales x_t en otras nuevas que conoceremos como variables de base. Para describir los efectos no lineales del exceso de temperatura sobre la salud, emplearemos funciones básicas como polinomios o splines.

La función s de la que hemos estado hablando se puede escribir de forma vectorial como sigue:

$$s(x_t, \beta) = \sum_{j=1}^J s_j(x_{tj}, \beta_j) = z_t^T \cdot \beta \quad (2.2)$$

donde z_t^T es la fila t -ésima de la matriz base Z de tamaño $n \times v_x$, y contiene las funciones de base que se han aplicado a las variables originales x_t .

Recordar que no solo se tienen que estudiar las relaciones lineales o no lineales de las variables, sino que además existe un cierto retardo. Las muertes que se producen diariamente debido a las altas temperaturas no se producen repentinamente. Notar que, en nuestro caso, no todas las muertes ocurren exactamente tras un desfase de T días. Dependiendo del individuo y de las circunstancias de calor a la que fue expuesto, este desfase puede ser mayor o menor. Por consiguiente, se tienen que imponer ciertas restricciones a los modelos, haciéndolos más complejos, pero a su vez más fiables.

Bajo efectos retardados, el resultado Y_t se puede explicar en términos de exposiciones pasadas x_{t-l} , donde l es el retardo. Dado un vector de exposiciones ordenadas x , le aplicamos una transformación para obtener una matriz Q de tamaño $n \times (L+1)$ cuyas columnas tienen la siguiente forma:

$$q_t = (x_t, x_{t-1}, \dots, x_{t-l}, \dots, x_{t-L})^T \quad (2.3)$$

donde x_t es la observación en un instante de tiempo t , y x_{t-k} es una observación pasada con un desfase k . Además, L se conoce como retraso máximo. Notar que la primera columna de la matriz Q es el vector original x sin desfases. Es decir, $q_{\cdot 1} = x$. Definimos también un vector de retrasos para cada columna de la matriz Q dado por $\mathbf{l} = (0, \dots, l, \dots, L)^T$.

Como ya hemos mencionado con anterioridad, para estudiar modelos lineales bajo efectos de retardo se suelen emplear los DLM. El modelo más sencillo es el que no posee ninguna restricción, y es el que solo tiene definido un parámetro para cada retardo. Como la estimación de

los parámetros no suele ser muy buena en estos casos debido a la alta correlación, se comienzan a imponer ciertas restricciones. Por ejemplo, se pueden pensar en modelos donde se asume un efecto constante en cada intervalo entre retardos. Un caso especial de DLM es aquel que tiene como predictor una media móvil de exposiciones de los últimos L días.

Para poder modelar correctamente los conceptos necesarios, tenemos que describir las relaciones no lineales existentes en el espacio del predictor a lo largo de los retardos. Lo primero de todo es definir lo que se conoce como base cruzada. Se trata de un espacio de funciones de dos dimensiones que describen tanto la forma de relación como los efectos de los retrasos. Para elegir una base cruzada hay que seleccionar dos conjuntos de funciones de base, que combinadas se conocerán como funciones de base cruzada.

Ahora, con la teoría vista hasta el momento, vamos a definir de manera general un DLM donde la forma de los efectos distribuidos a lo largo de los retardos se explica sobre una base adecuada. En forma matricial se tiene lo siguiente:

$$s(x_t, \eta) = z_t^T C \eta \quad (2.4)$$

donde C es una matriz de variables de base de obtenidas al aplicar las correspondientes funciones de base al vector de retardo l , de tamaño $(L + 1) \times v_l$, y η es un vector de parámetros desconocidos.

La dimensión adicional que hemos definido le otorga a la matriz C una estructura que le permite describir los efectos de las exposiciones retardadas. Los distintos tipos de DLM de los que hemos hablado se engloban en el modelo anterior, y se obtienen según C . Cuando C es un vector de unos se especifica el modelo con la media móvil. Cuando C es la matriz identidad, se habla del DLM sin restricciones. Si C se define como una serie de polinomios o splines, se define un DLM que describe el efecto de una curva suavizada a lo largo de los retardos.

De lo anterior, podemos definir también la siguiente matriz:

$$W = QC \quad (2.5)$$

Se trata de la matriz de las variables después de aplicar las transformaciones anteriores, y que nos permitirá estimar los parámetros η .

A partir de los parámetros estimados $\hat{\eta}$ podemos construir los efectos lineales β de cada desfase. Se hace lo siguiente:

$$\hat{\beta} = C \hat{\eta} \quad (2.6)$$

A continuación, vamos a modelar la forma de la relación en los dos espacios en los que estamos trabajando. Para ello, aplicamos las dos transformaciones que hemos descrito. En primer lugar, como en 2.2, elegimos las variables de base para definir la dependencia en el espacio del predictor. Con ello, obtendremos el vector z . Ahora, creamos la dimensión de retraso adicional. Hacemos lo mismo que en 2.3 pero para las variables que hemos almacenado en z . Como resultado, se obtiene un array \bar{R} de tamaño $n \times v_x \times (L + 1)$ que representa los retardos de todas las variables de base de x .

Si definimos la matriz C , que recordemos que es la matriz de variables base para l que hemos visto en 2.4, un DLNM se puede escribir como sigue:

$$s(x_t, \eta) = \sum_{j=1}^{v_x} \sum_{k=1}^{v_l} r_{tj}^T \cdot C \cdot k \eta_{jk} = w_t^T \eta \quad (2.7)$$

donde r_{tj} es el vector de exposiciones retardadas para el tiempo t transformado por la función base j . El vector w_t es el que resulta de aplicar las $v_x \cdot v_l$ funciones de la base cruzada a x_t , parecido a lo que ocurre en 2.5. Igualmente, también se podría llegar a una forma compacta para la matriz W como un producto tensorial, que requiere emplear productos de Kronecker y de Hadamard.

Capítulo 3

Modelo y aplicación

Para aplicar los modelos DLNM emplearemos RStudio. En concreto, la librería *dlnm* contiene todo lo necesario. Para crear la base cruzada se utiliza el comando *crossbasis*. Entre los argumentos que hay que incluir en esta función se encuentran las observaciones de los datos que se pretenden estudiar (en nuestro caso, la temperatura), el retardo a tener en cuenta, y la manera en la que modelamos el DLNM. Más adelante veremos distintas formas de hacer eso.

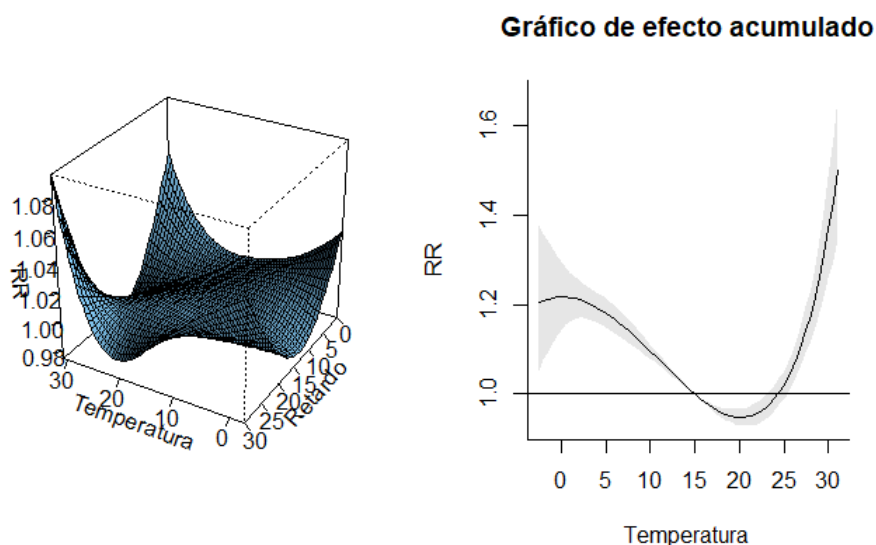
Después de definir la base cruzada, hay que crear el modelo que vamos a utilizar para estudiar la variable explicada (en nuestro caso, las defunciones observadas diariamente) en función de la base que se ha creado anteriormente. En general, se usará la regresión de quasi-Poisson ya que se tratan de datos contables y además existe una sobredispersión, puesto que hay varias defunciones en el mismo día.

Por último, se realizan las predicciones. Para ello se escribe la función *crosspred*. Dentro de sus argumentos hay que introducir el modelo y la base anteriores. Existen distintas formas de mostrar estas predicciones. Con la función *plot* se dibuja una superficie tridimensional que muestra la evolución de la relación entre la temperatura y el retardo. Además en esta función se puede introducir un argumento para cambiar el tipo de gráfico que sale. Si se pone *overall* se dibuja una curva que muestra el riesgo total acumulado a lo largo de todos los días de retardo. Con el argumento *slices* se pueden dibujar varias cosas. En particular, si añadimos un argumento *lag* se dibujan las curvas para distintos valores de la temperatura bajo días fijos. Por el otro lado, si se añade el argumento *var* se muestran curvas de distintos días bajo una temperatura fija. Por último, con el argumento *contour* se muestra lo mismo que la superficie tridimensional pero con un mapa de calor.

Una vez hemos visto los pasos a seguir para estudiar un modelo DLNM, vamos a crear el primero. Comenzamos leyendo los datos que son necesarios para el análisis. Los datos de mortalidad y defunciones diarias según las distintas comunidades autónomas o grupos de edad son sacados de la página web del Sistema de Monitorización de la Mortalidad Diaria (MoMo). Estos datos no contienen las temperaturas diarias, así que hay que conseguirlas de otras fuentes. Como vamos a estudiar lo que ocurre en la región de Navarra, de la página web de Meteo Navarra obtenemos las temperaturas máximas, mínimas y medias diarias. En concreto, los datos son de la estación automática de Pamplona GN. Esta estación no incluye los valores del 2021, y es por ello que los datos de ese año fueron sacados de la estación manual Pamplona MAN. Tampoco están incluidos los datos del año 2025.

Después de leer todos los datos, hacemos un primer análisis para evaluar si existen datos faltantes, o hay algo que no cuadra. Nos damos cuenta de que los datos de la temperatura media del 2020 no están, así que después de pasar todas las columnas a variables numéricas, la calculamos como la suma de la temperatura máxima y mínima de cada día dividido entre 2, que estos datos sí se conocen. Notamos que después de todo lo anterior, todavía tenemos una totalidad de 7 días que no conocemos sus valores.

Ya tenemos todos los datos requeridos para comenzar a estudiar los modelos, pero antes vamos a incluir todo dentro de la misma base de datos, para que no haya complicaciones más adelante. Con la función *filter* de la librería *dplyr* seleccionamos de los datos de MoMo las filas que sean de Navarra. Después, con la función *merge* de esa misma librería unimos bajo el mismo data frame los datos de Navarra y las temperaturas diarias. Ahora sí que podemos comenzar a hacer modelos. Para empezar, creamos la base cruzada, donde estudiamos la temperatura media diaria con un retardo de 30 días. El impacto de la temperatura se estudia mediante B-splines con 4 grados de libertad, y el retardo mediante splines naturales también con 4 grados de libertad. Como modelo, hacemos regresión de quasi-Poisson estudiando las defunciones observadas según la base cruzada anterior. Si hacemos las predicciones de este modelo se obtienen las siguientes gráficas:



En la primera imagen se muestra una superficie que muestra cómo aumenta o disminuye el Riesgo Relativo (RR) según el valor de la temperatura y de los días de retardo. Por ejemplo, vemos que en las esquinas, cuando la temperatura es muy alta o muy baja, el riesgo aumenta, aunque tampoco muy significativamente. El resto de la superficie es bastante plana, y el valor del riesgo se mantiene más o menos constante, incluso incluyendo algún punto donde el RR es menor que 1, luego hay un menor riesgo de mortalidad.

En la segunda gráfica se ve el efecto acumulado, esto es, cómo varía el Riesgo Relativo en función de la temperatura, acumulando el efecto a lo largo de todos los días de retardo. Además, en gris podemos apreciar un intervalo de confianza que indica posible variabilidad. Se llegan a las mismas conclusiones que hemos mencionado al visualizar la superficie. Cuando las temperaturas son altas o bajas, el riesgo es mayor, pero tampoco demasiado. El resto de la curva es próximo a 1, y cuando se rondan los 20 grados centígrados, el riesgo llega a disminuir.

Hemos elegido este primer modelo así porque sabemos que el tipo de relación entre la mortalidad y el calor no es lineal. Por tanto empleamos splines con la idea de modelar una curva. Tener 4 grados de libertad nos permite que el modelo encuentre un equilibrio entre el ajuste de los datos y la sencillez. Vamos a comprobar esto haciendo un modelo que todavía tenga menos grados de libertad, y otro que tenga 10, muchos más. Los gráficos obtenidos son:

Gráfico de efecto acumulado

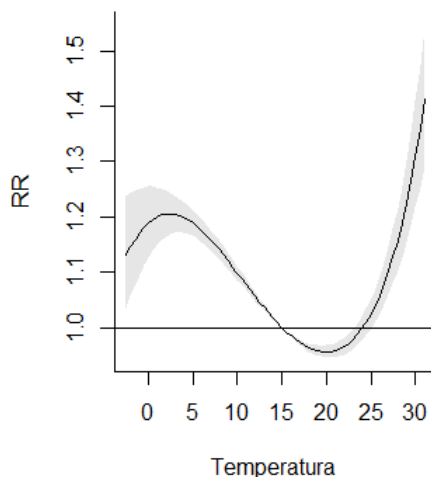
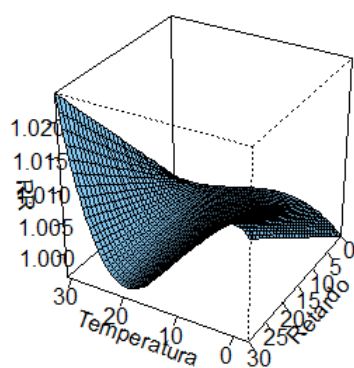
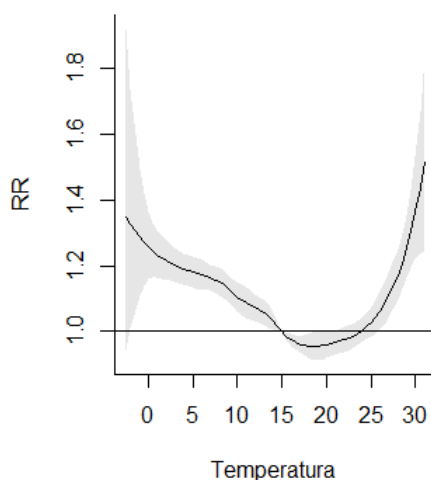
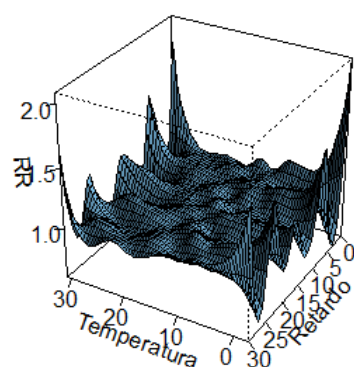


Gráfico de efecto acumulado



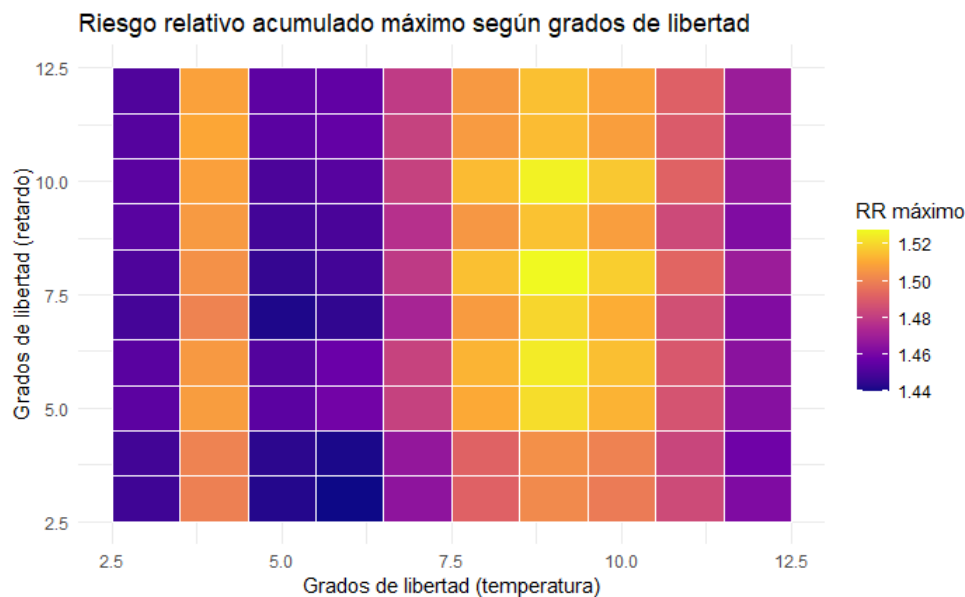
Los dos gráficos de arriba corresponden a un modelo con 3 grados de libertad para los splines que definen la temperatura y 2 grados de libertad para los splines naturales que modelan el retardo. Es el menor valor que pueden tomar. Vemos que ahora se tiene una superficie mucho más sencilla que la del modelo anterior. El Riesgo Relativo es ligeramente mayor en los valores extremos de la temperatura cuando el retardo es el máximo posible, de 30 días. La curva de efectos acumulados es bastante parecida a la que se ha analizado previamente.

Por otra parte, los gráficos de abajo son los de un modelo en el que los splines que modelan la temperatura y el retardo tienen 10 grados de libertad. Ahora la superficie no es tan plana. De hecho, es rugosa, con varios picos en las paredes. Además, notamos que ahora el Riesgo Relativo sí que aumenta mucho más en esas montañas mencionadas. El gráfico de efectos acumulados muestra una curva que toma los mayores valores en los extremos.

Notamos que el modelo con un menor número de grados de libertad es demasiado simple, y no es capaz de darnos buenas predicciones. Por otro lado, el modelo con muchos más grados de

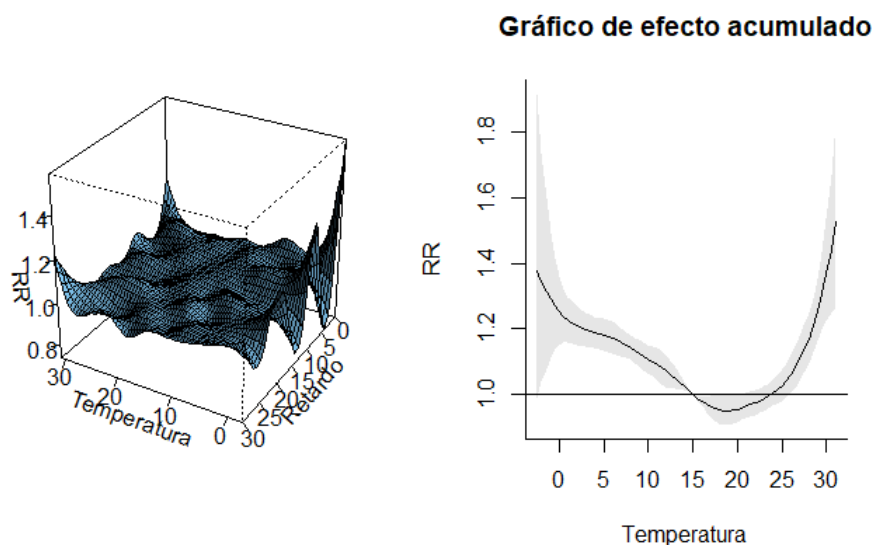
libertad es muy complejo, y se ajusta demasiado a los datos conocidos, luego tampoco nos sirve para una predicción adecuada. Por consiguiente, hay que encontrar un equilibrio, un modelo que no sea demasiado simple ni complejo.

A continuación vamos a proceder a hacer un análisis de sensibilidad en el que crearemos varios modelos como los anteriores, pero modificando tanto los grados de libertad correspondientes a la temperatura como los grados de libertad del retardo. Para cada uno de estos modelos se estudiará cómo evoluciona el riesgo relativo y de esta manera conocer cuáles son los posibles mejores modelos. Nos fijamos en el resultado, pero también hacemos un mapa de calor donde cada casilla representa un modelo con diferentes grados de libertad en la temperatura y el retardo, y cuanto más clara sea, mayor riesgo relativo acumulado tiene. Así, se obtiene lo siguiente:



Observamos que los mayores valores se obtienen con los modelos que poseen entre 8 y 10 grados de libertad para la temperatura, y entre 5 y 10 grados de libertad para el retardo. En particular, el mayor valor se obtiene con el modelo de 9 grados de libertad para la temperatura y 8 para el retardo. Por consiguiente, nos quedaremos con este modelo.

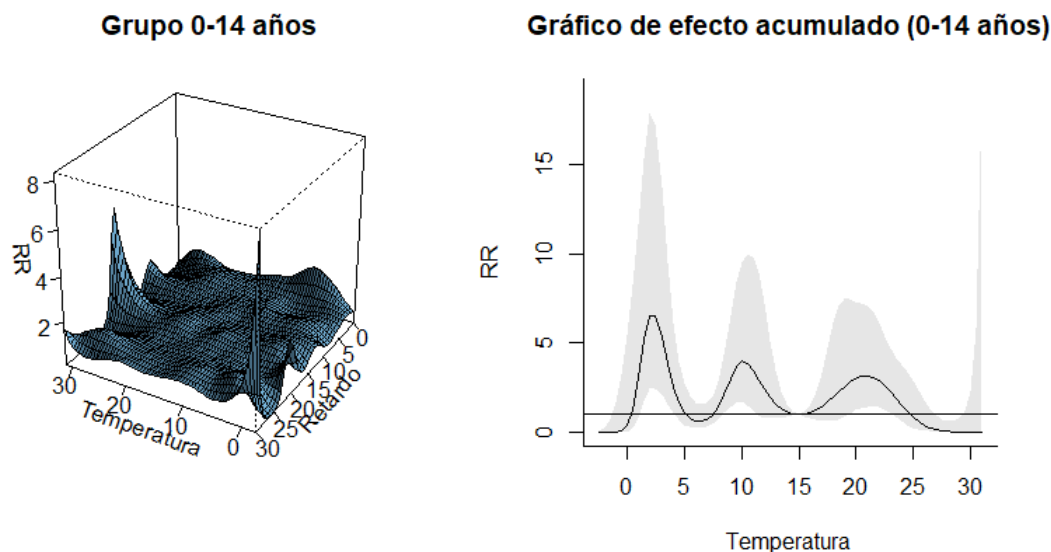
Si hacemos la predicción y los gráficos correspondientes se obtiene:



El análisis es el que ya hemos comentado en algún modelo previo. La superficie es rugosa, con altibajos, y los mayores valores se alcanzan en los extremos, no haciendo falta mucho retardo para ello. Si nos fijamos en la curva, el riesgo relativo es inferior a 1 cuando hay temperaturas suaves, entre los 15 y los 25 grados centígrados, y va aumentando conforme alcanzamos temperaturas más extremos. Cuando se llega a los 30 grados centígrados, el valor del riesgo relativo se acerca a 1.6, lo que ya es bastante.

Una vez tenemos un buen modelo que nos permite estudiar el efecto de la temperatura en la mortalidad correctamente, vamos a realizar un estudio según los grupos de población. En particular, dividiremos nuestros datos según la edad de las personas. La población se divide en los siguientes grupos: niños (0-14 años), jóvenes (15-44 años), adultos (45-64 años), mayores (65-74 años), más mayores (75-84 años) y abuelos (85 años o más). Los datos cuyo grupo de edad se corresponde con 'all' no se tienen en cuenta. Para cada uno de estos grupos vamos a crear la base y aplicar el modelo obtenido anteriormente para hacer la predicción. Dibujaremos la superficie y el gráfico de efecto acumulado y estudiaremos qué está ocurriendo.

Comenzando con el grupo de niños, se obtiene lo siguiente:



Si nos fijamos en la superficie vemos que, con un retardo de unos 20 días, el riesgo relativo que se llega a alcanzar en valores extremos de frío y de calor toma valores muy grandes, llegando incluso a superar el valor de 6 cuando la temperatura supera los 30 grados centígrados. Además, la curva del efecto acumulado para este grupo sube y baja de manera constante, alcanzando un riesgo de 0 en los extremos. Esto no parece muy coherente, y es posible que se deba a que no hay datos suficientes.

Si contamos el número de defunciones observadas según los grupos de edad obtenemos lo siguiente:

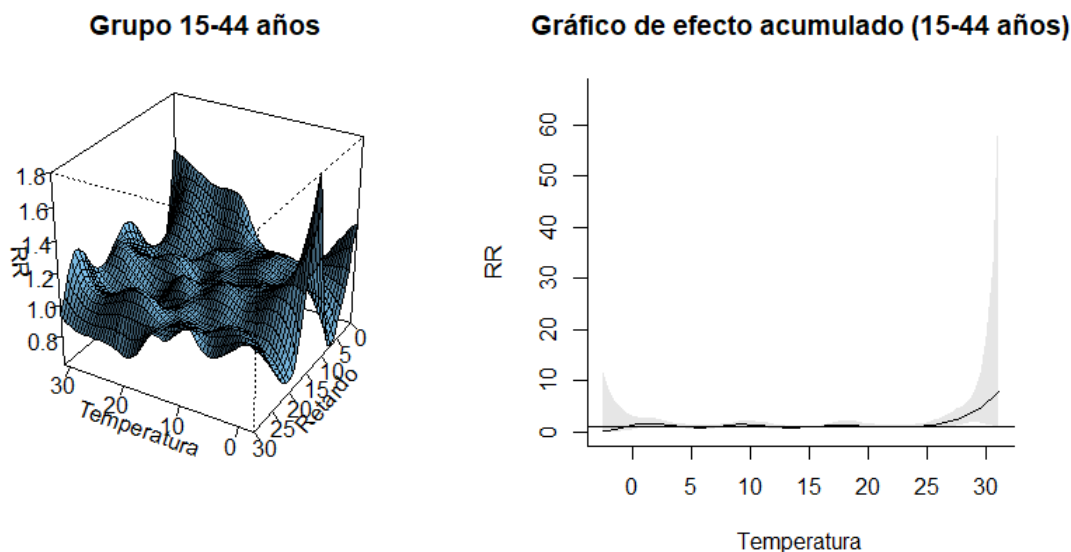
+85	0-14	15-44	45-64	65-74	75-84
62612.43	403.72	2010.13	12302.83	123006.61	29776.21

La fila de arriba corresponde a los distintos grupos, y la de abajo se trata de la suma de todas las defunciones observadas para cada grupo. Como intuíamos, solamente hay cerca de 404 defunciones para el grupo de niños entre los 0 y los 14 años, frente a las miles de defunciones que hay en otros grupos. El grupo de jóvenes que tienen entre 15 y 44 años es el siguiente que tiene menos defunciones, con unas 2010, pero ya son cuatro veces más que el de los niños. Por otro lado, el grupo con más defunciones es el de mayores (entre 65 y 74 años) con unas 123006.

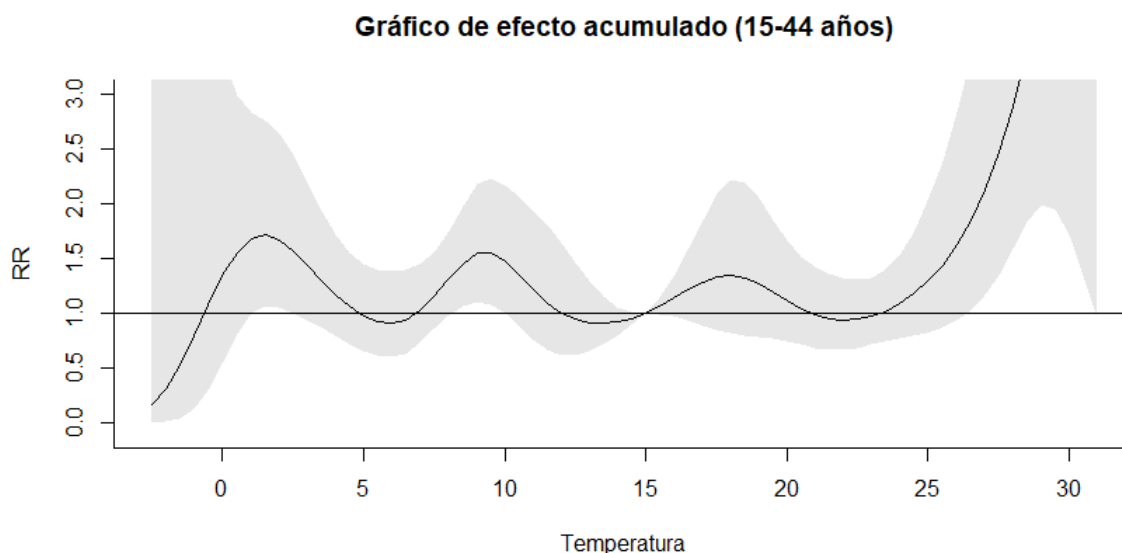
Lo que nos quiere decir esto es que el grupo con una mayor mortalidad, y al que habrá que prestar especial atención es el de los mayores. También podemos confirmar lo que ya sospechábamos, y es que los niños no son tan propensos a fallecer, y menos por el efecto de la temperatura. Por consiguiente, no existen datos suficientes, y tampoco tiene mucho sentido seguir estudiando este grupo.

Pasamos a continuación con el grupo de jóvenes de entre 15 y 44 años. Al igual que antes, tampoco contamos con muchas defunciones de este grupo de población, así que los gráficos que

se obtienen son:

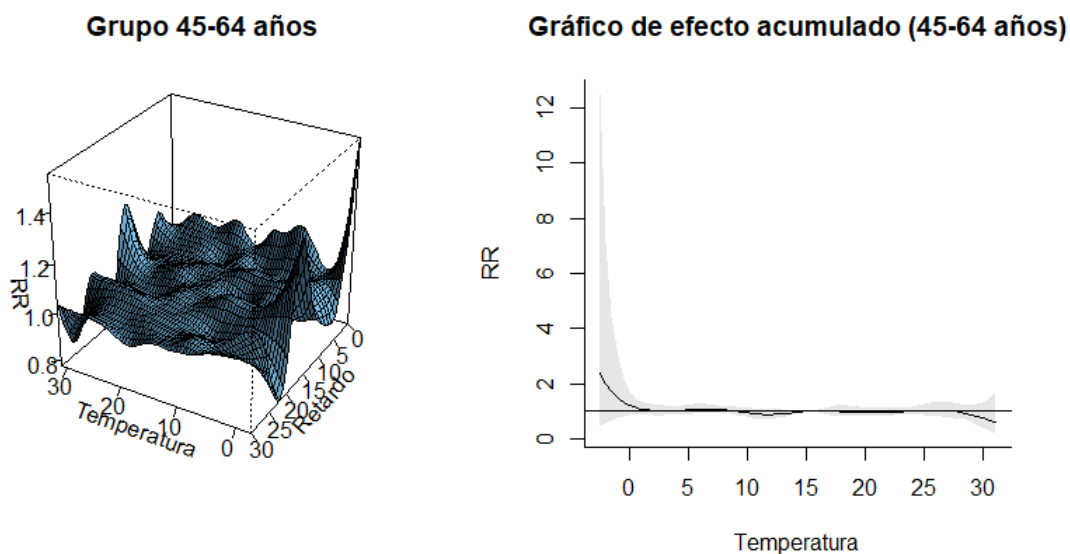


Observemos que la superficie resultante está conformada por muchos picos y valles, que ya nos indica que algo no va bien. Fijándonos en la curva de efecto acumulado, nos sorprende el eje correspondiente al riesgo relativo, que alcanza valores demasiado grandes. Además, la zona gris, que recordemos que se trata del intervalo de confianza, también es muy amplia. Por último, parece que la curva sube y baja constantemente, al igual que nos ocurría con el grupo anterior. Si hacemos una ampliación, se tiene:



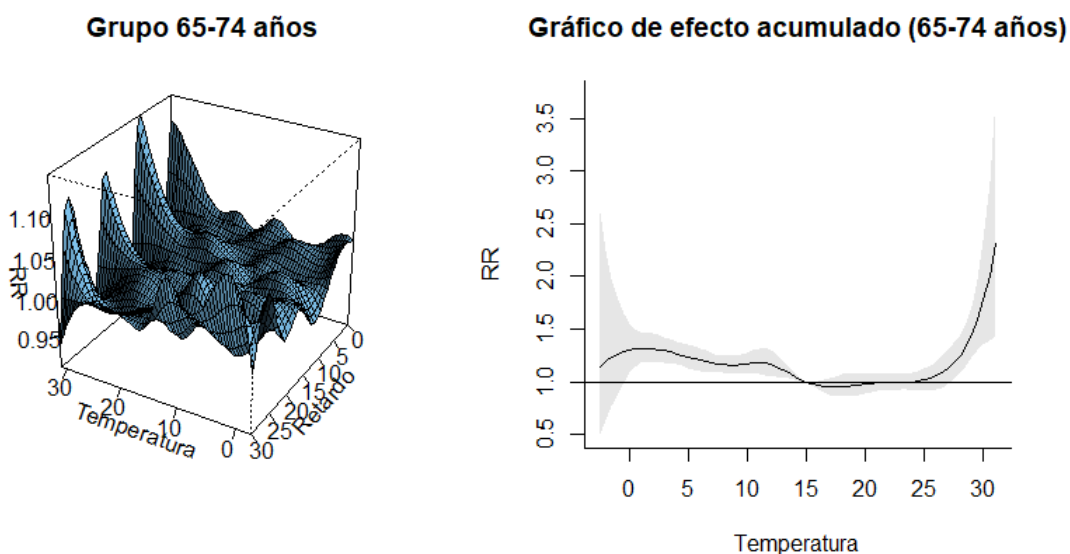
De manera semejante a lo que hemos visto con anterioridad, la curva se comporta de manera rara, con muchos altibajos. Se alcanzan valores del riesgo relativo exagerados y la zona gris es amplia. Todo esto nos lleva a pensar que, pese a que haya más datos que para el grupo de los niños, siguen sin ser suficientes defunciones para hacer un estudio. Para los siguientes grupos de población ya sí que se va a poder ver algo más. Además, tiene sentido que cuanto más mayores sean las personas, mayor riesgo tienen de fallecer, en particular, por el efecto de la temperatura.

Ahora, para el grupo de las personas que tienen entre 45 y 64 años, se tienen los siguientes gráficos:



Aunque parece que se obtiene una superficie un tanto irregular notemos que, como ya es habitual, el riesgo aumenta en las esquinas. Sin embargo, en las zonas centrales se mantiene más o menos igual. Observando el gráfico de efecto acumulado nos damos cuenta de que la curva es muy próxima a 1 constantemente, y al aproximarnos a temperaturas muy cálidas, el valor del riesgo relativo decrece. Esto es justo lo contrario a lo que se esperaría que ocurriera. Esto parece sugerir que no hay, en nuestros datos, suficientes muertes que se puedan atribuir al exceso de calor dentro de este grupo. Así, podemos pensar que este grupo de edad no es especialmente vulnerable al calor.

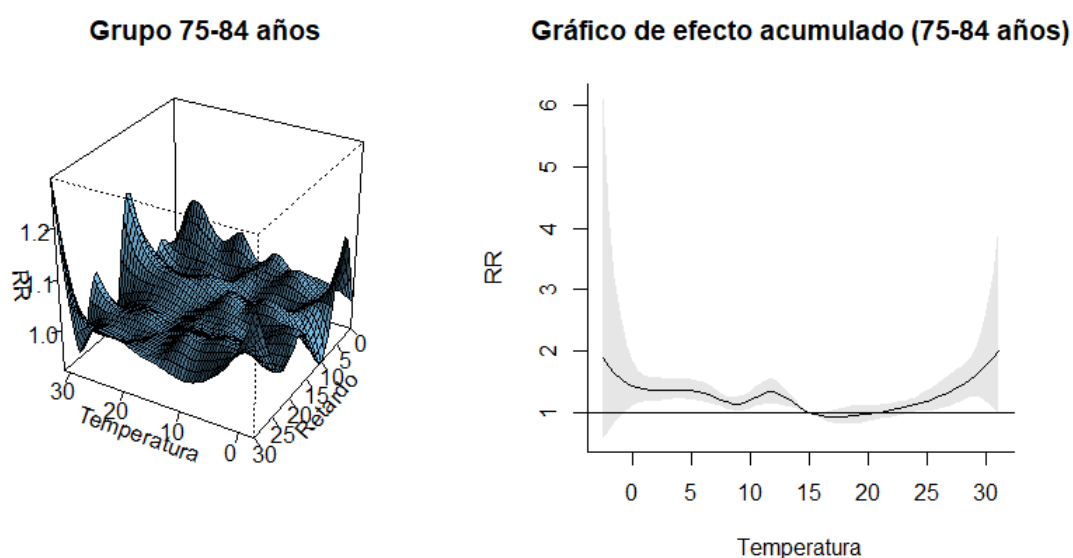
Pasemos ahora con el grupo de personas mayores, ubicado entre los 65 y los 74 años de edad. Las imágenes resultantes son:



En un primer vistazo se nota una superficie bastante irregular, con una totalidad de 4 picos con una temperatura de 30 grados y cada pocos días de retardo. Notemos que en dichos picos, el riesgo relativo tampoco es demasiado grande, se supera ligeramente el valor de 1.10. La curva que se visualiza del gráfico de efecto acumulado tiene una forma más esperable. En el centro, entre los 15 y los 20 grados de temperatura, el valor del riesgo relativo es ligeramente inferior a 1, lo que muestra que, con una temperatura normal, no existe ningún problema. Conforme la temperatura es más fría, el riesgo parece que es mayor, pero el hecho de que el intervalo de confianza, la zona gris, sea tan amplio, no nos confirma nada. Sin embargo, en la parte de las temperaturas altas, que es la que más nos interesa en este trabajo, se ve un evidente aumento del riesgo, incluso llegando cerca del 2.5. El intervalo de confianza no es muy amplio, luego es una medida en la que se puede confiar.

Por consiguiente, como ya veníamos esperando, este grupo, que es ya de una edad bastante avanzada es muy vulnerable a los efectos del exceso de calor, luego se tiene que tener cuidado.

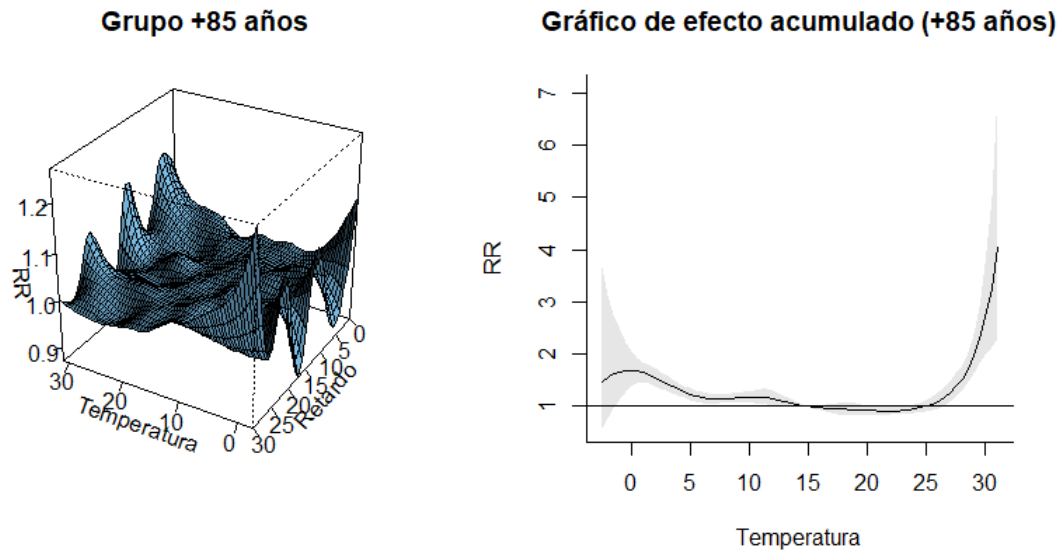
Sea ahora el penúltimo grupo, el de las personas de entre 75 y 84 años. Los gráficos que se obtienen son:



Se obtiene algo parecido al grupo anterior. Según la superficie obtenida, con 30 grados, el mayor valor del riesgo relativo se alcanza con 30 días de retardo, y el siguiente con unos 10 días, alcanzando el 1.2. Por otra parte, mediante el gráfico de efecto acumulado nos damos cuenta que, parecido a lo de antes, entre los 15 y los 20 grados no hay ningún problema, cerca del 0 parece que aumenta pero no podemos sacar conclusiones, y según nos acercamos a los 30 grados, el riesgo aumenta significativamente. Por tanto, este grupo también puede ser considerado como grupo vulnerable a los efectos del calor extremo.

Vamos a terminar este análisis con el grupo más mayor, compuesto por personas de más de

85 años. Las gráficas correspondientes son:



Se obtienen cosas bastante semejantes a los dos grupos anteriores. La superficie alcanza el mayor valor del riesgo relativo cerca del 1.2 cuando hay 30 grados de temperatura y unos 10 días de retardo. Sobre la curva, esta vez entre los 15 y los 25 grados se considera que no hay peligro. No se pueden sacar conclusiones del frío, y existe una evidente subida del riesgo cerca de los 30 grados. Notar que esta vez se supera la barrera del riesgo relativo de valor 2.5, llegando hasta 4. Esto nos quiere decir que el grupo es vulnerable y, de hecho, el más vulnerable de todos al exceso del calor.

Bibliografía

- [1] GASPARRINI, A., ARMSTRONG, B. AND KENWARD, M.G. (2010), *Distributed lag non-linear models*. *Statist. Med.*, 29: 2224-2234, <https://doi.org/10.1002/sim.3940>
- [2] GASPARRINI A. (2011), *Distributed Lag Linear and Non-Linear Models in R: The Package dlnm*. *Journal of statistical software*, 43(8), 1–20.
- [3] GASPARRINI, A., ARMSTRONG, B. (2013), *Distributed lag non-linear models in R: the package dlnm*. *London School of Hygiene and Tropical Medicine, UK. dlnm version*, 1(7), 05-16.
- [4] ACHEBAK, H., DEVOLDER, D., BALLESTER, J. (2018), *Heat-related mortality trends under recent climate warming in Spain: A 36-year observational study*. *PLoS medicine*, 15(7), e1002617.