RESEARCH PROJECT

Temperature-Mortality Association: Portuguese Extreme Weather Event Early Warning System

André Martins Brito[a,b]

[a] School of Science and Technology, NOVA University Lisbon
[b] National Institute of Health Doctor Ricardo Jorge

Abstract
Portugal is one of the European countries with a higher excess of mortality during winter, even though winters are considered relatively mild. There is also an excess of mortality during the summer. This excess mortality is associated with a larger vulnerability of the Portuguese population to non-optimal temperature exposure caused by poor housing conditions, with deficient insulation and weak energy efficiency, and an ageing population. More likely heat waves and cold spells due to climate change could result in excess deaths.

An update of the heat and cold health early warning systems is proposed for use in Mainland Portugal. The aim was to develop a risk indicator, active throughout the whole year, and easily understood by the entire population, with the highest possible spatial resolution. Daily data of all-cause mortality and maximum, minimum and mean temperatures was gathered from public data sources for the 1995-2020 time period. District-specific temperature-mortality associations were estimated using different quasi-Poisson regressions. Linear threshold Distributed Lag Models (DLM) were proposed and estimated for cold and warm semesters, where minimum temperatures were considered in autumn/winter and maximum temperatures in spring/summer, to identify worst case exposure scenarios. Influenza incidence was also included in the models to improve predictive performance. Model specification was selected per district independently based on goodness-of-fit criteria.

Models proposed here could serve as updates for heat and cold health early warning systems, as they provide the results to maintain a risk indicator, active throughout the whole year, and easily understood by the entire population, with the highest possible spatial resolution for the data available for Mainland Portugal. Differences between the optimum district-specific models completely justify the need for region-specific warnings.

KEYWORDS
Time series regression; Temperature; Mortality; Heat waves; Cold waves; Prevention plans; Distributed lag models; Early warning systems; Risk indicator;

1. Introduction

There is a clear shown association between temperature and human mortality, with both high and low ambient temperatures represent an increased risk to human health [1]. Throughout the world, several studies attribute excess mortality to heatwaves, these being extended periods of heat exposure. Additionally, low ambient temperatures also play an important role in excess mortality. In general, cold has a bigger impact on mortality than heat [1] and has also shown to have an impact on morbidity [2]. In addition, low or high temperatures may also have a lagged effect on mortality, increasing the number of recorded deaths only days after a possible exposure to unusual weather. On the other hand, sustained exposure to non-optimal temperatures over a certain period of time could also result in excess deaths [3]. There needs to be an understating

CONTACT André Brito. Email: anm.brito@campus.fct.unl.pt

of how mortality and morbidity are affected by temperature, in order to accurately warn the population of imminent danger to public health and implement preventive measures. Moreover, as global warming starts to have its effect on weather patterns, it is important to comprehend how increasingly likely extreme weather events impact mortality and morbidity and adjust measures to mitigate the effects of these events.

Extreme temperature events, included in Extreme Weather Events (EWE), are often referred to as heatwaves (for extreme heat) or as cold spells/waves (for extreme cold), however there is no consensual definition for these events, with different criteria being used in different countries [4, 5].

In regard to heatwaves, the WHO/Europe asserts that there is no standard definition for a heatwave. The EuroHEAT project carried out by WHO/Europe defined a heatwave as a period in which the maximum and minimum apparent temperatures are over the ninetieth percentile of the monthly temperature distribution for at least two days. Apparent temperature is a compound variable of temperature and humidity intended to measure individual's relative discomfort to a meteorological state [6]. Equally, Yang et al. [5] states that "*in general, heatwave is defined as several consecutive days (duration) with daily temperature measures exceeding specific threshold, including absolute threshold and relative threshold*".

According to the IPMA, that considers the definition provided by the WMO, a heatwave occurs when the maximum air temperature is 5°C higher than the average value of the daily maximum temperatures during a period of 6 consecutive days in the reference period (1961-1990) [7].

On the other end of the temperature spectrum, according to the WHO/Europe harsh cold weather spells happen regularly across Europe. These severe meteorologic conditions represent an aggravated health risk to the population, having grave impacts on mortality and morbidity, especially on vulnerable groups [8]. As it happens with heatwave definitions, at present, there is also no common description for a cold spell [9]. There is sufficient evidence that severe cold temperatures have an effect on mortality and morbidity and are a threat to individual's health [10, 11]. In accordance with the WMO, the IPMA considers that a coldwave occurs when the minimum air temperature is 5°C below the average value of the daily minimum temperatures during a period of of 6 consecutive days in the reference period (1961-1990) [12].

The effects of extreme cold and extreme heat on health have been studied extensively. However, extreme temperature events, these being heat or cold waves, are not those with the most attributable impact on mortality. In fact, days with moderate but not optimum temperatures, these being days with moderate heat or cold, are attributed a larger number of deaths than days with extreme temperatures, with moderate cold surpassing moderate heat in regards to the mortality burden attributable to it [1]. Thus, the effects of the moderate temperature range on mortality should not be ignored as evidence shows. These results could also affect how public health protection measures are implemented.

Several studies focus on the effects of ambient temperature on mortality. However, these investigations consider ambient temperature as a proxy for individuals' exposure to temperature with temperature data being extracted, in most cases, from weather stations. The assumption that ambient temperature is a good proxy for individuals' exposure to temperature fails to recognise the fact that most of our time is spent indoors rather than outdoors. Indoor temperature is seldom considered in studies whose objective is to estimate the associations between mortality and temperature mostly due to the lack of indoor temperature measurements [13].

The definition of indoor temperature thresholds that ensure the health of occupants, given high or low outside temperatures, remains a challenge [14, 15]. Increasing temperatures caused by global warming, possibly even exacerbated by urban heat island effects, represent a threat to human health, as individuals may struggle to feel comfortable inside their own homes. Energy inefficient homes are harder to keep at an acceptable temperature for occupants and efforts to do so may reveal to be far too expensive. An ageing population will further aggravate this issue due to their increased sensitivity to temperature. Anderson et al. [14] performed a literature review with the aim of summarising the groups vulnerable to heat exposure and identifying indoor heat thresholds for the United Kingdom. They identify a set of pre-existing health conditions that make sick individuals more vulnerable to heat exposure. These include obesity, the already mentioned cardiovascular and respiratory diseases, psychiatric and neurological disorders. In addition, older people, young children and people with impaired mobility, like those confined to bed, are also identified as vulnerable groups. Besides individual's characteristics, the authors also summarise environment and building features that increase vulnerability to heat, such as living on the upper floor of a building, residing in an urbanised areas, living in flat roofed buildings and lack of ventilation or air conditioning at home. Given the amount of variables that influence the vulnerability to heat exposure, indoor temperature thresholds are hard to define and fully justify. As such, the authors considered unfeasible to suggest any maximum indoor temperature thresholds for national use given current knowledge of the effects of indoor temperature on health, being the indoor temperature range that minimises risk to health defined by the WHO, as lying of between 18°C and 24°C, the largely accepted range. However, this range is not location, season or climate specific and is the default suggestion for worldwide application, not having in mind the adaptability and thermoregulation capabilities of individuals to local climates that evidence suggests to exist, following the judgement made by Thai et al. [13].

Evidence suggests that individuals are able to adapt to their local climate. The human thermoregulation and acclimatisation to ambient temperature abilities play a crucial role in how temperature affects health. Thus, optimal temperature range recommendations made by public health institutions should be sensible to that [13, 16]. In addition, behavioural adaptations, such as wearing light clothing in hot days or, on the other hand, turning on heating when its cold, has influence on the temperature experienced and, thus, also affect health. However, these behavioural changes should be adjusted to ensure that outcomes are, in fact, beneficial to health and do not represent actions that would increase the risk of unfavourable outcomes. This is where the importance of public health policies and institutions and subsequent warning systems prevails. There needs to be an understanding of what behaviours should be applied to given situations and which should not and communicate recommendations to the population in order to encourage the correct responses. Effective communication and rational recommendations are extremely important, especially when it comes to extreme weather events, like heatwaves and cold spells, where simple behavioural adjustments can dictate whether an individual survives or not [17].

The European summer heatwave of 2003 is an important moment in history that provides crucial information about how extreme weather events can be devastating. It is estimated that this particular heatwave caused 30000 excess deaths due to heat exposure all throughout Europe [18]. Particularly in France, 15000 excess deaths are attributed to this heatwave and the elderly were especially vulnerable. In an effort to understand what behaviours or health conditions increased the risk of mortality, a case-control study was performed by Vandentorren et al. [17] considering only people aged 65 or older who died during the 2003 heatwave in France. Results from this study showed that lack of mobility associated with poor state of health, cardiovascular, neurological and mental illnesses increased the risk of death. In respect to housing conditions, residing on the top floor of a building and deficient insulation were also risk increasing factors. Moreover, the bedroom's location and its exposure to sunlight were deemed to influence the risk of mortality. Bedroom's located right under the roof of buildings and a larger number of windows augmented the chances of death. In addition, individuals' behaviours directly impacted, positively and negatively, the risk of dying. For instance, people who attended air-conditioned or cooler venues, dressed lightly or used other tools at their disposal to remain cold were less likely to die. On the other hand, inadequate behaviours, like opening windows in the afternoon when outside temperatures were still high, showed to augment the risk of mortality. The key takeway from this study is that individuals who were able to adapt their behaviour accordingly were less vulnerable to heat. The authors suggest that recommendations from public health officials should encourage adequate behaviour changes like dressing lightly, attending cooler venues and opening windows in the hours of less heat, behaviours that clearly can make a difference when put into practice during extreme heat events.

An analysis performed by Fowler et al. [19] estimates the excess deaths attributable to cold temperatures through an EWDI during winter in several European countries. A pattern emerged in EWDI values per country, with countries in northern Europe having significantly lower EWDI values than southern European countries. Some of these southern countries are considered to have milder winters. Nevertheless, results show these countries have higher excess deaths in winter than those considered to have harsher winters. These findings suggest a higher vulnerability to lower temperatures of populations living in milder climates. Several factors may come into play when interpreting these results, for example, low income of individuals, which is directly correlated with national Gross Domestic Product (GDP), could impact their ability to remain warm due to the additional cost that represents. Moreover, building stock and adequate insulation could also play a role in the individual's exposure to cold temperatures. The authors also note that there is considerable heterogeneity in EWDI values between countries with identical climates, such as Mediterranean countries, and state that public health intervention can have a differentiating effect. With the objective of making sense of why some European countries suffer from excessively higher winter deaths than others, Healy [20] set to find out if other potential causative factors could explain this excess in mortality. Time series data from 14 European Countries were gathered containing information about all-cause mortality, macro- and socio-economic variables, health care and housing conditions, for the 1988-1997 time period. By calculating a coefficient of seasonal variation in mortality for the 14 countries, the obtained results show that Portugal had the highest seasonal variation in mortality within this set of European countries in the time period considered, with an increase of 28% above what would be the expected mortality rate, what translates to 8800 premature deaths during winter time per year. After Portugal, Ireland and Spain also present high seasonal variation in mortality. The macroeconomic state of a country was also shown to have an association with excess mortality during winter. Indeed, higher GDP per capita countries displayed lower mortality variation during winter, while countries with lower GDP per capita (e.g. Portugal and Spain) exhibit significantly higher mortality variations. Similarly, lower health care expenditure, high levels of income poverty and inequality and deficient building stock insulation are all associated with higher mortality variations. All of these factors combined could explain why Portugal shows the highest excess mortality during winter out of all the 14 countries throughout the study period. This research shows that non-climate related factors, such as macro- and socio-economic factors, have an impact on mortality. This means deaths can be prevented by a more favourable economic conjecture, enhanced public health services and improved building stock conditions.

As a result of these studies, the importance of public health institutions and actions is clear, as they can prevent the untimely death of several individuals by intervening accordingly to current circumstances. In Portugal, a Heat-Health Early Warning System - ICARO - has been in operation since 1999 [21] and was developed by the national public health institution INSA. The underlying model to the ICARO System is built upon knowledge obtained by analysing the effects of previously occurred heatwaves in Portugal. In 2003, during that year's summer heatwave, the underlying model of this system was able to successfully predict nefarious effects to mortality considering only air temperature values [21]. The 2003 heatwave episode revealed to be an opportunity to improve the model given the new data obtained regarding the effects extreme heat events. The Heat-Health Early Warning System is able to report the risk level related to heat for each of the 5 main health regions for mainland Portugal and is active between May and September. For the winter, another health warning system - FRIESA - was developed based on Poisson regression models and it is able to predict mortality associated with exposure to cold temperatures for only two districts of mainland Portugal, the Lisbon and Porto districts [11].

Portugal has 308 municipalities within its 18 districts and two autonomous regions, 278 of them are on the mainland, 11 in the Autonomous Region of Madeira, and 19 in the Autonomous Region of Azores. Mainland Portugal is located between latitudes 37° and 42°N which is a transitional region between the sub-tropical anticyclone and the subpolar depression zones. Its climate is characterised by a mostly mild Mediterranean climate highly conditioned by the proximity to the Atlantic Ocean and its orography, also exhibiting substantial variability between years and powerful seasonality. Due to its geographic location, mainland Portugal is also prone to experience extreme weather events like heatwaves, droughts and even floods. Precipitation and air temperature also vary largely across the territory. In the north of Portugal annual precipitation is far more abundant than in the south, while average air temperature in the south varies substantially more than in the central and northern parts of the country. In a relatively small country, several climate variables vary along north to south and east to west directions [22, 23]. These climate characteristics denote a high degree of variability to environment exposure within the country and, therefore, are sufficient to justify the existence of temperature related warning systems with the highest spatial resolution possible, that *tailors* the warning to each distinct area.

Besides climate variables, a region's socio-economic, demographic and building stock characteristics should also be considered when designing a warning system as they have an influence on the temperatures experienced by individuals and their thermal comfort and may vary significantly between regions, especially in Portugal.

About 80% of the Portuguese population resides in three out of seven regions, Norte, Área Metropolitana de Lisboa e Centro (NUTS.g II). The Alentejo region is consistently losing inhabitants and there is a clear trend of outflow of people from more inland areas to coastal ones which results in a concentration of older individuals in inland areas. There are also discrepancies of income per region, with Lisboa being where people earn, on average, the most and Norte where people earn, on average, the least [24]. These demographic and socio-economic traits combined with climate variability contribute to a heterogeneous vulnerability of individuals to temperature exposure within the country. Residential building stock may influence the temperature experienced by individuals when indoors [17] and, thus, has an impact on the risk of death, in spite of lack of research between indoor temperatures and mortality [13]. Indoor temperature is affected not only by outdoor temperature but also by construction material, proper insulation and energy efficiency of a building. An analysis performed by Oliveira et al. [25] attempted to provide an overview of the building stock in Portugal by considering energy certificates that can be associated with better indoor environments. Their results show a fairly heterogeneous characterisation of the building stock per region in Portugal. For instance, the authors found that the Lisbon Metropolitan Area is the area with the largest proportion of old buildings whereas the Algarve had the lowest proportion of old buildings. In regard to the energy certificates analysed, 30% of buildings in the Alentejo and Centro regions had the lowest rating possible. The results presented showed that a notable proportion of buildings in Portugal do not guarantee a healthy environment for their occupants, which combined with low average income and an ageing populations creates a worrying scenario for the Portuguese population where individuals may not be able to be comfortable inside their own homes. Hence, the argument for a temperature related warning system with a high spatial resolution is well established.

2. Objectives

The general objective of this work is to contribute to the understanding of how temperature exposures (Heat and Cold) influence mortality and morbidity, considering the highest spatial resolution in Mainland Portugal through an analysis of all-cause mortality (death counts). Hopefully, this knowledge will help to adjust public health actions to current circumstances and to specific locations given the risk they were estimated to represent.

The Portuguese National Public Health Institute (INSA), currently operates two different extreme temperature exposure early warning systems. One of these early warning systems, ICARO, operates during the summer months and it is based on a linear regression model, while the other early warning system, FRIESA, operates during the winter months and it is based on a quasi-Poisson regression model. Each early warning system provides daily risk indicators based on temperature forecasts which are then exchanged with civil protection agencies, public health officials, decision makers and other stakeholders. When it comes to the public health early warning systems currently operating in Portugal, two problems were identified. Firstly, heat and cold related warning systems are based on different analytical models and their results have different interpretations which could confuse those responsible of implementing appropriate prevention measures. Secondly, the warning systems have different coverage of the Portuguese territory, with the winter only providing warnings for 2 out of the 18 districts in Mainland Portugal. As stated before, the climate, socio-economic, demographic and building stock variability between geographical areas in Portugal completely justifies the need for a nationwide coverage with augmented spatial resolution. Hence, an update of the temperature-related warning systems is needed.

The main goal of this work is to develop an all year round indicator for the risk of mortality due to heat and cold related exposures for the whole country, with the highest spatial resolution possible, by resorting to state of the art modelling techniques. This indicator is to be presented to and understood by the general public and is not aimed only at public health officials, which represents a challenge. The objective is to provide a nationwide overview of the temperature related risks with specific information available for each of the territory subsections and to simplify targeted action on communities with higher vulnerability by reaching a wider audience.

The *RELIABLE* Project provided a wide range of data, making use of new public data sources, pertaining to temperature, mortality with several levels disaggregation. From these data it was possible to extract information about the evolution and characteristics of mortality per municipality and district of Mainland Portugal, and the intensity and duration of extreme temperature events during the study period, something that has become ever more important has the effects of climate change increase and climate becomes ever more volatile.

3. Data Sets

3.1. Data Presentation and Preprocessing

For this work, information about mortality, morbidity, temperature, demographic and influenza incidence was available. Focus was given to the 1995 - 2020 time period.

All-cause mortality data for Portuguese residents was provided by Statistics Portugal (INE) for the 1995 - 2020 time period. In the data set provided by INE, with 2 793 537 records, each row corresponds to an individual's death. About the individuals there is information about:

- `ANO_MORTE` - Year of death;
- `MES_MORTE` - Month of death;
- `DIA_MORTE` - Day of death;
- `DT_CC_RESID` - Municipality of residence;
- `SEXO` - Gender (1 - Male; 2 - Female; 3 - Unknown);
- `IDADE_ANO` - Age (years);
- `ANO_NASCIM` - Year of birth;

- LOCAL_MORTE - Place of death (1 - House; 2 - Medical facility; 3 - Other; 4 - Unknown);
- PAIS_RESID - Country of residence or birth country;
- MES_NASCIM - Month of birth;

Variables ANO_MORTE, MES_MORTE and DIA_MORTE were complete, with zero records for missing information. Concerning the DT_CC_RESID variable, there were 11,034 death records with invalid municipality codes. The records containing these invalid codes were aggregated together and set to resemble a non-existent municipality here coded as "INVALID". The SEXO (gender) variable has 5 records with unknown gender. Furthermore, there were 276 records with missing age. Unfortunately, the age cannot be computed through other variables (e.g. ANO_MORTE - ANO_NASCIM), because these are actually the records that have an invalid ANO_NASCIM (year of birth) of 9999. The records with missing age were deleted. MES_NASCIM (month of birth) were incomplete with 104 215 records with missing month of birth or with a nonexistent month. Month of birth is not a crucial piece of information, as the analysis performed here relies on the individual's age. PAIS_RESID may refer to country of residence or country of birth, thus, this information was disregarded and more importance was given to the municipality of residence. Records pertaining to the municipalities included in the Autonomous Region of Madeira and Azores were also discarded, as this work only focuses on the municipalities of Mainland Portugal. Mortality row data allows for flexible aggregations. Data can, thus, be aggregated to daily, weekly, monthly and/or yearly mortality counts by municipality, district and/or age group.

After mortality data preprocessing, there were 2 650 111 death records, in Mainland Portugal.

For each municipality, there are 9 497 records on temperature data. The data is complete and no outliers were detected. The Montijo municipality is physically divided in two and, thus, two temperature files were created for this municipality. To have only one temperature file for each of the municipalities, the Montijo files where combined in the following way:

- a minimum temperature variable was computed by selecting the lowest minimum temperature between the two files;
- a mean temperature variable was computed by calculating the mean between the mean temperature recordings in the two files;
- a maximum temperature variable was computed by selecting the highest maximum temperature between the two files.

By doing this, any extreme temperature, high or low, recordings were preserved.

Concerning the demographic information, annual average resident population (numbers) by municipality of residence and age group were exported from the data bases of INE for the 1995 - 2020 time period.

Weekly incidence of influenza like illness per 100 000 residents was provided by INSA for the 1995 - 2020 time period. No regional disaggregation was possible, hence, the incidence rate pertains to the whole country of Portugal. No other influenza indicator made available by the Transparency Portal of the Portuguese National Health Service [26] had the desired time span to include the whole study period.

3.2. Data Aggregation and Preparation

As mentioned before, many of the investigations regarding temperature-mortality associations rely on daily data for mortality counts and climate variables for specific cities or regions. Thus, the several data sets provided for this work needed preparation to comply with such structure.

Mortality and morbidity data were aggregated in order to obtain daily count time series for districts, municipalities and the whole Portuguese Mainland. Daily counts per age group could also be computed. For instance, deaths that occurred in the same district, in the same day and belonged to a specific age group were counted and a daily death count for that particular day, age group and district was obtained. However, some days had a zero death or hospitalisation count. For each district or municipality there was computed daily time series counts of death or hospitalisations per cause (all-, respiratory- and circulatory-causes). Besides the overall death and hospitalisation counts, age groups considered in this work were $[0, 65)$ and ≥ 65. The series start at 01/01/1995 and end at 31/12/2020, composed of 9 497 observations (days).

Moreover, following other investigations [16, 27], the year was classified into two periods, a warm semester and a cold semester. The warm semester included the months of May, June, July, August, September and October while the cold semester included January, February, March, November and December.

As mentioned before, temperature data relate to municipalities. If an analysis was to be performed at the district level, municipality temperature would not be representative of the whole district, as districts are composed of several municipalities. The temperature of only one municipality would not be representative of the temperature of an entire district. Hence, it was necessary to combine the temperature data from each municipality to have a temperature exposure for each district. The rationale was to formulate a temperature index for the district. This index was computed using the weighted average of the temperatures recorded in each municipality of a district, being the weights considered the resident population numbers in each municipality. The index i_temp_t is formulated as follows

$$i\_temp_t(D) = \frac{\sum_{M \in D} pop_t(M) \times temp_t(M)}{\sum_{M \in D} pop_t(M)} \tag{1}$$

where D refers to a specific district, M refers to the municipality that belongs to district D ($M \in D$), $i\_temp_t(D)$ refers to the temperature index in day t for district D, $temp_t(M)$ refers to the temperature recorded in day t in municipality M, $pop_t(M)$ refers to the estimated resident population in municipality M in the year to which day t belongs to. The temperature index was calculated as described for minimum, mean and maximum temperatures, thus, $temp_t$ will refer to minimum, mean and maximum temperature in day t.

Furthermore, influenza incidence data is a weekly rate estimated for the entire country. Thus, to be of use for this work, it was necessary to assume that this rate would be representative for the incidence of influenza in each district or municipality. A simple linear interpolation was used to transform weekly data into daily data.

The data preparation resulted in the construction of data sets for each municipality and district, each with 9 497 daily observations, for the time period between 01/01/1995 and 31/12/2020, composed by the following variables:

- `District` or `Municipality` - District or municipality;
- `deaths` - Number of deaths;
- `deaths_[0,65)` - Number of deaths for the $[0, 65)$ age group;
- `deaths_>=65` - Number of deaths for the ≥ 65 age group;
- `tmin` - Minimum temperature (°C) (index minimum temperature in the case of districts);
- `tmean` - Mean temperature (°C) (index mean temperature in the case of districts);
- `tmax` - Maximum temperature (°C) (index maximum temperature in the case of districts);
- `pop_total` - Annual estimated resident population for the specific municipality or district;
- `pop_[0,65)` - Annual estimated resident population aged between 0 and 64 for the specific municipality or district;
- `pop_>=65` - Annual estimated resident population aged 65 or over for the specific municipality or district;
- `Semester` - Classification between *warm* or *cold* according to the month;
- `ILI` - Computed influence like illness incidence.

All data manipulation and analysis was made using the `R` software [28] in the R Studio environment [29].

4. Methods

4.1. Generalised Linear Models

Most temperature-mortality association studies rely on Generalised Linear Models (GLMs) to estimate the relationships and make assessments. As the name implies, generalised linear models result from the generalisation of linear models and can be applied when the response variable does not follow a Normal distribution. The generalisation is achieved using what is called a link function. More specifically, in order to apply a GLM, the response variable needs to display a distribution belonging to a specific group of distributions called the exponential family [27]. Being regression models, GLMs are composed of a random component and a systematic component [30]. The random component is determined by the underlying distribution of the response variable, while the systematic component is determined by relationship between the predictors at the mean of the response variable μ. Most studies use as response variables death counts described as having a Poisson distribution. A GLM can be characterised by the following equation [31],

$$g\left(\mu_t\right) = g\left(E[Y_t|\boldsymbol{X}_t]\right) = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \cdots + \beta_N X_{N,t} \tag{2}$$

where $Y_t, X_{1,t}, X_{2,t}, \ldots, X_{N,t}$ are the random variables that constitute the model observed at time $t, (t = 1, \ldots, T)$, Y_t denotes the response variable and $X_{1,t}, X_{2,t}, \ldots, X_{N,t}$ represent the explanatory variables. $\beta_0, \beta_1, \beta_2, \ldots, \beta_N$ are the model parameters and $E[Y_t|\boldsymbol{X}_t]$ is the expected value of Y given the the explanatory variables $X_1, X_2, \ldots, X_N$ at time t. There is a linear relationship between the explanatory variables and the so-called link function $g(\cdot)$ applied to the conditional expectation of the dependent variable Y transformed by the link function $g(\cdot)$. Hence, GLM are regression models linear in the parameters.

As mentioned before, GLM rely on the assumption that the response variable's distribution belongs to the exponential family, also known as the EDM family. Based on the notation presented in Dunn and Smyth [30], the distributions belonging to the EDM family of distributions have a probability density function $f_Y(y;\theta,\phi)$, if Y is a continuous variable or a probability mass function $f_Y(y;\theta,\phi)$, if Y is discrete variable, that can be represented in the following form

$$f_Y(y;\theta,\phi) = a(y,\phi)\exp\left\{\frac{y\theta - \kappa(\theta)}{\phi}\right\} \tag{3}$$

where θ is the canonical parameter, $\kappa(\theta)$ is a function of θ called the cumulant function, ϕ is a dispersion parameter and $a(y,\phi)$ is a normalising function that ensures $f_Y(y;\theta,\phi)$ is, in fact, a probability function. Equation 3 represents the Exponential Dispersion Model (EDM) distribution's probability function in the canonical form. Other forms of representation are possible.

Most commonly in count data, frequently described has having a Poisson distribution, one may need to take into the account the context in which the counts were observed. For instance, when dealing with mortality data pertaining to different locations and at different points in time, one needs to account for the number of individuals that may die in each location at given time, meaning the exposed population [27, 30]. The population number varies between locations and points in time and, consequently, so does the expected number of deaths based on that exposed population. Hence, the inclusion of an offset in a Poisson regression converts our response variable from a count of events into a rate of events.

In this work, a particular focus will be given to the Poisson regression, as this will be the basis for the developed mortality-temperature models. As mentioned before, Poisson regression estimates the expected value of the response variable (event counts), in this study, death counts based on exposure to extreme temperatures.

The analysis of count data usually reveals what is termed as overdispersion. Overdispersion means that the observed variance of the count data significantly exceeds its observed mean. However, a Poisson-distributed variable has a variance equal to the mean and, thus, overdispersion could have severe consequences in the estimation of GLMs that it needs to be accounted for [32]. To solve this issue, a dispersion parameter is introduced into the Poisson model so that $V[Y_t|\boldsymbol{X}_t] = \phi\lambda_t$ where $\phi > 1$. This formulation implies that the conditional variance of Y_t increases at a higher rate than its mean, which poses another problem, as there are no distributions belonging to EDM that display this property. Hence, under the condition of overdispersion, there is no probability density function for Y_t and, therefore, the maximum likelihood method cannot be used to estimate the model's coefficients. Nonetheless, as stated by Fox [32], "*the maximum-likelihood estimation of a GLM yields the so-called quasi-likelihood estimators of the regression coefficients, which share many of the properties of maximum-likelihood estimators*".

There are various measures that can be used to quantify the health burden of certain exposures in order to convey them successfully to the public [33]. Most studies quantify temperature-mortality associations in terms of Relative Risk (RR). In a Poisson regression, the estimated parameters of the regression β_x can be used to calculate RR [27],

$$RR = \exp(\beta_x) \tag{4}$$

where β_x is a Poisson regression term related to exposure x, meaning that every unit increase of the associated exposure x would increase the probability of the outcome realisation by a value given by RR when compared to the probability of the outcome realisation with a reference value x_0.

Another way to measure RR using counts is to estimate the number of expected deaths due to exposure to extreme temperatures and compare them to the number of expected deaths when there is no exposure to extreme temperatures. Let A_1 be the number of expected deaths given an exposure to extreme temperatures and A_0 the number of expected deaths given no exposure to extreme temperatures, the estimate for RR, $\widehat{RR}$ is given by [33]

$$\widehat{RR} = \frac{A_1}{A_0}. \tag{5}$$

This RR can be interpreted as the rate at which the probability of death increases given an exposure to extreme temperatures compared to the probability of death given no exposure to extreme temperatures. Another way to look at the RR value is to consider $RR-1$ that gives the excess relative risk of deaths due to exposure to extreme temperatures and will be zero if there is no expected effect.

4.2. Distributed Lag Models

Temperature-mortality associations are characterised by their non-linear nature and delayed effects of an exposure. The Distributed Lag Models (DLM) were introduced in air pollution studies and aimed at detecting the delayed effect of an exposure on the response, albeit for linear associations.

In the case of the air pollution investigation performed by Schwartz et al. [34], the authors intended to evaluate the effects of exposure to PM10 concentration lagged in time. To achieve this, the authors defined the following linear model,

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \cdots + \beta_q X_{t-q} + \epsilon_t \tag{6}$$

where X_{t-l} is the PM10 measurements $l, l = 0, 1, \ldots, q$ days before, Y_t are deaths counts registered at time t and ϵ_t a random error. The regression coefficients $\beta_0, \beta_1 \ldots, \beta_q$ reflect the effects of PM10 concentration registered on days $t, t-1, \ldots, t-q$ on deaths registered on day t, respectively.

The defined DLM is

$$f(\mathbf{X}_{t-\Delta t} \mid \boldsymbol{\beta}) = \sum_{l=0}^{q} \beta_l X_{t-l} \tag{7}$$

where $\Delta t \geq 0$ and $l = 0, 1, \ldots, q$ is the number of lag days considered.

In the case of temperature-mortality investigations, Armstrong [35] formulated the following quasi-Poisson regression

$$E[Y_t] = \exp\left\{\alpha + f\left(\text{temp}_t \mid \boldsymbol{\beta}\right) + \texttt{seasonality control} + \texttt{covariates}\right\} \tag{8}$$

where Y_t is the death count for day t, temp_t is temperature (in their case, mean temperature) registered for day t, $\boldsymbol{\beta}$ are the regression coefficients related to the temperature terms, $f(\cdot)$ is a function that links the response with the exposure, `seasonality control` represents a function or a set of variables to capture seasonality patterns and long-term trends of the outcome variable and `covariates` represent any other variables that can influence the response. The author applied a DLM to a linear-thresholds model to capture the delayed effects of temperature on deaths. The linear-thresholds model is defined by

$$f\left(\text{temp}_t \mid \boldsymbol{\beta}\right) = \beta_C \; T_{C,t} + \beta_H \; T_{H,t} \tag{9}$$

where $T_{C,t} = \max\left\{(\tau_C - \text{temp}_t), 0\right\}$ and $T_{H,t} = \max\left\{(\text{temp}_t - \tau_H), 0\right\}$ represent the excess of temperature below a defined cold-effect threshold τ_C and the excess of temperature above a defined hot-effect threshold τ_H, respectively, and, the coefficients β_C and β_H their respective regression coefficients. The application of the DLM on the linear-thresholds model resulted in the functional form of $f\left(\text{temp}_t \mid \boldsymbol{\beta}\right)$ changing to

$$f\left(\text{temp}_{t-\Delta t} \mid \boldsymbol{\beta}\right) = \sum_{l=0}^{L} \beta_{C,l} \; T_{C,t-l} + \sum_{l=0}^{L} \beta_{H,l} \; T_{H,t-l} \tag{10}$$

where $T_{C,t}$ and $T_{H,t}$ have the same definition as before, while the $\beta_{C,l}$ and $\beta_{H,l}$ coefficients now represent the effects of cold and heat lagged by $l, l = 0, 1, \ldots, L$ days, respectively, with L is the maximum number of lag days with $\Delta t \geq 0$. In this particular case, this DLM is considered an unconstrained model, as the coefficients are independently estimated from each other.

The $\beta_{C,l}$ and $\beta_{H,l}$ have simple interpretations [35, 36]. From a forward perspective, an increase above the heat threshold on day t will impact mortality risk in each of the $t + l$ days. On the other hand, a decrease below the cold threshold on day t will impact mortality risk in each of the days $t + l$ days, where $l = 0, 1, \ldots, L$. The terms $\beta_{H,l}$ or $\beta_{C,l}$ are the contributions from the exposure at time t to the risk at times $t + 0, \ldots, t + L$

4.3. Proposed Models

In this section, the temperature-mortality models developed for this work will be presented. The challenges of the temperature-mortality associations identified on the aforementioned Chapter, the non-linearity of the relationship and the delayed effects of the exposure on the response, will be addressed using the several approaches presented. An attempt to also estimate temperature-morbidity associations will be made. One of the main goals of this work is to develop a risk indicator with the highest spatial resolution possible for Mainland Portugal, in an attempt to issue risk warnings related to disaggregated areas of territory, taking into to account the effects of temperature on the number of deaths and hospitalisations. The risk indicator will be based on the RR that can be computed by the estimates provided by GLMs. The models will estimate the temperature-mortality and temperature-morbidity associations related to each of the adequate areas of territory chosen.

As it was mentioned before, Poisson time series regressions are the go-to models when the response variable is count data, like death and hospitalisation counts. Most of the times these kinds of data suffer from overdispersion, something that needs to be accounted for when estimating the regression parameters. Thus, the response variables will be considered to have a quasi-Poisson distribution, turning the regression into a quasi-Poisson regression. The usual model formulation employed in the literature [35] takes the following form, presented in equation 11:

$$E[Y_t] = \exp\left\{\alpha + f\left(\text{temp}_{t-\Delta t} \mid \boldsymbol{\beta}\right) + \texttt{seasonality control} + \texttt{covariates}\right\} \tag{11}$$

where Y_t is the death/hospitalisation count for day t, $f\left(\text{temp}_{t-\Delta t} \mid \boldsymbol{\beta}\right)$ is the functional form of the temperature terms registered on day $t - \Delta t$ with $\Delta t \geq 0$, allowing for delayed effect of temperature, $\boldsymbol{\beta}$ are the regression coefficients related to the temperature terms, `seasonality control` represents a function or a set of variables to capture seasonality patterns long-term trends of the outcome variable and `covariates` represents any other variables that can influence the response. Additionally, Y_t follows a quasi-Poisson distribution with $E[Y_t] = \mu$ and $V[Y_t] = \phi\ \mu,\ \phi > 1$.

In this work, Model 1 will be based on a more simple and direct DLM framework. The objective is to analyse and evaluate the performance of the framework on its predictive ability and risk warning accuracy.

Given the several formulations for the temperature-mortality, the models will include a function of time to capture seasonality and long term trends, temperature terms and influenza incidence. Yearly resident population data will be used as an offset to account for population variations over time. The following Directed Acyclic Graph (DAG), displayed in Figure 1, intends to represent the structure of the relationships assumed between the available variables considered important to the modelling of the temperature-mortality and temperature-morbidity associations. The relationships between between time (t), ambient temperature (T), influenza incidence (I) and mortality/morbidity (M) are represented.

It is assumed that time (t) influences temperature and influenza incidence, as both variables show seasonal patterns, hence the arrows directed from t to T and I. Furthermore, the direction of the arrow between T and I implies that a variation in ambient temperature has an effect on influenza incidence. Temperature was shown to have a negative correlation with influenza incidence [37], i.e., lower temperatures are correlated with higher influenza incidence. As mentioned before, temperature exposure and influenza incidence have both harmful impacts on mortality

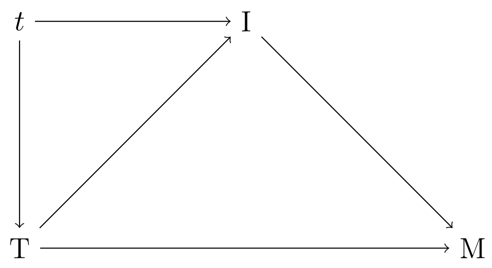


Figure 1. DAG diagram for the structure of the relationships between time (t), temperature (T), influenza incidence (I) and mortality/morbidity (M) considered in the proposed models.

and morbidity, hence the arrows directed at M from I and T [4, 35].

It is important to include influenza incidence in these types of models, as influenza epidemics are associated with excess mortality due to respiratory complications that may occur [19]. Its seasonality is also a crucial aspect to have into consideration due to the peaks in influenza incidence during the winter time. The non inclusion of influenza incidence in the model would result in a biased estimation of the effects of lower temperatures, given that both register a similar yearly seasonality. On the other hand, the inclusion of a influenza incidence variable in the models will impute part of the deaths to observed influenza peaks, hopefully improving the predictive performance of the models and making the estimated effects of colder temperatures more precise and closer to the reality. However, it is important to note that influenza incidence is not a confounding factor for the temperature-mortality/morbidity association. It would only be a confounding factor in such associations if influenza incidence had both an effect on mortality/morbidity and , simultaneously, an effect on temperature as well, which is not a very plausible assumption. Rather, the assumption that temperature has an effect on influenza incidence is far more plausible [37]. For that reason, influenza incidence does not confound the effects of temperature on mortality/morbidity. However, temperature and influenza incidence have an interaction effect, as both exposures have an effect on the outcome, mortality/morbidity [38]. The joint effect of both exposures is expected to be higher than the effect expected by the sum of their individual effects.

The model will have the following form

$$E[Y_t] = \exp\left\{\alpha + \texttt{seasonality control} + f(\text{temp}_{t-\Delta t}|\boldsymbol{\beta}) + \texttt{influenza incidence} + \texttt{offset}\left[\log(\text{pop}_t)\right]\right\} \tag{12}$$

where Y_t is the death/hospitalisation count for day t, α is the intercept coefficient, Δt is a time variation $\Delta t = 0, 1 \ldots$, $\text{temp}_{t-\Delta t}$ represents temperature variables registered in day $t - \Delta t$, $\boldsymbol{\beta}$ are the regression coefficients related to the temperature terms, `seasonality control` represents a set of variables to capture seasonality patterns long-term trends of the outcome variable and `influenza incidence` represents a set of variables to control for the effects of influenza in the response variable.

Regarding the seasonality control, most studies control for day of the week. This is to avoid systematic errors that could underlie count data. For instance, death registrations could be more prevalent in work days when compared to the weekend days due to administrative issues, while hospitalisations could be more frequent on the weekend when people are free to seek medical services. Day of the week will be included in the model as a categorical variable defined by

$$\sum_{d=1}^{7} \delta_d \texttt{dow}_d(t) \quad \text{where} \quad \texttt{dow}_d(t) = \begin{cases} 1 & \text{if day } t \text{ is weekday } d \\ 0 & \text{otherwise} \end{cases}, \tag{13}$$

where δ_d is the coefficient related to the day of the week d and $d = 1, 2, 3, 4, 5, 6, 7$ which correspond, respectively, to Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday.

Additionally, most studies also control for seasonality and trends with the use of spline functions of time. Spline functions are more complex when compared to other approaches like Fourier terms or Serfling Models and result in an increased number of variables in the model. Moreover, spline functions are interpolation methods used in a retrospective fashion in time series studies, not allowing for the prediction of the effects of time on mortality in the short-term outside the study period. On this work, the objective is to predict, in the short-term, mortality and morbidity counts and assess the risk in the following days using forecasted temperatures. Spline functions do not enable this. Hence, a Serfling Model [39] was employed. The polynomial terms allow for the capture of long term trends and the Fourier pair allows for the capture of yearly seasonality in mortality and morbidity. This model makes it possible to establish a mortality and morbidity baseline that will be the reference point from which to measure risk. Thus, seasonality was controlled for in the model by a categorical variable of day of of the week and a Serfling Model defined as

$$\begin{aligned} \texttt{seasonality control} := {} & \alpha_1 \times t + \alpha_2 \times t^2 + \\ & + \alpha_3 \times \cos\left(\frac{2\pi}{365.25} t\right) + \alpha_4 \times \sin\left(\frac{2\pi}{365.25} t\right) + \sum_{d=1}^{7} \delta_d \texttt{dow}_d(t) \end{aligned} \tag{14}$$

where δ_d is the coefficient related to the day of the week d, α_1 and α_2 the coefficients related to the polynomial trend and α_3 and α_3 the coefficients related to the Fourier pair.

As mentioned before, the national weekly incidence of ILI per 100 000 residents was transformed into a national daily indicator with the use of a simple linear interpolation. By doing this, there are no jumps in the influenza incidence between days and there is a smoother evolution of influenza incidence values through time. In order to include influenza incidence in the models, the hypothesis that the national incidence values would be representative for the district and municipality incidence had to be considered true. Influenza incidence has a delayed effect on health [35, 40], and similarly to other studies this was accounted for in the models presented here. A linear relationship between mortality/morbidity and influenza incidence is assumed with delayed effects accounted for with distributed lags. The following equation shows how influenza incidence was added to the model

$$\texttt{influenza incidence} := \sum_{l=0}^{L_{ILI}} \gamma_l \texttt{ILI}_{t-l} \tag{15}$$

where L_{ILI} is the maximum number of lag days considered for influenza incidence, $\texttt{ILI}_{t-l}$ is the daily estimated influenza incidence in day $t - l$ and γ_l the respective coefficient for lag l.

All the models will be area specific, meaning that each adequate areas of territory chosen will have their own model and all of the data used to estimate these models will pertain to each of the areas.

In the following subsections, the underlying rationale for the definition of the functional form of the temperature terms will be presented, along with the complete model definition.

4.3.1. Model 1

Model 1 is based on the DLM framework to account for the delayed effects of temperature. In formulating this model, the objective was to obtain a simple to understand and interpretable model. Given that the intention was to use the estimates provided by this model to assist on the issuing of warnings to the population in the event of extreme temperature, a conservative perspective was needed and, thus, the focus turned to the most extreme exposures to which the population could be exposed to, these being the minimum and maximum temperatures. Meaning that, in this model, the minimum and maximum temperatures were considered, since they are, in the worst possible scenarios, the most extreme temperatures to which the population will be exposed to. Thus, for both minimum and maximum temperatures, temperature thresholds will be defined for which below or above, respectively, there is an expected effect of temperature on mortality/morbidity. The temperature thresholds will correspond to quantiles in the distribution of the minimum and maximum temperatures. In the case of cold during the winter, one wants to quantify by how many degrees the minimum temperature observed on day t is below the threshold for cold. On the other hand, in the case of heat, during the summer, one wants to quantify by how many degrees the maximum temperature on day t is above the threshold for heat. These temperature differentials will be the base of the DLM framework to estimate the effects of temperatures on mortality/morbidity. Additionally, the number of lag days to be considered needs to be defined for both cold and hot days. Many investigations report a different lag structure for low and high temperatures. Thus, different lag days will be considered for the minimum and maximum temperature differentials. Cold is know to have a longer effect on mortality/morbidity risk while heat has a more immediate effect. Thus, a higher lag will be needed for cold when compared to heat.

Having explained the rationale behind the formulation of Model 1, the functional form of the temperature terms is the following

$$f(\text{temp}_{t-\Delta t}|\boldsymbol{\beta}) = \sum_{l=0}^{L_C} \beta_{C,l}\ T_{C,t-l} + \sum_{l=0}^{L_H} \beta_{H,l}\ T_{H,t-l} \tag{16}$$

where $L_C, L_C = 0, 1, \ldots$ is the maximum lag days for cold considered, $L_H, L_H = 0, 1, \ldots$ is the maximum lag days for heat considered, $\beta_{C,l}$ are the regression coefficients for the effect of cold lagged $l = 0, \ldots, L_C$ days, $\beta_{H,l}$ are the regression coefficients for the effect of heat lagged $l = 0, \ldots, L_H$ days. The temperature differentials for cold and heat, $T_{C,t}$ and $T_{H,t}$, respectively, are measured as follows

$$T_{C,t} = \max\left\{(\tau_C - \text{temp}_t^{min}), 0\right\} \times \mathcal{I}_C(t);\ T_{H,t} = \max\left\{(\text{temp}_t^{max} - \tau_H), 0\right\} \times \mathcal{I}_H(t) \tag{17}$$

where τ_C is the minimum temperature (temp^{min}) value that occurs with a relative frequency p_C considering the minimum temperature distribution during the cold semester, i.e. τ_C is the cold threshold in a specific area, τ_H is the maximum temperature (temp^{max}) value value that occurs with a relative frequency p_H considering the maximum temperature distribution during the warm semester, i.e. τ_H is the heat threshold, in a specific area. Additionally, temp_t^{min} and temp_t^{min} refer to the minimum and maximum temperatures observed on day t. In an effort not to have a heat differential during the winter and vice-versa, the temperature differentials for cold and heat are restricted to their specific semesters, thus, they are allowed to be different from zero in the November, December, January to April months (cold semester) for the cold differential and May until October (hot semester) for heat differential. This is achieved with the use of the indicator functions, $\mathcal{I}_C(t)$ and $\mathcal{I}_H(t)$, defined as follows

$$\begin{aligned} \mathcal{I}_C(t) &= \begin{cases} 1 & \text{if day } t \text{ belongs to the cold semester} \\ 0 & \text{otherwise} \end{cases} \\ \mathcal{I}_H(t) &= \begin{cases} 1 & \text{if day } t \text{ belongs to the hot semester} \\ 0 & \text{otherwise} \end{cases} . \end{aligned} \tag{18}$$

In other words, $T_{C,t}$ and $T_{H,t}$ measure the excess of cold during the cold semester and the excess of heat during the hot semester based on defined cold and heat thresholds, respectively.

One can also experiment with constraints in the DLM. Given the advantages that constrained models provide, the following coefficient constraints will be considered

$$\begin{aligned} \beta_{C,4} = \beta_{C,5},\ \beta_{C,6} = \beta_{C,7},\ \beta_{C,8} = \cdots = \beta_{C,11},\ \beta_{C,12} = \cdots = \beta_{C,L_C} \\ \beta_{H,4} = \beta_{H,5},\ \beta_{H,6} = \beta_{H,7} \end{aligned} \tag{19}$$

The structure of $f(\text{temp}_{t-\Delta t}|\boldsymbol{\beta})$ will allow the estimation of the overall cumulative risk and the analysis of the lag structure of the effects of low and high temperatures and the estimation of the number of deaths or hospitalisations per day, given a set of temperature observations.

Each parameter ($\tau_C,\ \tau_H,\ L_C,\ L_H,\ L_{ILI}$) used in this model will be specific to each area of territory. The values of the parameters will be chosen based on which set of parameter values improve a certain fit criterion. The procedure for this sensitivity analysis to the parameter values chosen will be explained in a later section.

In the case of Model 1, the number of expected deaths given no exposure to extreme temperatures A_0 for a given day t, mentioned in the Section 4.1 will be given by the model by replacing the values of the temperature differentials for cold and heat, $T_{C,t}$ and $T_{H,t}$, by 0 and the daily influenza incidence values by 0, as well. Thus, achieving a model estimate without considering any exposure to extreme temperatures and providing a baseline daily time series for mortality.

4.3.2. Model Specification Choice

Having specified the fit criterion to measure the goodness-of-fit of the models, the QAIC, this sections aims to detail the model choices that are going to be considered and evaluated, so as to reach the model with the best fit for each area of territory. This will be made for each area of territory and for Model 1. A grid search approach will be considered.

Regarding Model 1, for each area of territory, the following listing presents the parameters for which a value needs to be chosen and the values indeed considered for each one,

- L_C maximum number of lag days for cold, $L_C \in \{8, 10, 12, 14\}$,
- L_H maximum number of lag days for heat, $L_H \in \{4, 5, 6, 7\}$,
- L_{ILI} the maximum number of lag days for influenza incidence rate, $L_{ILI} \in \{8, 10, 12, 14\}$
- τ_C the cold threshold, with a relative frequency p_C considering the minimum temperature distribution during the cold semester, $p_C \in \{0.40, 0.35, 0.30, 0.25, 0.20, 0.15, 0.10, 0.05, 0.01\}$,
- τ_H the hot threshold, with a relative frequency p_H considering the maximum temperature distribution during the warm semester, $p_H \in \{0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 0.99\}$.

A grid search will be used to extensively evaluate each combination possible for each area of territory and its goodness-of-fit. According to other research [35], the Akaike Information Criterion for overdispersed models (QAIC) will be the fit criterion used.

4.4. Results

Some of the results and plots presented in this section rely on the `R` software package *dlnm* [41], developed by Gasparrini.

The models estimated refer only to temperature-mortality associations and were estimated only for data pertaining to the years between 1995 and 2019. Moreover, the areas of territory chosen in which to perform model estimation were the 18 districts of Mainland Portugal. A disaggregation of data to the municipality level would entail very few daily counts per municipality, which would have consequences on precision of the models [42]. Thus, the decision was made to use district-aggregated death count data to estimate the models and, subsequently, use the estimated district models and use the municipality temperature data as inputs. To do this, the assumption that the characteristics of the effects of temperature on mortality estimated for the district would apply and generalise to each of the district's municipality.

Figure 2 displays daily time series for deaths, maximum and minimum temperatures and influenza incidence at a national level for the 2012 and 2013 years. During this period, days 2012-02-23 and 2013-07-08 (identified with dashed vertical lines) were the deadliest.

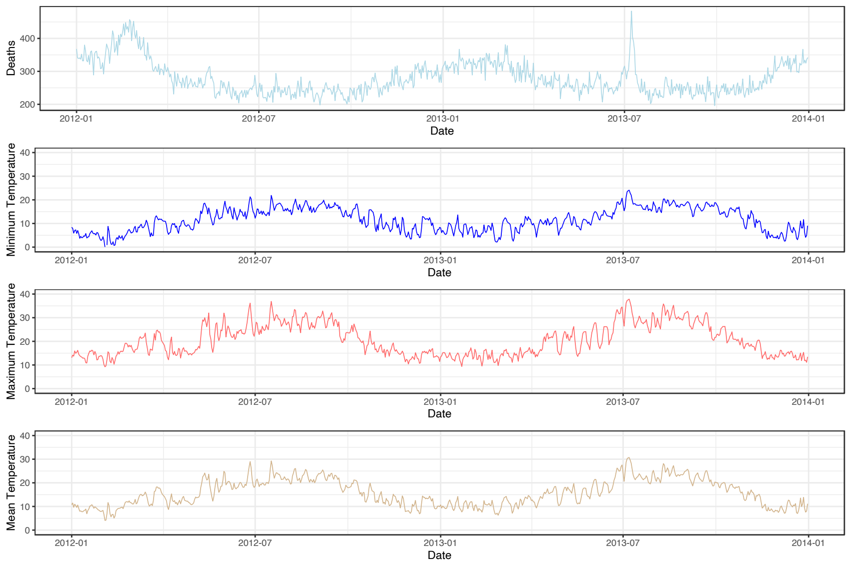


Figure 2. Time series representation for deaths, maximum and minimum temperatures and influenza incidence at a national level for the 2012 and 2013 years.

In this section, the specification choice results for Model 1 will be presented. Additionally, the results for the Lisboa District will be presented and used as examples of what could be presented for all of the other districts.

4.4.1. Model 1

The grid search for the assessment of the best set of parameter choices for the model were performed on a district basis. The QAIC was used to compare the model's goodness-of-fit given a set of parameter choices. Table 1 summarises the results for the parameter choices per District that resulted in the best quality models.

First of all, it can be seen that the optimum values for L_C are equal across most of the range of districts, with up to 14 days of lag for cold. Only two of the districts don't have $L_C = 14$, Beja and Castelo Branco, with $L_C = 8$ and $L_C = 10$, respectively. Regarding the maximum number of lag days for the influenza incidence rate, lags with the better model fit varied between districts, with a third of districts having a $L_{ILI} = 8$. Unlike L_C, optimum values for L_H differ between districts, with 5 days of lag being the most common. Secondly, regarding the resulting optimum temperature thresholds, the probabilities associated with the cold threshold, p_C, are equal for a great majority of the of districts, settling at larger quantiles, suggesting that cold starts to have an effect on mortality at milder temperatures. On the other hand, the probabilities associated with the hot threshold p_H have a more diverse and extreme selection. In fact, with the cold thresholds

Table 1. Results for the best parameter choices per District, for Model 1.

District	L_C	L_H	L_{ILI}	p_C	p_H	τ_C (°C)	τ_H (°C)	QAIC
Aveiro	14	4	8	0.40	0.85	8.67	28.48	52481.17
Beja	14	4	8	0.40	0.60	8.19	30.73	43206.05
Braga	14	6	8	0.40	0.85	7.04	29.08	52187.64
Bragança	14	5	8	0.35	0.70	3.70	28.81	41293.55
Castelo Branco	14	6	8	0.25	0.75	4.64	31.41	44853.25
Coimbra	14	4	8	0.40	0.80	8.23	28.41	50463.01
Évora	14	7	8	0.40	0.85	8.05	35.42	42286.90
Faro	14	6	8	0.40	0.65	10.45	28.73	49586.64
Guarda	14	7	8	0.40	0.85	4.94	31.25	43983.33
Leiria	14	5	8	0.40	0.70	9.69	25.56	49929.82
Lisboa	14	7	8	0.40	0.85	11.75	26.38	64797.67
Portalegre	14	4	8	0.35	0.60	7.03	30.68	40896.99
Porto	14	4	8	0.40	0.85	8.64	27.95	60857.06
Santarém	14	7	8	0.40	0.85	8.48	32.89	51266.56
Setúbal	14	5	8	0.40	0.85	10.56	29.94	54745.48
Viana do Castelo	14	4	8	0.35	0.75	6.08	25.70	45062.07
Vila Real	14	4	8	0.40	0.60	4.83	25.32	43834.12
Viseu	14	6	8	0.40	0.85	5.43	29.91	49400.88

the results showed that 16 out of the 18 districts have an optimum p_C of 40%, associated with the milder quantile admissible in the grid search performed, never exceeding the 30% quantile of minimum temperatures, while the optimum p_H range from 60% to 85% with 7 districts having the 85% quantile as the optimum heat threshold, which is a more extreme position in the distribution of maximum temperatures. The placement of the optimum temperatures quantiles suggests that the Portuguese population has a lower tolerance to cold temperatures.

Figure 3 shows a graphical representation of the temperature thresholds τ_C and τ_H organised by decreasing latitude, considering the district's geographical centroid. In this graphic, it is clear that some districts have a larger range between temperature thresholds, $\tau_H - \tau_C$. This could be seen as a measure of thermal comfort per district, as some populations could be considered more resilient than others. For the same temperature exposure, some populations have already exceed their temperature thresholds while others haven't.

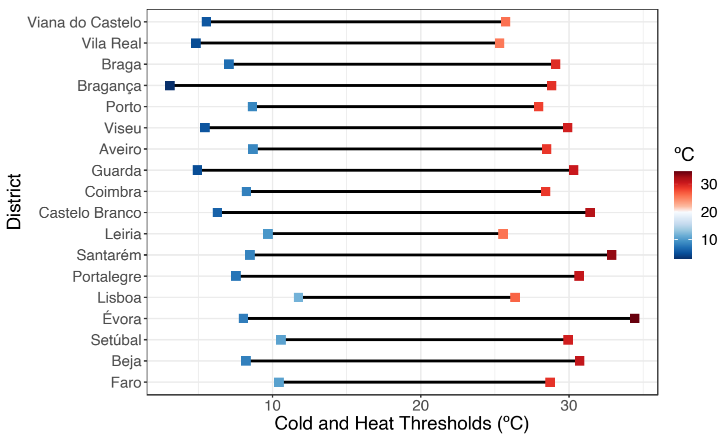


Figure 3. Graphical representation of the temperature thresholds, τ_C and τ_H, organised by decreasing district centroid latitude, for Model 1.

Just a reminder that, the temperature values differ per district with the same p_C or p_H because each model is estimated using that district's own temperature data, hence, own temperature distribution.

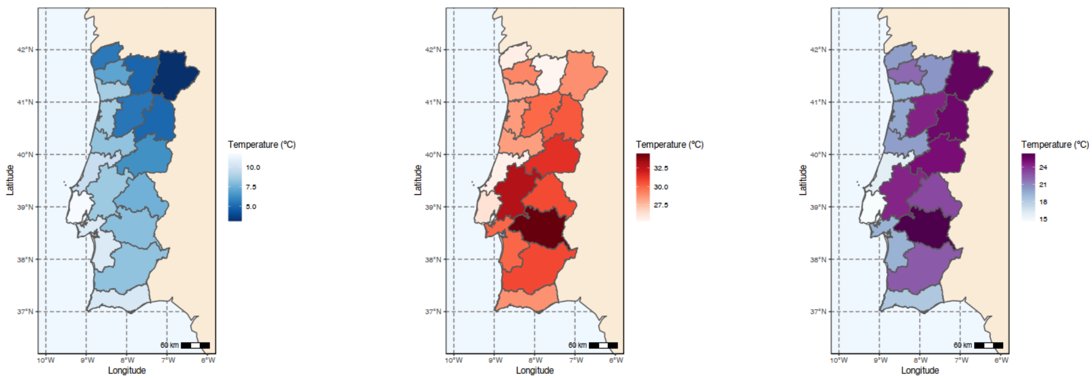
Table 2 shows some summary statistics for τ_C, τ_H and $\tau_H - \tau_C$.

Figures 4a and 4b show the results for the cold and heat thresholds, τ_C and τ_H, respectively,

Table 2. Summary statistics of τ_C, τ_H and $\tau_H - \tau_C$.

Variable	Min.	1st Q.	Median	Mean	3rd Q.	Max.	SD
τ_C	3.08	5.71	8.12	7.63	8.66	11.75	2.28
τ_H	25.32	28.07	28.95	29.15	30.59	34.42	2.47
$\tau_H - \tau_C$	14.63	19.49	21.26	21.52	24.46	26.37	3.40

and 4c shows the respective threshold range, $\tau_H - \tau_C$, represented on the map of Mainland Portugal.



(a) Cold thresholds. (b) Heat thresholds. (c) Threshold range.

Figure 4. Results for the best parameter choices for temperatures thresholds per district.

A quick visual inspection of the maps suggests that the cold threshold values decrease as latitude increases, with lower thresholds being recorded in the northern part of the country, and with cold threshold values also decreasing with longitude, with lower thresholds being recorded in the more inland parts of the country (related to lower longitudes). In regards to the heat thresholds, another pattern appears. Northern districts have lower heat thresholds whilst more inland regions seem to have higher heat thresholds. Moreover, threshold range could also be associated with longitude, as closeness to the sea provides milder climates.

To support the aforementioned associations, simple linear regressions to explain the variations τ_C and τ_H per latitude, longitude individually or at the same time were tested. The geographical location occupied by each district was summarised into the district's geographical centroids. Overall, the cold threshold τ_C revealed better associations with latitude and longitude than the τ_H, this could be due to the the fact that p_C is the same for a great majority of the districts and the correlation with latitude and longitude could come only from the district's temperature distributions and its relation with latitude and longitude. Nevertheless, these are interesting associations.

In regards to cold, populations seem to acclimatise to their specific temperature exposure distribution and not to a fixed value minimum temperature threshold. The argument is not so strong when heat is concerned. In fact, τ_H and p_H have a larger variation than their cold counterparts. Half of the districts share the same p_H of 85%, suggesting some stabilisation on the same relative threshold, but not as strong an evidence as in the case of cold. Additionally, thresholds range show to have an association with longitude, with districts closer to the Western coastline being associated with a lower thresholds range.

Model estimates make it possible to calculate daily values for $\widehat{RR}$ per district to be of use as a risk measure. Table 3 displays some summary statistics for the computed $\widehat{RR}$.

Table 4 summarises the cumulative effects RR over the lag period and exposure considered. The temperature values corresponding to the to the district's 1% minimum and maximum tem-

Table 3. Summary statistics of daily $\widehat{RR}$ per district.

District	Min.	1st Q.	Median	Mean	3rd Q.	95% Q.	99% Q.	99.9% Q.	Max.	SD
Aveiro	0.84	0.94	0.98	1.00	1.02	1.20	1.36	1.54	1.70	0.09
Beja	0.86	0.94	0.97	1.00	1.03	1.19	1.32	1.51	1.67	0.09
Braga	0.86	0.95	0.98	1.00	1.01	1.19	1.35	1.54	1.79	0.09
Bragança	0.87	0.94	0.96	1.00	1.02	1.21	1.36	1.47	1.58	0.09
Castelo Branco	0.86	0.94	0.96	1.00	1.03	1.22	1.39	1.62	1.77	0.10
Coimbra	0.86	0.95	0.98	1.00	1.02	1.20	1.35	1.51	1.58	0.09
Évora	0.88	0.95	0.97	1.00	1.02	1.19	1.36	1.53	1.78	0.09
Faro	0.89	0.93	0.96	1.00	1.03	1.21	1.39	1.58	1.81	0.10
Guarda	0.83	0.94	0.97	1.00	1.02	1.22	1.41	1.63	1.83	0.10
Leiria	0.85	0.95	0.98	1.00	1.02	1.16	1.29	1.43	1.53	0.08
Lisboa	0.87	0.95	0.98	1.00	1.01	1.17	1.33	1.51	1.81	0.08
Portalegre	0.87	0.93	0.96	1.00	1.03	1.21	1.35	1.51	1.66	0.10
Porto	0.88	0.95	0.98	1.00	1.01	1.18	1.31	1.49	1.71	0.08
Santarém	0.87	0.95	0.97	1.00	1.01	1.20	1.39	1.62	1.93	0.09
Setúbal	0.88	0.95	0.98	1.00	1.01	1.18	1.33	1.56	2.05	0.09
Viana do Castelo	0.87	0.94	0.97	1.00	1.02	1.21	1.38	1.58	1.74	0.10
Vila Real	0.85	0.94	0.97	1.00	1.03	1.19	1.31	1.49	1.59	0.09
Viseu	0.87	0.94	0.97	1.00	1.02	1.21	1.40	1.80	2.03	0.10

perature quantiles and a 90% CI are also presented. Considering these temperatures for the cold and heat cases, heat presents a higher risk of mortality to the population when compared to cold. It is interesting to see that, in the case of Évora and Faro, an exposure of 40°C in Évora presents less of a risk than an exposure of 35°C in Faro, in terms of cumulative RR. This is clearly due to the higher temperatures that the Évora population is exposed and their ability to adapt to these temperatures. A similar situation can be named for the cold case. Let's consider as examples the Guarda and Lisboa districts. For a temperature exposure of -1°C, Guarda's population is subjected to less risk of mortality than the Lisboa population at an exposure of 7°C. A fairly different temperature exposure. This highlights the temperature variations and discrepancies in population acclimatisation within Mainland Portugal.

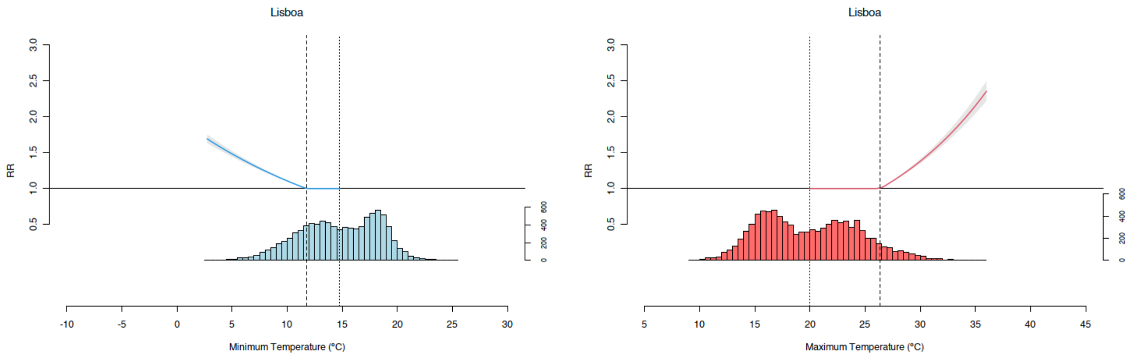
Table 4. Cumulative temperature effects per district and type of exposure, cold or heat.

	Cold			Heat		
District	1% Q. $temp^{min}(°C)$	RR	90% CI	99% Q. $temp^{max}(°C)$	RR	90% CI
Aveiro	2.56	1.37	1.32 - 1.43	33.45	1.41	1.36 - 1.46
Beja	1.93	1.23	1.16 - 1.31	39.08	1.39	1.32 - 1.46
Braga	0.46	1.34	1.29 - 1.40	33.72	1.41	1.36 - 1.47
Bragança	-1.54	1.33	1.24 - 1.43	35.02	1.36	1.29 - 1.44
Castelo Branco	0.16	1.27	1.20 - 1.34	37.43	1.56	1.48 - 1.64
Coimbra	1.81	1.34	1.27 - 1.40	34.72	1.41	1.35 - 1.47
Évora	1.99	1.37	1.28 - 1.48	40.05	1.45	1.36 - 1.54
Faro	5.00	1.28	1.22 - 1.34	34.54	1.54	1.49 - 1.60
Guarda	-1.15	1.20	1.13 - 1.28	35.42	1.51	1.43 - 1.59
Leiria	3.69	1.29	1.23 - 1.35	33.59	1.30	1.25 - 1.35
Lisboa	6.66	1.35	1.31 - 1.38	30.23	1.41	1.38 - 1.44
Portalegre	1.45	1.24	1.14 - 1.33	39.56	1.50	1.42 - 1.58
Porto	2.82	1.35	1.31 - 1.39	32.70	1.38	1.35 - 1.42
Santarém	1.91	1.32	1.26 - 1.38	38.34	1.51	1.45 - 1.57
Setúbal	5.03	1.37	1.32 - 1.42	34.13	1.40	1.35 - 1.44
Viana do Castelo	0.31	1.44	1.36 - 1.53	31.94	1.45	1.38 - 1.52
Vila Real	-1.30	1.23	1.15 - 1.30	33.00	1.33	1.27 - 1.39
Viseu	-0.79	1.33	1.27 - 1.40	34.12	1.64	1.57 - 1.72

4.4.2. Lisboa District

The Lisboa district will be the use case for the several results that can be more comprehensively presented and performed for each one of the other districts. This specific district was chosen on the basis that is the most populated district and provided larger daily mortality counts than the other districts, thus, providing more robust results.

Figure 5 combines the exposure-response curves for the Lisboa district provided by Model 1. The lines represent the estimated cumulative RR when only temperature exposure is concerned. Moreover, it is an overall cumulative RR because the risk is cumulated over the lag period considered. The solid vertical lines represent the district's average minimum or maximum temperature during the study period and dashed lines represent the district's cold or heat thresholds, respectively. The grey shadowed area around the coloured blue or red exposure–response lines represent their respective RR 90% CI. For these risk-temperature associations, the computed RR has as its baseline the number of deaths when the respective population is exposed to the average minimum or maximum temperature of its district, accordingly. When minimum temperatures are concerned, exposures below the cold threshold are associated with an increase in RR. Meanwhile, when maximum temperatures are concerned, exposures above the heat threshold are associated with an increase in RR.

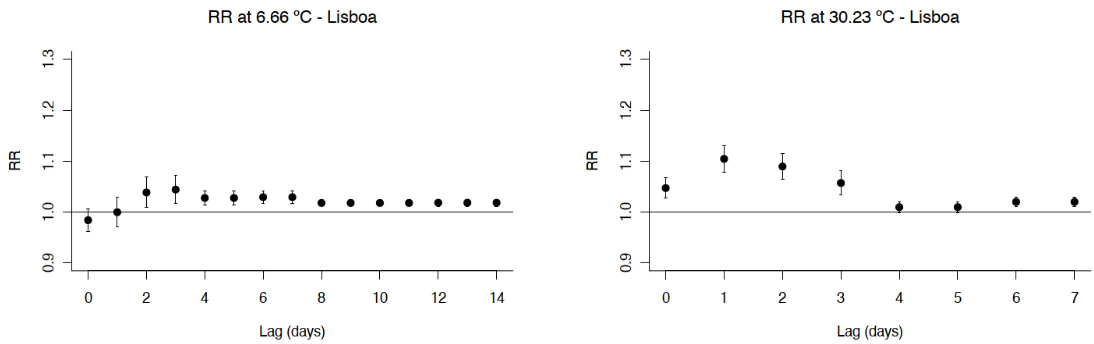


(a) Model 1 - Minimum Temperatures.

(b) Heat thresholds.

Figure 5. Overall cumulative RR associated with minimum (left) and maximum (right) temperatures for Model 1 for the Lisboa district.

Besides the exposure-response curves, there are lag-response structures. Figures 6a and 6b combine the lag-response structures for the Lisboa district, considering an exposure to cold and heat, respectively.



(a) Model 1 - Cold Effect.

(b) Model 1 - Heat Effect.

Figure 6. Lagged effects, RR, considering an exposure equal to the district's 1% minimum temperature quantile and an exposure equal to the district's 99% maximum temperature quantile, with their respective 90% CI.

In the case of cold, peak risk is reached at the third day after exposure, with a controlled decrease of risk in the following days. As expected per the literature, cold effects seem to linger over the lag period, resisting to reach a non-significant risk. At the same time, heat effects reach a peak risk at one day after exposure and are followed by a more accentuated descend of risk over the next couple of days after the peak. From the fourth day onwards, heat effects are fairly negligible. The cold effects are, thus, characterised by a delayed peak risk and a more lasting effect while heat effects are more immediate and don't last nearly as long.

5. Discussion

At the municipality level, a difficulty that was encountered was the small number of daily death counts at the municipality level, as there were several days in several municipalities that had null or very low count of deaths. According to Bhaskaran et al. [42], an average of tens of events per day is needed for credible precision of the estimated models. This problem was solved by moving on to the next level of aggregation, the districts, consequently reducing the spatial resolution. Nevertheless, it was a trade that was forcibly necessary in order to complete what was set out to be done.

In this work, temperature-mortality associations were estimated for each one of the 18 districts in Mainland Portugal based on one model. The results of the model were presented and a comparison was made between each district and each model. The models provide mortality estimates for the whole year and make it possible to establish an all-year round mortality baseline which, subsequently, makes it possible to estimate a RR of mortality. As it was explained, the model not only accounts for the effects of temperature exposure but also, the effects of influenza incidence as illustrated in the DAG in Figure 1. The $\widehat{RR}$ values provided reflect this relationship and, given the similarity between the seasonality of cold temperatures and of influenza incidence, early warning systems should also take this into account [19, 37].

Model 1 describes a simple and conservative approach to the challenge of estimating temperature-mortality associations. It is based on the DLM framework and receive as inputs time of the year, maximum and minimum temperature values, influenza incidence to provide a daily mortality estimate. The use of maximum and minimum temperature values determine that the model takes into the account the most extreme temperature exposure possible registered in the study period and provide a risk measure. As the effects of climate change increase the frequency of extreme temperature events [43], mean average daily temperature may not provide the full picture of exposure. Thus, the use of maximum and minimum temperatures could provide a better framework for public health prevention and early warning systems in the very near future. This was the rationale behind the formulation of Model 1. Furthermore, the analysis of the optimum temperature thresholds has provided evidence of population acclimatisation to their district's climate [44]. Optimum cold and heat thresholds revealed to be both different at the real temperature value and at the relative temperature scale, indicating that the population acclimatises to their own environment and perform behavioural adaptations. Were this not to happen, a clear evidence of non-conformity to a specific climate would be very similar real temperature threshold values, which is clearly not the case given the results presented in this work. Evidence of population acclimatisation to their specific climates has been frequently presented in the literature [1, 44–46]. Additionally, it is also interesting to note that optimum cold thresholds are, in most cases, relatively mild temperatures when compared to optimum heat thresholds, which lay in the more extreme side of the maximum temperature distribution. This suggests that the effects of *cold* temperatures on mortality start at fairly milder temperatures and continue stronger as temperatures decrease. This is corroborated by the findings of Gasparrini et al. [1] which concluded that the effect of days of extreme temperature was substantially less than that attributable to milder but non-optimum weather with more temperature-attributable deaths caused by cold. Moreover, optimum cold thresholds showed to have a fairly significant relationship with latitude and longitude, which is in line with what has been reported [1, 44].

In terms of limitations, the use of a district temperature index results in a loss of information because municipality-specific temperatures are summarised by their weighted averages. The more spatial specific is a temperature exposure, the more representative it is. However, due to the aforementioned challenges of low mortality counts this was a necessary step. Secondly, different population sizes and population structures, like ageing populations, per district result in different mortality counts. Despite models being adjusted for population size, lower or higher mortality counts have impacts on model precision.

References

[1] Antonio Gasparrini, Yuming Guo, Masahiro Hashizume, Eric Lavigne, Antonella Zanobetti, Joel Schwartz, Aurelio Tobias, Shilu Tong, Joacim Rocklov, Bertil Forsberg, Michela Leone, Manuela De Sario, Michelle L Bell, Yue-Liang Leon Guo, Chang fu Wu, Haidong Kan, Seung-Muk Yi, Micheline de Sousa Zanotti Stagliorio Coelho, Paulo Hilario Nascimento Saldiva, Yasushi Honda, Ho Kim, and Ben Armstrong. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet*, 386(9991):369–375, 2015. ISSN 0140-6736. . URL `https://www.sciencedirect.com/science/article/pii/S0140673614621140`.

[2] Xuping Song, Shigong Wang, Yuling Hu, Man Yue, Tingting Zhang, Yu Liu, Jinhui Tian, and Kezheng Shang. Impact of ambient temperature on morbidity and mortality: An overview of reviews. *Science of the Total Environment*, 586(222):241–254, 2017. ISSN 18791026. . URL `http://dx.doi.org/10.1016/j.scitotenv.2017.01.212`.

[3] Antonio Gasparrini and Ben Armstrong. Reducing and meta-analysing estimates from distributed lag non-linear models. *BMC Medical Research Methodology*, 13(1):1–10, 2013. ISSN 14712288. .

[4] K. Laaidi, A. Economopoulou, V. Wagner, M. Pascal, P. Empereur-Bissonnet, A. Verrier, and P. Beaudeau. Cold spells and health: Prevention and warning. *Public Health*, 127(5):492–499, 2013. ISSN 00333506. . URL `http://dx.doi.org/10.1016/j.puhe.2013.02.011`.

[5] Jun Yang, Peng Yin, Jimin Sun, Boguang Wang, Maigeng Zhou, Mengmeng Li, Shilu Tong, Bohan Meng, Yuming Guo, and Qiyong Liu. Heatwave and mortality in 31 major Chinese cities: Definition, vulnerability and implications. *Science of the Total Environment*, 649:695–702, 2019. ISSN 18791026. . URL `https://doi.org/10.1016/j.scitotenv.2018.08.332`.

[6] Who/europe - heat threatens health: key figures for europe. `https://www.euro.who.int/en/health-topics/environment-and-health/Climate-change/archive/heat-threatens-health-key-figures-for-europe`, . Accessed: 10-03-2022.

[7] Ipma - Área educativa - onda de calor. `https://www.ipma.pt/pt/educativa/glossario/meteorologico/index.jsp?page=glossario_op.xml&print=true`, . Accessed: 10-03-2022.

[8] Who/europe - how cold weather affects health. `https://www.euro.who.int/en/health-topics/environment-and-health/Climate-change/news/news/2013/02/how-cold-weather-affects-health`, . Accessed: 10-03-2022.

[9] Chengzhen Meng, Fang Ke, Yao Xiao, Suli Huang, Yanran Duan, Gang Liu, Shuyuan Yu, Yingbin Fu, Ji Peng, Jinquan Cheng, and Ping Yin. Effect of Cold Spells and Their Different Definitions on Mortality in Shenzhen, China. *Frontiers in Public Health*, 9(January):1–9, 2022. ISSN 22962565. .

[10] R. Carmona, J. Díaz, I. J. Mirón, C. Ortíz, I. León, and C. Linares. Geographical variation in relative risks associated with cold waves in Spain: The need for a cold wave prevention plan. *Environment International*, 88:103–111, 2016. ISSN 18736750. . URL `http://dx.doi.org/10.1016/j.envint.2015.12.027`.

[11] Liliana Antunes, Susana Pereira Silva, Jorge Marques, Baltazar Nunes, and Sílvia Antunes. The effect of extreme cold temperatures on the risk of death in the two major Portuguese cities. *International Journal of Biometeorology*, 61(1):127–135, 2017. ISSN 00207128. . URL `http://dx.doi.org/10.1007/s00484-016-1196-x`.

[12] Ipma - Área educativa - onda de frio. `https://www.ipma.pt/pt/educativa/glossario/meteorologico/index.jsp?page=glossario_op.xml`, . Accessed: 10-03-2022.

[13] Phong K. Thai, Christhina Cândido, Akwasi Asumadu-Sakyi, Adrian Barnett, and Lidia Morawska. Variation of indoor minimum mortality temperature in different cities: Evidence of local adaptations. *Environmental Pollution*, 246:745–752, 2019. ISSN 18736424. .

[14] Mindy Anderson, Catriona Carmichael, Virginia Murray, Andy Dengel, and Michael Swainson. Defining indoor heat thresholds for health in the UK. *Perspectives in Public Health*, 133(3):158–164, 2013. ISSN 17579139. .

[15] R. Jevons, C. Carmichael, A. Crossley, and A. Bone. Minimum indoor temperature threshold recommendations for English homes in winter – A systematic review. *Public Health*, 136:4–12, 2016. ISSN 14765616. . URL `http://dx.doi.org/10.1016/j.puhe.2016.02.007`.

[16] Jian Cheng, Zhiwei Xu, Hilary Bambrick, Hong Su, Shilu Tong, and Wenbiao Hu. Impacts of heat, cold, and temperature variability on mortality in Australia, 2000–2009. *Science of the Total Environment*, 651:2558–2565, 2019. ISSN 18791026. . URL `https://doi.org/10.1016/j.scitotenv.2018.10.186`.

[17] S. Vandentorren, P. Bretin, A. Zeghnoun, L. Mandereau-Bruno, A. Croisier, C. Cochet, J. Ribéron, I. Siberan, B. Declercq, and M. Ledrans. August 2003 heat wave in France: Risk factors for death of elderly people living at home. *European Journal of Public Health*, 16(6):583–591, 2006. ISSN 11011262. .

[18] J. Díaz, R. García-Herrera, R. M. Trigo, C. Linares, M. A. Valente, J. M. De Miguel, and E. Hernández. The impact of the summer 2003 heat wave in Iberia: How should we measure it? *International Journal of Biometeorology*, 50(3):159–166, 2006. ISSN 00207128. .

[19] Tom Fowler, Rosamund J. Southgate, Thomas Waite, Ruth Harrell, Sari Kovats, Angie Bone, Yvonne Doyle, and Virginia Murray. Excess winter deaths in Europe: A multi-country descriptive analysis. *European Journal of Public Health*, 25(2):339–345, 2015. ISSN 1464360X. .

[20] J. D. Healy. Excess winter mortality in Europe: A cross country analysis identifying key risk factors. *Journal of Epidemiology and Community Health*, 57(10):784–789, 2003. ISSN 0143005X. .

[21] Paulo Nogueira and Eleonora Paixão. Models for mortality associated with heatwaves: update of the Portuguese heat health warning system. *International Journal of Climatology*, 28(4):545–562, 2008. . URL `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1546`.

[22] Pedro M A Miranda, F átima E S Coelho, António R Tomé, Maria A Valente, Anabela Carvalho, Carlos Pires, Henrique O Pires, Vanda C Pires, and Carlos Ramalho. 2 20 th Century Portuguese Climate and Climate Scenarios. pages 23–83, 2002.

[23] M. Isabel P. De Lima, Fátima Espírito Santo, Alexandre M. Ramos, and João L.M.P. de Lima. Recent changes in daily precipitation and surface air temperature extremes in mainland Portugal, in the period 1941-2007. *Atmospheric Research*, 127:195–209, 2013. ISSN 01698095. . URL `http://dx.doi.org/10.1016/j.atmosres.2012.10.001`.

[24] Instituto Nacional de Estatística. *Censos 2011 Resultados Definitivos - Portugal*. Instituto Nacional de Estatística, 2011.

[25] Rosa Oliveira, Carlos C Antonio, Hugo Santos, and Eduardo O Fernandes. the Portuguese Housing Stock: an Initial Analysis of the National Energy Certification Database. *Proceedings of the 7Th International Conference on Mechanics and Materials in Design (M2D2017)*, (June):1595–1600, 2017.

[26] Portal da transparência do serviço nacional de saúde - enquadramento. `https://www.sns.gov.pt/transparencia/enquadramento/`. Accessed: 21-04-2022.

[27] Adrian G. Barnett and Annette J. Dobson. *Analysing Seasonal Health Data*. Statistics for Biology and Health. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-10747-4. . URL `http://link.springer.com/10.1007/978-3-642-10748-1`.

[28] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL `https://www.R-project.org/`.

[29] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2019. URL `http://www.rstudio.com/`.

[30] Peter K Dunn and Gordon K Smyth. *Generalized Linear Models With Examples in R by Peter K. Dunn, Gordon K. Smyth (z-lib.org).pdf*. ISBN 9781441901170.

[31] Manuel L Esquível and Pedro P Mota. *Inferência e Modelação Estística*. Nova FCT Editorial, 2020. ISBN 9789899952898.

[32] John Fox. *Applied Regression Analysis and Generalized Linear Models*. SAGE Publications, 2008. ISBN 9780761930426. URL `http://books.google.ch/books?id=GKkn3LSSHFsC`.

[33] Kyle Steenland and Ben Armstrong. An overview of methods for calculating the burden of disease due to specific risk factors. *Epidemiology*, 17(5):512–519, 2006. ISSN 10443983. .

[34] Joel Schwartz, Source Epidemiology, No May, and Joel Schwartz. The Distributed Lag between Air Pollution and Daily Deaths. *Epidemiology*, 11(3):320–326, 2000.

[35] Ben Armstrong. Models for the relationship between ambient temperature and daily mortality. *Epidemiology*, 17(6):624–631, 2006. ISSN 10443983. .

[36] Antonio Gasparrini and Michela Leone. Attributable risk from distributed lag models. *BMC Medical Research Methodology*, 14(1):1–8, 2014. ISSN 14712288. .

[37] Ji Eun Park, Woo Sik Son, Yeonhee Ryu, Soo Beom, Okyu Kwon, and Insung Ahn. Effects of temperature , humidity , and diurnal temperature range on influenza incidence in a temperate region. (April 2018):11–18, 2020. .

[38] Priscila Corraini, Morten Olsen, Lars Pedersen, Olaf M. Dekkers, and Jan P. Vandenbroucke. Effect modification, interaction and mediation: An overview of theoretical insights for clinical investigators. *Clinical Epidemiology*, 9:331–338, 2017. ISSN 11791349. .

[39] Robert E Serfling. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public health reports (Washington, D.C. : 1896)*, 78(6):494–506, 1963. ISSN 0094-6214. URL https://pubmed.ncbi.nlm.nih.gov/19316455https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1915276/.

[40] Juan Carlos Alberdi, Julio Díaz, Juan Carlos Montero, and Isidro Mirón. Daily mortality in Madrid community 1986–1992: Relationship with meteorological variables. *European Journal of Epidemiology*, 14(6):571–578, 1998. ISSN 1573-7284. . URL https://doi.org/10.1023/A:1007498305075.

[41] Antonio Gasparrini. Distributed lag linear and non-linear models in R: The package dlnm. *Journal of Statistical Software*, 43(8):2–20, 2011. ISSN 15487660. .

[42] Krishnan Bhaskaran, Antonio Gasparrini, Shakoor Hajat, Liam Smeeth, and Ben Armstrong. Time series regression studies in environmental epidemiology. *International Journal of Epidemiology*, 42 (4):1187–1195, 2013. ISSN 03005771. .

[43] Michael Sanderson, Katherine Arbuthnott, Sari Kovats, Shakoor Hajat, and Pete Falloon. *The use of climate information to estimate future mortality from high ambient temperature: A systematic literature review*, volume 12. 2017. ISBN 1111111111. .

[44] Christopher R. de Freitas and Elena A. Grigorieva. Role of acclimatization in weather-related human mortality during the transition seasons of autumn and spring in a thermally extreme mid-latitude continental climate. *International Journal of Environmental Research and Public Health*, 12(12): 14974–14987, 2015. ISSN 16604601. .

[45] Yuming Guo, Antonio Gasparrini, Ben G. Armstrong, Benjawan Tawatsupa, Aurelio Tobias, Eric Lavigne, Micheline de Sousa Zanotti Stagliorio Coelho, Xiaochuan Pan, Ho Kim, Masahiro Hashizume, Yasushi Honda, Yue Leon Guo, Chang Fu Wu, Antonella Zanobetti, Joel D. Schwartz, Michelle L. Bell, Ala Overcenco, Kornwipa Punnasiri, Shanshan Li, Linwei Tian, Paulo Saldiva, Gail Williams, and Shilu Tong. Temperature variability and mortality: A multi-country study. *Environmental Health Perspectives*, 124(10):1554–1559, 2016. ISSN 15529924. .

[46] Yuming Guo, Antonio Gasparrini, Ben Armstrong, Shanshan Li, Benjawan Tawatsupa, Aurelio Tobias, Eric Lavigne, Micheline De Sousa Zanotti Stagliorio Coelho, Michela Leone, Xiaochuan Pan, Shilu Tong, Linwei Tian, Ho Kim, Masahiro Hashizume, Yasushi Honda, Yue Liang Leon Guo, Chang Fu Wu, Kornwipa Punnasiri, Seung Muk Yi, Paola Michelozzi, Paulo Hilario Nascimento Saldiva, and Gail Williams. Global variation in the effects of ambient temperature on mortality: A systematic evaluation. *Epidemiology*, 25(6):781–789, 2014. ISSN 15315487. .