Statistical methods in studies on temperature-health associations

Antonio Gasparrini

A thesis submitted to the University of London for the degree of Doctor of Philosophy

London School of Hygiene and Tropical Medicine, 2011

To that tiny little thing,
which always reminds me that my job is not
the most important issue in my life

Declaration

Statement of Own Work

All students are required to complete the following declaration when submitting their thesis. A shortened version of the School's definition of Plagiarism and Cheating is as follows (the full definition is given in the Research Degrees Handbook):
The following definition of plagiarism will be used:
Plagiarism is the act of presenting the ideas or discoveries of another as ones own. To copy sentences, phrases or even striking expressions without acknowledgement in a manner which may deceive the reader as to the source is plagiarism. Where such copying or close paraphrase has occurred the mere mention of the source in a biography will not be deemed sufficient acknowledgement; in each instance, it must be referred specifically to its source. Verbatim quotations must be directly acknowledged, either in inverted commas or by indenting. (University of Kent)
Plagiarism may include collusion with another student, or the unacknowledged use of a fellow student's work with or without their knowledge and consent. Similarly, the direct copying by students of their own original writings qualifies as plagiarism if the fact that the work has been or is to be presented elsewhere is not clearly stated.
Cheating is similar to plagiarism, but more serious. Cheating means submitting another student's work, knowledge or ideas, while pretending that they are your own, for formal assessment or evaluation.
Supervisors should be consulted if there are any doubts about what is permissible.

Declaration by candidate

I have read and understood the School's definition of plagiarism and cheating given in the Research Degrees Handbook. I declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

Signed: [redacted]

Date: **August 12, 2011**

Full name: **Antonio Gasparrini**

Abstract

Research on the health effects of temperature has expanded greatly in recent years, mainly due to the occurrence of extreme weather events and predicted climate change scenarios. The development of appropriate statistical methodology has been an important component of this research, and standard approaches, primarily based on multi-city time series regression analysis, are now well established. However, particular aspects of temperature-health associations, such as the non-linear and delayed relationship and the joint handling of multi-city data, still pose important methodological problems. During my PhD research, I have contributed to the development of statistical methods that address two particular limitations of traditional approaches, focusing on the development of two modelling frameworks: distributed lag non-linear models and multivariate meta-analysis. The former is a class of models that specify simultaneously non-linear and delayed exposure-response relationships in time series data, while the latter is an extension of traditional meta-analysis for the combination of multiple correlated outcomes across studies, that is also applicable to multi-parameter associations. These methods are placed within the traditional two-stage approach that is adopted in temperature-health studies. The first stage is city-specific, with analyses deriving the estimated relationship within each city. The second-stage is meta-analytical procedure for combining the results from the first stage. I have implemented these modelling frameworks in two packages within the statistical environment R. In this PhD thesis I present a series of publications which summarize my research work. Their content focuses on three key aspects: the development of the statistical methodology, the implementation of the software, and the application of the methods to real data. The papers are preceded by an epidemiological and statistical introduction to the topic, and followed by a final discussion where I illustrate potential future developments and provide some conclusions. These methodological advancements contribute several improvements over standard methods that are applied to investigate temperature-health associations in time series data, and may be easily extended to other research fields and study designs.

Contents

List of Tables

List of Figures

Acknowledgements

I am grateful to my supervisor, Prof. Mike Kenward, for his support and clever supervision over the last three years. I will always remember his enthusiasm and shining eyes when we discuss challenging statistical problems.

I would also like to thank my co-supervisor, Prof. Ben Armstrong, for his wise guidance on the methodological and substantive issues related to my research topics. I am extremely indebted to him for the patience and wisdom he showed every time he let me take wrong decisions in order to learn from my mistakes.

The past three years would not have been so productive without the exceptional researchers I have had the chance to meet. In particular, my colleagues Paul Wilkinson, Zaid Chalabi, Cathryn Tonne and Shakoor Hajat, who I am sure I have bothered with my silly questions. I am also grateful to Bianca De Stavola for her advice.

My sincere gratitude also goes to my family and friends, who have being so close to me during these years. In particular, Giorgio, Sara, and Francesco for their company, Costanza for hundreds of reasons, my parents for their support, and my friends and colleagues James Grellier and Sotiris Vardoulakis for welcoming me at the very beginning.

The final thank you is to Genny, for all the things we have shared together, and for being there every time I needed it.

Preface

This PhD thesis consists of a collection of research papers, commentaries and software documentation. Although these publications are focused on the same research topic, they have been published, or submitted, as independent research contributions. Inevitably, concepts and definitions are often repeated in different papers, and, more importantly, their content is not uniformly linked and standardized. The thesis is therefore divided into three main parts, where the selected publications are preceded by an introduction and followed by a final discussion. The aim is to "tell the story" of my research activity during the PhD project, presenting my contribution to the topic as a coherent body of work.

The introduction in Part I contains two main chapters. The epidemiological and statistical context of temperature-health studies is illustrated in Chapter 1, focusing in particular on the methodological aspects of the association under study. Chapter 2 offers a summary of the publications, also introducing the main statistical developments and the related software implementation. In Part III, Chapter 12, I provide a final discussion and describe potential directions for future research.

Part II includes the selected nine publications in different chapters. The order has been chosen to reflect the progressive research steps of the PhD project. The first publication in Chapter 3 is an example of standard analysis of temperature-health associations. The paper in Chapter 4 discusses the methodological advancements and limitations of the studies in this field, and anticipates the two main statistical developments I wish to present, described in two blocks of papers in Chapters 5-7 and 8-9, respectively. The last Chapters 10 and 11 include two publications where the two statistical frameworks are applied for substantive analyses.

Part I

Introduction

Chapter 1

Epidemiological and statistical context

1.1 Temperature and health

Exposure to extreme temperatures has long been recognized as a threat to health, and, in the last decade, this association has been intensely scrutinized by the scientific community. This growing interest has been stimulated by specific episodes of extreme weather, characterized by an exceptional increase in mortality and other health outcomes. Particularly infamous events have been reported as public health disasters, for example the heat waves in Chicago during July 1995 (Semenza *et al.*, 1996, 1999) or in France during August 2003 (Le Tertre *et al.*, 2006; Poumadere *et al.*, 2005).

More generally, the need to deepen our understanding of the relationship between extreme temperature and health is motivated by the mounting evidence about climate change. Over the past century, the overall global surface temperature has increased by 0.4-0.8°C, and the global sea level has risen 10-25 cm from the melting of the polar ice caps (National Reasearch Council (NRC), 2000). The scenarios for the next decades predict an increased intensity and frequency of extreme temperature events, in particular heat waves (Luber and McGeehin, 2008; Meehl *et al.*, 2000), and several studies have foreseen an increased health impact (Kalkstein and Greene, 1997; O'Neill and Ebi, 2009; Patz *et al.*, 2005).

Thus, it is no surprise if, during recent years, the epidemiological literature assessing the health effects of temperature has greatly expanded: Gosling *et al.* (2009) provides a critical review of the literature on this issue, acknowledging the inter-disciplinary nature of the topic and examining the findings presented in epidemiological, environmental and climatological journals. A comprehensive review of the epidemiological evidence on the effects of hot temperatures published between 1970 and 2008, together with a thorough discussion on methodological issues, is offered by Basu and Samet (2002a) and Basu (2009). Kovats and Hajat (2008) have performed a similar assessment, focusing on the public health implications of temperature-related mortality and morbidity. Studies on the the health effects of cold temperatures are instead less common (Analitis *et al.*, 2008; Carder *et al.*, 2005; Eurowinter Group, 1997; Gorjanc *et al.*, 1999; Hajat and Haines, 2002; Wilkinson *et al.*, 2004).

Temperature indexes

The assessment of temperature-health associations commonly relies on ambient temperatures and other meteorological variables measured at central weather stations, characterizing the average exposure experienced by individuals living in the same city or region (Basu and Samet, 2002a; Gosling *et al.*, 2009). The exposure is usually collected at equally-spaced times, commonly on a daily basis, describing the temporal variability of temperatures. Different exposure indexes have been proposed, from maximum (Armstrong *et al.*, 2010) or minimum temperature (Schwartz, 2005) measured within each day, or most frequently mean temperature, defined as the average between maximum and minimum or between the hourly measurements (Hajat *et al.*, 2002; Pattenden *et al.*, 2003). Composite indexes with a combination of dry-bulb temperature and measures of humidity (relative humidity or dew-point temperatures) have been also proposed, for example apparent temperature (Michelozzi *et al.*, 2007; Stafoggia *et al.*, 2006) or the humidex (Conti *et al.*, 2005, 2007). Each measure has a specific characterization: composite indexes are commonly build as surrogates of the thermal stress of the body, a *perceived* temperature which depends also on humidity. In addition, in the ecological approach based on aggregated exposure measurements, the estimated effects of maximum and minimum temperature are frequently interpreted as the specific risk for exposures experienced by individuals during day and night time, respectively, and their comparison fosters additional hypotheses on the causal pathway

linking temperature and health outcomes. Barnett *et al.* (2010) have provided an assessment of the predictive ability of different exposure indexes in a large dataset including 107 cities in the USA, concluding that no measure shows a significant and coherent improved performance.

Some investigations have instead applied micro-environmental models to determine the personal exposure as the time-weighted average concentration in the different locations where participants in a study spend their time (Basu and Samet, 2002b), but this method is considered expensive as well impractical to be broadly implemented in epidemiological studies. Recently, more sophisticated studies have assessed the exposure at a finer geographical grid, using model-based predicted temperature in small-area analysis (Aylin *et al.*, 2001). Such approaches allow the inspection of further issues such as the *urban heat island* or effect modification by small-area characteristics, for example deprivation indexes, building type and land use characteristics.

Adaptation and susceptibility

The assessment of temperature as a risk factor for human health needs to accommodate the additional complexity of the adaptation of populations to their own climate. A suitable index to measure this phenomenon is the *comfort range*, the temperature band at which a specific population experiences the lowest risk. The limits of this range are commonly interpreted as the threshold values beyond which the risk increases above the baseline level for both cold and hot temperatures (Gosling *et al.*, 2009). Several studies on all-cause mortality have reported that the range of minimum effects vary in different climates, with populations living in colder and hotter regions showing lower and higher thresholds, respectively, for both cold and heat effects (Baccini *et al.*, 2008; Eurowinter Group, 1997; Keatinge *et al.*, 2000; McMichael *et al.*, 1998). Moreover, some evidence suggests that urban populations are more susceptible to extreme heat events, if compared to people living in non-urbanized areas (McGeehin and Mirabelli, 2001; Smoyer *et al.*, 2000).

The health risk associated with extreme temperature is modified by several variables. Kovats and Hajat (2008) categorized these risk factors in *intrinsic* and *extrinsic*: the former relates to features of individuals, while the latter refers to environmental and behavioural aspects. Intrinsic characteristics which have been assessed as potential modifiers of the relationship between temperature and health

are aging (Filleul *et al.*, 2004; Hajat *et al.*, 2007), sex (Bell *et al.*, 2008; Stafoggia *et al.*, 2006) and clinical of patho-physiological factors (Medina-Ramn *et al.*, 2006; Stafoggia *et al.*, 2008a; Wilkinson *et al.*, 2004). The evidence for a differential effect of race (Basu and Ostro, 2008; Medina-Ramn *et al.*, 2006) might be explained in terms of different patterns by racial groups of extrinsic factors which have been found to modify the association, such as socio-economic characteristics (Borrell *et al.*, 2006; Gouveia *et al.*, 2003; Rey *et al.*, 2009) and housing (Vandentorren *et al.*, 2006). Several studies also reported an interaction with air pollution, although the evidence is not conclusive (Carder *et al.*, 2008; Rainham and Smoyer-Tomic, 2003; Stafoggia *et al.*, 2008b).

The adaptation of individuals to different climates is paramount for the prediction of the future burden of climate change, and several studies have assessed geographical and temporal variations in the temperature-mortality association (McGeehin and Mirabelli, 2001; Medina-Ramon and Schwartz, 2007; Michelozzi *et al.*, 2006). Interestingly, some studies reported a progressive reduction in heat-related mortality along the last century, despite the aging of populations (Barnett, 2007; Carson *et al.*, 2006). This trend is likely to reflect improvements in social, environmental, behavioural, and health-care factors: in particular, the increased prevalence of air conditioning seems to play an important role in decreasing heat-related deaths (Davis *et al.*, 2003; ONeill *et al.*, 2005). In addition, some investigators have reported that heat waves occurring early in the summer are associated with a higher mortality risk than extreme events assessed later in the hot season (Baccini *et al.*, 2008; Hajat *et al.*, 2002), suggesting a short-term adaptation of population to changing climate.

1.2 Study designs and analytical approaches

The vast majority of studies assessing the health effects of temperature rest upon ecological study designs that use aggregated data. Most studies examine the relationship between ambient temperature and the number of cases in a defined period, frequently all-cause and cause-specific mortality or morbidity. The outcome is usually routinary collected for administrative purposes on a daily basis, and associated with the exposure averaged on the same temporal scale, as described in Section 1.1.

The preferred approach to investigate temperature-mortality associations is through time series regression analysis, a popular analytical tool in environmental epidemiology (Bell *et al.*, 2004; Touloumi

et al., 2004; Zeger *et al.*, 2006). In this approach the series of daily counts and ambient levels of temperature are compared, while controlling for potential confounding variables such as long-term and seasonal trends, air pollution and influenza epidemics. The standard model assumes Poisson responses, allowing for overdispersion. This analytical framework has been widely used in recent years, and has been thoroughly reviewed and extended by Armstrong (2006). A brief overview is provided in Section 1.3.

Alternative methods proposed for studying the effects of temperature are the case-crossover (Basu and Ostro, 2008; Bell *et al.*, 2008; Medina-Ramn *et al.*, 2006; Stafoggia *et al.*, 2006), case-only (Armstrong, 2003; Medina-Ramon and Schwartz, 2007; Schwartz, 2005), and case-control (Naughton *et al.*, 2002; Semenza *et al.*, 1996). Descriptive studies or simple analyses of specific heat waves have been published as well (Basu and Samet, 2002a).

Many studies on temperature-health associations are focused on the analysis of specific *episodes*, generally referred to as periods characterized by unusual weather patterns, such as heat waves and cold spells. In these analyses, the excess risk during the event can be estimated through contrast with comparable referent periods, usually chosen as the same days in the previous years (Conti *et al.*, 2007; Huynen *et al.*, 2001) or the other non-event days in the same month (Hoffmann *et al.*, 2008; Knowlton *et al.*, 2009). Alternatively, the health impact is computed including an indicator variable for specific weather events in a time series regression models (Hajat *et al.*, 2002; Huynen *et al.*, 2001; Rey *et al.*, 2007). Although a common definition has not yet been reached (Gosling *et al.*, 2009), the extreme episodes are commonly identified in terms of both *intensity* and *duration*, labelling as heat wave days those with temperature above a threshold for a minimum number of days (Robinson, 2001).

Other investigators have treated temperature as a continuous risk factor, including daily values in the time series regression model and thus estimating the proper exposure-response relationship (Anderson and Bell, 2009; Armstrong *et al.*, 2010). This approach provides a more detailed assessment of the association between temperature and various health outcomes, allowing the simultaneous estimation of cold and heat effects and the inspection of the comfort range or point of minimum effect described in Section 1.1 (Curriero *et al.*, 2002). A few studies have also investigated the association between heat, heat waves and mortality, by including both a heat wave indicator and a continuous term for

temperature in a regression model (Anderson and Bell, 2009; Hajat *et al.*, 2006; Rocklov and Forsberg, 2008). The rationale behind this methodology is that the effect of heat may be described as the sum of two contributions: an increased risk due to the independent effects of daily temperature levels, and an additional risk due to duration of heat sustained for several consecutive days.

1.3 Time series analysis in environmental epidemiology

Time series analysis is a common analytical tool applied in many different fields, from econometrics and ecology to physics and engineering. Commonly in these subject-areas, the main purpose is to predict future outcomes given the sequence of past observations: from a statistical perspective, the focus of the analysis is on specifying the proper auto-correlation structure of the series (Diggle, 1990), in order to provide reliable predictions. In contrast, in many applications in biomedical research, and particularly in environmental epidemiology, the scope of time series analysis is shifted from prediction to estimation, and the analysis is commonly carried out through standard regression techniques. The aim is not to predict future occurrences given an observed series, but to provide an appropriate description of the association between the exposure and outcome series, controlling for potential confounders (Zeger *et al.*, 2006). Still, the ordered temporal structure of the observations needs to be accounted for.

In applications in temperature-health studies, the time series regression model usually contains the series of daily measurements of health outcome and temperature as dependent and independent variables, respectively. Control for potential confounders is achieved by including series for additional variables. In the standard formulation, the model includes terms for controlling for seasonal and long-term trends, which remove the effect of measurable or un-measurable confounding factors acting in long temporal frames (Dominici *et al.*, 2003a; Peng and Dominici, 2008). Recently, the use of a spline function of time to model the trend components has been favoured: the main concern is the selection of the appropriate amount of smoothing, corresponding to the selection of the optimal number of (effective) degrees of freedom (df) per year (Peng *et al.*, 2006). Residual confounding effects are controlled by adding other variables showing short-term variation, such as air pollution, humidity and day of the week.

The analysis may focus on the whole continuous series, or be restricted to seasonal data. In the former,

a single spline function of time is usually preferred, producing an irregular seasonal trend which is believed to control for additional confounding effects operating at medium timescales. Commonly, in seasonal analyses, the seasonal and long-term trends are instead controlled with separate terms.

The standard design of modern time series studies in environmental epidemiology is based on a two-stage procedure involving multiple populations, commonly cities (Katsouyanni *et al.*, 1997; Samet *et al.*, 2000b). This approach guarantees against un-representativeness of single-city results, and offers a method to investigate heterogeneity and effects modification (Dominici *et al.*, 2003b). The analytical framework adopted in this context has been described as a two-stage hierarchical model, with a first-stage analysis which provides city-specific estimates to be then pooled in a second-stage meta-analysis (Dominici *et al.*, 2000). A single-step analysis, based on hierarchical models, is considered unfeasible given the complexity of the first-stage model, with a high number of parameters used to specify, for example, seasonal and long-term trends.

1.4 Methodological aspects in temperature-health studies

As mentioned earlier (Section 1.2), time series regression analysis has gained popularity in recent years for the assessment of the short-term effects of environmental stressors: in particular, this approach has been frequently used in order to estimate the health effects of air pollution. Almost 15 years ago, Schwartz *et al.* (1996) offered a comprehensive overview of the methodological problems, and since then, standard, well-grounded statistical methods have been developed for the application of time series analysis in this field (Dominici, 2004).

Temperature is usually included as a potential confounder in models assessing the health effects of air pollution. However, new and challenging methodological issues arise when temperature becomes the exposure of interest. In spite of this, many of the studies on health effects of temperature have applied the same methodology, and only few methodological works have systematically re-assessed it for the new context (Armstrong, 2006). These issues are briefly described in the next sections.

Non-linearity

While the exposure-response association between all-cause and cause-specific mortality and different air pollution indexes is commonly assumed as linear (Bell *et al.*, 2006; Daniels *et al.*, 2000; Dominici *et al.*, 2002; Schwartz *et al.*, 2002), the relationship with temperature is usually described as U, V or J-shaped (Braga *et al.*, 2001a; Curriero *et al.*, 2002; Hajat *et al.*, 2007).

A simple approach to deal with non-linear effects is to assume a threshold-type association, with the risk increasing linearly beyond a specific temperature values: this method includes the definition of *bathtub*-shaped relationships, with two distinct thresholds for cold and heat and a flat region in between (Pattenden *et al.*, 2003), or the so-called V-model, with a specific point of minimum effect (Ballester *et al.*, 1997; Huynen *et al.*, 2001; Nafstad *et al.*, 2001). Season-specific analyses simplify the relationship further, assuming a linear relationship within the season (Analitis *et al.*, 2008; Zanobetti and Schwartz, 2008), or an *hockey-stick* model with a single threshold (Armstrong *et al.*, 2010; Baccini *et al.*, 2008; Eurowinter Group, 1997). Muggeo (2008) proposed an interesting algorithm to simultaneously estimate both threshold and slopes, recently implemented in a statistical software (Muggeo, 2010).

Threshold-type models show some advantages in terms of interpretation and communication, with the effect being summarized by a single estimate of linear increase in risk, and the threshold value informing about the adaptation of populations to different climates. However, these approaches are based on strong assumptions on the shape of the exposure-response relationship, which could generate important biases, if wrong (Hajat *et al.*, 2006). An alternative options is to apply polynomial (Braga *et al.*, 2001a, 2002) or spline functions (Anderson and Bell, 2009; Armstrong, 2006; Curriero *et al.*, 2002), in order to model flexibly a smooth non-linear association.

Delayed effects and harvesting

Typically, exposure to environmental stressors generates effects delayed in time: a specific exposure events produces a risk which lasts well beyond the exposure period itself. Several time series analyses of temperature have reported that the exposure to extreme temperatures affects health for a period lasting days or weeks for exposures to extremely hot and cold days, respectively (Anderson and Bell, 2009; Braga *et al.*, 2001a,b; Carder *et al.*, 2005). The simplest methods to allow for lagged effects in a

time series analysis is to include the moving average of the exposure on the same and previous days, up to a maximum *lag* period. This method has been frequently used in air pollution studies (Dominici, 2004), and also applied to estimate the delayed effect of temperature (Anderson and Bell, 2009).
Furthermore, the complexity increases in the presence of so-called *harvesting* effect (or *mortality displacement*): the phenomenon that arises when a stressor affects mainly a pool of frail individuals, whose events are only brought forward by a brief period of time by the effect of exposure (Rabl, 2005; Schwartz, 2001). For non-recurrent outcomes, the depletion of the pool following a stress results in some reduction of cases few days later, thereby reducing the overall long-term impact. Some investigators assessed the presence of harvesting in temperature-health associations, but the evidence is still unclear (Goodman *et al.*, 2004; Hajat *et al.*, 2005).
A detailed analysis of delayed effects and harvesting would require the specification the distribution of the effects at different times after the event, modelling the relationship between an exposure occurrence and a sequence of future outcomes.

Pooling the results

The health effects of environmental factors are often assessed through multi-site studies. In the context of air pollution, the usual framework consists in a two-stage hierarchical analysis with a common site-specific model and then the application of meta-analytic techniques to pool the results (Dominici *et al.*, 2000; Samoli *et al.*, 2008). This approach ensures that the heterogeneity between different locations is properly accounted for, allowing model parameters to vary across sites, but at the same time avoiding additional variability and potential biases due to differential modelling choices (Dominici *et al.*, 2003b; Touloumi *et al.*, 2004). Meta-regression methods are commonly applied to assess the effect modification of site-level characteristics.
However, the non-linear association between temperature and health outcomes poses further problems for the pooling of more complex relationships estimated from the first-stage model, which cannot be adequately summarized by a single parameter. More sophisticated methods are needed to combine estimates of complex non-linear and delayed associations between temperature and health outcomes.

Chapter 2

Contribution of selected publications

In this chapter I provide a summary of my contribution to the research on statistical methods in studies on temperature-health associations, illustrating the content of my publications on the topic. In the next Section 2.1, I will first introduce two important methodological developments which may be applied to extend the ordinary methods, addressing the problems described in Section 1.4. The implementation of these methods on the statistical environment R is illustrated in Section 2.2. Finally, Section 2.3 will provide an overview of the publications, also highlighting my role in the various steps from study planning to article publication.

2.1 Methodological developments

The methodological content of the publications included in Part II focuses on two main statistical techniques: *distributed lag non-linear models* (DLNMs) and *multivariate meta-analytic methods*. These two analytical frameworks are described in detail in the related publications: here I provide a simple summary to introduce the main concepts.

Distributed lag non-linear models

Delayed effects may be defined following two different, but nevertheless complementary, perspectives: an exposure in a given day exerts an increased risk in multiple future days (forward), or the risk in a given day is caused by multiple exposures occurring in multiple previous days (backward). The basic idea of distributed lag models (DLM) for modelling delayed effects in time series data is therefore quite simple: in a regression model, the outcome for a given day is related to a linear predictor which includes multiple terms for current and lagged exposures, up to a maximum lag period. For long lag periods, the effects at different lags are constrained through a pre-specified function, whose coefficients are the only parameters estimated by the regression model.

The first formulation, based on a polynomial function, was originally adopted in econometrics to model capital expenditures (Almon, 1965), and recently re-proposed for epidemiological time series data (Schwartz, 2000). Although elegant and simple, this approach may nevertheless be applied only to linear dependencies. The extension to model non-linear and delayed exposure-response relationships brought to the development of DLNMs, firstly conceived and applied by Armstrong (2006). Although the algebraic definition is relatively complex, requiring the use of tensor product matrices, the idea is straightforward: the association is modelled through two independent functions, describing the relationship in the dimension of predictor and lags, respectively.

Multivariate meta-analysis and meta-regression

The standard design adopted in environmental epidemiology is based on a two-stage analysis, as described in Section 1.3. City-specific estimates are obtained from first-stage regression models following a common specification, and are then pooled in a second-stage meta-analysis. However, the association between temperature and health outcomes is usually non-linear and with delay. Previous investigations have proposed approaches to simplify the relationships, as described in Section 1.4, or to summarize it, for example computing average slopes (Curriero *et al.*, 2002), or estimating a single relative risk measure for specific absolute or relative (distribution percentiles) temperatures (Anderson and Bell, 2009; Stafoggia *et al.*, 2006).

However, these methods may not be appropriate for describing the complex pattern of health effects

of temperature. A solution is to retain the complexity of relationship as estimated in the first-stage model, describing the association with multiple parameters. The estimates can be then combined using a multivariate meta-analysis, a method originally proposed to pool multiple correlated outcomes in randomized clinical trials (Berkey *et al.*, 1996, 1998). This approach allows the synthesis of complex multi-parameter associations, producing average relationships across cities and providing a way to examine heterogeneity and effect modification through multivariate meta-regression.

The methodology of multivariate meta-analysis is the object of current research (Jackson *et al.*, 2011), although its statistical development can be easily described within the framework of linear mixed models (Verbeke and Molenberghs, 2000). Even if this technique has been already applied in the context of environmental epidemiology for multi-parameter associations (Analitis *et al.*, 2008; Dominici *et al.*, 2002; Samoli *et al.*, 2003, 2005, 2009), a methodological overview has not been presented yet.

2.2 Software implementation

The statistical methods described in Section 2.1 have been implemented within R (R Development Core Team, 2011), a free programming language and software environment for statistical computing and graphics. R was created by Ross Ihaka and Robert Gentleman in 1996 at the University of Auckland, New Zealand (Ihaka and Gentleman, 1996). The software is now maintained and developed by the R Development Core Team. The basic distribution may be extended via specific *packages*, a structured collection of functions built to produce specific statistical computations. Packages are usually documented through help pages and optionally *vignettes*, documents which include a detailed description, references and code examples. The packages may be included in the Comprehensive R Archive Network (CRAN), and then downloadable through R.

The choice to create the two R packages has been motivated by several considerations. First, the two approaches are based on relatively complex statistical methods, and on routines which require non-trivial computing skills in order to provide stable and trustworthy results. The availability of fully-documented packages in a freely-available software can promote the application of the methods by other research teams. Second, the implementation has required a long series of tests on the original scripts, and the analysis and results are hopefully less prone to errors and bugs. Finally, the production

of the packages involves the generalization of the methodologies beyond the specific data and models I have used in my research. The packages are therefore expected to be applicable in a wider range of analyses and potentially easier to improve and extend.

The R package dlnm

Distributed lag non-linear models have been implemented in the R package dlnm (`http://cran.r-project.org/web/packages/dlnm/index.html`). The package was first released on CRAN on the 1st of July 2009. The current version is 1.4.1, after 17 updates.

The package dlnm contains functions for building basis matrices to specify DLMs and DLNMs, and then to predict and plot the results for a fitted model. The first function, `crossbasis()`, is used to define the two basis matrices to model the relationship in the dimensions of predictor and lags, respectively, then combining them in a so-called cross-basis matrix obtained by a tensor product. Different models may be defined by choosing different functions for the 2 dimensions. The model is fitted using standard regression functions which include the cross-basis matrix, and the results are predicted over a set of values using the function `crosspred()`. A method function `plot()` is used to graph the estimated association.

A dataset with the time series data of mortality, temperature and air pollution for Chicago in the period 1987-2000 (Samet *et al.*, 2000a) is included in the package and used in the examples. Documentation of the package is provided through the help pages of the functions, and the package vignette included in Chapter 7.

The R package mvmeta

Multivariate meta-analysis and meta-regression have been implemented in the R package mvmeta (`http://cran.r-project.org/web/packages/mvmeta/index.html`). The package was first released on CRAN on the 9th of May 2011. The current version is 0.2.3, after 4 updates.

The main function in the package is `mvmeta()`, which performs fixed and random-effects multivariate meta-analysis and meta-regression. This regression-type function contains a formula which specifies the outcomes and linear predictors, and calls internal functions to compute maximum likelihood and

restricted maximum likelihood estimates through a Quasi-Newton algorithm. Additional functions are used, among other purposes, to obtain predictions and best linear unbiased predictions, to run a multivariate heterogeneity test and to compute fit statistics.
The dataset `berkey98`, used in the examples, contains the results of 5 published trials comparing surgical and non-surgical treatments for medium-severity periodontal disease (Berkey *et al.*, 1998). Documentation of the package is provided through the help pages of the functions.

2.3 Overview of the publications

The nine publications included as chapters in Part II summarize my research activity within the PhD project. They include six research papers, two commentaries and a package vignette. I am the first-author on eight, and the sole author on two. Seven contributions have already been published (one as on-line version only), one has been accepted for publication and the last one has been submitted.
The order of the publications has been carefully chosen to best describe a coherent research project. However, the manuscripts have been published or submitted as independent contributions, and the text included in the different chapters is not consistently linked. The purpose of this section is to provide the reader with a summary of each publication, progressively illustrating the conceptual and methodological steps undertaken during my research.

Research paper I

The first research paper, originally published as Gasparrini *et al.* (2011) and included in Chapter 3, is a thorough assessment of the effects of high temperatures on a long list of different causes of death, in 9 regions of England and in Wales. This investigation adopts a standard time series analysis on multiple regions, assuming a linear threshold relationship above region-specific percentiles to summarize the linear increase in log-relative risk per 1°C increase in summer temperature. The estimates of this single parameter are then pooled across regions using a traditional univariate meta-analysis. The relatively simple analytical approach is consistent with the research question, and appropriate to provide a comparative assessment of the effects of temperature by different causes of death. However, the analysis does not exploit any of the methodology I developed during my PhD research, and is

included here as an example of simple approach, which may be extended by the methods described in Section 2.1 and in the following publications.

As the first author of the paper, I took an active role in discussing the study design, research question, and the relevant epidemiological and public health issues. I independently conducted the analysis, discussing the analytical approaches and results with the second author, and the interpretation and conclusions with the research team. I took the lead on writing the manuscript and acted as corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Commentary I

This publication originally appeared in Gasparrini and Armstrong (2010) and is included in Chapter 4. After the release on CRAN of the dlnm package in July 2009 and the on-line publication of the methodological paper on DLNMs (Gasparrini *et al.*, 2010) in May 2010, I was offered by the editor of the journal to write a commentary on the article by Barnett *et al.* (2010). This paper presented a sophisticated analysis on the comparative assessment of different temperature indexes, using both the DLNM methodology and the R package. The aim of the commentary is to elucidate the development and application of these analytical tools, and more generally to offer a critical overview on the statistical issues and potential advancements in studies on temperature-health associations. In the context of my PhD research, this publication takes stock of the statistical approaches available for assessing the health effects of temperature and indicates potential research directions to improve the standard models proposed in the first research paper presented above. The developments are then described in the two blocks of methodological papers presented below.

As the first author of the commentary, I decided upon the structure and contents, in agreement with the co-author, and took the lead in writing up and submitting the manuscript.

Research paper II

This research paper was originally published as Gasparrini *et al.* (2010) and included in Chapter 5. Together with the fourth research paper included in Chapter 9, it represents the main methodologi-

cal output included in my PhD thesis. It is the first of a block of three publications illustrating the methodology and software implementation of DLNMs. In this contribution, I aim to extend and integrate the development of this class of models firstly proposed by Armstrong (2006). In particular, I endeavour to establish a general methodological framework to describe non-linear and delayed associations in time series data, also providing a consistent algebraic definition of DLNMs. This step offers some advantages in clarifying the development of the models and interpretation of the results, and is of paramount importance for future extensions, as described in Section 12.1.

As the first author of the paper, I structured the methodological description of DLNMs, in agreement with the co-authors, and developed the algebra for this model family. I independently choose the example included in the paper and performed the analysis. I was the lead author of the manuscript and acted as corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Research paper III

This article was originally published as Gasparrini (2011) and included in Chapter 6. In this contribution I illustrate the software implementation of DLMs and DLNMs in the R package dlnm. Although it may represent a peer-reviewed version of the documentation of the package included in the vignette in Chapter 7, the paper actually offers additional insights, reconsidering the main conceptual and practical steps to define this modelling framework and linking the algebraic formulation with the use of the functions in the package. The manuscript is written using the R function `Sweave()`, a tool that allows embedding the R code and associated results within the text in the document. Given the lapse of time since the publication of the methodological paper on DLNMs presented above (Gasparrini *et al.*, 2010), I also took the opportunity to present further research advancements, discussing the conceptual framework of DLNMs and addressing relevant issues such as modelling strategies and research directions.

As the sole author of the paper, I autonomously conceived the outline and contents of the article. I wrote the software and package and exploited it in the example illustration. I also drafted the manuscript and dealt with all the submission issues.

Package vignette

This document accompanies the R package dlnm, and represents its main documentation. The version for the release 1.4.1 of the package is included in Chapter 7, and an updated version is available at `http://cran.r-project.org/web/packages/dlnm/vignettes/dlnmOverview.pdf`. The content of the document is partly similar to what presented in the third research paper in Chapter 6, but adopting a more practical perspective. After some information on the installation of the package and a brief summary of the theory of DLNMs, the vignette illustrates in details the use of the functions, thoroughly describing their arguments and specificities, and provides multiple examples on the use of the package for increasingly complex analyses. The vignette is updated at each release of a new version of the package.

As the sole author of the vignette, I autonomously conceived the outline and contents of the document. I wrote the software and package and applied it in the example illustration. I also drafted the manuscript and dealt with all the publication issues.

Commentary II

This document has been accepted for publication, and is included in Chapter 8. It is the first of a block of two publications on multivariate meta-analytical techniques. It represents a commentary to the article by Jackson *et al.* (2011), which arose from an event organized at the Royal Statistical Society in London in January 2010 on multivariate meta-analysis, where I was invited as discussant together with the co-author of the commentary. Our contribution focused on the application of multivariate meta-analysis to combine estimates of non-linear associations specified by multiple parameters. The commentary provides a brief overview of methodological issues and discusses advantages and limitations if compared to alternative traditional approaches, and may be considered a preamble to the fourth research paper included in Chapter 9.

As the first author of the commentary, I structured the contents of the manuscript, in agreement with the co-author, and performed the analysis for the example illustration. I also drafted the manuscript and dealt with all the submission issues.

Research paper IV

This paper has been submitted for publication, and is included in Chapter 9, together with the document added in the on-line appendix. Along with the second research paper included in Chapter 5, this represents the main methodological output included in my PhD thesis. In this contribution, I offer a methodological overview on the use of multivariate meta-analysis and meta-regression to synthesize multi-parameter associations, formalizing and extending previous applications of the method. I also take advantage to present the second R package mvmeta, which is used to perform the analysis for the example illustration, involving a multi-city time series dataset. The paper illustrates the development of the methodology within the framework of linear mixed models, focusing on maximum likelihood and restricted maximum likelihood estimation, and discusses methodological issues and future developments. The R code is included as an on-line appendix, so the analysis may be completely reproduced and extended.

As the first author of the paper, I outlined the description of the methodology of multivariate meta-analysis and meta-regression, in agreement with the co-authors, and extensively revised the existing literature. I independently choose the examples included in the paper and performed the analysis. I autonomously produced the R routine and implemented it in the related R package. I was the lead author of the manuscript and submitted the draft, acting as corresponding author.

Research paper V

This article was originally published as Goldberg *et al.* (2011) and included in Chapter 10. It is the first of a block of two publications exploiting the two proposed methodologies for substantive analyses on temperature-health associations. This paper applies distributed lag non-linear models to investigate the relationship between cold and hot temperatures with mortality, using the dataset for a single city in temperate climate. The analysis assesses the relationship with mortality for all-cause, cardiovascular and respiratory mortality, exploring both overall and lag-specific effects.

In contrast to the other publications included in this PhD thesis, I am not the first author of this article. My contribution focused on discussing the modelling approaches and the interpretation of the results, given my expertise in the methodology of distributed lag non-linear models and on the epidemiological

issues related to temperature-health associations. Moreover, as author of the R package dlnm, I acted as an expert on the use of the software, providing routines with examples and checking the final code.

Research paper VI

This paper was originally published as Gasparrini and Armstrong (2011) and included in Chapter 11, together with the document added in the on-line appendix. It illustrates an analysis on a multi-city dataset, and offers some substantive evidence on the association between high temperature, heat waves and all-cause mortality. It represents an excellent example of my contribution to the research in the field, employing at the same time the two methodologies of distributed lag non-linear models and multivariate meta-analysis. The paper clearly shows the advantages of the application of these more sophisticated statistical approaches, if compared to simpler traditional methods. The R code is included in the on-line appendix, so the analysis may be completely reproduced and extended.
As the first author of the paper, I took an active role in discussing the study design, research question, and the modelling approaches, in agreement with the co-author, also extensively revising the existing literature. I independently conducted the analysis, discussing the analytical issues, results, interpretation and conclusions with the co-author. I took the lead on writing the manuscript and acted as corresponding author during the submission process, drafting the responses to reviewers and changes to the various versions.

Part II

Selected publications

Chapter 3

Research paper I

Title: The effect of high temperatures on cause-specific mortality in England and Wales.

Author(s): Antonio Gasparrini, Ben Armstrong, Sari Kovats, Paul Wilkinson.

Journal/Publisher: Occupational and Environmental Medicine.

Type of publication: Research paper.

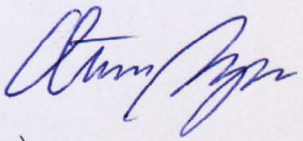
Stage of publication: Published online on March 9, 2011 as doi:10.1136/oem.2010.059782.

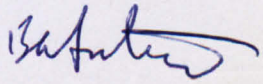
URL: `http://oem.bmj.com/content/early/2011/03/08/oem.2010.059782.short`.

Academic peer-reviewed: Yes.

Copyright: Permission not needed for inclusion in thesis/dissertation.

Candidate's role: See Section 2.3.

Candidate's signature: 

Senior author: (Prof. Ben Armstrong) 

OEM Online First, published on March 9, 2011 as 10.1136/oem.2010.059782

The effect of high temperatures on cause-specific mortality in England and Wales

Antonio Gasparrini, Ben Armstrong, Sari Kovats, Paul Wilkinson

► Additional figure and tables are published online only. To view these files please visit the journal online (http://oem.bmj.com).

Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, UK

Correspondence to
Antonio Gasparrini, London School of Hygiene and Tropical Medicine, 15–17 Tavistock Place, London WC1H 9SH, UK; antonio.gasparrini@lshtm.ac.uk

Accepted 10 February 2011

ABSTRACT

Objectives Several observational studies have suggested an association between high temperatures and all-cause mortality. However, estimates on more specific mortality outcomes are sparse, and frequently assessed in studies using different analytical methods.

Methods A time series analysis was performed on 10 regions in England and Wales during the summers (June–September) of 1993–2006. Average percentage linear increases in risk for a 1°C increase in temperature above region-specific thresholds and attributable deaths were computed by cause-specific mortality and age groups (0–64, 65–74, 75–84, 85+).

Results There was evidence of increased mortality with heat for almost all cause-of-death groups examined, with an overall increase in all-cause mortality of 2.1% (95% CI 1.6% to 2.6%) for a 1°C rise above the regional heat threshold. Among main causes, the steepest increase in risk was for respiratory mortality (+4.1% (3.5% to 4.8%) per 1°C). It was much smaller for cardiovascular causes (+1.8% (1.2% to 2.5%)) and myocardial infarction (+1.1% (0.7% to 1.5%)), but comparatively high for arrhythmias (+5.0% (3.2% to 6.9%)) and pulmonary heart disease (+8.3% (2.7% to 14.3%)). Among non-cardiorespiratory causes, the strongest effects were for genitourinary (+3.8% (2.9% to 4.7%)) and nervous system (+4.6% (3.7% to 5.4%)) disorders. 33.9% of heat deaths were attributable to cardiovascular causes, 24.7% to respiratory causes and 41.3% to all other causes combined.

Conclusions These results suggest that the risk of heat-related mortality is distributed across a wide range of different causes, and that targeting of preventative actions based on pre-existing disease is unlikely to be efficient.

What this paper adds

- A number of observational studies have suggested an association between hot temperatures and all-cause, cardiovascular and respiratory mortality.
- Estimates on the effect of heat on more specific mortality outcomes are sparse, and frequently assessed in studies relying on different analytical methods.
- Analysis of the association between heat and an extensive list of mortality outcomes, based on a common analytical approach, indicates a very widely distributed risk in relation to contributing cause.
- This suggests that targeting preventative actions based on assessment of existing diseases is unlikely to be an efficient strategy.

INTRODUCTION

Periods of high temperature in England and Wales are likely to increase in frequency and intensity as a result of climate change,[1 2] and, unless protective measures are taken,[3] so too their attendant impact on mortality and morbidity.[4]

In England, the strategy for the prevention of heat-related health effects is outlined in the Department of Health's Heatwave Plan, launched in 2004 and subsequently revised.[5] The plan includes both a weather-based warning system, and advice to primary and social care professionals and to the general public. The identification of individuals at high risk of heat-related mortality or morbidity is a key part of the plan. It defines at risk people as those with serious chronic conditions (especially heart or breathing problems), mobility problems (eg, Parkinson's disease or a previous stroke) and serious mental health problems. In addition, people on certain medications and those who misuse alcohol or drugs are considered at risk. However, to date, evidence concerning the degree to which risk is concentrated in these groups has been very limited, and it is unclear, even in theory, whether targeting these groups would prevent a large number of heat deaths.

The aim of this contribution is to assess heat-related mortality in relation to a wide range of causes using data for England and Wales in the period 1993–2006. The analysis may help define high-risk groups and estimate the effectiveness of preventive actions undertaken so far. In addition, the comparative assessment of multiple causes of deaths could help formulate or assess hypotheses on the underlying pathophysiological mechanisms in the association between high temperature and human health.

METHODS

Weather and mortality data

In order to account for the different climates within the UK, a meteorological data series was created for each Government Office Region in England and Wales. Data on minimum and maximum dry-bulb and dew point temperatures were obtained from the British Atmospheric Data Centre. Relative humidity was calculated using the average from dew point and dry-bulb temperatures at 9:00 h and 15:00 h. For each measure we used only data from stations reporting on 75% of days between 1993 and 2006, with a mean of 29 stations contributing to each regional series (range: 7 in

London to 44 in Wales). The temperature series are highly correlated within regions (mean r=0.95, range 0.94–0.98) and station means vary little within region (mean SD 0.7°C, range 0.3–1.1°C). Missing values in day i in station j were imputed using a combination of period average of the station j and a weighted average of the other regional stations. We then combined the values for day i from all monitoring stations in each region using a weighted average with weights equal to the populations residing closest to each station. Details of these methods have been previously described.[6 7]

Individual death record data were obtained from the Office for National Statistics, including date of death, age, underlying cause of death and postcode of residence at time of death. A broad list of 33 causes of death potentially associated with heat was selected, based on published epidemiological and physiological evidence regarding chronic diseases previously reported to be affected by hot weather. ICD-9 codes were used for the period 1993–2000 and ICD-10 for 2001–2006 (see online eTable 1 for the full list). Data were also disaggregated into five age groups: all, 0–64, 65–74, 75–84, 85+ years of age). The data were collapsed to series of daily counts for each region, for the specified age/cause-of-death subgroups.

Statistical analysis

The region and age-specific association between temperature and each outcome was assessed by time series analysis. Given the focus of the analysis on the effect of heat, we restricted the period of observation to the summer months (June–September). The different methods used to express the relationship between temperature and health outcomes have been extensively reviewed in previous work,[8] and modelling choices in this contribution are based on an extensive assessment performed on the same dataset for all-cause mortality.[7] We rely on a simple model, choosing a parameterisation where the effect of the average of the same and previous day's maximum temperatures (lag 0–1) is assumed to follow a log-linear increase in risk above a heat threshold, suggested as a reasonable and transparent approximation to more complex non-linear models. Lag choices are based on previous research in the UK,[9 10] while maximum temperature was chosen as the index providing the best fit in an analysis of all-cause mortality. In order to achieve comparable estimates for different regions and outcomes, we fixed the threshold to the 93rd percentile of region-specific year-round distribution of lag 0–1 maximum temperature (reported in table 1), the value showing the best fit for overall mortality.

The model follows a standard form for time series regression of season-specific data.[11 12] Here the expected number of deaths $E(Y_{ir})=\mu_{ir}$ in day i for each region r, assumed to follow an overdispersed Poisson distribution, is described by the formula:

$$\log(\mu_{ir}) = \alpha + \beta_r(t_{ir} - t_{0r})_+ + \sum_{k=1}^{K} g_k(x_{kir})$$

where α is an intercept and g_k functions are modelling the effects of confounders x_k. These terms included indicator variables for day of the week, natural cubic splines with 4 degrees of freedom (equally spaced knots) of day of the year in order to control for within-summer seasonal variation, and linear and quadratic terms of time to describe the long time trend. The effect of relative humidity is included with natural cubic splines with 3 degrees of freedom (knots at equally spaced quantiles of distribution). The region-specific coefficient β_r describes the log-linear increase in deaths for a unit increase in lag 0–1 temperature t above the threshold t_{0r}, with $(t-t_0)_+$ as a threshold term assuming value $(t-t_0)$ if $t>t_0$ and 0 otherwise. The analyses were repeated separately for each outcome and age category in the 10 regions. Coefficients β_r were pooled to derive the average β^m across regions, estimated using a random effect meta-analysis through restricted maximum likelihood.[13] The results are reported as pooled relative risk $\exp(\beta^m)$ or percentage increase $(\exp(\beta^m)-1)\times100$ related to a 1°C increase above the region-specific heat thresholds, and as numbers and fractions of deaths attributable to days with temperatures exceeding such thresholds. Heterogeneity is measured with the I^2 statistic, measuring the proportion of total variation due to difference between regions.[14] Attributable deaths n_{ir} in each day i for region r were computed through the formula $n_{ir}=N_{ir}\cdot(RR_{ir}-1)/RR_{ir}$, with $RR_{ir}=\exp(\beta_r\cdot(t_{ir}-t_{0r})_+)$ and N_{ir} as the total number of deaths. Daily attributable deaths are then summed over days and regions. We do not present means where estimates did not converge in any region, due to the small number of death in some subgroups. A sensitivity analysis on modelling choices for controlling seasonal and long time trends was carried out. All analyses were performed with R software, v 2.12.0.[15]

RESULTS

Descriptive statistics for each region are reported in table 1. During the period considered in the analysis (summer months in 1993–2006), a total of 2 285 519 deaths occurred in England and Wales, with an average of 134 each day per region. Regions show some differences in the distribution of maximum temperature, with the 93rd centiles (thresholds) increasing from 20.9°C for the North East to 24.7°C for Greater London.

The main results are plotted in table 2, which shows the pooled estimates of effect (percentage increase in risk for

Table 1 Descriptive statistics for overall mortality and maximum temperature by each region (June–September)

Region	Daily mortality		Maximum temperature (°C)		
	Mean	Range	Mean	Range	Threshold*
North East	72.0	41–104	18.4	8.9–29.4	20.9
North West	189.0	138–249	19.3	11.5–32	21.6
Yorkshire and Humber	131.8	90–180	19.5	10.5–30.3	22.2
West Midlands	136.2	93–186	20.3	9.9–33.8	23.0
East Midlands	107.0	70–156	20.3	9.7–32.3	23.0
Wales	83.8	45–119	19.4	11.8–31.7	21.6
East	132.2	87–187	21.2	10.3–34.4	23.9
South East	200.4	147–308	21.0	10.2–34.0	23.5
South West	136.1	98–187	20.1	12.1–31.2	22.3
Greater London	149.7	102–280	21.8	10.7–37.3	24.7

*93rd percentile of year-round (January–December) lag 0–1 maximum temperature distribution.

Table 2 Total number of deaths, pooled estimates of effect (% increase, with 95% CI) related to a 1°C increase above the region-specific heat thresholds and attributable burden for each cause of death

		Percent increase in mortality per 1°C increase above heat threshold		Attributable deaths	
	n	%	95% CI	n	%
All causes	2 285 519	2.1	1.6 to 2.6	23 617	100
Cardiovascular diseases	904 131	1.8	1.2 to 2.5	8005	33.9
Stroke	235 681	2.5	1.6 to 3.4	2864	12.1
Ischaemic heart diseases	475 235	1.7	1.2 to 2.2	3725	15.8
Myocardial infarction	230 343	1.1	0.7 to 1.5	1121	4.7
Chronic ischaemic heart diseases	237 973	2.3	1.6 to 3.0	2598	11.0
Atrial fibrillation	10 001	4.5	2.7 to 6.3	210	0.9
Atrio-ventricular conduction disorders	469	5.8	−2.8 to 15.1	9	0.0
Arrhythmias	5226	5.0	3.2 to 6.9	132	0.6
Pulmonary embolism	18 679	1.4	0.0 to 2.7	118	0.5
Heart failure	38 611	3.6	2.4 to 4.8	658	2.8
Sudden cardiac death	218	3.6	−8.5 to 17.2	1	0.0
Pulmonary heart diseases	1015	8.3	2.7 to 14.3	37	0.2
Respiratory	289 516	4.1	3.5 to 4.8	5841	24.7
Chronic obstructive pulmonary disease	87 980	4.3	3.6 to 5.1	1821	7.7
Asthma	5307	5.5	2.8 to 8.3	133	0.6
Respiratory infections	157 206	4.2	3.5 to 5.0	3194	13.5
Other	1 091 872	1.8	1.4 to 2.2	9764	41.3
Endocrine diseases	32 437	2.9	1.7 to 4.2	446	1.9
Diabetes	25 554	3.0	1.8 to 4.2	360	1.5
Genitourinary system	37 327	3.8	2.9 to 4.7	723	3.1
Urinary system	33 300	4.3	3.3 to 5.3	726	3.1
Mental diseases	48 022	3.1	1.7 to 4.6	776	3.3
Organic mental disorders	38 019	3.5	2.0 to 5.1	695	2.9
Psychoactive substance use	2542	3.1	−2.6 to 9.1	12	0.1
Schizophrenia	3856	0.9	−3.1 to 5.2	1	0.0
Nervous system	49 549	4.6	3.7 to 5.4	1118	4.7
Extra-pyramidal disorders	14 105	5.5	4.0 to 7.0	382	1.6
Other disorders of the nervous system	14 514	3.3	1.8 to 4.9	236	1.0
External causes	72 844	3.0	2.3 to 3.6	1037	4.4
Accidents/injuries	46 199	2.7	1.9 to 3.6	614	2.6
Intentional self-harm	15 935	2.8	1.4 to 4.3	202	0.9
Sudden infant death	967	−0.8	−6.8 to 5.5	−8	−0.0

each 1°C increase above the regions-specific heat threshold) and attributable deaths for each mortality outcome (see online eTable 1 for ICD-9 and ICD-10 codes). The age-specific estimates are included in figure 1 (see online eTable 2 for a complete list of age-specific values). The analysis by main causes shows the typical trend of risk by age, with an estimated increase for overall mortality of 2.1%, ranging from 1.3% in 0–64 year olds to 3.0% in 85+ year olds. Thus, ambient heat is responsible for 1.03% of the overall mortality that occurred in the summer months during the study period, equating to approximately 23 617 deaths. The results confirm the higher risk for respiratory as compared to cardiovascular mortality, with an increased risk of 4.1% and 1.8%, respectively. However, the attributable burden is higher for the latter, with 8005 attributable cases compared to 5841 for the former. While the risk for cardiovascular mortality increases with increasing age, the effect by respiratory causes is consistent across age groups. Non-cardiorespiratory ('other') deaths were also associated with heat, with similar slopes and age-specific effects as cardiovascular causes.

Within cardiovascular causes, the highest estimated risk was found for atrial fibrillation (4.5%), arrhythmias (5.0%) and pulmonary heart disease (8.3%). We found a much lower additional risk per 1°C increase in heat for myocardial infarction (1.1%) and ischaemic heart diseases (1.7%), although these causes are associated with large absolute risks. Stroke and heart failure show a strong pattern by age. The increases for respiratory causes are higher and less heterogeneous, both between causes and age groups, than for cardiovascular causes, ranging from 4.2% for respiratory infections to 5.5% for asthma.

There was also a significant heat-related risk for most non-cardiorespiratory outcomes, with increases of 2.9% for endocrine, nutritional and metabolic disorders, 3.8% for diseases of the genitourinary system, 3.1% for mental and behavioural disorders, and 4.6% for diseases of the nervous system. External causes showed an increase of 3.0%, with a slightly higher risk for deaths among those aged 0–64 than for older age groups.

Table 2 and online eFigure 1 show that there is no clear relationship between cause-specific heat-mortality and the number of heat attributable deaths. Most notably, ischaemic heart disease and myocardial infarction both had comparatively small temperature–mortality relationships (smaller point estimates than for all non-cardiorespiratory causes combined) but account for sizeable fractions of the overall burden of heat deaths, while some of the causes with strong temperature–mortality relationships, such as arrhythmias, asthma and nervous system disorders, account for a comparatively small fraction of the overall heat mortality burden.

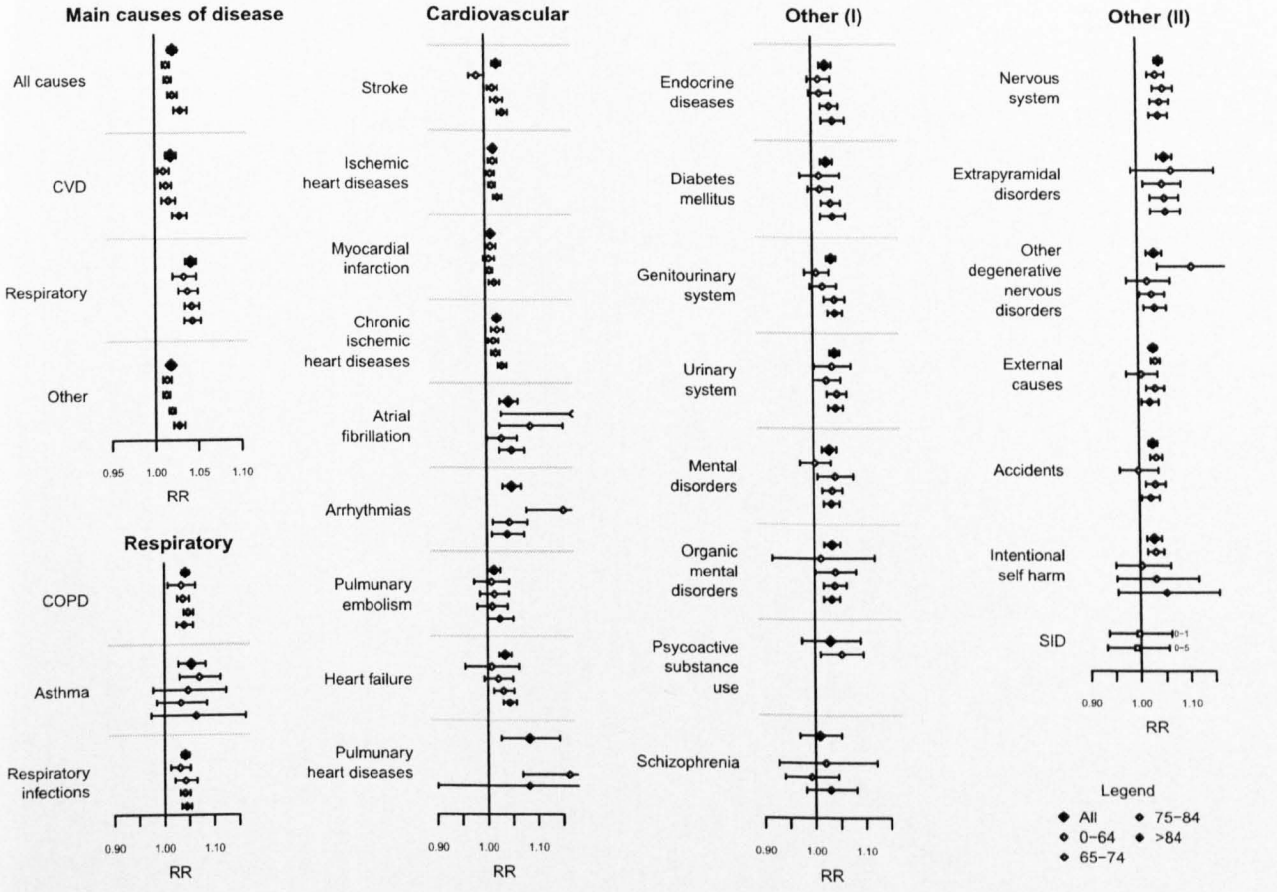


Figure 1 Pooled estimates of relative risk (with 95% CI) related to a 1°C increase above the region-specific heat thresholds for each cause-of-death and age group.

Region-specific estimates for broad causes (all-cause, cardiovascular disease, respiratory, other) are illustrated in figure 2. Estimates of the increase in risk for each 1°C above the region-specific heat threshold are substantially heterogeneous across regions, with I^2 statistics (the proportion of variance estimated to be true between-region variation) ranging between 67% and 92% for these causes of deaths. However, although heterogeneous, the cause-specific heat slopes show broadly similar relationships to each other across regions, with the gradients being generally steeper in London and other warmer regions, as previously reported for all-cause mortality.[7]

In order to assess the sensitivity of the results to the modelling choices, we repeated the analysis on broad causes increasing to 6 the degrees of freedom of the spline for seasonality and/or including a cubic term for long time trend. The results are robust to these choices, with the percentage change in the effects (measured as percentage increase) for the four broad causes ranging in the order of 0.1–7.5%.

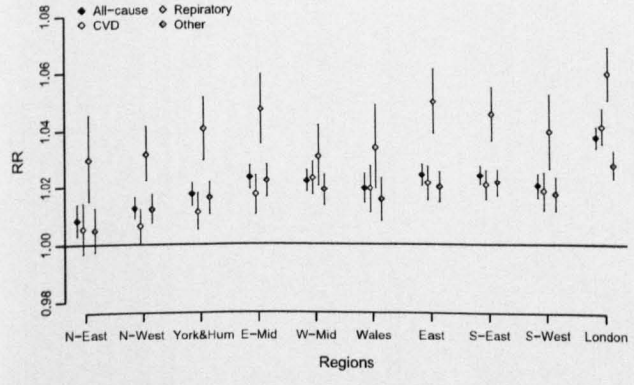


Figure 2 Region-specific estimates of relative risk (with 95% CI) related to a 1°C increase above the region-specific heat thresholds for broad causes of death. CVD, cardiovascular disease.

DISCUSSION

This study provides detailed evidence from England and Wales on the relationship between high temperature and mortality by cause-of-death and age. Comparability of the results by cause is enhanced by the application of a common analytical framework and modelling choices.

One of the notable observations of this analysis is the apparently widespread effect of heat, with evidence of a heat-related increase in mortality for almost all of the cause-of-death and age groups analysed. The effect of heat generally increases by age, as reported by other investigators.[16 17] Ageing induces physiological changes in thermoregulation and homeostasis, together with the increased prevalence of chronic conditions and use of medication, which are likely to increase vulnerability to heat.[18 19] There were some variations in the pattern of risk with age, with relatively steep gradients for stroke and heart failure, for example, and flatter slopes for respiratory diseases.

Although heat risk is often thought of mechanistically in terms of its effects on the cardiovascular system, it is noteworthy that relative risks for cardiovascular causes in general were no higher than those of most other causes of death and appreciably lower than those of respiratory causes. Of particular note is the relatively low relative risk for myocardial infarction, which has an underlying thrombotic genesis. This argues against changes in the coagulation properties of the blood as a major pathway for heat-related mortality risk, although numerically, myocardial infarction still contributes substantially to the excess burdens of deaths. However, the excess risks appear somewhat higher for stroke, which also is partly a thrombotic phenomenon.

Although the estimates are very imprecise, it appears that some of the highest cardiovascular excess risks are for pulmonary heart disease. Pathophysiologically, this may tie in with the relatively high excess risks for respiratory categories in general, and perhaps suggests critical exacerbation of right heart failure or other circulatory decompensation in the context of increased demand for cardiac output (for cooling) but limited reserve. Heart failure in general showed relatively high excess risks. Also of note is the comparatively high excess risks for arrhythmias and atrial fibrillation in particular, which have been noted previously.[17 20] The reasons for the large excess relative risk for this cause-of-death group are not clear, but such arrhythmias may contribute to cardiovascular compromise.

We found an important risk for chronic diseases such as diabetes mellitus, as well as for diseases of the genitourinary system, which may well reflect adverse effects on fluid and electrolytic balance, especially in those on medication.[4 20] The higher risk for nervous system diseases and mental disorders is likely to be related to impaired perception of environmental conditions and impaired ability to take actions to protect health.[18 19] Note the relatively high excess risks for extra-pyramidal disorders, which includes Parkinson's disease.

Our results on the overall relationship between heat and all-cause mortality are broadly compatible with evidence from the USA,[16 21–23] the UK,[7 24 25] and elsewhere. In particular, two recent studies summarise the effects for 107 U.S.A. communities and 15 European cities, and report an average increase of around 3% for 1°C increase in temperature.[12 26] Where cause-of-death has been examined, mainly for cardiovascular and respiratory deaths as broad groupings, the evidence has generally shown larger effect on respiratory causes,[21 27 28] while some studies of other causes have found a marked increase also for nervous system diseases and mental illness.[17 19 29] However, a comprehensive analysis of specific causes of death in relation to high temperatures has rarely been reported.

Our analytical approach is based on a simple linear-threshold model with cut-offs at percentiles of region-specific distributions. This choice is coherent with the findings of the systematic assessment by Armstrong and colleagues for all-cause mortality, performed on the same data.[7] In their analysis, the linear threshold model performed only marginally less well compared to more complex models with non-linear terms, although some evidence of non-linearity for extremely hot temperatures was also reported. Here, given our focus on the relative impacts of heat on different causes of death, we favoured interpretability over flexibility, but the presence of bias in the analysis of cause-specific mortality due to this approximation cannot be entirely ruled out. The same applies to other complexities in the temperature–mortality association, such as harvesting and time-varying effects.

The adoption of region-specific thresholds assumes the (partial) adaptation of populations to their own climates, a phenomenon previously reported.[12 22] Nonetheless, a substantial heterogeneity still remains, and we measured a North–South gradient in the supra-threshold linear effect, reported previously in detail for all-cause mortality.[7] However, given the relative similarity in these patterns across different causes, as illustrated in figure 2, we do not expect important biases in our comparative assessment.

The analyses do not account for the potential mediating, confounding or modifying effects of air pollution, in particular ozone, as these measures were not available for regional data. While some studies have reported some evidence of effect modification by particulates and ozone in the temperature–mortality association,[30 31] other investigators have measured a relatively small confounding effect.[26 32] A recent study exploring the heat–ozone interaction in 15 British conurbations, a subset of the data used here, has found that the effect of temperature was robust to the confounding or modification effect of ozone.[33] In particular, the effect of ozone seems to disappear when maximum temperature is used, the same temperature metric adopted in our analysis. However, these results are not conclusive and the chance that the estimated effects are partly due to unmeasured air pollution should be taken into account.

Although vulnerability to heat in those with recognised chronic diseases is not directly estimated in these cause-specific mortality analyses, the widespread heat-related increased mortality in many cause-of-death groups suggests that many groups of individuals with recognised disease conditions would need to be targeted if preventative actions were to cover most of the population-level risk of heat mortality. Indeed, in the case of cardiovascular disease, for example, which numerically contributes the largest group of excess heat deaths despite a comparatively low relative risk, it is reasonable to assume that pre-existing disease may often not be even recognised. Thus, heat death may be little more predictable than death in general. And, while of course it is appropriate in clinical settings to pay attention to those who may be at relatively greater risk during periods of heat, the observed increase in heat mortality in a wide range of cause-of-death groups suggests the need to pursue a broad-based population strategy for prevention as well as targeted strategies. This has evident implications for the emphasis in the Heatwave Plan in England and elsewhere.

Funding This study was funded by the Medical Research Council (grant reference G0501810). The funder had no role in the planning of the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication. All authors are independent of the funder.

Competing interests None.

Contributors AG was involved in the analysis, design and writing of the study and is guarantor. SK and PW were involved in the design and writing of the study. BA was involved in the design and writing of the study and in the analysis of the results. All authors had full access to all of the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the analysis.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **National Research Council.** *Reconciling Observations of Global Temperature Changeed*. Washington, DC: National Academy Press, 2000.
2. **Intergovernmental Panel on Climate Change (IPCC).** *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC*. Cambridge and New York: Cambridge University Press, 2007.
3. **Kovats RS,** Hajat S. Heat stress and public health: a critical review. *Annu Rev Public Health* 2008;**29**:41–55.
4. **Basu R,** Samet JM. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiol Rev* 2002;**24**:190–202.
5. **Department of Health.** *Heatwave: Plan for England—Protecting Health and Reducing Harm From Extreme Heat and Heatwave*. London: Department of Health, 2009.

6. **Rückerl R,** Greven S, Ljungman P, *et al*. Air pollution and inflammation (interleukin-6, C-reactive protein, fibrinogen) in myocardial infarction survivors. *Environ Health Perspect* 2007;**115**:1072–80.
7. **Armstrong BG,** Chalabi Z, Fenn B, *et al*. The association of mortality with high temperatures in a temperate climate: England and Wales. *J Epidemiol Community Health*. Published Online First: 3 May 2010. doi:10.1136/jech.2009.093161.
8. **Armstrong B.** Models for the relationship between ambient temperature and daily mortality. *Epidemiology* 2006;**17**:624–31.
9. **Hajat S,** Armstrong BG, Gouveia N, *et al*. Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. *Epidemiology* 2005;**16**:613–20.
10. **Pattenden S,** Nikiforov B, Armstrong BG. Mortality and temperature in Sofia and London. *J Epidemiol Community Health* 2003;**57**:628–33.
11. **Hajat S,** Armstrong B, Baccini M, *et al*. Impact of high temperatures on mortality: is there an added heat wave effect? *Epidemiology* 2006;**17**:632–8.
12. **Baccini M,** Biggeri A, Accetta G, *et al*. Heat effects on mortality in 15 European cities. *Epidemiology* 2008;**19**:711–19.
13. **van Houwelingen HC,** Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002;**21**:589–624.
14. **Higgins JP,** Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;**21**:1539–58.
15. **R Development Core Team.** *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2010.
16. **Basu R,** Ostro BD. A multicounty analysis identifying the populations vulnerable to mortality associated with high ambient temperature in California. *Am J Epidemiol* 2008;**168**:632–7.
17. **Stafoggia M,** Forastiere F, Agostini D, *et al*. Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. *Epidemiology* 2006;**17**:315–23.
18. **Basu R,** Samet JM. An exposure assessment study of ambient heat exposure in an elderly population in Baltimore, Maryland. *Environ Health Perspect* 2002;**110**:1219–24.
19. **Conti S,** Masocco M, Meli P, *et al*. General and specific mortality among the elderly during the 2003 heat wave in Genoa (Italy). *Environ Res* 2007;**103**:267–74.
20. **Medina-Ramón M,** Zanobetti A, Cavanagh DP, *et al*. Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environ Health Perspect* 2006;**114**:1331.
21. **Braga AL,** Zanobetti A, Schwartz J. The effect of weather on respiratory and cardiovascular deaths in 12 U.S. cities. *Environ Health Perspect* 2002;**110**:859–63.
22. **Curriero FC,** Heiner KS, Samet JM, *et al*. Temperature and mortality in 11 cities of the eastern United States. *Am J Epidemiol* 2002;**155**:80–7.
23. **Zanobetti A,** Schwartz J. Temperature and mortality in nine US cities. *Epidemiology* 2008;**19**:563–70.
24. **Hajat S,** Kovats RS, Atkinson RW, *et al*. Impact of hot temperatures on death in London: a time series approach. *J Epidemiol Community Health* 2002;**56**:367–72.
25. **Hajat S,** Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk? *Occup Environ Med* 2007;**64**:93–100.
26. **Anderson BG,** Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology* 2009;**20**:205–13.
27. **Ishigami A,** Hajat S, Kovats RS, *et al*. An ecological time-series study of heat-related mortality in three European cities. *Environ Health* 2008;**7**:5.
28. **Michelozzi P,** Accetta G, De Sario M, *et al*. High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities. *Am J Respir Crit Care Med* 2009;**179**:383–9.
29. **Rey G,** Jougla E, Fouillet A, *et al*. The impact of major heat waves on all-cause and cause-specific mortality in France from 1971 to 2003. *Int Arch Occup Environ Health* 2007;**80**:615–26.
30. **Ren C,** Williams GM, Morawska L, *et al*. Ozone modifies associations between temperature and cardiovascular mortality: analysis of the NMMAPS data. *Occup Environ Med* 2008;**65**:255–60.
31. **Ren C,** Williams GM, Tong S. Does particulate matter modify the association between temperature and cardiorespiratory diseases? *Environ Health Perspect* 2006;**114**:1690–6.
32. **Medina-Ramon M,** Schwartz J. Temperature, temperature extremes, and mortality: a study of acclimatization and effect modification in 50 United States cities. *Occup Environ Med* 2007;**64**:827–33.
33. **Pattenden S,** Armstrong B, Milojevic A, *et al*. Ozone, heat and mortality: acute effects in 15 British conurbations. *Occup Environ Med* 2010;**67**:699.

The effect of high temperatures on cause-specific mortality in England and Wales

Antonio Gasparrini, Ben Armstrong, Sari Kovats, et al.

Occup Environ Med published online March 9, 2011
doi: 10.1136/oem.2010.059782

Chapter 4

Commentary I

Title: Time series analysis on the health effects of temperature: advancements and limitations.

Author(s): Antonio Gasparrini, Ben Armstrong.

Journal/Publisher: Environmental Research.

Type of publication: Commentary.

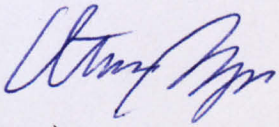
Stage of publication: Published in Volume 110, Issue 6, August 2010, Pages 633-638.

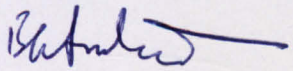
URL: `http://www.sciencedirect.com/science/article/pii/S0013935110000976`.

Academic peer-reviewed: No.

Copyright: Permission obtained from the publisher.

Candidate's role: See Section 2.3.

Candidate's signature: 

Senior author: (Prof. Ben Armstrong) 

Environmental Research 110 (2010) 633–638

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Commentary

Time series analysis on the health effects of temperature: Advancements and limitations ☆

Antonio Gasparrini *, Ben Armstrong

Department of Social and Environmental Health Research, London School of Hygiene & Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK

ARTICLE INFO

Article history:
Received 8 June 2010
Accepted 9 June 2010
Available online 23 June 2010

Keywords:
Temperature
Time series
Distributed lag
Harvesting
Multivariate meta-analysis

ABSTRACT

The association between extreme temperatures and health outcomes has been frequently investigated during the last few years. This assessment is usually based on a time series design, a framework which has gained a substantial development in the last two decades. In this contribution we offer an overview of the recent methodological advancements which provide new statistical tools to examine the health effects of temperature in a time series setting, highlighting at the same time the main limitations that still affect this research area.

1. Introduction

The increase in frequency and intensity of extreme weather events predicted in the near future (Luber and McGeehin, 2008) is arousing a growing interest, in the scientific and public health communities (Basu, 2009; Basu and Samet, 2002; Gosling et al., 2009). Several studies have investigated the association between mortality and both hot and cold temperatures, reporting increased risks in populations exposed to a wide range of climates (Analitis et al., 2008; Anderson and Bell, 2009; Baccini et al., 2008; McMichael et al., 2008). These studies are usually based on a time series design, where the series of daily counts of death or hospitalisations and ambient levels of temperature are compared, while controlling for potential confounding variables such as long-term and seasonal trends, air pollution and influenza epidemics. The purpose of these studies is to estimate the change in the counts of events associated with ambient temperature on the same day and on previous days (so-called *lagged effects*). Statistical approaches focus on regression methods within the generalized linear or additive modelling frameworks (GLM or GAM, respectively), assuming a Poisson distribution of the response (daily counts), and usually accounting for overdispersion (arising when the observed variance is greater than the expected number of daily events, differently from the standard Poisson assumption). Time series studies on temperature have benefitted from the remarkable statistical developments, achieved in the last two decades, to quantify the short-term effects of air pollution (Bell et al., 2004; Dominici, 2004; Schwartz et al., 1996; Touloumi et al., 2004).

In a paper published in this issue of the Journal, Barnett et al. (this issue) performed an analysis to examine the relationship between mortality and different temperature indexes, using a large dataset from 107 cities in the USA over a 14 years period. The analytical approach proposed by the authors highlights the flexibility and effectiveness of time series methods to attain sophisticated inferential deductions about complex associations. However, this complexity requires elaborate statistical tools that might appear obscure to many readers inexperienced with time series methods. In this contribution we attempt to review and elucidate recent advances in and limitations of this study design when applied to examine temperature–health associations, focussing mostly on the statistical issues.

2. Temporal decomposition

The time series design is characterized by a distinctive temporal structure of the data, with observations collected at ordered and equally spaced time points. In applications in environmental epidemiology, these time periods usually correspond to days, the smallest unit of time for which health outcome data are collected routinely. The main feature of the analytical methods is the *temporal decomposition* of the outcome and exposure series, where the variability is partitioned into contributions related to different timescales (Dominici et al.,

☆ *Funding:* Medical Research Council (UK); contract/grant number G0707030.
* Corresponding author. Fax: +44 20 75804524.
E-mail address: antonio.gasparrini@lshtm.ac.uk (A. Gasparrini).

0013-9351/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.envres.2010.06.005

2003a; Zeger et al., 2006). From an epidemiological perspective, the temporal partition of contributions to the exposure–response function addresses different issues. First, an exposure may lead to multiple physiological mechanisms operating at different timescales, whose effects can be disentangled by decomposing the series. In addition, specific confounding factors can act at different temporal frames; hence, the decomposition may produce virtually unbiased estimates at specific timescales in the presence of unmeasured confounders, if such factors act on longer temporal frames.

In the first methods that were proposed, the partition of both the response and exposure series was obtained by Fourier series decomposition (Zeger et al., 1999) or seasonal-trend decomposition using LOESS functions (Schwartz, 2000b; Schwartz, 2001), and then the correlations between components at corresponding timescales were estimated. In current applications, the decomposition is directly achieved through regression models, applying functions to describe seasonality and long-term trends, thus *filtering out* the effects of unmeasured factors that change slowly in time (Peng and Dominici, 2008). This approach leaves only the residual shorter-term variation to be explained by other factors that have day-to-day variability, like temperature. Originally, harmonic functions based on pairs of sine-cosine terms of day of the year were used to model the cyclic seasonal component (Hunsberger et al., 2002; Stolwijk et al., 1999), with non-linear functions of time like polynomial terms to describe the long-term trend. Recently, the use of a single spline function of time has been favoured, producing an irregular seasonal trend which is believed to control for additional confounding effects operating at medium timescales. The main choices are based on regression splines within GLM or penalized splines within GAM (Ruppert et al., 2003; Schimek, 2009).

Independently of the type of spline and modelling framework, the main concern is the selection of the appropriate amount of smoothing in order to avoid residual confounding, but at the same time leaving a temporal window with enough variability to be explained by temperature. This choice corresponds to the selection of the optimal number of (effective) degrees of freedom (df) per year in the spline for time.

3. Exposure–response relationship

Temperature usually shows a typical association with health outcomes, characterized by *non-linear* and *delayed* effects. Empirically, risk may increase at both hot and cold temperatures, with the exposure–response relationship being described as *U*, *V* or *J*-shaped (Curriero et al., 2002; Hajat et al., 2007; Pattenden et al., 2003). In addition, the effect of an exposure to extreme temperatures is not limited to the same day, but persists for a period of time, typically from a few days for heat to some weeks for cold (Anderson and Bell, 2009; Braga et al., 2001). When assessing non-recurrent outcomes, an additional complexity is given by the *harvesting* effect: if temperature mainly affects a pool of susceptible individuals who would have otherwise experienced the outcome a few days later, the depletion of the pool after an extreme event will result in a decreased occurrence in those days (Hajat et al., 2005; Schwartz, 2000b). This anticipation (*displacement*) of the outcome will be measured as an increase in risk in the very first days, followed by a decrease some time later, with a smaller net effect.

These aspects require the application of subtle statistical approaches to accurately express in a regression model the exposure–response relationship for temperature effects, whose estimates usually require careful interpretation. The issue of non-linearity has been addressed in different ways, using a threshold parameterization to describe linear effects of cold and heat below and above specific cut-off temperatures, or alternatively relying on spline functions within GLM or GAM, as those previously described (Armstrong, 2006). The problem of delayed and harvesting effects has been tackled in air pollution studies, proposing the so-called *distributed lag models* (DLMs), where the linear but delayed effects were modelled including multiple lagged exposures (Schwartz, 2000a). In practice, the health effect in day t of the series is explained in terms of exposures at days $t-\ell$, with ℓ as the lag in the interval $0,\ldots,L$, and L as the maximum lag period. If the lag period considered is long, the distribution of effects can be modelled through a mathematical function; for example, strata (Welty and Zeger, 2005), polynomials (Goodman et al., 2004) or splines (Zanobetti et al., 2000) can be used to avoid collinearity in the model.

Despite the availability of well-developed methods to describe flexible but un-lagged exposure–response relationships, or alternatively flexible distributed lag models for simple linear dependencies, these two issues have been rarely addressed together. Extensions to DLMs have been proposed, applying distributed lag functions to each term of polynomial (Braga et al., 2001), linear piecewise (Roberts and Martin, 2007) or threshold functions (Muggeo, 2008). Nonetheless, these methods remain somewhat limited in their ability to describe complex dependencies. Recently, we have proposed a new modelling framework which can describe flexible relationships both in the space of the predictor and the lag dimension, leading to the family of *distributed lag non-linear models* (DLNMs) (Armstrong, 2006; Gasparrini et al., 2010). The core of this methodology is the specification of two independent functions to model the relationship along the two dimensions of predictor (temperature) and lags, respectively, given a menu of available choices. These two functions are then combined to form *cross-basis* variables to be included in the regression model, whose estimated parameters describe the bi-dimensional effect.

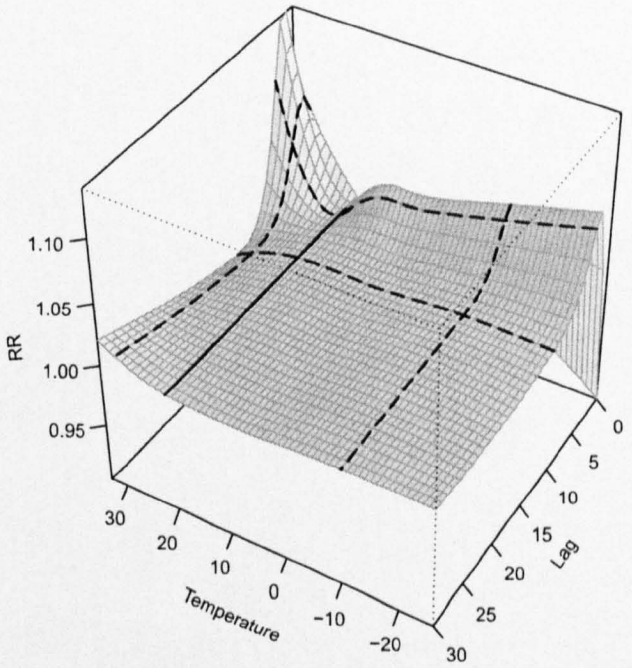


Fig. 1. 3-D plot of RR surface by temperature and lags. Highlighted are the reference at 21 °C (continuous line), the RR by lag at −10 and 30 °C, and the RR by temperature at lag 3 and 10 (dashed lines). Chicago 1987–2000.

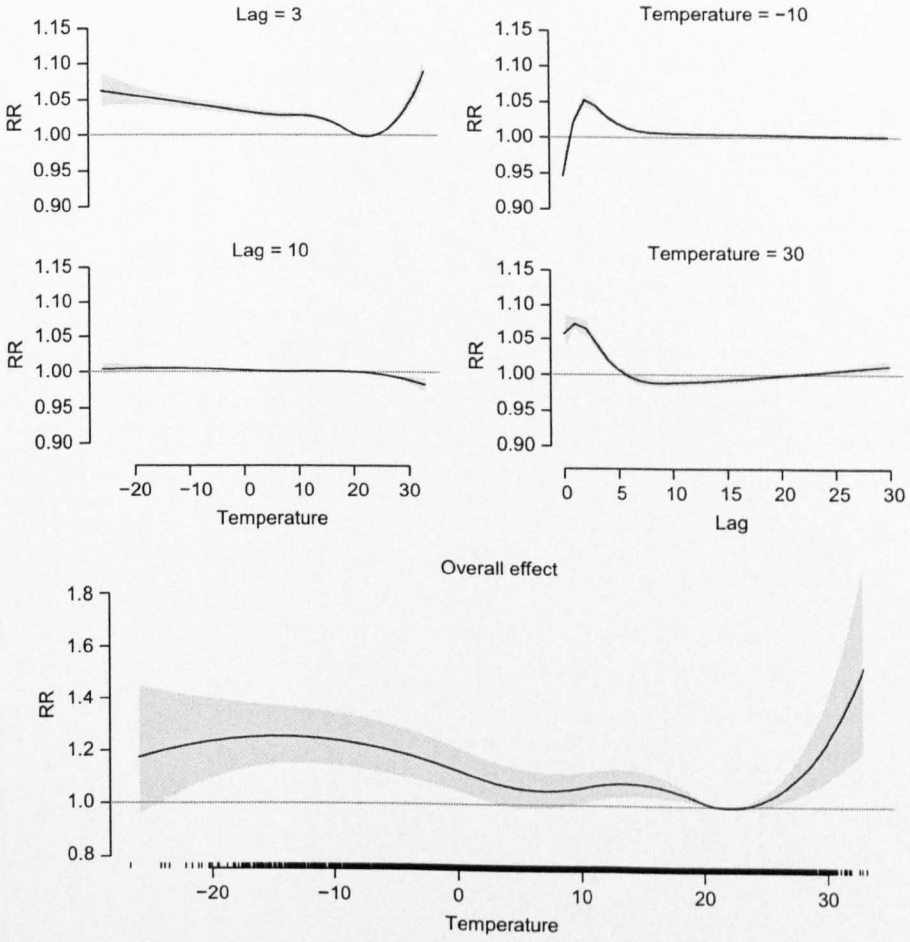


Fig. 2. Plot of RR by temperature at lag 3 and 10 (top left), RR by lag at −10 and 30 °C (top right) and overall RR (below). Reference at 21 °C. Chicago 1987–2000.

Figs. 1 and 2 show the application of the DLNM methodology to a time series of non-accidental deaths in Chicago, Illinois, during the period 1987–2000, using the same database that Barnett and colleagues analysed (Samet et al., 2000a, 2000b). In this example, we defined a cross-basis for temperature choosing a quadratic spline with 5 df for the space of temperature, and a natural cubic spline with 5 df for the space of lags, with 25 total parameters included. Fig. 1 shows the bi-dimensional relative risk (RR) surface for non-accidental mortality using a reference value of 21 °C, the empirical point of minimum mortality: the dashed lines represent the effects by lags for specific temperatures (−10 and 30 °C), and conversely the effects by temperature at specific lags (3 and 10 days). These effects are also reported, together with confidence intervals, in the top of Fig. 2. Lag-specific effects have a two-fold interpretation: each of them represents the increase in risk in a day t given a unit increase in temperature at day $t-\ell$ (*backward* interpretation, from outcome to exposure), or alternatively the increase in risk related to a unit increase in temperature at the day t during the following $t+\ell$ day (*forward* interpretation, from exposure to outcome). The overall effects are computed by the sum of lag contributions, and are illustrated in the bottom of Fig. 2. The results from this model suggest a strong and immediate effect of heat in the first 5 days, followed by a decrease after around 10 days, potentially interpreted as harvesting; cold temperatures display a more delayed effect, lasting up to 15 days.

The DLNM modelling framework is implemented within the software R (R Development Core Team, 2010) in the package 'dlnm' (Gasparrini and Armstrong, 2010). Further information about the analysis for Chicago and the package can be found at http://cran.r-project.org/web/packages/dlnm/vignettes/dlnmOverview.pdf.

4. Pooling the results

The health effects of environmental factors are assessed often through multi-site studies, using a two-stage hierarchical analysis with a common site-specific model and then the application of meta-analytic techniques to pool the results (Dominici et al., 2000; Samoli et al., 2008), the same strategy used by Barnett and colleagues. This approach ensures that the heterogeneity between different locations is properly accounted for, allowing model parameters to vary across sites, but at the same time avoiding additional variability and potential biases due to differential modelling choices (Dominici et al., 2003b; Touloumi et al., 2004). Meta-regression methods are commonly applied to assess the effect modification of site-level characteristics.

Air pollution studies are consistent with a linear exposure–response relationship, summarizing the effect with a single coefficient estimating the log-RR for a unit increase in exposure. The non-linear effect of temperature poses additional challenges, and several solutions have been proposed. First, the exposure–response relationship may be simplified assuming linear dependencies beyond site-specific thresholds (Baccini et al., 2008;

McMichael et al., 2008), or alternatively restricting the analysis to specific seasons, where strong deviations from linearity are not expected (Analitis et al., 2008; Zanobetti and Schwartz, 2008). An alternative solution is to produce a summary measure of the estimated non-linear relationship, for example computing average slopes (Curriero et al., 2002), or estimating a single RR for specific absolute or relative (distribution percentiles) temperatures (Anderson and Bell, 2009; Stafoggia et al., 2006). The use of site-specific thresholds or relative temperatures is usually preferred, in order to take into account the adaptation of populations to their own climate. A standard meta-analysis is then carried out for these single parameters.

The methods illustrated above have limitations: constraining the exposure–response to a simple shape could generate biased results, especially when assessing lag-specific effects. Even if strong assumptions are not formulated on the single-site models, pooling only simple summary measures of these might lose important features of a complex association. More sophisticated approaches rely on multivariate meta-analytical techniques, applied to relationships described by multiple parameters (Jackson et al., 2010; van Houwelingen et al., 2002), which are simultaneously pooled while accounting for their correlations within each site-specific model. These methods have been used to investigate dose–response functions (Baccini et al., 2008; Dominici et al., 2002; Samoli et al., 2005) or distributed lag curves (Analitis et al., 2008; Samoli et al., 2009) in multi-site studies. However, this approach is suitable only for associations expressed by a limited number of parameters. It is currently computationally infeasible, for example, to pool surfaces as the one illustrated in Figs. 1 and 2. An interesting compromise is the *meta-smoothing* method (Schwartz and Zanobetti, 2000), where simple univariate meta-analyses can be applied to pool the effects for any combinations of temperature values and lags, without accounting for correlations. Further research is needed to assess the presence and extent of biases in this type of approach for point estimates and standard errors if compared to proper multivariate methods (Riley, 2009), and to develop approaches to investigating heterogeneity in temperature–health association over sites (multivariate meta-regression).

As it is, investigators must balance the advantages of keeping information from the site-specific model with the need to reduce the number of parameters (summary measures) to make second-stage meta-analytical methods feasible. This choice also depends on the aim of the investigation and the availability of data.

5. Model selection

In contrast to analyses performed in many other subject-areas, the regression models applied in time series data for environmental factors are based on a limited number of predictors, such as day of the week, indicators for holiday periods and influenza epidemics, weather and pollution variables. The predictors to be included in the model are typically defined a priori, in particular in multi-site studies. In air pollution research, the critical choice to select the final model thus commonly focuses on the specification of the functions to account for seasonal and long-term trend, as discussed above in Section 2. Several contributions have addressed this issue, comparing alternative selection criteria (Baccini et al., 2007; Burnett et al., 1997; Peng et al., 2006; Touloumi et al., 2006): the main options are based on Akaike or Bayesian information criteria (AIC and BIC, respectively), (generalized) cross-validation techniques, minimization of the partial auto-correlation function of the residuals (PACF) or the related white noise test. The first 3 statistics aim to maximize the ability of the model to predict new observations arising from the same phenomenon which produced the data, while the last two intend to minimize the correlation between residuals from proximate observations in the series, to match the standard assumption of uncorrelated residuals. While these models fit statistics and residual analyses provide helpful insight, as criteria none can guarantee control of confounding (Peng et al., 2006). More complex approaches remain in the domain of statistical theory (Crainiceanu et al., 2008; Dominici et al., 2004). This leaves this aspect of model choice controversial, and makes analyses of sensitivity of key findings to variations in model choices very important.

When temperature is the focus of the analysis, similar considerations apply but additional issues should be taken into account. First, given the stronger association of temperature than of pollution with season, the optimal amount of smoothing to control for time may not be the same as that applied when air pollution is the exposure of interest. Second, given the longer lag often suggested for temperature (Anderson and Bell, 2009; Braga et al., 2001) there is more tension between the need to control confounding by unmeasured factors causing medium-term fluctuations in mortality (favoring many degrees of freedom in the time smooth) and the need to leave variation from which effects of interest can be estimated (favoring fewer degrees of freedom). Finally, when the objective is to investigate temperature–mortality relationships in their own right, there is usually a trade-off between complete description of all patterns not explicable by noise (many criteria often select quite complex models) and simplicity of interpretation. Specific study purposes may suggest different trade-offs.

The analysis performed by Barnett and colleagues illustrates some of these issues. A cross-validation procedure is specifically justified for the purpose of comparing the predictive ability of different temperature indexes, which might turn out to be useful, for example, to assess the future burden of climate change. In their analysis, the performance of the selection criteria is thus consistent with the research question. In other circumstances, for example when the goal is to obtain unbiased estimates of the exposure–response relationship or to control for confounding effects, the choice of selection criteria may be different (Dominici et al., 2008; Peng et al., 2006).

6. Implication of the ecological design

Ecological studies are defined as those in which the unit of the analysis is represented by aggregated or grouped observations (Last, 2001): the evidence from these research designs is interpreted with caution, given the inherent risk of biases due to the lack of information about individual characteristics (Greenland and Robins, 1994). Two of the main limitations that have been emphasized are the presence of unmeasured confounders, and error due to measures being collected from monitors at central sites, which do not represent personal exposures, which vary.

However, for investigating acute effects of environmental stressors (ambient temperature or air pollution) that change over time, the limitations inherent in ecological designs are offset by advantages. Many individual factors such as genetic make-up do not vary over time so cannot confound. Others, such as diet or smoking, vary only slowly so their effects are filtered out by the smooth function of time as discussed in Section 2. Moreover, variation in exposure across individuals is not the problem that it first appears to be. While there may be large variation in temperature or air pollution over a city, changes in daily population averages are usually much better reflected by the central monitor. This implies that the error in assigning

individuals to central site levels is primarily of a Berkson-type, which in linear models does not lead to bias in estimates of effect in linear models, though it reduces the precision (Armstrong, 1998; Zeger et al., 2000). Thus, results from time series studies in this context are considered more robust than those achieved from other ecological designs.

Nevertheless, there are some reasons for caution in interpreting ecological time series studies of the effects of temperature. Berkson error may be a more relevant problem for studies of temperature than of air pollution effects, because of the non-linearity of most temperature–health relationships. For example, if a threshold-linear model pertained at the individual level, variation of temperatures across the city would blur the threshold, even if the between-day time fluctuations in individuals were perfectly correlated with those of the central site measurements. There are also implications of the ecological design for assessment of harvesting and, more broadly, of the lag pattern of the effect. Elegant conceptual frameworks have been proposed to characterize this phenomena (Rabl, 2005; Schwartz, 2000b), but it may not be appreciated that the apparent post-exposure protective effect in the presence of harvesting is not a real effect acting at the individual level, but an artefact of the ecological nature of the design. The measured decrease in risk is explained by the change in the structure of the underlying population after the depletion of the pool of frail individuals. More generally, lag curves such as those depicted in Fig. 2 are likely to be the results of the sum of delayed positive effects and harvesting, and should not be automatically interpreted as the temporal representation of some physiological mechanism linking temperature and mortality.

7. Conclusions

Time series analysis represents a valuable tool to assess the acute health effects of environmental factors that fluctuate over time. The recent developments described above address some of the main problems regarding its application in temperature–health studies, providing flexible methods to investigate this complex association. These investigations are relatively simple to conduct because of the routinely collected data, available in most locations. However, it is important to consider that this ecological design still has some limitations, which need to be kept in mind when planning a study or interpreting analytical results.

In addition, these new approaches face new challenges related to the complexity of the analytical methods, mainly due to the need to select a model from a large number of alternatives. As highlighted by Barnett and colleagues, the estimate of the association is particularly sensitive to the choice of functions, lag period and other model parameters, and available selection criteria are still limited to reliably identify a "best" model. An extensive sensitivity analysis on the various modelling choices is therefore always recommended.

References

Analitis, A., Katsouyanni, K., Biggeri, A., Baccini, M., Forsberg, B., Bisanti, L., et al., 2008. Effects of cold weather on mortality: results from 15 European cities within the PHEWE Project. American Journal of Epidemiology. 168, 1397.

Anderson, B.G., Bell, M.L., 2009. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. Epidemiology 20, 205–213.

Armstrong, B., 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occupational and Environmental Medicine 55, 651.

Armstrong, B., 2006. Models for the relationship between ambient temperature and daily mortality. Epidemiology 17, 624–631.

Baccini, M., Biggeri, A., Accetta, G., Kosatsky, T., Katsouyanni, K., Analitis, A., et al., 2008. Heat effects on mortality in 15 European cities. Epidemiology 19, 711–719.

Baccini, M., Biggeri, A., Lagazio, C., Lertxundi, A., Saez, M., 2007. Parametric and semi-parametric approaches in the analysis of short-term effects of air pollution on health. Computational Statistics and Data Analysis 51, 4324–4336.

Barnett, A.G., Tong, S., Clements, A.C.A. What measure of temperature is the best predictor of mortality? Environmental Research, this issue, doi:10.1016/j.envres.2010.05.006.

Basu, R., 2009. High ambient temperature and mortality: a review of epidemiological studies from 2001 to 2008. Environmental Health 8, 40.

Basu, R., Samet, J.M., 2002. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. Epidemiologic Reviews 24, 190–202.

Bell, M.L., Samet, J.M., Dominici, F., 2004. Time-series studies of particulate matter. Annual Review of Public Health 25, 247–280.

Braga, A.L., Zanobetti, A., Schwartz, J., 2001. The time course of weather-related deaths. Epidemiology 12, 662–667.

Burnett, R.T., Brook, J.R., Yung, W.T., Dales, R.E., Krewski, D., 1997. Association between ozone and hospitalization for respiratory diseases in 16 Canadian cities. Environmental Research 72, 24–31.

Crainiceanu, C.M., Dominici, F., Parmigiani, G., 2008. Adjustment uncertainty in effect estimation. Biometrika 95, 635.

Curriero, F.C., Heiner, K.S., Samet, J.M., Zeger, S.L., Strug, L., Patz, J.A., 2002. Temperature and mortality in 11 cities of the eastern United States. American Journal of Epidemiology 155, 80–87.

Dominici, F., 2004. Time-series analysis of air pollution and mortality: a statistical review. Research report—Health Effects Institute, 3–27; discussion, pp. 29–33.

Dominici, F., Daniels, M.J., Zeger, S.L., Samet, J.M., 2002. Air pollution and mortality: estimating regional and national dose-response relationships. Journal of the American Statistical Association 97, 100–111.

Dominici, F., McDermott, A., Hastie, T.J., 2004. Improved semiparametric time series models of air pollution and mortality. Journal of the American Statistical Association 99, 938–949.

Dominici, F., McDermott, A., Zeger, S.L., Samet, J.M., 2003a. Airborne particulate matter and mortality: timescale effects in four US cities. American Journal of Epidemiology 157, 1055–1065.

Dominici, F., Samet, J.M., Zeger, S.L., 2000. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. Journal of the Royal Statistical Society: Series A 163, 263–302.

Dominici, F., Sheppard, L., Clyde, M., 2003b. Health effects of air pollution: a statistical review. International Statistical Review 71, 243–276.

Dominici, F., Wang, C., Crainiceanu, C., Parmigiani, G., 2008. Model selection and health effect estimation in environmental epidemiology. Epidemiology 19, 558–560.

Gasparrini, A., Armstrong, B., 2010. dlnm: Distributed Lag Non-linear Models in R. R package version 1.2.3.

Gasparrini, A., Armstrong, B., Kenward, M.G., 2010. Distributed lag non-linear models. Statistics in Medicine. Epub ahead of print May 7. 10.1002/sim.3940.

Goodman, P.G., Dockery, D.W., Clancy, L., 2004. Cause-specific mortality and the extended effects of particulate pollution and temperature exposure. Environmental Health Perspectives 112, 179–185.

Gosling, S.N., Lowe, J.A., McGregor, G.R., Pelling, M., Malamud, B.D., 2009. Association between elevated athmospheric temperatures and human mortality: a critical review of the literature. Climate Change 92, 299–341.

Greenland, S., Robins, J., 1994. Invited commentary: ecologic studies—biases, misconceptions, and counterexamples. American Journal of Epidemiology 139, 747.

Hajat, S., Armstrong, B.G., Gouveia, N., Wilkinson, P., 2005. Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. Epidemiology 16, 613–620.

Hajat, S., Kovats, R.S., Lachowycz, K., 2007. Heat-related and cold-related deaths in England and Wales: Who is at risk? Occupational and Environmental Medicine 64, 93–100.

Hunsberger, S., Albert, P.S., Follmann, D.A., Suh, E., 2002. Parametric and semiparametric approaches to testing for seasonal trend in serial count data. Biostatistics 3, 289–298.

Jackson, D., White, I.R., Thompson, S.G., 2010. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. Statistics in Medicine 29, 1282–1297.

Last, J.M., 2001. A dictionary of epidemiology. Oxford University Press, New York.

Luber, G., McGeehin, M., 2008. Climate change and extreme heat events. American Journal of Preventive Medicine 35, 429–435.

McMichael, A.J., Wilkinson, P., Kovats, R.S., Pattenden, S., Hajat, S., Armstrong, B., et al., 2008. International study of temperature, heat and urban mortality: the 'ISOTHURM' project. International Journal of Epidemiology 37, 1121.

Muggeo, V.M., 2008. Modeling temperature effects on mortality: multiple segmented relationships with common break points. Biostatistics 9, 613–620.

Pattenden, S., Nikiforov, B., Armstrong, B.G., 2003. Mortality and temperature in Sofia and London. Journal of Epidemiology and Community Health 57, 628–633.

Peng, R.D., Dominici, F., 2008. Statistical Methods for Environmental Epidemiology with R—A Case Study in Air Pollutioon and Health. Springer-Verlag, New York.

Peng, R.D., Dominici, F., Louis, T.A., 2006. Model choice in time series studies of air pollution and mortality. Journal of the Royal Statistical Society: Series A 169, 179–203.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing.

Rabl, A., 2005. Air pollution mortality: harvesting and loss of life expectancy. Journal of Toxicology and Environmental Health: Part A 68, 1175–1180.

Riley, R.D., 2009. Multivariate meta-analysis: the effect of ignoring within-study correlation. Journal of the Royal Statistical Society: Series A 172, 789–811.

Roberts, S., Martin, M.A., 2007. A distributed lag approach to fitting non-linear dose–response models in particulate matter air pollution time series investigations. Environmental Research 104, 193–200.

Ruppert, D., Wand, M., Carroll, R., 2003. Semiparametric Regression. Cambridge University Press.

Samet, J.M., Zeger, S.L., Dominici, F., Curriero, F., Coursac, I., Dockery, D.W., 2000a. The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 2, Morbidity and Mortality from Air Pollution in the United States. Health Effects Institute.

Samet, J.M., Zeger, S.L., Dominici, F., Dockery, D., Schwartz, J., 2000b. The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 1, Methods and Methodological Issues. Health Effects Institute.

Samoli, E., Analitis, A., Touloumi, G., Schwartz, J., Anderson, H.R., Sunyer, J., et al., 2005. Estimating the exposure–response relationships between particulate matter and mortality within the APHEA multicity project. Environmental Health Perspectives 113, 88–97.

Samoli, E., Peng, R., Ramsay, T., Pipikou, M., Touloumi, G., Dominici, F., et al., 2008. Acute effects of ambient particulate matter on mortality in Europe and North America: results from the APHENA study. Environmental Health Perspectives 116, 1480–1486.

Samoli, E., Zanobetti, A., Schwartz, J., Atkinson, R., Le Tertre, A., Schindler, C., et al., 2009. The temporal pattern of mortality responses to ambient ozone in the APHEA project. Journal of Epidemiology and Community Health 63, 960–966.

Schimek, M., 2009. Semiparametric penalized generalized additive models for environmental research and epidemiology. EnvironMetrics 20, 699–717.

Schwartz, J., 2000a. The distributed lag between air pollution and daily deaths. Epidemiology 11, 320–326.

Schwartz, J., 2000b. Harvesting and long term exposure effects in the relation between air pollution and mortality. American Journal of Epidemiology 151, 440–448.

Schwartz, J., 2001. Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? Epidemiology 12 55–61.

Schwartz, J., Spix, C., Touloumi, G., Bacharova, L., Barumamdzadeh, T., le Tertre, A., et al., 1996. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. Journal of Epidemiology and Community Health 50 (Suppl 1), S3–S11.

Schwartz, J., Zanobetti, A., 2000. Using meta-smoothing to estimate dose–response trends across multiple studies, with application to air pollution and daily death. Epidemiology 11, 666–672.

Stafoggia, M., Forastiere, F., Agostini, D., Biggeri, A., Bisanti, L., Cadum, E., et al., 2006. Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. Epidemiology 17, 315–323.

Stolwijk, A., Straatman, H., Zielhuis, G., 1999. Studying seasonality by using sine and cosine functions in regression analysis. Journal of Epidemiology and Community Health 53, 235–238.

Touloumi, G., Atkinson, R., Le Tertre, A., Samoli, E., Schwartz, J., Schindler, C., et al., 2004. Analysis of health outcome time series data in epidemiological studies. EnvironMetrics 15, 101–117.

Touloumi, G., Samoli, E., Pipikou, M., Le Tertre, A., Atkinson, R., Katsouyanni, K., 2006. Seasonal confounding in air pollution and health time-series studies: effect on air pollution effect estimates. Statistics in Medicine 25, 4164–4178.

van Houwelingen, H.C., Arends, L.R., Stijnen, T., 2002. Advanced methods in meta-analysis: multivariate approach and meta-regression. Statistics in Medicine 21, 589–624.

Welty, L.J., Zeger, S.L., 2005. Are the acute effects of particulate matter on mortality in the National Morbidity, Mortality, and Air Pollution Study the result of inadequate control for weather and season? A sensitivity analysis using flexible distributed lag models. American Journal of Epidemiology 162, 80–88.

Zanobetti, A., Schwartz, J., 2008. Temperature and mortality in nine US cities. Epidemiology 19, 563–570.

Zanobetti, A., Wand, M.P., Schwartz, J., Ryan, L.M., 2000. Generalized additive distributed lag models: quantifying mortality displacement. Biostatistics 1, 279–292.

Zeger, S.L., Dominici, F., Samet, J., 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10, 171–175.

Zeger, S.L., Irizarry, R., Peng, R.D., 2006. On time series analysis of public health and biomedical data. Annual Review of Public Health 27, 57–79.

Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwartz, J., Dockery, D., et al., 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environmental Health Perspectives 108, 419–426.

Chapter 5

Research paper II

Title: Distributed lag non-linear models.

Author(s): Antonio Gasparrini, Ben Armstrong, Michael G. Kenward.

Journal/Publisher: Statistics in Medicine.

Type of publication: Research paper.

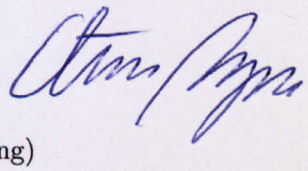
Stage of publication: Published in Volume 29, Issue 21, September 2010, Pages 2224-34.

URL: `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2998707`.

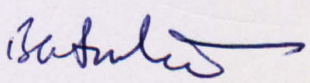
Academic peer-reviewed: Yes.

Copyright: Permission obtained from the publisher.

Candidate's role: See Section 2.3.

Candidate's signature: 

Senior author: (Prof. Ben Armstrong)



Research Article

Statistics in Medicine

Received 4 November 2009, Accepted 18 March 2010 Published online 7 May 2010 in Wiley Interscience

(www.interscience.wiley.com) DOI: 10.1002/sim.3940

Distributed lag non-linear models

A. Gasparrini[a*†], B. Armstrong[a] and M. G. Kenward[b]

Environmental stressors often show effects that are delayed in time, requiring the use of statistical models that are flexible enough to describe the additional time dimension of the exposure–response relationship. Here we develop the family of distributed lag non-linear models (DLNM), a modelling framework that can simultaneously represent non-linear exposure–response dependencies and delayed effects. This methodology is based on the definition of a 'cross-basis', a bi-dimensional space of functions that describes simultaneously the shape of the relationship along both the space of the predictor and the lag dimension of its occurrence. In this way the approach provides a unified framework for a range of models that have previously been used in this setting, and new more flexible variants. This family of models is implemented in the package `dlnm` within the statistical environment R. To illustrate the methodology we use examples of DLNMs to represent the relationship between temperature and mortality, using data from the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) for New York during the period 1987–2000. Copyright © 2010 John Wiley & Sons, Ltd.

Keywords: distributed lag; time series; smoothing; delayed effects

1. Introduction

Sometimes the effect of a specific exposure event is not limited to the period when it is observed, but it is *delayed* in time. This introduces the problem of modelling the relationship between an exposure occurrence and a sequence of future outcomes, specifying the distribution of the effects at different times after the event (defined *lags*). Ultimately, this step requires the definition of the additional lag dimension of an exposure–response relationship, describing the *time structure* of the effect.

This situation occurs frequently when assessing the short-term effects of environmental stressors: several time-series studies have reported that the exposure to high levels of air pollution or extreme temperatures affect health for a period lasting some days after its occurrence [1, 2]. Furthermore, the complexity increases in the presence of so-called 'harvesting': the phenomenon that arises when a stressor affects mainly a pool of frail individuals, whose events are only brought forward by a brief period of time by the effect of exposure [3, 4]. For non-recurrent outcomes, the depletion of the pool following a stress results in some reduction of cases few days later, thereby reducing the overall long-term impact. For both these reasons, the estimate of the effect depends on the appropriate specification of the lag dimension of the dependency, defining models flexible enough to represent simultaneously the exposure–response relationship and its temporal structure.

Among the various methods that have been proposed to deal with delayed effects, a major role is played by *distributed lag models* (DLM), recently used to quantify the health effect and assess the presence of harvesting in air pollution and temperature studies [2, 5, 6]. The main advantage of this method is that it allows the model to contain a detailed representation of the time-course of the exposure–response relationship, which in turn provides an estimate of the overall effect in the presence of delayed contributions or harvesting.

While conventional DLMs are suitable for describing the lag structure of linear effects, they show some limitations when used to represent non-linear relationships. We propose a solution, to relax further the assumptions on the shape of the relationship and extend this methodology to *distributed lag non-linear models* (DLNM), a family of models which can describe, in a flexible way, effects that vary simultaneously both along the space of the predictor and in the lag

[a] *Public Health and Policy Department, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1C 7HT, U.K.*
[b] *Epidemiology and Population Health Department, London School of Hygiene and Tropical Medicine, London, U.K.*
* *Correspondence to: A. Gasparrini, Public Health and Policy Department, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1C 7HT, U.K.*
† *E-mail: antonio.gasparrini@lshtm.ac.uk*

dimension of its occurrence. In this way the class of DLNMs also provides a unifying framework for existing simpler methods.

DLNMs have been previously described only briefly in epidemiological terms [7]: the aim of this paper is to develop this method rigorously, and to describe implementation in the specifically written package `dlnm` included in the statistical software R [8], providing an illustrative example of its application using a real data set. In Section 2 we briefly describe the basic model used in time series analysis and introduce the idea of basis as a general way to describe a non-linear relationship between a predictor and a response. In Section 3 we outline the additional complexity of effects delayed in time and provide a general representation of simple DLMs. In Section 4 we use the results obtained in the previous sections to define the general framework of DLNMs which includes all the models previously described as special cases. An application of this methodology to modelling the effect of temperature on mortality for New York is illustrated in Section 5. In Section 6 we provide some discussion and propose possible further developments.

2. The basic model

2.1. A general representation

A general model representation to describe the time series of outcomes Y_t with $t=1,\ldots,n$ is given by

$$g(\mu_t)=\alpha+\sum_{j=1}^{J} s_j(x_{tj};\boldsymbol{\beta}_j)+\sum_{k=1}^{K}\gamma_k u_{tk}, \tag{1}$$

where $\mu\equiv E(Y)$, g is a monotonic link function and Y is assumed to arise from a distribution belonging to the exponential family [9, 10]. The functions s_j denote smoothed relationships between the variables x_j and the linear predictor, defined by the parameter vectors $\boldsymbol{\beta}_j$. The variables u_k include other predictors with linear effects specified by the related coefficients γ_k. The functions s_j might be also specified through non-parametric methods based on generalized additive models [11, 12]. However, in the present development we rely on a completely parametric approach.

In time series analyses of environmental factors the outcomes Y_t are commonly daily counts, assumed to originate from a so-called overdispersed Poisson distribution with $E(Y)=\mu$, $V(Y)=\phi\mu$, and a canonical log-link in (1). These studies have taken advantage of the substantial improvements, during the last years, of statistical methods to quantify the short-term effects of air pollution [13, 14]. Usually these include a smooth function of time to capture the effect of confounders changing slowly over time, expressed as seasonality or long-time trends. Non-linear effects of metereological factors such as temperature and humidity are included as well. Categorical variables such as days of the week or age groups are modelled as factors. Although air pollution is commonly described by a linear relationship, this assumption may be relaxed in order to assess non-linear effects.

Here we focus on a general function s specifying the potentially non-linear and delayed effect of the predictor x, often referring, without loss of generality, to air pollution or temperature.

2.2. Basis functions

The relationship between x and $g(\mu)$ is represented by $s(x)$, which is included in the linear predictor of a generalized linear model as a sum of linear terms. This can be done through the choice of a *basis*, a space of functions of which we believe s to be an element [12]. The related *basis functions* comprise a set of completely known transformations of the original variable x that generate a new set of variables, termed *basis variables*. The complexity of the estimated relationship depends on the type of basis and its dimension.

Several different basis functions have been used to describe the potentially non-linear health effects of environmental factors, the choice depending on the assumptions about the shape of the relationship, the degree of approximation required by the specific purposes of the investigation, and interpretational issues. Among completely parametric methods, the main choices typically rely on functions describing smooth curves, such as polynomials or spline functions [15], or on the use of a linear threshold parameterization, represented by a truncated linear function $(x-\kappa)_+$ which equals $(x-\kappa)$ when $x>\kappa$ and 0 otherwise [16].

A general representation of the simple models described above is given by

$$s(x_t;\boldsymbol{\beta})=\mathbf{z}_t^{\mathrm{T}}\boldsymbol{\beta} \tag{2}$$

with $\mathbf{z}_{t\cdot}$ as the tth row of the $n\times v_x$ basis matrix $\mathbf{Z}$, obtained by the application of the basis functions to the original vector of exposures $\mathbf{x}$. $\mathbf{Z}$ can be then included in the design matrix of the model in (1) in order to estimate the related unknown parameters $\boldsymbol{\beta}$ defining the shape of the relationship.

3. Delayed effects

3.1. An additional dimension

In the presence of delayed effects, the outcome at a given time t may be explained in terms of past exposures $x_{t-\ell}$, with ℓ as the *lag*, representing the period elapsed between the exposure and the response. A comparatively simple approach is to apply a transformation to the original vector of ordered exposures $\mathbf{x}$, deriving the $n \times (L+1)$ matrix $\mathbf{Q}$, such as

$$\mathbf{q}_{t\cdot} = [x_t, \ldots, x_{t-\ell}, \ldots, x_{t-L}]^{\mathrm{T}} \tag{3}$$

with L defining the *maximum lag* and $\mathbf{q}_{\cdot 1} \equiv \mathbf{x}$ (the first column of $\mathbf{Q}$). We can also define $\boldsymbol{\ell} = [0, \ldots, \ell, \ldots, L]^{\mathrm{T}}$ as vector of lags corresponding to the $L+1$ columns of $\mathbf{Q}$.

This step specifies the additional lag dimension of the exposure–response relationship. Ultimately, the aim of the modelling framework proposed here is to simultaneously describe the dependency along two dimensions: the usual predictor space and in the new lag dimension.

3.2. Distributed lag models

When a linear relationship is assumed, the delayed effects can be naturally described by distributed lag models (DLM). This methodology allows the effect of a single exposure event to be distributed over a specific period of time, using several parameters to explain the contributions at different lags. These models have been extensively used to assess the lagged effects of environmental factors. The simplest formulation is an *unconstrained* DLM, specified by the inclusion of a parameter for each lag [5, 17]. Unfortunately, the precision of the estimates for the effects at specific lags is often very poor, due to the high correlation between exposures in adjacent days and the resulting collinearity in the model [1]. To gain more precision in the estimate of the distributed lag curve, it is possible to impose some constraints, for example assuming a constant effect within lag intervals [18], or describing a smooth curve using continuous functions such as polynomials [5, 19] or splines [6]. A simple model with the moving average of the exposures in the previous L days as a predictor can be considered as a special case of a DLM: such a model has been extensively used in the field of air pollution epidemiology [20] and sometimes used as well to quantify the effects of temperatures [21].

The algebraic notation for this class of models has only been given previously for polynomial DLMs [5]. Using the development provided in Sections 2.2 and 3.1, it is possible to formulate a simpler and general definition of DLM, in which the shape of the distributed effects along lags is specified by a proper basis. In matrix notation

$$s(x_t; \boldsymbol{\eta}) = \mathbf{q}_{t\cdot}^{\mathrm{T}} \mathbf{C} \boldsymbol{\eta}, \tag{4}$$

where $\mathbf{C}$ is an $(L+1) \times v_\ell$ matrix of basis variables derived from the application of the specific basis functions to the lag vector $\boldsymbol{\ell}$, and $\boldsymbol{\eta}$ a vector of unknown parameters. The addition of the supplementary dimension in (3) provides a structure for the application of the basis matrix $\mathbf{C}$, in order to describe the effects of lagged exposures. All the different DLMs described above can be derived from (4), by specifying the correspondent basis matrix: $\mathbf{C} \equiv \mathbf{1}$ (a vector of ones) for the moving average model, $\mathbf{C} \equiv \mathbf{I}$ (an identity matrix) for the unconstrained DLM, or $\mathbf{C}$ defined as a series of polynomial or splines functions of $\boldsymbol{\ell}$ for DLMs describing the effect as a smoothed curve along lags.

From (4) we can define

$$\mathbf{W} = \mathbf{QC} \tag{5}$$

with $\mathbf{W}$ the matrix of the v_ℓ transformed variables that are included in the design matrix to allow estimation of the parameters $\boldsymbol{\eta}$. The interpretation of the estimated parameters $\hat{\boldsymbol{\eta}}$ is aided by construction from them of the implied linear effects $\boldsymbol{\beta}$ at each lag, following:

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= \mathbf{C}\hat{\boldsymbol{\eta}}, \\ \mathrm{V}(\hat{\boldsymbol{\beta}}) &= \mathbf{C}V(\hat{\boldsymbol{\eta}})\mathbf{C}^{\mathrm{T}}, \end{aligned} \tag{6}$$

Here the choice of the basis to derive $\mathbf{C}$ can be considered as the application of a constraint to the shape of the distributed lag curve described by $\hat{\boldsymbol{\beta}}$.

Despite the specification of the basis functions in (4) being slightly different to that in (2), i.e. being applied to the vector $\boldsymbol{\ell}$ instead of the exposure series $\mathbf{x}$ itself, their goal is conceptually similar to describe the shape of the relationship, the former along distributed lags and the latter in the space of x.

4. Distributed lag non-linear models

As described in Sections 2 and 3, there are well-developed methods to describe flexible exposure–response relationships for simple lag models, or alternatively flexible DLMs for simple linear effects, but rarely are these two components modelled simultaneously. Extensions to describe non-linear effects have been proposed, using a piecewise parameterization or polynomials, for which a DLM can be constructed by applying the constraint matrix $\mathbf{C}$ to each term of a threshold [22] or piecewise function [23] or to the linear and quadratic terms [2], respectively. Nonetheless, these methods remain somewhat limited in their ability to describe this complex dependency.

A useful generalization is achieved through the generation of a new model framework which can describe non-linear relationships both in the space of the predictor and along lags, leading to the family of DLNM.

4.1. The concept of cross-basis

While the algebraic notation of DLNMs can be quite complex, involving three-dimensional arrays, the basic concept, which rests on the definition of a *cross-basis*, is straightforward. Extending the idea of basis described in Section 2, a cross-basis can be pictured as a bi-dimensional space of functions describing simultaneously the shape of the relationship along x and its distributed lag effects. Choosing a cross-basis amounts to choosing two sets of basis functions, which will be combined to generate the *cross-basis functions*.

4.2. The algebra of DLNM

To model the shape of the relationship in each of the two spaces we are considering, we need to apply simultaneously the two transformations described in Sections 2 and 3. First, as in (2), we choose a basis for $\mathbf{x}$ to define the dependency in the space of the predictor, specifying $\mathbf{Z}$. Then we create the additional lag dimension, as in (3), for each one of the derived basis variables of $\mathbf{x}$ stored in $\mathbf{Z}$. This produces a $n \times v_x \times (L+1)$ array $\dot{\mathbf{R}}$, which represents the lagged occurrences of each of the basis variables of $\mathbf{x}$. The construction is symmetric, in the sense that the order of the two transformations can be reversed, applying the basis functions directly to each column of the matrix $\mathbf{Q}$.

Defining $\mathbf{C}$, the matrix of basis variables for $\boldsymbol{\ell}$ seen in (4), a DLNM can then be specified by

$$s(x_t;\boldsymbol{\eta}) = \sum_{j=1}^{v_x}\sum_{k=1}^{v_\ell} \mathbf{r}_{tj\cdot}^{\mathrm{T}}\mathbf{c}_{\cdot k}\eta_{jk} = \mathbf{w}_{t\cdot}^{\mathrm{T}}\boldsymbol{\eta}, \tag{7}$$

where $\mathbf{r}_{tj\cdot}$ is the vector of lagged exposures for the time t transformed through the basis function j. The vector $\mathbf{w}_{t\cdot}$ is obtained by applying the $v_x \cdot v_\ell$ cross-basis functions to x_t, similarly to (5). We keep the same notation to emphasize the fact that the DLM specified in (4) is a special case of the more general DLNM in (7). To reach a compact formula for $\mathbf{W}$ of a similar form to (5), we need to present it as a tensor product. Defining $P_{i,j}$ as the operator permuting the indexes i and j of an array and assuming a generic $i \times j$ matrix as a $i \times j \times 1$ array, it follows that

$$\dot{\mathbf{A}} = (\mathbf{1}^{\mathrm{T}} \otimes \dot{\mathbf{R}}) \odot (\mathbf{1} \otimes P_{1,3}(\mathbf{C}) \otimes \mathbf{1}^{\mathrm{T}}) \tag{8}$$

with $\mathbf{1}$ indicating vectors of ones with appropriate dimensions. The symbols $\otimes$ and $\odot$ represent the Kronecker and Hadamard products, respectively. The $n \times (v_x \cdot v_\ell) \times (L+1)$ array $\dot{\mathbf{A}}$ is then re-arranged, summing along the third dimension of lags to obtain the final matrix of cross-basis functions $\mathbf{W}$. The equation in (8) is a modified version of the formula used to implement smoothing on a multidimensional grid through tensor product bases [24, 25]. The main difference in the cross-basis approach lies in the dimensions considered in the model. While the original method provides a framework to describe a smooth surface in the space of two distinct variables, the DLNM expresses simultaneously the effects in the space of a variable and in its lag dimension.

4.3. Interpreting a DLNM

Despite its complex parameterization, estimation of and inference about the parameters of a DLNM raise no more problems than any other generalized linear model, and can be carried out with common statistical softwares after the cross-basis variables have been specified. Nonetheless, while the interpretation of the simpler DLM in (4) is straightforward, consisting in reporting the estimated linear effects $\hat{\boldsymbol{\beta}}$ in (6) for each lag, the results of a more complex DLNM with smoothed non-linear dependencies are harder to summarize. One solution is to build a grid of predictions for each lag and for suitable values of exposure, using three-dimensional plots to provide an overall picture of the effects varying along the two dimensions.

Given a vector $\mathbf{x}^p$ of the m exposure values used for prediction and the resultant $m \times v_x$ matrix $\mathbf{Z}^p$, the corresponding $m \times v_x \times (L+1)$ array $\dot{\mathbf{R}}^p$ can be derived by repeating the matrix $\mathbf{Z}^p$ $L+1$ times along the dimension of the lags. The

computation of $\dot{\mathbf{R}}^p$ is slightly different than for the array $\dot{\mathbf{R}}$ used in the estimation process in (7). In this case the interest lies in the prediction of the effects at each lag given an exposure, not in the temporal sequence of the exposures themselves. The final array $\dot{\mathbf{A}}^p$ follows simply substituting $\mathbf{r}_{tj\cdot}$ with $\mathbf{r}^p_{tj\cdot}$ in (7) or $\dot{\mathbf{R}}$ with $\dot{\mathbf{R}}^p$ in (8).

The prediction grid, expressed with the $m \times (L+1)$ matrix of predicted effects $\mathbf{E}$ and related matrix of associated standard errors $\mathbf{E}^{sd}$, can be derived using the vector of estimated coefficients $\hat{\boldsymbol{\eta}}$, computed from the model fitted including the matrix of cross-basis functions $\mathbf{W}$. For each lag ℓ

$$\mathbf{e}_{\cdot\ell} = \mathbf{A}^p_{\cdot\cdot\ell}\hat{\boldsymbol{\eta}} \tag{9}$$

and, given $V(\hat{\boldsymbol{\eta}})$ the variance–covariance matrix of the estimated coefficients

$$\mathbf{e}^{sd}_{\cdot\ell} = \sqrt{\text{diag}(\mathbf{A}^p_{\cdot\cdot\ell} V(\hat{\boldsymbol{\eta}}) \mathbf{A}^{pT}_{\cdot\cdot\ell})}. \tag{10}$$

This grid is useful to compute the estimates of the effects by exposure at lag ℓ_p or by lag at exposure x_p, simply taking $\mathbf{e}_{\cdot\ell_p}$ and $\mathbf{e}_{x_p\cdot}$, respectively.

Finally, an estimate of the overall effect can be computed by summing all the contributions at different lags. The vector $\mathbf{e}_{\text{tot}}$ and associated standard errors $\mathbf{e}^{sd}_{\text{tot}}$, obtained summing the contributions at each lag, specify the effects by exposure over the whole lag period. They are obtained from

$$\mathbf{e}_{\text{tot}} = \mathbf{W}^p\hat{\boldsymbol{\eta}} \tag{11}$$

and

$$\mathbf{e}^{sd}_{\text{tot}} = \sqrt{\text{diag}(\mathbf{W}^p V(\hat{\boldsymbol{\eta}}) \mathbf{W}^{pT})}. \tag{12}$$

5. An application

5.1. Data and model choices

We apply DLNMs to investigate the effect of temperature on overall mortality for the city of New York, during the period 1987–2000. The data set is taken from the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) [26], available publicly through the Internet-based Health and Air Pollution Surveillance System website (http://www.ihapss.jhsph.edu). It includes 5114 daily observations of overall and cause-specific mortality, weather and pollution data.

The analysis is based on the model in (1), fitted through a generalized linear model with quasi-Poisson family, with the following choices regarding the control of confounders: natural cubic splines of time with 7 degrees of freedom (df) per year to describe long-time trends and seasonality; indicator variables for day of the week; natural cubic splines with 3 df at equally spaced quantiles for the average of dew point temperature at lag 0–1; linear terms for the average of ozone and CO at lag 0–1. These choices are motivated by several methodological and substantive papers on time-series analyses [21, 26, 27].

The effect of mean temperature has been investigated through the choice of two bases to describe the relationship in the space of temperature and lags; we illustrate a flexible model with natural cubic splines to describe the relationship in each dimension. The knots were placed at equally spaced values in the range of temperature, to allow enough flexibility in the tails, and at equal intervals in the logarithmic scale of lags, to allow more flexibility in the first part of the distributed lag curve, where more variability is expected [22, 28]. The maximum lag L was set to 30 days. Simpler models with the moving average of temperature in previous days have been fitted for comparison.

We have based the choice of the number of knots, which defines the df in each dimension, on modified Akaike and Bayesian information criteria for models with overdispersed responses fitted through quasi-likelihood [11, 27], given by:

$$\text{QAIC} = -2\mathscr{L}(\hat{\boldsymbol{\theta}}) + 2\hat{\phi}k \quad \text{and} \quad \text{QBIC} = -2\mathscr{L}(\hat{\boldsymbol{\theta}}) + \log(n)\hat{\phi}k, \tag{13}$$

where $\mathscr{L}$ is the log-likelihood of the fitted model with parameters $\hat{\boldsymbol{\theta}}$ and $\hat{\phi}$ the estimated overdispersion parameter, whereas k and n are the number of parameters and number of observations, respectively. The best model is chosen that minimizes the criteria above.

All the analyses were performed with the software R, version 2.10.1 [8], using the package `dlnm`, version 1.1.1, developed by the first two authors and publicly available on the R comprehensive archive network (CRAN). The code of the main analysis is included in Appendix A.

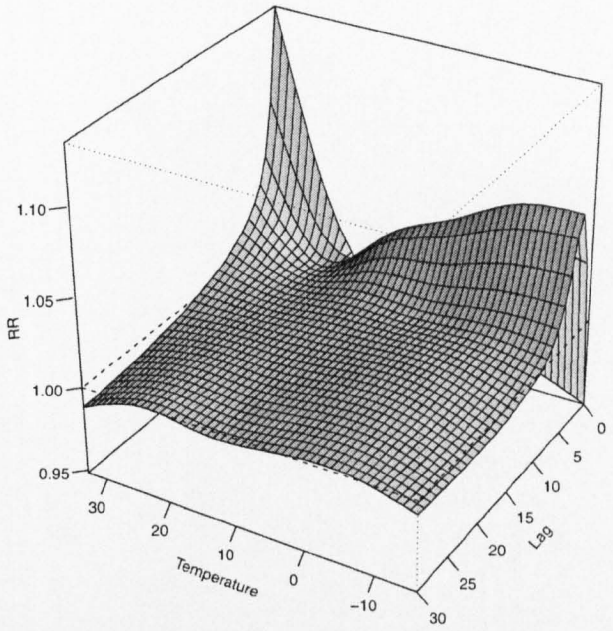


Figure 1. 3-D plot of RR along temperature and lags, with reference at 21°C.

5.2. Results

When used to compare different modelling choices, the QAIC led to a comparatively complex model, with 11 df for the space of the predictor and 5 df for the lag dimension, and a total of 55 parameters used to define the relationship. In contrast, the QBIC indicated a 5×5 df model, with 25 df spent to describe the overall effect. In the absence of any knowledge about the performances of these criteria within the DLNM framework, we chose the latter as our final model on the grounds of parsimony.

An overall picture of the effect of temperature on mortality is provided in Figure 1, showing a 3-D graph of the relative risk (RR) along temperature and lags compared with a reference value of 21°C, the point of overall minimum mortality. The plot shows a very strong and immediate effect of heat, and suggests a more delayed effect for extremely hot temperatures. The maximum effect of cold temperatures is reached approximately at lag 2–3. Inspection of the graph at longer lags suggests some harvesting for extreme temperatures.

Although the 3-D plot is a useful tool for summarizing the overall relationship in the two dimensions, uncertainty in the estimates cannot be included. In order to provide a more specific assessment of the relationship, we can plot the effects for specific temperatures or lags. Figure 2 shows the RR by temperature at specific lags (0, 5, 15 and 28) and by lag at specific temperatures (−10.8, −2.4, 26.5 and 31.3°C), corresponding approximately to 0.1th, 5th, 95th and 99.9th percentiles of temperature distribution (termed as moderate and extreme cold and heat). The overall effect of temperature, summing up the contributions for the 30 days of lag considered in the analysis, is included below. The shape of the temperature–mortality relationship seems to change along lags, with a different points of minimum mortality for lag 0 and 5 (first two graphs on top left). This plot confirms the more delayed effect of extreme heat if compared with moderate hot temperatures, with a significant risk lasting up to 10 and 3 days, respectively (third and fourth graphs from top right). Nonetheless, only extreme hot temperatures suggest a possible harvesting effect, starting after 15 days of lag. The overall estimated RR versus 21°C is 1.24 (95 per cent CI: 1.13–1.36) and 1.07 (95 per cent CI: 1.03–1.11) for extreme and moderate heat, respectively. Cold temperatures show a completely different pattern, with the effect of moderate cold sustained up to 25 days of lag (first two graphs on top right). In addition, the effect of cold seems to level off, with a slightly higher overall RR of 1.30 (95 per cent CI: 1.20-1.40) for moderate cold, compared to 1.20 (95 per cent CI: 1.04-1.39) for extreme cold (graph below).

To compare this DLNM with simpler alternatives, models with the moving average of lag 0–1 and lag 0–30 and the same spline functions for the space of temperature have been fitted. The former provides similar estimates for the effect of heat, but shows a weaker effect of low temperatures, with an estimated RR of 1.06 (95 per cent CI: 1.03–1.09) for moderate cold. This difference is probably due to underestimation, given the fact that low temperatures exert effects lasting longer than 2 days. Conversely, the moving average model with lag 0–30 shows similar effects for cold, but lower estimates for hot temperatures, with a RR of 1.01 (95 per cent CI: 0.97–1.04) and 1.06 (95 per cent CI: 0.97–1.17)

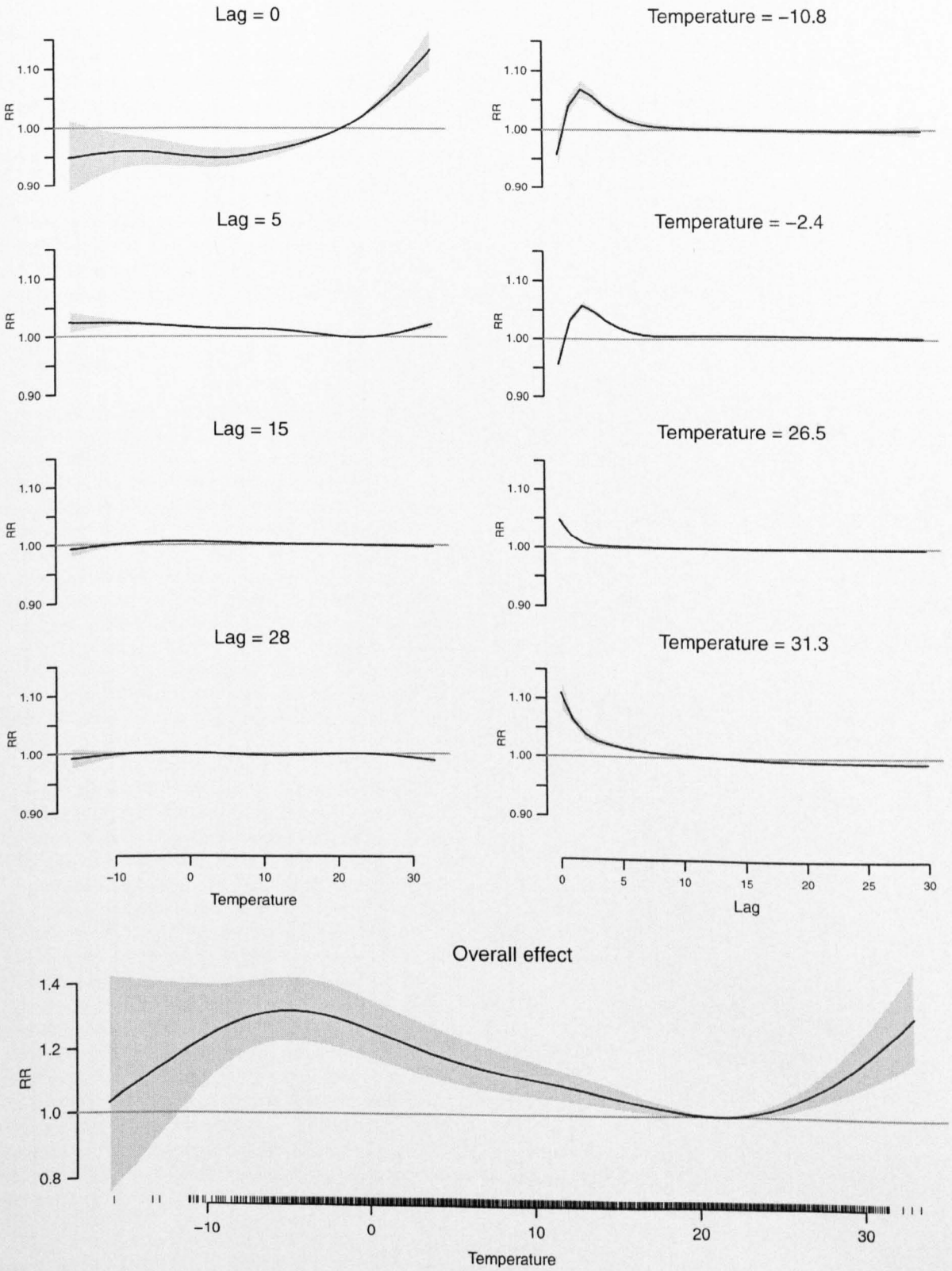


Figure 2. Plot of RR by temperature at specific lags (top left), RR by lag at 0.1th, 5th, 95th and 99.9th percentiles of temperature distribution (top right) and overall RR (below). Reference at 21°C.

for moderate and extreme heat, respectively. It is plausible that averaging over 31 days could cause some bias in the estimates, considering that each previous exposure within the lag period is assumed to provide the same contribution to the effect on each day. The criteria above indicate a better fit of DLNM, with a difference of 571 and 517 for QAIC and of 468 and 445 for QBIC if compared with lag 0–1 and 0–30 moving average models, respectively.

A sensitivity analysis has been carried out to assess the impact of model choices. In particular, we evaluate changes in the estimated overall effect (as described in the bottom of Figure 2) associated with varying the df used to specify the cross-basis functions (along both dimensions) and the seasonal and long-term trend component. Increasing the number of knots in the space of temperature produces a much less smoothed curve, probably due to overfitting, while no appreciable change is noted with different choices for spline in the lag dimension. Using more df to control for season and long-time trend does not affect the estimates, apart from a less pronounced decrease in the temperature–mortality curve at very low temperature. In addition, the inspection of lag and temperature-specific curves reveals that the supposed negative effect of heat at long lags, attributed to harvesting, completely disappears when increasing the seasonal control. This is plausible, given that the effects of models with an extended lag periods are more sensitive to the seasonal component.

6. Discussion

In this paper we have described the class of DLNMs, the members of which can be used to model the effect of factors showing at the same time non-linear dependencies and delayed effects. The specification of a DLNM is conceptually simple but flexible enough to allow a wide range of models including simple previously used as well as more complex new variants. The conceptual simplicity has allowed construction of an R package to fit this wide range of models.

One difficulty highlighted by this abundance of choice (basis types, number and placement of knots, maximum lag) is what criteria can be used to chose between alternatives. In the example above we used information criteria to guide choice of number of knots, but *a priori* arguments for choice of basis types and maximum lag. A previous discussion on choice of DLNM from an epidemiological perspective emphasized compromise between sufficient complexity to capture detail and sufficient simplicity to allow interpretation [7]. Because there is no consensus on what comprises an 'optimal' model, sensitivity analyses are particularly important, allowing dependence of key conclusions on model choice to be assessed. The broad range of DLMNs facilitates this. Regression diagnostics, such as residuals and partial autocorrelation plots, may also be helpful. In addition, we have discussed choice of DLNM assuming that it focuses on the variable of interest (temperature in our example). There is also a problem of model selection for covariates, some parts of which might also be DLNMs. This problem, sometimes referred to as adjustment uncertainty, has received some attention in time series studies of pollution [29, 30] as well as generally [31]. Again no consensus has emerged on what approach is optimal, and analyses of sensitivity to this component of model choice is also important.

The current implementation of DLNMs as illustrated in Section 5 is based on a completely parametric method, where the cross-basis dimension $v_x \times v_l$ equals the number of df spent to describe the relationship. Recently, interesting alternatives based on penalized regression with low-rank smoothers have been proposed to deal with non-linear effects [32, 33], and also applied to describe the distributed lag curve [6, 22]. Although completely parametric approaches seems to be preferred to control for season and long-term trend in time series data [27, 34, 35], the penalized methods could show some advantage in the bi-dimensional framework of DLNM. This issue represents an opportunity for further development, and could benefit from the research already carried out on penalized tensor-product smoothers [25, 36]. In addition, the algebraic definition in Section 4 is still valid in this new context, and only the estimation algorithm to derive $\hat{\boldsymbol{\eta}}$ and $V(\hat{\boldsymbol{\eta}})$ actually changes.

The development of DLNMs described in Section 4 involves only a comparatively complex parameterization of the lagged exposure series, as expressed by (7). Although our application has involved the use of an overdispersed Poisson log-linear model, we do emphasize through the development and notation in (1), that this framework has very general applicability, for example to time series data with other outcome distributions. More importantly, the main concept is fairly general, and can be easily translated in other study design and regression models.

The analysis of the data for New York during the period 1987–2000 offers some evidence for the potential of this framework to highlight complex dependencies of environmental factors, which would be largely obscured when using simpler models. We believe this approach represents a useful tool to gain understanding of phenomena investigated in environmental studies and other scientific fields.

Appendix A: R code

The following code reproduces the main analysis and graphs included in Section 5. The packages `dlnm` and `NMMAPSlite` may be downloaded directly through R from the CRAN.

A detailed overview of the capabilities of the package `dlnm` is illustrated in the vignette included in the implementation, available typing `vignette("dlnmOverview")` in R.

```
require(dlnm);require(NMMAPSlite)

################################
# LOAD AND PREPARE THE DATASET
################################

initDB()
data <- readCity("ny", collapseAge = TRUE)
data <- data[,c("city","date","dow","death","tmpd","dptp","rhum",
 "o3tmean","o3mtrend","cotmean","comtrend")]

# TEMPERATURE: CONVERSION TO CELSIUS
data$temp <- (data$tmpd-32)*5/9
# POLLUTION: O3 AND CO AT LAG-01
data$o3 <- data$o3tmean + data$o3mtrend
data$co <- data$cotmean + data$comtrend
data$o301 <- filter(data$o3,c(1,1)/2,side=1)
data$co01 <- filter(data$co,c(1,1)/2,side=1)
# DEW POINT TEMPERATURE AT LAG 0-1
data$dp01 <- filter(data$dptp,c(1,1)/2,side=1)

################################
# CROSSBASIS SPECIFICATION
################################

# FIXING THE KNOTS AT EQUALLY SPACED VALUES
range <- range(data$temp,na.rm=T)
ktemp <- range[1] + (range[2]-range[1])/5*1:4
# CROSSBASIS MATRIX
ns.basis <- crossbasis(data$temp,varknots=ktemp,cenvalue=21,
 lagdf=5,maxlag=30)

################################
# MODEL FIT AND PREDICTION
################################

ns <- glm(death ~ ns.basis + ns(dp01,df=3) + dow + o301 + co01 +
 ns(date,df=14*7),family=quasipoisson(), data)
ns.pred <- crosspred(ns.basis,ns,at=-16:33)

################################
# RESULTS AND PLOTS
################################

# 3-D PLOT (FIGURE 1)
crossplot(ns.pred,label="Temperature")
# SLICES (FIGURE 2, TOP)
percentiles <- round(quantile(data$temp,c(0.001,0.05,0.95,0.999)),1)
ns.pred <- crosspred(ns.basis,ns,at=c(percentiles,-16:33))
crossplot(ns.pred,"slices",var=percentiles,lag=c(0,5,15,28),
 label="Temperature")
# OVERALL EFFECT (FIGURE 2, BELOW)
crossplot(ns.pred,"overall",label="Temperature",
  title="Overall effect of temperature on mortality
```

```
New York 1987-2000")
# RR AT CHOSEN PERCENTILES VERSUS 21C (AND 95%CI)
ns.pred$allRRfit[as.character(percentiles)]
cbind(ns.pred$allRRlow,ns.pred$allRRhigh)[as.character(percentiles),]

##############################

# THE MOVING AVERAGE MODELS UP TO LAG x (DESCRIBED IN SECTION 5.2)
# CAN BE CREATED BY THE CROSSBASIS FUNCTION INCLUDING THE
# ARGUMENTS lagtype="strata", lagdf=1, maxlag=x
```

Acknowledgements

The authors are thankful to the two anonymous reviewers for their interesting and constructive comments on the manuscript.

References

1. Zanobetti A, Schwartz J, Samoli E, Gryparis A, Touloumi G, Atkinson R, Le Tertre A, Bobros J, Celko M, Goren A, Forsberg B, Michelozzi P, Rabczenko D, Aranguez Ruiz E, Katsouyanni K. The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement. *Epidemiology* 2002; **13**(1):87–93.
2. Braga AL, Zanobetti A, Schwartz J. The time course of weather-related deaths. *Epidemiology* 2001; **12**(6):662–667.
3. Schwartz J. Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? *Epidemiology* 2001; **12**(1):55–61.
4. Rabl A. Air pollution mortality: harvesting and loss of life expectancy. *Journal of Toxicology and Environmental Health: Part A* 2005; **68**(13–14):1175–1180.
5. Schwartz J. The distributed lag between air pollution and daily deaths. *Epidemiology* 2000; **11**(3):320–326.
6. Zanobetti A, Wand MP, Schwartz J, Ryan LM. Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics* 2000; **1**(3):279–292.
7. Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology* 2006; **17**(6):624–631.
8. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria 2009. Available from: http://www.R-project.org. ISBN 3-900051-07-0.
9. McCullagh P, Nelder JA. *Generalized Linear Models* (2nd edn). Chapman & Hall/CRC Press: London/Boca Raton, 1989.
10. Dobson A, Barnett A. *An Introduction to Generalized Linear Models* (3rd edn). CRC Press/Chapman & Hall: Boca Raton/London, 2008.
11. Hastie T, Tibshirani R. *Generalized Additive Models* (2nd edn). Chapman & Hall/CRC Press: London/Boca Raton, 1990.
12. Wood S. *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC Press: London/Boca Raton, 2006.
13. Dominici F, Sheppard L, Clyde M. Health effects of air pollution: a statistical review. *International Statistical Review* 2003; **71**(2):243–276.
14. Touloumi G, Atkinson R, Le Tertre A, Samoli E, Schwartz J, Schindler C, Vonk J, Rossi G, Saez M, Rabszenko D. Analysis of health outcome time series data in epidemiological studies. *EnvironMetrics* 2004; **15**(2):101–117.
15. Dominici F, Daniels M, SLand Samet ZJ. Air pollution sand mortality: estimating regional and national dose–response relationships. *Journal of the American Statistical Association* 2002; **97**:100–111.
16. Daniels M, Dominici F, Samet J, Zeger S. Estimating particulate matter-mortality dose–response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. *American Journal of Epidemiology* 2000; **152**(5):397.
17. Hajat S, Armstrong BG, Gouveia N, Wilkinson P. Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. *Epidemiology* 2005; **16**(5):613–620.
18. Welty LJ, Zeger SL. Are the acute effects of particulate matter on mortality in the National Morbidity, Mortality, and Air Pollution Study the result of inadequate control for weather and season? A sensitivity analysis using flexible distributed lag models. *American Journal of Epidemiology* 2005; **162**(1):80–88.
19. Rondeau V, Berhane K, Thomas DC. A three-level model for binary time-series data: the effects of air pollution on school absences in the Southern California Children's Health Study. *Statistics in Medicine* 2005; **24**(7):1103–1115.
20. Dominici F. Time-series analysis of air pollution and mortality: a statistical review. *Research Report*, Health Effects Institute 2004; (123):3–27. Discussion 29–33.
21. Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology* 2009; **20**(2):205–213.
22. Muggeo VM. Modeling temperature effects on mortality: multiple segmented relationships with common break points. *Biostatistics* 2008; **9**(4):613–620.
23. Roberts S, Martin MA. A distributed lag approach to fitting non-linear dose–response models in particulate matter air pollution time series investigations. *Environmental Research* 2007; **104**(2):193–200.
24. Eilers PHC, Currie ID, Durbn M. Fast and compact smoothing on large multidimensional grids. *Computational Statistics and Data Analysis* 2006; **50**(1):61–76.
25. Eilers PH, Gampe J, Marx BD, Rau R. Modulation models for seasonal time series and incidence tables. *Statistics in Medicine* 2008; **27**(17):3430–3441.
26. Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. Revised analyses of the National Morbidity, Mortality, and Air Pollution Study: mortality among residents of 90 cities. *Journal of Toxicology and Environmental Health: Part A* 2005; **68**(13–14):1071–1092.

27. Peng RD, Dominici F, Louis TA. Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society, Series A* 2006; **169**(2):179–203.
28. Peng RD, Dominici F. A Bayesian hierarchical distributed lag model for estimating the time course of risk of hospitalization associated with particulate matter air pollution. *Journal of the Royal Statistical Society, Series A* 2009; **58**(1):3–24.
29. Dominici F, Wang C, Crainiceanu C, Parmigiani G. Model selection and health effect estimation in environmental epidemiology. *Epidemiology* 2008; **19**(4):558–560.
30. Crainiceanu C, Dominici F, Parmigiani G. Adjustment uncertainty in effect estimation. *Biometrika* 2008; **95**(3):635.
31. Ritov Y, Bickel P. Achieving information bounds in non and semiparametric models. *Annals of Statistics* 1990; **18**(2):925–938.
32. Ruppert D, Wand M, Carroll R. *Semiparametric Regression*. Cambridge University Press: Cambridge, 2003.
33. Schimek M. Semiparametric penalized generalized additive models for environmental research and epidemiology. *EnvironMetrics* 2009; **20**(6):699–717.
34. Baccini M, Biggeri A, Lagazio C, Lertxundi A, Saez M. Parametric and semi-parametric approaches in the analysis of short-term effects of air pollution on health. *Computational Statistics and Data Analysis* 2007; **51**(9):4324–4336.
35. He S, Mazumdar S, Arena V. A comparative study of the use of GAM and GLM in air pollution research. *EnvironMetrics* 2006; **17**(1):81–93.
36. Wood S. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics* 2006; **62**(4):1025–1036.

Chapter 6

Research paper III

Title: Distributed lag linear and non-linear models in R: the package dlnm.

Author(s): Antonio Gasparrini.

Journal/Publisher: Journal of Statistical Software.

Type of publication: Research paper.

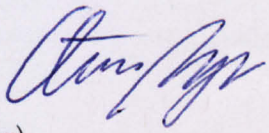
Stage of publication: Published in Volume 43, Issue 8, July 2011, Pages 1-20.

URL: http://www.jstatsoft.org/v43/i08.

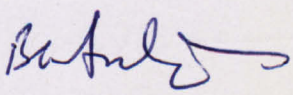
Academic peer-reviewed: Yes.

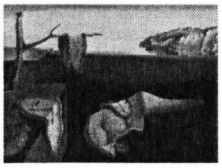
Candidate's role: See Section 2.3.

Candidate's signature:



Co-supervisor: (Prof. Ben Armstrong)





Journal of Statistical Software

July 2011, Volume 43, Issue 8. *http://www.jstatsoft.org/*

Distributed Lag Linear and Non-Linear Models in R: The Package dlnm

Antonio Gasparrini
London School of Hygiene and Tropical Medicine

Abstract

Distributed lag non-linear models (DLNMs) represent a modeling framework to flexibly describe associations showing potentially non-linear and delayed effects in time series data. This methodology rests on the definition of a *crossbasis*, a bi-dimensional functional space expressed by the combination of two sets of basis functions, which specify the relationships in the dimensions of predictor and lags, respectively. This framework is implemented in the R package **dlnm**, which provides functions to perform the broad range of models within the DLNM family and then to help interpret the results, with an emphasis on graphical representation. This paper offers an overview of the capabilities of the package, describing the conceptual and practical steps to specify and interpret DLNMs with an example of application to real data.

Keywords: distributed lag models, time series, smoothing, delayed effects, R.

1. Introduction

The main purpose of a statistical regression model is to define the relationship between a set of predictors and an outcome, and then to estimate the related effect. A further complexity arises when the dependency shows some *delayed effects*: in this case, a specific occurrence of a predictor (let us call it an *exposure event*) affects the outcome for a lapse of time well beyond the event period. This step requires the definition of more complex models to characterize the association, specifying the temporal structure of the dependency.

1.1. Conceptual framework

The specification of suitable statistical models for delayed effect, and the interpretation of their results, is aided by the development of a proper conceptual framework. The key feature of this framework is the definition of an additional dimension to characterize the association,

which specifies the temporal dependency between exposure and outcome on the scale of *lag*. This term, borrowed by the literature on time series analysis, represents the time interval between the exposure event and the outcome when evaluating the delay of the effect. In case of protracted exposures, the data can be structured by the partition in equally-spaced time periods, defining a series of exposure events and outcomes realizations. This partitioning also defines lag units. Within this time structure, the exposure-response relationship can be described with either of two opposite perspectives: we can say that a specific exposure events produces effects on multiple future outcomes, or alternatively that a specific outcome is explained in terms of contributions by multiple exposure events in the past. The concept of lag can then be used to describe the relationship either *forward* (from a fixed exposure to future outcomes) or *backward* in time (from a fixed outcome to past exposures).

Ultimately, the main feature of statistical models for delayed effects is their bi-dimensional structure: the relationship is simultaneously described both along the usual space of the predictor and in the additional dimension of the lags.

1.2. Distributed lag models

The issue of delayed effects has been recently addressed in studies assessing the short term effects of environmental stressors: several time series studies have reported that the exposure to high levels of pollution or extreme temperatures affects health for a period lasting some days after the its occurrence (Braga *et al.* 2001; Goodman *et al.* 2004; Samoli *et al.* 2009; Zanobetti and Schwartz 2008).

The time series study design offers several advantages in order to deal with delayed effects, given the defined temporal structure of the data and the straightforward definition of the lag dimension, where the time partitioning is directly specified by the equally-spaced and ordered time points. In this setting, delayed effects are elegantly described by *distributed lag models* (DLMs), a methodology originally developed in econometrics (Almon 1965) and recently been used to quantify health effects in studies on environmental factors (Schwartz 2000; Zanobetti *et al.* 2000; Muggeo and Hajat 2009). This methodology allows the effect of a single exposure event to be distributed over a specific period of time, using several parameters to explain the contributions at different lags, thus providing a comprehensive picture of the time-course of the exposure-response relationship.

Conventional DLMs rely on the assumption of a linear effect between the exposure and the outcome. Some attempts to relax this assumption and explore delayed effects of factors showing non-linear relationships have been proposed (Roberts and Martin 2007; Braga *et al.* 2001). In particular, Muggeo (2008) introduced a methodology based on constrained segmented parameterization, assuming distributed lag linear effects of hot and cold temperatures beyond two thresholds. This methodology is implemented in an R package presented in a previous issue of this Journal (Muggeo 2010).

More recently, a general approach has been proposed to further relax the linearity assumption, and flexibly describe simultaneously non-linear and delayed effects. This step has lead to the generation of the new modeling framework of *distributed lag non-linear models* (DLNMs) (Gasparrini *et al.* 2010; Armstrong 2006), implemented in the R package **dlnm** (Gasparrini and Armstrong 2010).

1.3. Aim of the paper

The package **dlnm** within the statistical environment R (R Development Core Team 2011) offers a set of tools to specify and interpret the results of DLNMs. The aim of this paper is to provide a comprehensive overview of the capabilities of the package, including a detailed summary of the functions, with an example of application to real data. The example refers to the effects on all-cause mortality of two environmental factors, air pollution (ozone) and temperature, in the city of Chicago during the period 1987-2000. A thorough methodological description of DLNMs, together with the complete algebraical development, have been given elsewhere (Gasparrini *et al.* 2010). In this paper I reconsider the main conceptual and practical steps to define a DLNM, predict the effects and interpret the results with the aid of graphical features. The description of the functions included in the package and the related code for each step will be presented. The code is also available as supplemental material.

The paper is structured as follows: Section 2 considers the general problem of modeling non-linear or delayed effects, with an overview of the statistical approaches proposed so far. In the next three Sections, the development of the methodology is illustrated in details, showing the specification (Section 3), effect prediction (Section 4) and representation (Section 5) of DLNMs. Section 6 shows an example of alternative modeling approaches and the issue of model selection, while Section 7 discusses specific data requirements. Section 8 describes potential future developments. Final comments are provided in Section 9.

The package **dlnm** (current version 1.4.1) is expected to be loaded in the session, by typing:

```
R> library("dlnm")
```

Complementary information on the capabilities of the package, together with additional examples of application to real data, can be found in the vignette `dlnmOverview`. This document is included in the implementation of the package, and can be visualized by typing:

```
R> vignette("dlnmOverview", package = "dlnm")
```

2. Non-linear and delayed effects

In this section I present the basic formulation for a time series model, then introducing the methods to describe non-linear and then delayed effects, the latter through the specification of simple DLMs. The developtment will be formulated in such a way to facilitate the introduction of the DLNM framework in Sections 3–5.

2.1. The basic model

A model for time series data may be generally represented by:

$$g(\mu_t) = \alpha + \sum_{j=1}^{J} s_j(x_{tj}; \boldsymbol{\beta}_j) + \sum_{k=1}^{K} \gamma_k u_{tk} \tag{1}$$

where $\mu_t \equiv \mathrm{E}(Y_t)$, g is a monotonic link function and Y_t is a series of outcomes with $t = 1, \ldots, n$, assumed to arise from a distribution belonging to the exponential family (Dobson and Barnett 2008). The functions s_j specify the relationships between the variables x_j and

the linear predictor, defined by the parameter vectors $\boldsymbol{\beta}_j$. The variables u_k include other predictors with linear effects specified by the related coefficients γ_k.

In the illustrative example on Chicago data described in Section 1.3, the outcome Y_t is daily death counts, assumed to originate from a so-called overdispersed Poisson distribution with $\mathrm{E}(Y) = \mu$, $\mathrm{V}(Y) = \phi\mu$, and a canonical log-link in (1). The analysis follows a conventional approach used in time series studies on environmental epidemiology (Dominici 2004; Touloumi *et al.* 2004), where the association between daily ozone and temperature levels on mortality is controlled for other confounding factors like seasonal and long time trend and day of the week. However, the framework is general and applies to every outcome and predictors measures collected as time series data.

The non-linear and delayed effects of ozone and temperature are modeled through as particular functions s_j which define the relationship along the two dimensions of predictor and lags.

2.2. Non-linear exposure-response relationships

The first step in the development of DLNMs is to define the relationship in the space of the predictor. Generally, non-linear exposure-response dependencies are expressed in regression models through appropriate functions s. Within completely parametric approaches, several different functions have been proposed, each of them characterized by different assumptions and degree of flexibility. The main choices typically rely on functions describing smooth curves, like polynomials or spline functions (Braga *et al.* 2001; Dominici *et al.* 2004); on the use of a linear threshold parameterization (Muggeo 2010; Daniels *et al.* 2000); or on the simple stratification through dummy parameterization.

All of these functions apply a transformation of the original predictor to generate a set of transformed variables included in the model as linear terms. A useful generalization is achieved introducing the concept of *basis*: a space of functions of which we believe s to be an element (Wood 2006). The related *basis functions* comprise a set of completely known transformations of the original variable x that generate a new set of variables, termed *basis variables*. An algebraic representation may be given by:

$$s(x_t; \boldsymbol{\beta}) = \mathbf{z}_{t\cdot}^{\top} \boldsymbol{\beta} \tag{2}$$

with $\mathbf{z}_{t\cdot}$ as the t^{th} row of the $n \times v_x$ basis matrix $\mathbf{Z}$. In the parametric approach adopted here, the basis dimension v_x equals the degrees of freedom (df) spent to define the relationship in this space, and is proportional to the degree of flexibility of the function. The unknown parameters $\boldsymbol{\beta}$ can be estimated including $\mathbf{Z}$ in the design matrix of the model in (1) .

This first step in the definition of DLNMs is performed in the package **dlnm** with the function `mkbasis()`, used to create the basis matrix $\mathbf{Z}$. The purpose of this function is to provide a general way to include non-linear effects of x, with different choices specified as different arguments of `mkbasis()`. As an example, I build a basis matrix applying the selected basis functions to the vector $\mathbf{x} = [1, \ldots, 5]^{\top}$:

```
R> mkbasis(1:5, type = "bs", df = 4, degree = 2, cenvalue = 3)

$basis
           b1        b2        b3     b4
[1,] -0.12500 -0.75000 -0.12500 0.0000
```

```
[2,]  0.53125 -0.46875 -0.12500 0.0000
[3,]  0.00000  0.00000  0.00000 0.0000
[4,] -0.12500 -0.46875  0.53125 0.0625
[5,] -0.12500 -0.75000 -0.12500 1.0000

$type
[1] "bs"

$df
[1] 4

$degree
[1] 2

$knots
33.33333% 66.66667%
 2.333333  3.666667

$bound
[1] 1 5

$int
[1] FALSE

$cen
[1] TRUE

$cenvalue
[1] 3
```

The result is a list object storing the basis matrix and the arguments defining it. In this case, the chosen basis is a quadratic spline with 4 df, defined by the arguments `type`, `df` and `degree`. The basis variables are centered to the value of 3.

Different types of basis may be chosen through the second argument `type`. The available options are natural cubic or simple B-splines (`type = "ns"` or `"bs"`, through a call to the related functions in the package **splines**); strata through dummy variables (`"strata"`); polynomials (`"poly"`); threshold-type functions such as low, high or double threshold or piecewise parameterization (`"lthr"`, `"hthr"`, `"dthr"`); and simply linear (`"lin"`). The argument `df` defines the dimension of the basis (the number of its columns, basically the number of transformed variables). This value may depend on the argument `knots` (which overcomes `df`), specifying the position of the internal knots for types `"ns"` and `"bs"` (with boundary knots specified by `bound`), the cut-off points for `"strata"` (defining right-open intervals) and the thresholds/cut-off points for `"lthr"`, `"hthr"` and `"dthr"`. If not defined (as in the example above), the knots are placed at equally-spaced quantiles by default, and the boundary knots at the range of the predictor values. The argument `degree` select the degree of polynomial for `"bs"` and `"poly"`.

The arguments `cen` and `cenvalue` are used to center the basis for continuous functions (types `"ns"`, `"bs"`, `"poly"`, and `"lin"`), with default to the mean of the original variable if `cenvalue` is not provided. An "intercept" can be included with the argument `int`, set by default at `FALSE` to avoid identifiability problems. The concept of intercept is different between bases: types `"ns"` and `"bs"` apply a complex parameterization where the intercept is implicitly built within the basis variables (see the related help pages typing `?ns` and `?bs`); in type `"strata"`, the intercept corresponds to the dummy variable for the baseline stratum (the first one by default), which is excluded if `int = FALSE`; the intercept is the usual vector of 1's in the other types. See the help page (typing `?mkbasis`) for additional information.

2.3. Delayed effects

The second step to define a DLNM is to specify the function to model the relationship in the additional dimension of lags, allowing for delayed effects. In this situation, the outcome Y_t at a given time t may be explained in terms of past exposures $x_{t-\ell}$, with ℓ as the lag. Given a maximum lag L, the additional lag dimension can be expressed by the $n \times (L+1)$ matrix $\mathbf{Q}$, such as:

$$\mathbf{q}_{t\cdot} = [x_t, \ldots, x_{t-\ell}, \ldots, x_{t-L}]^\top \tag{3}$$

with $\mathbf{q}_{t\cdot}$ as the t^{th} row of $\mathbf{Q}$. The vector of lags $\boldsymbol{\ell} = [0, \ldots, \ell, \ldots, L]^\top$ corresponds to the scale of this additional dimension.

Simple DLMs allow for delayed effects of linear relationships using a function to describe the dependency between the outcome and lagged exposures. Several alternatives have been proposed, from an *unconstrained* DLM (simply a parameter for each $x_{t-\ell}$ with $\ell \in \boldsymbol{\ell}$) (Hajat *et al.* 2005), to the use of strata (Pattenden *et al.* 2003), polynomials (Schwartz 2000) or splines (Zanobetti *et al.* 2000; Armstrong 2006). A compact and general algebraic definition of a DLM is given by (see Gasparrini *et al.* 2010, Section 3.2):

$$s(x_t; \boldsymbol{\eta}) = \mathbf{q}_{t\cdot}^\top \mathbf{C}\boldsymbol{\eta} \tag{4}$$

where $\mathbf{C}$ is an $(L+1) \times v_\ell$ matrix of basis variables derived from the application of the specific basis functions to the lag vector $\boldsymbol{\ell}$, and $\boldsymbol{\eta}$ a vector of unknown parameters. This basis matrix is used to define the relationship along the lag dimension. All the DLMs described above differ only in the choice of the basis to derive the matrix $\mathbf{C}$.

This second step is carried out in **dlnm** through the function `mklagbasis()`, which calls `mkbasis()` in order to build the basis matrix $\mathbf{C}$. For example:

```
R> mklagbasis(maxlag = 5, type = "strata", knots = c(2, 4))

$basis
     b1 b2 b3
lag0  1  0  0
lag1  1  0  0
lag2  0  1  0
lag3  0  1  0
lag4  0  0  1
lag5  0  0  1
```

```
$type
[1] "strata"

$df
[1] 3

$knots
[1] 2 4

$int
[1] TRUE

$maxlag
[1] 5
```

In this example, after the maximum lag is fixed at 5 through the first argument `maxlag`, the lag vector `0:maxlag`, corresponding to $\boldsymbol{\ell} = [0, \ldots, 5]^\top$, is automatically created and the chosen function applied to it. In this case, a dummy parameterization with strata defined by the cut-off points 2 and 4 (right-open intervals) included in `knots`. The available functions, and the arguments to specify them, are essentially the same illustrated above for `mkbasis()`. The only difference is that the basis matrix is never centered and by default includes an intercept (`int = TRUE`, see `?mklagbasis`). In addition, the knots (if not specified) are placed by default at equally-spaced values in the log scale, allowing more flexibility in the first lag period. The specific argument `type = "integer"` produces strata variables for each integer values, defining $\mathbf{C}$ as an identity matrix, and may be used to specify unconstrained DLMs.

3. Specifying a DLNM

The last step in the specification of a DLNM involves the simultaneous definition of the relationship in the two dimensions of predictor and lags, as described in Sections 2.2 and 2.3. In spite of the different terminology of non linearity and delayed effects, the two procedures are conceptually similar: to define a basis which expresses the relationship in the related space. This similarity is highlighted by the analogy of the two functions `mkbasis()` and `mklagbasis()`.

DLNMs are then specified by the definition of a *cross-basis*, a bi-dimensional functional space describing at the same time the dependency along the range of the predictor and in its lag dimension. Algebraically, this reduces to concurrently apply the two transformations explained in (2) and (3). First, choosing a basis for $\mathbf{x}$ to derive $\mathbf{Z}$, then creating the additional lag dimension for each one of the derived basis variables of $\mathbf{x}$, producing a $n \times v_x \times (L+1)$ array $\dot{\mathbf{R}}$. With $\mathbf{C}$ defined in (4), a DLNM can be represented by:

$$s(x_t; \boldsymbol{\eta}) = \sum_{j=1}^{v_x} \sum_{k=1}^{v_\ell} \mathbf{r}_{tj.}^\top \mathbf{c}_{.k} \eta_{jk} = \mathbf{w}_{t.}^\top \boldsymbol{\eta} \tag{5}$$

with $\mathbf{w}_{t.}$ as the t^{th} row of the cross-basis matrix $\mathbf{W}$. Additional details are given in Gasparrini *et al.* (2010, Section 4.2).

Choosing a cross-basis amounts to selecting two sets of basis functions as described above, which will be combined to generate *cross-basis functions*. This is carried out by the function `crossbasis()`, which calls the functions `mkbasis()` and `mklagbasis()` to generate the two basis matrices **Z** and **C**, respectively, than combining them through a tensor product to produce **W** following (5). This function can be applied to specify the two cross-bases for ozone and temperature in the example described in Section 2.1. The related code is:

```
R> basis.o3 <- crossbasis(chicagoNMMAPS$o3, vartype = "hthr",
+    varknots = 40.3, lagtype = "strata", lagknots = c(2, 6), maxlag = 10)
R> basis.temp <- crossbasis(chicagoNMMAPS$temp, vartype = "bs",
+    vardegree = 3, vardf = 6, cenvalue = 25, lagdf = 5, maxlag = 30)
```

The result is an object of class '`crossbasis`', corresponding to the cross-basis matrix **W** in (5) and the related arguments as attributes. The first argument `x` of `crossbasis()` is the predictor series, in this case `chicagoNMMAPS$o3` and `chicagoNMMAPS$temp` available in the dataset included in the package (see `?chicagoNMMAPS`). In the current implementation, the values in `x` are expected to represent an equally-spaced and ordered series, with the interval defining the lag unit. The series must be complete, although missing values are allowed (see Section 7). The argument `maxlag` defines the maximum lag.

The other arguments are similar to those enumerated in Sections 2.2 - 2.3. The function `crossbasis()` passes the arguments with prefix `var-` to `mkbasis()`, in order to specify **Z**, and the arguments with prefix `lag-` to `mklagbasis()`, producing **C**. In this example, the cross-basis for ozone comprises a threshold function for the space of the predictor, with a linear relationship beyond 40.3 μgr/m^3, and a dummy parameterization assuming constant distributed lag effects along the strata of lags 0-1, 2-5 and 6-10. In contrast, the options for temperature are a cubic spline with 6 df (knots at equally-spaced percentiles by default) centered at 25°C, and a natural cubic spline (`lagtype = "ns"` by default) with 5 df (knots at equally-spaced values in the log scale of lags by default), up to a maximum of 30 lags.

As explained in Section 2.2, the basis variables for the space of the predictor are centered by default for continuous functions. The default centering point is the predictor mean, if not set with `cenvalue` (for example at 25°C for the cross-basis of temperature above). This value represents the reference for predicted effects from a DLNM (see Section 4). The choice of the reference value does not affect the fit of the model, and different values may be chosen depending on interpretational issues. The reference in non-continuous functions is automatically set to the first interval in `strata` and `integer`, or to the flat region in `lthr`, `hthr`, `dthr`. As suggested in Section 2.2, it is strongly recommended to avoid the inclusion of an intercept in the basis for the predictor space (`varint` must be `FALSE`, as default), otherwise a rank-deficient cross-basis matrix will be specified, causing some of the cross-variables to be excluded in the regression model. A complete overview of the available options is given in the help page (typing `?crossbasis`).

These choices may be checked by the function `summary.crossbasis()`. For example:

```
R> summary(basis.temp)

CROSSBASIS FUNCTIONS
observations: 5114
range: -26.66667 , 33.33333
```

```
total df: 30
maxlag: 30

BASIS FOR VAR:
type: bs with degree 3
df: 6 , knots at: 1.666667 10.55556 19.44444
boundary knots at -26.66667 33.33333
centered on 25

BASIS FOR LAG:
type: ns
df: 5 , knots at: 1.105502 3.322105 9.983144
boundary knots at 0 30
with intercept
```

The cross-basis matrices can be included in the model formula of a common regression function in order to estimate the corresponding parameters $\boldsymbol{\eta}$ in (5). In the example, the final model includes also a natural cubic spline with 7 df/year to model the seasonal and long time trend components and a factor for day of the week, specified by the function `ns()` in the package **splines**, which needs to be loaded in the session. The code is:

```
R> library(splines)
R> model <- glm(death ~ basis.temp + basis.o3 + ns(time, 7 * 14) + dow,
+    family = quasipoisson(), chicagoNMMAPS)
```

4. Predicting from a DLNM

As shown in Section 3, the specification of a DLNM involves a complex parameterization of the exposure series, but the estimation of the parameters $\boldsymbol{\eta}$ is carried out with common regression commands. However, the meaning of such parameters, which define the relationship along two dimensions, is not straightforward. Interpretation can be aided by the prediction of lag-specific effects on a grid of suitable exposure values and the $L + 1$ lags. In addition, the overall effects, predicted from exposure sustained over lags L to 0, can be computed by summing the lag-specific contributions. The algebraic details to derive such estimates have been described elsewhere (see Gasparrini *et al.* 2010, Section 4.3).

Predicted effects are computed in **dlnm** by the function `crosspred()`. The following code computes the prediction for ozone and temperature in the example:

```
R> pred.o3 <- crosspred(basis.o3, model, at = c(0:65, 40.3, 50.3))
R> pred.temp <- crosspred(basis.temp, model, by = 2)
```

The first two arguments passed to `crosspred()` are the object of class '`crossbasis`' and the model object used for estimation. The vector of exposure values for which the effects must be predicted may be directly specified by the argument `at`, as in the first example above. Here I chose the integers from 0 to 65 μgr/m^3 in ozone, plus the value of the chosen threshold and 10 units above (40.3 and 50.3 μgr/m^3, respectively). The values are automatically ordered

and made unique. Alternatively, the vector may be selected through the arguments `by`, `from`, `to`, as in the second example above. In this case I simply chose rounded values within the temperature range with an increment of 2°C. The function `crosspred()` extracts from `model` the parameters (coefficients and (co)variance matrix) corresponding to the cross-basis variables through method functions `coef()` and `vcov()`. For model classes for which such methods are not available, the parameters must be manually extracted and included in the arguments `coef` and `vcov`. The function then calls `crossbasis()` to build a prediction cross-basis and to generate the predicted effects and standard errors given the parameters in `model`.

The result is a list object of class '`crosspred`' which stores the predicted effects. It includes a matrix of lag-specific effects and a vector of overall effects, with corresponding matrix and vector of standard errors. If `model` includes a log or logit link, exponentiated effects and confidence intervals are returned as well. The confidence level of the intervals is defined by the argument `ci.level`, with default `0.95`. The argument `cumul` (default to `FALSE`) adds the matrices of cumulative effects and standard errors along lags.

The results stored in the '`crosspred`' object can be directly accessed to obtain specific figures or customized graphs other than those produced by **dlnm** plotting functions, illustrated in Section 5. For example, the overall effect for the 10-unit increase in ozone, expressed as RR and 95% confidence intervals, can be derived by:

```
R> pred.o3$allRRfit["50.3"]

   50.3
1.05387

R> cbind(pred.o3$allRRlow, pred.o3$allRRhigh)["50.3",]

[1] 1.003354 1.106930
```

See the help page (typing `?crosspred`) for additional information.

5. Representing a DLNM

The bi-dimensional exposure-response relationship estimated by a DLNM may be difficult to summarize. A general description is provided by the graphical representation of the association. The method functions `plot()`, `lines()` and `points()` for class '`crosspred`' offer flexible plotting tools to aid the interpretation of results. The method `plot()` calls high-level functions `plot.default()`, `persp()` and `filled.contour()` to produce scatter plots, 3-D and contour plots of overall and lag-specific effects. These methods allow the user to specify the whole range or arguments of the plotting functions above, providing complete flexibility in the choice of colours, axes, labels and other graphical parameters. Methods `lines()` and `points()` may be used as low-level plotting functions to add lines or points to an existing plot.

For example, the association between ozone and mortality can be summarized by the RR for an increase of 10 μgr/m^3 above the threshold at each lag. This plot, illustrated in Figure 1 (left), is obtained by:

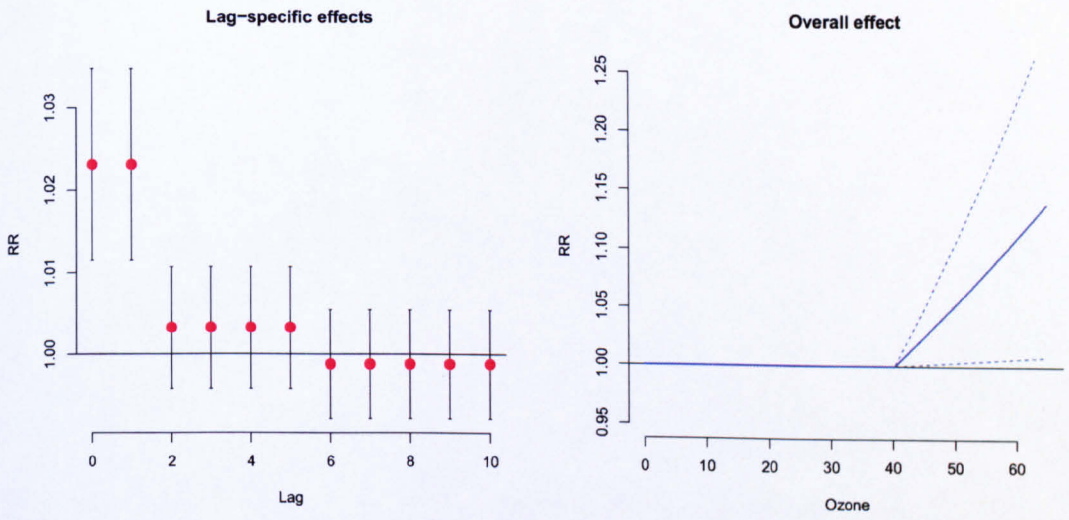


Figure 1: Lag-specific (left) and overall (right) effects on all-cause mortality for a 10-unit increase in ozone above the threshold (40.3 μgr/m^3). Chicago 1987–2000.

```
R> plot(pred.o3, "slices", type = "p", pch = 19, cex = 1.5, var = 50.3,
+     ci = "bars", ylab = "RR", main = "Lag-specific effects")
```

The first argument `x` of the method function `plot()` indicates the object of class '`crosspred`' where the results are stored. The second argument `ptype = "slices"` specifies the type of plot, in this case a *slice* of the matrix of predicted effect along the space of the lag at the predictor value `var=50.3`, corresponding to the 10-unit increase above the threshold set at 40.3 μgr/m^3. The argument `ci` indicates the plot type for confidence intervals. Exponentiated effects are automatically returned for models with log or logit links, or forced by the argument `exp`. Cumulative effects may be plotted with `cumul=TRUE`, if this option has been previously set when generating the prediction with `crosspred()`. Additional parameters are passed to the high-level plotting function (`plot.default()` in this example) to define points, title and the axis labels. See the help of the original high-level functions for additional details and a complete list of the arguments.

Following the conceptual definition described in Section 1.1, the left plot in Figure 1 can be read using two different perspectives: it represents the increase in risk in each $t + \ell$ future day following a single exposure at 50.3 μgr/m^3 in ozone at day t (*forward* interpretation), or otherwise the contributions of each $t - \ell$ past day with ozone at 50.3 μgr/m^3 to the increase in risk at day t (*backward* interpretation).

Alternatively, it is possible to plot the overall effect, computed by summing the lag-specific contributions via the argument `ptype = "overall"`:

```
R> plot(pred.o3, "overall", ci = "lines", ylim = c(0.95, 1.25), lwd = 2,
+     col = 4, xlab = "Ozone", ylab = "RR", main = "Overall effect")
```

The plot is shown in Figure 1 (right). Note the different representation of confidence intervals obtained by the argument `ci`, and non-default colour and line type.

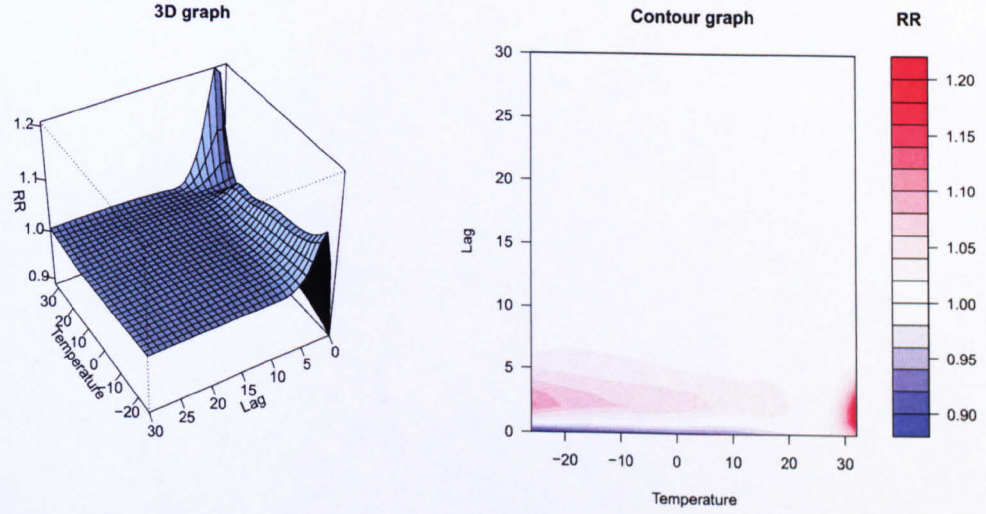


Figure 2: Three-dimensional graphs of the exposure-response relationship between temperature and all-cause mortality, with reference at 25°C. Chicago 1987–2000.

A more detailed approach is instead required to represent the smooth relationship between temperature and mortality, where splines functions have been used to define the dependency in both dimensions. A general description of this complex dependency may be given using 3-D and contour graphs (the default `ptype = "3d"` or `ptype = "contour"`), which illustrates the effect surface given by the whole grid of predicted effects. The graphs, shown in Figure 2, are obtained by:

```
R> plot(pred.temp, xlab = "Temperature", theta = 240, phi = 40,
+    ltheta = -185, zlab = "RR", main = "3D graph")
R> plot(pred.temp, "contour", plot.title = title(xlab = "Temperature",
+    ylab = "Lag", main = "Contour graph"), key.title = title("RR"))
```

The reference point (here 25°C) is the value at which the crossbasis functions have been centered in `crossbasis()`. Arguments `theta`, `phi`, `ltheta` and `plot.title`, `key.title` are used to modify the perspective and lighting in the 3-D plot and the labels in the contour plot, respectively. Other additional parameters may be specified as well (see `?persp` and `?filled.contour`).

Tri-dimensional or contour plots offer a comprehensive summary of the relationship, but are limited in their ability to inform on effects at specific values of predictor or lags. In addition, they are also limited for inferential purposes, as the uncertainty of the estimated effects is not reported. A more comprehensive picture is given by Figure 3, obtained by:

```
R> plot(pred.temp, "slices", var = -20, ci = "n", ylim = c(0.95, 1.22),
+    lwd = 1.5)
R> for(i in 1:2) lines(pred.temp, "slices", var = c(0, 32)[i], col = i + 2,
+    lwd = 1.5)
R> legend("topright", paste("Temperature =", c(-20, 0, 32)), col = 2:4,
```

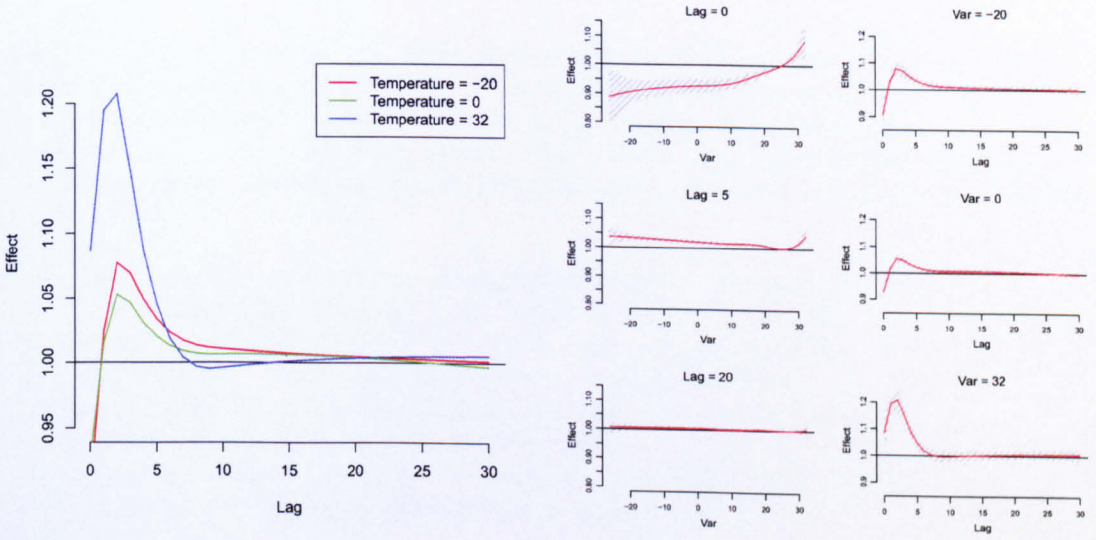


Figure 3: Lag-specific effects at different temperatures (left panel, and right column in right panel) and temperature-specific effects at different lags (left column in right panel) on all-cause mortality, with reference at 25°C. The right panel also shows 99% confidence intervals. Chicago 1987–2000.

```
+     lwd = 1.5)
R> plot(pred.temp, "slices", var = c(-20, 0, 32), lag = c(0, 5, 20),
+     ci.level = 0.99, xlab = "Temperature",
+     ci.arg = list(density = 20, col = grey(0.7)))
```

Figure 3 (left) shows predicted lag-specific effects for temperature values selected by the argument `var` in `plot()` and `lines()`. Alternatively, Figure 3 (right) illustrates a multiple plot of predicted effects along temperature for specific lags (left), and the same lag-specific effect plotted in Figure 3 (right), together with 99% confidence intervals. The arguments `var` and `lag` define the values in the two dimensions, while `ci.level` specifies the confidence level of the intervals. The argument `ci.arg` includes a list of arguments to be passed to low-level plotting functions, which draw confidence intervals. In this case, the default `ci = "area"` calls the function `polygon()`, and the arguments in `ci.arg` are used to select a shading area with increased grey contrast. However, plotting features such as labels and titles may not be included in this automatic multi-plot representation.

These graphs suggest different patterns for the effects of hot and cold temperatures, with a very strong and immediate effect of heat and a more delayed association with cold, negative in the very first lags. This analytical level is not obviously reached with simpler models.

6. Modeling strategies

The DLNM framework offers the opportunity to specify a wide selection of models through the choice of the basis functions for each of the two dimensions of predictor and lags. The example illustrated in the previous sections represents one of the potential modeling alternatives. In

order to discuss the flexibility of the methodology, and the related problems with model selection, a comparison with different models to estimate the association with temperature is shown below. Specifically, polynomial and strata functions are selected for the space of the predictor, while keeping the same natural cubic spline to model the distributed lag curve up to 30 days of lag. The code to specify the cross-basis, run the models and predict the effect is:

```
R> basis.temp2 <- crossbasis(chicagoNMMAPS$temp, vartype = "poly",
+    vardegree = 6, cenvalue = 25, lagdf = 5, maxlag = 30)
R> model2 <- update(model, .~. - basis.temp + basis.temp2)
R> pred.temp2 <- crosspred(basis.temp2, model2, by = 2)
R> basis.temp3 <- crossbasis(chicagoNMMAPS$temp, vartype = "dthr",
+    varknots = 25, lagdf = 5, maxlag = 30)
R> model3 <- update(model, .~. - basis.temp + basis.temp3)
R> pred.temp3 <- crosspred(basis.temp3, model3, by = 2)
```

The first alternative proposes, for the predictor dimension, a polynomial function with the same degrees of freedom as the original cubic spline in Section 5. The second model is based on a simpler double threshold function with a single threshold placed at 25°C, previously identified as the point of minimum mortality. This choice also facilitates the comparison of the models, as this is the centering point for the other two continuous functions. The overall effect estimated by the three models is displayed in Figure 4 (left), produced by the code:

```
R> plot(pred.temp, "overall", ylim = c(0.5, 2.5), ci = "n", lwd = 1.5,
+    main = "Overall effect")
R> lines(pred.temp2, "overall", col = 3, lty = 2, lwd = 2)
R> lines(pred.temp3, "overall", col = 4, lty = 4, lwd = 2)
R> legend("top", c("natural spline", "polynomial", "double threshold"),
+    col = 2:4, lty = c(1:2, 4), lwd = 1.5, inset = 0.1, cex = 0.8)
```

As expected, the alternative models produce different results. In particular, the polynomial model estimates a "wiggly" relationship for cold temperatures, if compared to the original cubic spline with equally-spaced knots. Instead, the two functions provide very close estimates for the effect of hot temperatures. Conversely, while the linearity assumption of the double threshold model seems adequate to model the dependency for cold, there is some evidence that this approach tends to underestimate the effect of heat. A second comparison of the estimated distributed lag curves is illustrated in Figure 4 (right), following:

```
R> plot(pred.temp, "slices", var = 32, ylim = c(0.95, 1.22), ci = "n",
+    lwd = 1.5, main = "Lag-specific effect")
R> lines(pred.temp2, "slices", var = 32, col = 3, lty = 2, lwd = 2)
R> lines(pred.temp3, "slices", var = 32, col = 4, lty = 4, lwd = 2)
R> legend("top", c("natural spline", "polynomial", "double threshold"),
+    col = 2:4, lty = c(1:2, 4), inset = 0.1, cex = 0.8)
```

Although exactly the same function for the space of lag was selected in all the three models, a different choice for the predictor dimension provides different estimates of the distributed lag curve, representing the effect at 32°C compared to the common reference point of 25°C.

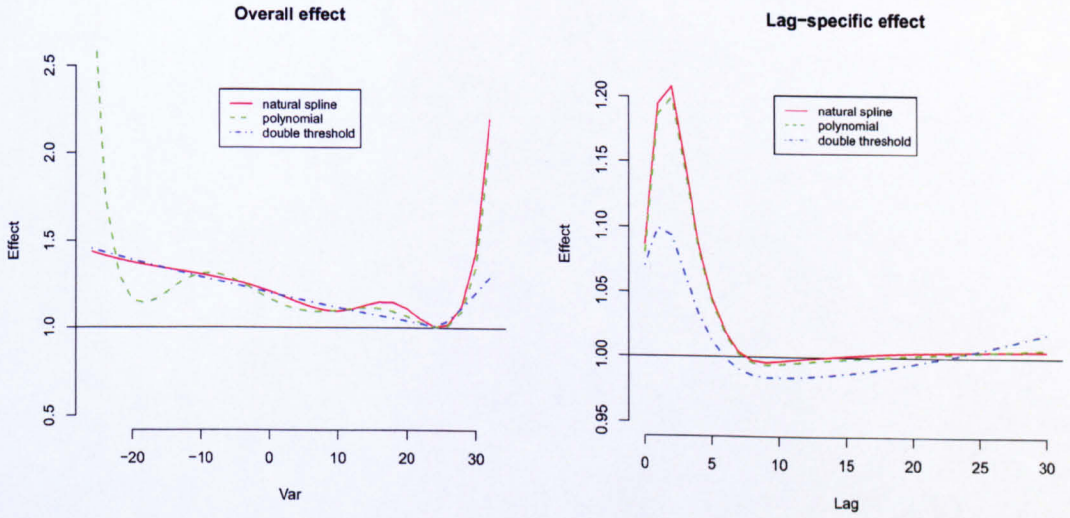


Figure 4: Overall effect (left) and lag-specific effect at 32°C (right) of temperature on all-cause mortality for 3 alternative models, with reference at 25°C. Chicago 1987-2000.

In particular, the spline and polynomial models produce very similar effects (as expected, given the almost identical fit in the other dimension for the hot tail), while the curve for the double threshold models shows quite a different shape. Specifically, the suggestion of an harvesting effect (the negative estimate at longer lags) may represent an artifact due to the lack of flexibility of this model.

Such richness in the specification of different alternatives is tempered by the lack of general criteria to select, among the available choices, the *best* model to summarize the association. In the example above, I showed a clear preference for the spline model. This choice is based both on knowledge of the properties of the function, such as flexibility and stability, and on reasonable arguments given the results plotted in Figure 4. However, this conclusion is questionable, and not grounded on solid and general statistical selection criteria. Moreover, the conclusion is based on several a-priori choices, just like the threshold location or the number of knots or polynomial degree.

Generally, within DLNMs, two different levels of selection may be described. The first one pertains to the specification, in both dimensions, of different functions. As illustrated above, this choice should be based both on the plausibility of the assumed exposure-response shape, and on a compromise between complexity, generalizability and ease of interpretation. The second level focuses on different choices within a specific function, such as the number and location of knots for the definition of a spline basis. The latter is more difficult to address, although not inherent to DLNM development. Several researchers have investigated this issue within time series analysis, proposing methods based on information criteria (Akaike, Bayesian and other variants), partial autocorrelation or (generalized) cross-validation (Peng *et al.* 2006; Baccini *et al.* 2007). The user may apply the same methods within DLNMs, but he should bear in mind that the bi-dimensional nature of these models brings along additional complexities, such as the definition of the maximum lag. Moreover, the evidence on the performance of different criteria is not conclusive, and this represents an issue of current

debate (Dominici *et al.* 2008). Further research is needed to provide some guidance on model selection within DLNMs.

Alternative approaches may be suggested. Muggeo (2008) proposed a model with a constrained segmented parameterization for the space of the predictor, and a doubly penalized spline-based distributed lag parameterization. This methods includes an automatic selection for the threshold(s) and for the smoothness of the distributed lag curve, and it is fully implemented in the R package **modTempEff** (Muggeo 2010). The comparison of such an approach with flexible DLNMs which relax the assumptions on the shape in the dimension of the predictor may provide some additional insights on the relationship.

7. Data requirements

The DLNMs framework introduced in this paper is developed for time series data. The general expression of the basic model in (1) allows this methodology to be applied to any family distribution and link function within (generalized) linear models (GLM), with extensions to generalized additive models (GAM) or models based on generalized estimating equations (GEE). However, the current implementation of DLNMs requires single series of equally-spaced, complete and ordered data.

Each value in the series of transformed variables is computed also using previous observations included in the selected lag period. Therefore, the first `maxlag` observations in the transformed variables are set to `NA`. Missing values in `x` are allowed, but, for the same reason, the same and the next `maxlag` transformed values will be set to `NA`. Although correct, this could generate computational problems for DLNMs with long lag periods in the presence of scattered missing observations. Some imputation methods may be considered in this case.

One of the main advantages of the **dlnm** package is that the user can perform DLNMs with standard regression functions, simply including the cross-basis matrix in the model formula. Its use is straightforward with the functions `lm()`, `glm()` or `gam()` (package **mgcv**, see Wood 2006). However, the user can apply different regression functions, compatibly with the time series structure of the data. These functions should have methods for `coef()` and `vcov()`, or alternatively the user must extract the parameters and include them in the arguments `coef` and `vcov` of `crosspred()` (see Section 4).

8. Future developments

The conceptual framework depicted in Section 1.1 is general, and may be applied to other study designs and data structures other than time series. This idea is hidden by the ordered nature of the time series approach, where each observation is naturally included in a temporal sequence specified by the index t. This represents the unique temporal scale of the study design, and the lag dimension, which lies on the same scale, is automatically defined as $t - \ell$. The temporal structure of different study designs may be more complex, implying multiple time scales. However, the lag dimension can be still expressed through *exposure histories* for each observation, defining an additional temporal scale. This step involves a slightly different definition of the matrix $\mathbf{Q}$ in (3), where each $\mathbf{q}_i$. represents the exposure history for the observation i from the exposure vector $\mathbf{x}$, which does not express anymore a series of observations ordered in time. Interestingly, the conceptual and algebraical process outlined above,

concerning the definition, prediction and representation of DLNMs, still applies. Preliminary tests on the application of the functions included in the package **dlnm** in case-control, cohort and longitudinal data are promising. Further development may lead to a general framework to describe delayed effects, which spans different study designs.

The current implementation of **dlnm** only comprises completely parametric methods to specify the model in (1). A potential alternative is offered by generalized additive models (GAM) based on penalized splines (Wood 2006). Specification and estimation methods for tensor product bases for bivariate smoothing, closely related to the DLNM definition, have been already developed in this framework, and well implemented in the R package **mgcv**. This methodology show clear advantages, primarily the higher flexibility and automatic smoothness selection. Interestingly, the algebraic development of cross-basis described in (5) is still valid, and the actual problem reduces to define suitable penalization methods for the parameters of the cross-basis functions. An extension of DLNMs with penalized splines is currently under development.

9. Final comments

The class of DLNMs represents a unified framework to describe phenomena showing both non-linear and delayed effects. The main advantage of this model family is to unify many of the previous methods to deal with delayed effects in a unique framework, also providing more flexible alternatives regarding the shape of the relationships. The specification of a DLNM involves only the choice of two bases to generate the cross-basis functions in (5), including, for example, linear thresholds, strata, polynomials, and spline transformations.

This flexibility is retained in the implementation of the methodology in the **dlnm** package, which provides functions to specify the model, predict the effects and plot the results. Several different models with an increasing level of complexity can be performed using a simple and general procedure. The example included in this paper illustrates the application of these functions to describe the association between two environmental stressors and mortality, although the framework is easily generalized to other applications. The package includes a thorough documentation of the functions. An overview of its capabilities, together with an update of the last advancements, is provided in the vignette `dlnmOverview` accompanying the implementation.

The separation of cross-basis specification and parameters estimation offers several advantages. First, as illustrated in the example, more than one variable showing delayed effects can be transformed through cross-basis functions and included in the model. Second, standard regression commands can be used for estimation, with the default set of diagnostic tools and related functions. More importantly, this implementation provides an open platform where additional models specified with different regression commands can be implemented, aiding the development of the methodology in other contexts or study designs.

Acknowledgments

Distributed lag non-linear models were originally conceived and applied by Armstrong (2006), who also provided useful comments on this contribution. The author is grateful to the two anonymous reviewers for their interesting and constructive reviews, which noticeably contributed to improve the manuscript.

References

Almon S (1965). "The Distributed Lag between Capital Appropriations and Expenditures." *Econometrica*, **33**, 178–196.

Armstrong B (2006). "Models for the Relationship Between Ambient Temperature and Daily Mortality." *Epidemiology*, **17**(6), 624–31.

Baccini M, Biggeri A, Lagazio C, Lertxundi A, Saez M (2007). "Parametric and Semi-Parametric Approaches in the Analysis of Short-Term Effects of Air Pollution on Health." *Computational Statistics and Data Analysis*, **51**(9), 4324–4336.

Braga AL, Zanobetti A, Schwartz J (2001). "The Time Course of Weather-Related Deaths." *Epidemiology*, **12**(6), 662–7.

Daniels MJ, Dominici F, Samet JM, Zeger SL (2000). "Estimating Particulate Matter-Mortality Dose-Response Curves and Threshold Levels: An Analysis of Daily Time-Series for the 20 Largest US Cities." *American Journal of Epidemiology*, **152**(5), 397.

Dobson AJ, Barnett AG (2008). *An Introduction to Generalized Linear Models*. 3rd edition. CRC Press/Chapman & Hall.

Dominici F (2004). "Time-Series Analysis of Air Pollution and Mortality: A Statistical Review." *Research Report 123*, Health Effects Institute.

Dominici F, McDermott A, Hastie TJ (2004). "Improved Semiparametric Time Series Models of Air Pollution and Mortality." *Journal of the American Statistical Association*, **99**(468), 938–949.

Dominici F, Wang C, Crainiceanu C, Parmigiani G (2008). "Model Selection and Health Effect Estimation in Environmental Epidemiology." *Epidemiology*, **19**(4), 558–60.

Gasparrini A, Armstrong B (2010). ***dlnm**: Distributed Lag Non-Linear Models*. R package version 1.4.1, URL `http://CRAN.R-project.org/package=dlnm`.

Gasparrini A, Armstrong B, Kenward MG (2010). "Distributed Lag Non-Linear Models." *Statistics in Medicine*, **29**(21), 2224–2234.

Goodman PG, Dockery DW, Clancy L (2004). "Cause-Specific Mortality and the Extended Effects of Particulate Pollution and Temperature Exposure." *Environmental Health Perspectives*, **112**(2), 179–85.

Hajat S, Armstrong BG, Gouveia N, Wilkinson P (2005). "Mortality Displacement of Heat-Related Deaths: A Comparison of Delhi, Sao Paulo, and London." *Epidemiology*, **16**(5), 613–20.

Muggeo VM (2008). "Modeling Temperature Effects on Mortality: Multiple Segmented Relationships with Common Break Points." *Biostatistics*, **9**(4), 613–620.

Muggeo VM, Hajat S (2009). "Modelling the Nonlinear Multiple-Lag Effects of Ambient Temperature on Mortality in Santiago and Palermo: A Constrained Segmented Distributed Lag Approach." *Occupational Environmental Medicine*, **66**(9), 584.

Muggeo VMR (2010). "Analyzing Temperature Effects on Mortality within the R Environment: The Constrained Segmented Distributed Lag Parameterization." *Journal of Statistical Software*, **32**(12), 1–17. URL http://www.jstatsoft.org/v32/i12/.

Pattenden S, Nikiforov B, Armstrong BG (2003). "Mortality and Temperature in Sofia and London." *Journal of Epidemiology and Community Health*, **57**(8), 628–33.

Peng RD, Dominici F, Louis TA (2006). "Model Choice in Time Series Studies of Air Pollution and Mortality." *Journal of the Royal Statistical Society A*, **169**(2), 179–203.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Roberts S, Martin MA (2007). "A Distributed Lag Approach to Fitting Non-Linear Dose-Response Models in Particulate Matter Air Pollution Time Series Investigations." *Environmental Research*, **104**(2), 193–200.

Samoli E, Zanobetti A, Schwartz J, Atkinson R, Le Tertre A, Schindler C, Perez L, Cadum E, Pekkanen J, Paldy A, Touloumi G, Katsouyanni K (2009). "The Temporal Pattern of Mortality Responses to Ambient Ozone in the APHEA Project." *Journal of Epidemiology and Community Health*, **63**, 960–966.

Schwartz J (2000). "The Distributed Lag between Air Pollution and Daily Deaths." *Epidemiology*, **11**(3), 320–6.

Touloumi G, Atkinson R, Le Tertre A, Samoli E, Schwartz J, Schindler C, Vonk JM, Rossi G, Saez M, Rabszenko D (2004). "Analysis of Health Outcome Time Series Data in Epidemiological Studies." *EnvironMetrics*, **15**(2), 101–117.

Wood SN (2006). *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC.

Zanobetti A, Schwartz J (2008). "Mortality Displacement in the Association of Ozone with Mortality: An Analysis of 48 Cities in the United States." *American Journal of Respiratory and Critical Care Medicine*, **177**(2), 184–9.

Zanobetti A, Wand MP, Schwartz J, Ryan LM (2000). "Generalized Additive Distributed Lag Models: Quantifying Mortality Displacement." *Biostatistics*, **1**(3), 279–92.

Affiliation:

Antonio Gasparrini
Department of Social and Environmental Health Research

London School of Hygiene and Tropical Medicine
15-17 Tavistock Place, London WC1H 9SH, United Kingdom
E-mail: antonio.gasparrini@lshtm.ac.uk
URL: http://www.lshtm.ac.uk/people/gasparrini.antonio/

Journal of Statistical Software
published by the American Statistical Association
Volume 43, Issue 8
July 2011

http://www.jstatsoft.org/
http://www.amstat.org/
Submitted: 2010-06-03
Accepted: 2011-07-13

Chapter 7

R package vignette

Title: Distributed lag non-linear models in R: the package dlnm.

Author(s): Antonio Gasparrini.

Journal/Publisher: R Foundation for Statistical Computing.

Type of publication: Package documentation.

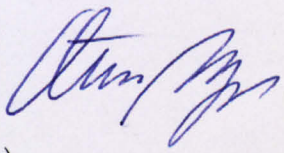
Stage of publication: Published online on 20 July 2011.

URL: `http://cran.r-project.org/web/packages/dlnm/index.html`.

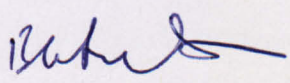
Academic peer-reviewed: No.

Copyright: Retained by the author.

Candidate's role: See Section 2.3.

Candidate's signature: 

Co-supervisor: (Prof. Ben Armstrong)



Distributed lag non-linear models in R: the package dlnm

Antonio Gasparrini and Ben Armstrong
London School of Hygiene and Tropical Medicine, UK

dlnm version 1.4.0 , 2011-07-20

Contents

[1]This document is included as a vignette (a LaTeX document created using the R function `Sweave()`) of the package dlnm. It is automatically dowloaded together with the package and can be accessed through R typing *vignette("dlnmOverview")*.

1 Preamble

The R package dlnm offers some facilities to run *distributed lag non-linear models* (DLNM's), a modelling framework to describe simultaneously non-linear and delayed effects between predictors and an outcome in time-series data. This document complements the description provided in Gasparrini (2011) (freely available at `http://www.jstatsoft.org/v43/i08/`), which represents the main reference to the package.

The aim of this contribution is to provide an extended overview of the capabilities of the package, together with additional examples of application with real data. Some information on installation procedures and on the data included in the package are given in Section 2. The theory underlying the DLNM methodology is briefly illustrated in Section 3, while the functions included in the package are described in Section 4. Some examples of applications are provided in Section 5: users mainly interested in the application can skip the previous Sections and and start with these examples. Finally, Section 6 offers some conclusions.

The DLNM's methodology has been previously described in Gasparrini et al. (2010), together with a detailed algebraical development. This framework was originally conceived and proposed to investigate the health effect of temperature by Armstrong (2006).

Type `citation("dlnm")` in R to cite the dlnm package after installation (see Section 2). A list of changes included in the current and previous versions can be found typing `file.show(system.file("ChangeLog", package = "dlnm"))`.

Please send comments or suggestions and report bugs to `antonio.gasparrini@lshtm.ac.uk`.

2 Installation and data

2.1 Installing the package dlnm

The dlnm package is installed in the standard way for CRAN packages from version 2.9.0 onwards, for example typing `install.packages("dlnm")` or directly through the menu in R, clicking on *Packages* and then on *Install package(s)....* The package can be alternatively installed using the .zip file containing the binaries, via *Packages* and then *Install package(s) from local zip files....*

The functionalities of dlnm depend on other packages whose commands are called to specify the dlnm functions. This hierarchy is ruled by the field *Imports* of the file `description` included in the package. The functions are imported from the packages splines (functions `ns()` and `bs()`) and tsModel (function `Lag()`). The former must be independently installed if a .zip file is used.

2.2 Data

Until the version 0.4.1, the package dlnm did not contain any data, and used the datasets stored in the package NMMAPSlite.

In this version the package contains its own dataset `chicagoNMMAPS`, with daily mortality (all causes, CVD, respiratory), weather (temperature, dew point temperature, relative humidity) and pollution data (PM10 and ozone) for Chicago in the period 1987-2000. The data were assembled from publicly available data sources as part of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) sponsored by the Health Effects Institute (Samet et al., 2000a,b). They are downloadable from the Internet-based Health and Air Pollution Surveillance System (iHAPSS) website (`http://www.ihapss.jhsph.edu`) or through the packages NMMAPSdata or NMMAPSlite. See `?chicagoNMMAPS` for additional information on the variables included.

3 Distributed lag non-linear models (DLNM's)

The aim of this Section is to provide a methodological summary of the DLNM framework. A detailed description of this methodology and the algebraical development have been published elsewhere (Armstrong, 2006; Gasparrini, 2011; Gasparrini et al., 2010).

3.1 The issue

The main purpose of a statistical regression model is to define the relationship between a predictor and an outcome, and then to estimate the related effect. A further complexity arises when the dependency shows some *delayed effects*: in this case, a specific occurrence of the predictor (let us call it an *exposure event*) affects the outcome for a certain period in the future. This step requires the definition of more complex models to characterize the association, specifying the temporal structure of the dependency. The main feature of DLNM's is their *bi-dimensional* structure: the model describes simultaneously the potentially non-linear relationship in the space of the predictor and along the new temporal dimension.

3.2 The concept of basis

Several different methods have been adopted to specify non-linear effects in a regression models. A simple solution is to generate strata variables, applying specific cut-off points along the range of the predictor in order to define specific intervals, and then specifying new variables through a dummy parameterization.

Other types of manipulations of the original variable are applied when there are specific assumptions on the shape of the relationship, for example when the effect is likely to exist and be linear only above or below a specific threshold (*hockey-stick* model). An extension of this model assumes two distinct linear effects below a first threshold and above a second threshold, with a null effect in between them.

An alternative to the strata or threshold approaches is to include in the model some terms allowing a true non-linear relationship, describing a smooth curve between the predictor and the outcome. The traditional methods include a quadratic term or higher degree polynomials. Recently, spline functions have been favoured, especially through a natural cubic parameterization.

A generalization may be provided assuming that all the approaches above imply the choice of a *basis*, defined as a *space of functions* used to define the relationship (Wood, 2006). The choice of the basis defines the related *basis functions*, completely known transformations of the original predictor generating a new set of transformed variables, defined *basis variables*. Independently from the basis chosen, the final result will be a matrix of transformed variables which can be included in the design matrix of a model in order to estimate the related parameters. The choice of different bases leads to the specification of different matrices, but the mechanism is common.

3.3 Delayed effect: DLM's

In the specific context of time series analysis, given the ordered series of the predictor values, a delayed (or lagged) effect is present when the outcome in a specific time is influenced by the level of the predictor in previous times, up to a maximum lag. Therefore, the presence of delayed effects requires to take into account the *time dimension* of the relationship, specifying the additional virtual dimension of the *lags*.

A very simple model to deal with delayed effects considers the moving average of the predictor up to a certain lag, specifying a transformed predictor which is the average of the values in that specific lag

period. Although simple, this model is limited if the purpose is to assess the temporal structure of the effects.

These limitations have been addressed using a more elegant approach based on distributed lag models (DLM's). The main advantage of this method is the possibility to depict a detailed description of the time-course of the relationship. Originally developed in econometrics (Almon, 1965), this method has recently been used to quantify the health effect in studies on environmental factors (Braga et al., 2001; Schwartz, 2001; Welty and Zeger, 2005; Zanobetti et al., 2000).

In the basic formulation, a DLM is fitted by the inclusion of a parameter for each lagged predictor occurrence. An estimate of the overall effect is given by the sum of the single lag effects upon the whole lag period considered (Hajat et al., 2005; Schwartz, 2000).

This *unconstrained* version of DLM does not require any assumption on the shape of the effect along lags, and consequently on the relationship between parameters. In order to define a more parsimonious model, it is possible to specify some assumptions on the shape of the distributed effect, applying some constraint. The simplest solution is to group the lags in different strata (Pattenden et al., 2003; Welty and Zeger, 2005), while a more complex option is to force the curve along lags to follow a specific smooth function, for example polynomials (Baccini et al., 2008; Schwartz et al., 2004; Zanobetti and Schwartz, 2008) or splines (Zanobetti et al., 2000).

Following the general approach used in Section 3.2, it may be shown that all the different DLM's above can be described by the same equation, where different models are specified through different basis functions to be applied to the vector of lags, building a new basis matrix (see Gasparrini et al., 2010, Eq. 4). Again, the choice of different bases generates different matrices, but the mechanism is general.

3.4 The extension to DLNM's

A general approach to specify non-linear but un-lagged effects has been introduced in Section 3.2, while the methods to define distributed lag functions for simple linear effects have been presented in Section 3.3. An obvious extensions is to combine these approaches to define distributed lag non-linear models (DLNM's), a family of models which can deal at the same time with non-linear and delayed effects.

The different issues of non-linearity and delayed effects share a common feature: in both cases the solution is to choose a basis to describe the shape of the relationship in the relative dimension. This step leads to the concept of *cross-basis*: following the idea of basis in 3.2, a cross-basis can be imagined as a bi-dimensional space of functions describing on the same time the shape of the relationship and the distributed lag effects. The algebraic notation to define the cross-basis and then the DLNM can be quite complex, involving tensor products of 3-dimensional arrays, and has been presented elsewhere (Gasparrini et al., 2010, Section 4.2). Nonetheless, the basic concept is straightforward: choosing a cross-basis amounts to choosing two independent set of basis functions, which will be combined to generate the specific cross-basis functions. The DLM's described in 3.3 can be considered as special cases of DLNM's with a simple linear function in the dimension of the predictor.

The result of a DLNM can be interpreted building a grid of predictions for each lag and for suitable values of the predictor, using three dimensional plots to provide an overall picture of the effects varying along the two dimensions. In addition, it is possible to compute the effects for single predictor levels or lags, simply cutting a "slice" of the grid along specific values of predictor or lags, respectively. Finally, an estimate of the overall effect can be computed by summing all the contributions at different lags. The effects are usually reported versus a reference value of the predictor, centering the basis functions for this space to their corresponding transformed values (Cao et al., 2006).

The choice of the two set of basis functions for each space is perfectly independent, and should be based on a-priori assumptions or on a compromise between complexity and generalizability. Linear,

threshold, strata, polynomial or splines functions can be used to define the relationship along the space of predictor, while unconstrained, strata, polynomial or splines functions can be applied to specify the shape along lags.

4 The functions in the package dlnm

This section describes the main functions included in the package dlnm. Here we provide a description of all the stages involved in the definition, estimation and interpretation of DLNMs, summarizing the conceptual and analytical steps. In addition, we illustrate the structure of the functions and discuss specific issues about their usage. Examples of applications to real time series data are described in Section 5. Additional information is provided in Gasparrini (2011).

4.1 Internal functions: `mkbasis()` and `mklagbasis()`

These functions build the basis matrices for the dimension of the predictor and lags, respectively. In concrete terms, they apply a transformation to the vector of predictor and to the vector of lags, and store the transformed variables and information about the chosen basis in list objects. These functions are called by `crossbasis()` (see Section 4.2) and are not expected to be directly run by the user in order to specify DLNMs. Their first arguments are `x` and `maxlag`, respectively, representing the original predictor and the maximum lag. The latter is used by `mklagbasis()` to generate the lag vector `0:maxlag`.

Different types of basis may be chosen through the argument `type`: the possible options are natural cubic or simple B-splines (`type="ns"` or `"bs"`), strata through dummy variables (`"strata"`), polynomials (`"poly"`), threshold-type functions such as low, high or double threshold or piecewise parameterization (`"lthr"-"hthr"-"dthr"`), strata variables for each integer values (`"integer"`, used in unconstrained DLMs) and simply linear (`"lin"`).

The argument `"df"` defines the dimension of the basis (the number of its columns, basically the number of transformed variables), which, in completely parametric models, corresponds to the number of degrees of freedom spent to define the relationship in the regression model including the basis. This value may depend on the argument `knots` (which overcomes `df`), specifying the position of the internal knots for `"ns"` and `"bs"` (with boundary knots specified in `bound`), the cut-off points for `"strata"` (defining right-open intervals) and the thresholds/cut-off points for `"lthr"`, `"hthr"` and `"dthr"`. The argument `degree` select the degree of polynomial for `"bs"` and `"poly"`.

The arguments `cen` and `cenvalue` state if the basis must be centered and the centering value to be used. The presence of an intercept in the basis matrix is determined by the argument `int`. Actually, the concept of intercept is different between bases: types `"ns"` and `"bs"` apply a complex parameterization where the intercept is implicitly built within the basis variables (see the related help pages typing `?ns` and `?bs`); in type `"strata"` the intercept corresponds to the dummy variable for the baseline stratum (the first one by default), which is excluded if `int=F`; the intercept is the usual vector of 1's in the other types. See Section 4.2 for additional information.

The value returned by `mkbasis()` and `mklagbasis()` is a list object, whose first component `basis` is the matrix created by the application of the chosen basis functions to `x` or `0:maxlag`, respectively. Additional values corresponding to the arguments above are returned in the other components of the list object.

4.2 The function `crossbasis()`

This is the main function in the package **dlnm**. It calls the internal functions `mkbasis()` and `mklagbasis()` and combines the two basis matrices through a tensor product in order to create the cross-basis, which specifies the dependency simultaneously in the two dimensions. See Gasparrini et al. (2010, Section 4.1 - 4.2) for details. Its first argument is `x`, assumed to represent an equally-spaced, complete and ordered series of observations, in order for the function to be coherently applied.

The function uses arguments `df-knots-bound-degree-int-cen-cenvalue-maxlag`, with specific (optional) prefix `var-` or `lag-` to pass them to `mkbasis()` or `mklagbasis()`, respectively (see Section 4.1, and type `?crossbasis` for a complete list of the arguments). The additional argument `group` defines groups of observations to be considered as individual unrelated series, and may be useful for example in seasonal analyses (see Section 5.3). In this case, each series must be consecutive, complete and ordered.

The function returns an object of class *"crossbasis"*, together with attributes defining the choices for the two basis functions. The arguments are set to some default values, and can be automatically changed for nonsensical combinations, or set to null if not required. Meaningless combinations of arguments (for example knots defined outside the predictor range) could lead to collinear variables, with identifiability problems in the model. The function applies some coherence checks and fix some specific problem (for example discarding strata intervals where no observation lies), but other problem may arise. The user is advised to test the result with the function `summary.crossbasis()`, which provides a summary of the choices made for the two bases and the final cross-basis.

The values in `x` are expected to be equally-spaced (with the interval defining the lag unit) and ordered in time. The series must be complete. Each value in the series of transformed variables is computed also using previous observations included in the lag period considered: therefore, the first `maxlag` observations in the transformed variables are set to `NA`. Missing values in `x` are allowed, but, for the same reason, the same and the next `maxlag` transformed values will be set to `NA`. Although correct, this could generate computational problems for DLNMs with long lag periods in the presence of scattered missing observations.

The basis variables for the space of the predictor are centered by default for continuous functions (types `"ns"`, `"bs"`, `"poly"` and `"lin"`). The default centering point is the predictor mean, if not set with `cenvalue`. This value will represent the reference for predicted effects from a DLNM (see Section 4.3). The choice of the reference value does not affect the fit of the model, and should be based on interpretational issues. The reference in non-continuous functions is automatically set to the first interval in `strata` and `integer`, or to the flat region in `lthr-hthr-dthr`.

An intercept is included by default only in the basis defining the lag dimension. It is strongly recommended to avoid the inclusion of an intercept in the basis for `x`, otherwise a rank-deficient cross-basis matrix will be specified, causing some of the cross-variables to be excluded in the regression model.

4.3 The function `crosspred()`

The cross-basis matrix produced by `crossbasis()` need to be included in a regression model formula in order to run a DLNM. The interpretation of the estimated related parameters, specifying a bi-dimensional relationship, is virtually impossible in complex DLNMs. The association is summarized through the function `crosspred()`, which predicts the effects for a set of values of the original predictor, and return the results for each combination of predictor values and lags. The function creates the same cross-basis functions for the chosen predictor values, based on the attributes of the original cross-basis matrix, and generates estimated effects and standard errors by extracting the related parameters estimated in the model (see Gasparrini et al. (2010, Section 4.3) for details).

The first two arguments of the function are `basis` (the matrix object of class *"crossbasis"*) and `model` (the regression model object which includes `basis`). The function extracts the information about the cross-basis from the attributes of the former, and links each cross-basis variables with the estimated parameters in the latter through their names. Multiple cross-basis matrices associated with different predictors may be included in `model`: in this case, the user must specify different names for the cross-basis objects.

One of the main advantages of the dlnm package is that the user can perform DLNMs with standard regression functions, simply including the cross-basis matrix in the model formula. The current implementation only works with time series data, basically involving an equally-spaced and ordered predictor series, and its use is straightforward with the functions `lm()`, `glm()` or `gam()` (package mgcv). However, the user can apply different regression functions, compatibly with the time series structure of the data. Alternative use beyond time series analysis, such as in case-control or cohort designs, is in development. The function `crosspred()` exploits `coef()` and `vcov()` methods to extract the coefficients and related (co)variance matrix from `model`, respectively: for classes of regression functions without these methods, the user needs to manually extract the parameters and include them in the arguments `coef` and `vcov`. In this case, their dimensions and order must match the variables included in `basis`.

The predictor values used for prediction are selected with the argument `at`, or alternatively with `from-to-by`. If specified by `at`, the values are automatically ordered and made unique. If `at` and `by` are not provided, approximately 50 equally-spaced rounded values are returned using `pretty()`.

The function returns an object of class *"crosspred"*, simply a list of components including the vector of prediction values, matrices of lag-specific effects and standard errors for combinations of each prediction value and lag, plus vectors of overall effects (summed up along lags) and standard errors. Matrices of cumulative effects and standard errors are included for `cumul=T` (default to `FALSE`), which represent the sum of the lag-specific effects at each lag. Exponentiated effects are added if the link of the regression model is equal to log or logit, together with confidence intervals computed using a normal approximation and a confidence level selected by `ci.level`. The model link is automatically selected from model for classes `"lm"`, `"glm"`, `"gam"` (package mgcv) and `"clogit"` and `"coxph"` (package survival), but needs to be provided for different classes or if arguments `coef-vcov` are used to input the parameters.

4.4 Plotting functions

Interpretation of the bi-dimensional predicted effects are aided by graphical representation. High and low-level plotting functions are provided through the methods `plot()`, `lines()` and `points()`. The method `plot()` calls high-level functions `plot.default()`, `persp()` and `filled.contour()` to produce scatter plots, 3-D and contour plots of overall and lag-specific effects. These methods have replaced the old function `crossplot()` since version 1.3.0, providing the user to specify the whole range or arguments of the plotting functions above, allowing complete flexibility in the choices of colours, axes, labels and other graphical parameters. See the help of the original high-level functions for additional details and a complete list of the arguments. Methods `lines()` and `points()` may be used as low-level plotting functions to add lines or points to an existing plot.

The first argument of the functions is `x`, a list object of class *"crosspred"*. The argument `ptype` specifies the type of plot, choosing among `"3d"`, `"contour"`, `"overall"` and `"slices"`, the latter selecting effects along lags at specific predictor values and effects along the predictor at specific lags. These are chosen through the additional arguments `var-lag`, respectively. Cumulative effects along lags are reported if `cumul=TRUE`: in this case, the same option must have been set to obtain the prediction saved in `x` (see Section 4.3). Confidence intervals are optionally plotted for `"overall"` and `"slices"`. The type is chosen by the argument `ci` among `"area"`, `"bars"` and `"lines"`. Low-level plotting functions

`polygon()`, `segments()` and `lines()` are called, respectively, whose arguments are passed by a list specified with the argument `ci.arg`. See the help of these low-level functions for additional details and a complete list of the arguments.

All the effects are reported versus a reference value. For continuous functions, this is specified by the centering point defined in the `crossbasis` object (see Section 4.2). Exponentiated effects are automatically returned if the component `model.link` of `x` is equal to log or logit, or forced with the argument `exp=TRUE`.

5 Some examples

This Section provides some examples of the use of the functions included in the dlnm package, described in Section 4. In spite of the specific application on the health effects of air pollution and temperature, these examples are easily generalized to different topics. The results included in this Section are not meant to represent scientific findings, but are reported with the only purpose to illustrate the capabilities of the dlnm package.

First, some simple examples of the internal functions are showed in Section 5.1. Then, 3 different examples of the application of DLNM's are illustrated in the Sections 5.2 - 5.4, using the NMMAPS dataset for the city of Chicago in the period 1987-2000 included in the package, which has been described in Section 2.2. These different cases cover most of the functionalities of the package, providing a detailed overview of its capabilities and a basis to perform analyses on this dataset or on other data sources.

The package is assumed to be present in the R library (see Section 2.1) and loaded in the session, typing:

```
> library(dlnm)
```

5.1 Examples for internal functions

As a first step, we provide an example of the use of the function `mkbasis()`. We build different basis matrices applying the selected basis functions to the vector of integers going form 1 to 5. In the first example we leave many of the arguments at their default values, apart from the selection of the degrees of freedom `df`:

```
> basis.var <- mkbasis(1:5, knots=3)
> basis.var

$basis
             b1          b2
[1,] -0.56626284  0.21084190
[2,] -0.20921622 -0.00635585
[3,]  0.00000000  0.00000000
[4,] -0.03716777  0.37894518
[5,] -0.22216593  0.98144395

$type
[1] "ns"

$df
```

```
[1] 2

$knots
[1] 3

$bound
[1] 1 5

$int
[1] FALSE

$cen
[1] TRUE

$cenvalue
[1] 3
```

The result is list object with the basis matrix and other components returning the chosen arguments. Here the basis is a natural cubic B-splines (default `type="ns"`) with 1 knot and `df=2` (`df` is equal to `length(knots)+1+int` for `type="ns"`). Apart from the fact that the basis variables are centered at `cenvalue=3` (the mean of the predictor values, the default for this argument), the same results could be obtained by the command *`ns(1:5, knots=3)`*.

Alternative choices may be specified through the following code (results not shown, the user can try to run the commands):

```
> mkbasis(1:5, type="bs", df=4, degree=2)
> mkbasis(1:5, type="lin", cenvalue=4)
```

In the first case the result is a quadratic spline where the number and location of `knots` are chose automatically, and fixed to 2 (`df` is `length(knots)+degree+int` for this `type`) and at equally spaced quantiles, respectively. The second line returns a simple linear function, where the only transformation is the centering at the value of 4.

The function `mklagbasis()` calls `mkbasis()` to create a basis matrix for the space of the lag. The basis functions are applied to the vector `0:maxlag` expressly created by the function. This is an example of application:

```
> mklagbasis(maxlag=5, type="poly", degree=3)

$basis
     b1 b2 b3  b4
lag0  1  0  0   0
lag1  1  1  1   1
lag2  1  2  4   8
lag3  1  3  9  27
lag4  1  4 16  64
lag5  1  5 25 125

$type
[1] "poly"
```

```
$df
[1] 4

$degree
[1] 3

$int
[1] TRUE

$maxlag
[1] 5
```

The statement specifies a 3rd degree polynomial. Differently from the bases for the space of the predictor build above, this matrix contains an intercept (`int=TRUE` by default), in this case a vector of 1's (see Section 4.2), and is never centered. `df` is equal to `degree+1` when an intercept is included. In this case, for a polynomial basis, the argument `knots` is not included.

Other examples (results not shown):

```
> mklagbasis(maxlag=5, type="integer")
> mkbasis(1:5, type="dthr", knots=c(2,3))
```

In the first line, the function applies a specific transformation in the space of lags in order to define unconstrained distributed lag effects (see Section 3.3), simply returning an identity matrix. The second choice returns a double threshold basis which can be applied to describe linear effects below 2 and above 3, with a null effect in between them.

A basis matrix of `type="strata"` with and without intercept is created by (results not shown):

```
> mklagbasis(maxlag=10, type="strata", knots=c(4,7))
> mklagbasis(maxlag=10, type="strata", knots=c(4,7), int=F)
```

In this case, the intercept is represented by the dummy variable for the first stratum (see Section 4.2). The values in `knots` specify the cut-off point for the strata, and represent the lower boundaries for the right-open intervals.

The effect of centering is illustrated below (results not shown):

```
> mkbasis(0:10, type="poly", degree=3)
> mkbasis(0:10, type="poly", degree=3, cen=F)
```

Each basis function is centered on the relative transformation of `cenvalue`, which is placed at the mean of the predictor values by default, or defined by the user.

5.2 Example 1: a simple DLM

In this first example, we specify a simple DLM, assessing the effect of PM_{10} on overall mortality, while adjusting for the effect of temperature. In order to do so, we first build two cross-basis matrices for the two predictors, and then include them in a model formula of a regression function. The effect of PM_{10} is assumed linear in the dimension of the predictor, so, from this point of view, we can define this as a simple DLM even if it estimates also the distributed lag function for temperature, which is included

as a non-linear term. As highlighted above, the data are assumed to be composed by equally-spaced, complete and ordered series.

First, we run `crossbasis()` to build the two cross-basis matrices, saving them in two objects. The names of the two objects must be different in order to predict the effects separately for each of them (see Section 4.3). This is the code:

```
> basis.pm <- crossbasis(chicagoNMMAPS$pm10, vartype="lin", lagtype="poly",
          lagdegree=4, cen=F, maxlag=15)
> basis.temp <- crossbasis(chicagoNMMAPS$temp, vardf=5, lagtype="strata",
          lagknots=1, cenvalue=21, maxlag=3)
```

In this case, we assume that the effect of PM_{10} is linear (`vartype="lin"`), while we model the relationship with temperature through a natural cubic spline with 5 degrees of freedom (`vartype="ns"`, chosen by default). In this space, the internal knots (if not provided) are placed by default at equally spaced quantiles, while the boundary knots are located at the range of the observed values, so we need to specify only `vardf`. We did not center PM_{10}, in order to compute the predicted effects versus a reference value of 0 $\mu gr/m^3$ (the same results could be obtained setting `cen=TRUE` and `cenvalue=0`). The reference value for temperature is set to 21°C.

The basis for the space of the lags is chosen through the same arguments but with prefix `lag-`. We specify the lagged effect of PM_{10} up to 15 days of lag with a 4^{th} degree polynomial function (setting `lagdegree=4`). The delayed effect of temperature are defined by two lag strata (0 and 1-3), assuming the effects as constant within each stratum. The argument `varknots=1` defines the lower boundary of the second interval.

An overview of the specifications for the cross-basis (and the related bases in the two dimensions) is provided by the function `summary.crossbasis`, which calls the attributes of the `crossbasis` object:

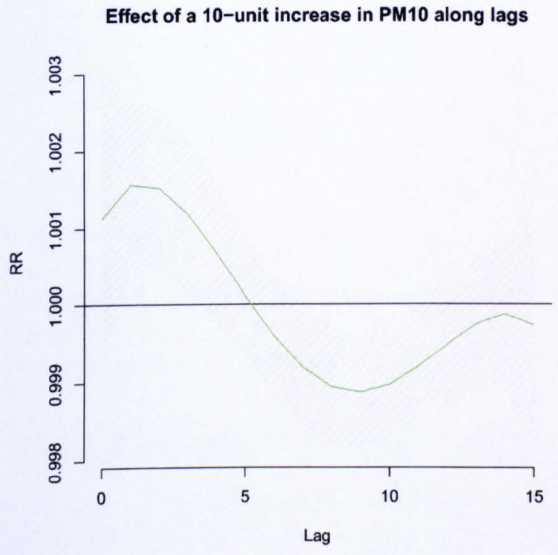
```
> summary(basis.pm)

CROSSBASIS FUNCTIONS
observations: 5114
range: -3.049835 , 356.1768
total df: 5
maxlag: 15

BASIS FOR VAR:
type: lin
df: 1

BASIS FOR LAG:
type: poly with degree 4
df: 5
with intercept
```

Now the two `crossbasis` objects can be included in a model formula in order to fit the DLM. The packages `splines` is loaded, as it is needed in the examples. In this case we model the effect assuming an overdispersed Poisson distribution, including a smooth function of time with 7 df/year (in order to correct for seasonality and long time trend) and day of the week as factor:

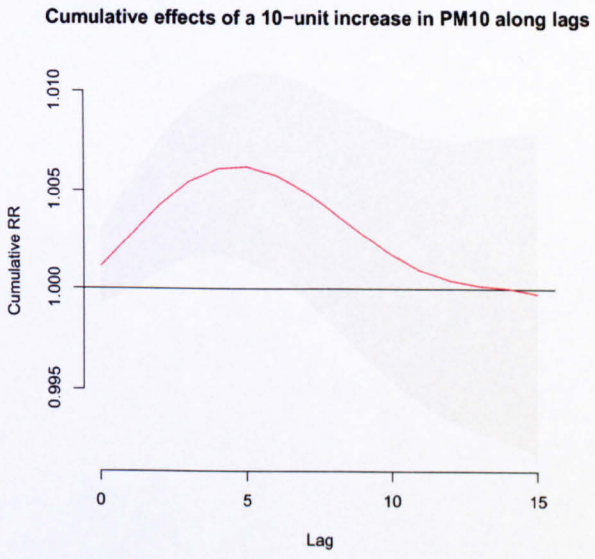
```
> library(splines)
> model <- glm(death ~ basis.pm + basis.temp + ns(time, 7*14) + dow,
          family=quasipoisson(), chicagoNMMAPS)
```

Figure 1

(a)



(b)



The effects of specific levels of PM_{10} on overall mortality, predicted by the model above, can be computed by the function `crosspred()` and saved in an object with the same class:

```
> pred.pm <- crosspred(basis.pm, model, at=0:20, cumul=T)
```

The functions include the `basis.pm` and `model` objects used to estimate the parameters as the first two arguments, while `at=0:20` states that the prediction must be computed for each integer value from 0 to 20 μgr/m^3. The argument `cumul` (default to `FALSE`) indicates that also cumulative effects along lags must be included. Now that the predicted effects have been stored in `pred.pm`, they can be plot by the methods functions described in Section 4.4. For example:

```
> plot(pred.pm, "slices", var=10, col=3, ylab="RR", ci.arg=list(density=15,lwd=2),
    main="Effect of a 10-unit increase in PM10 along lags")
> plot(pred.pm, "slices", var=10, cumul=TRUE, ylab="Cumulative RR",
          main="Cumulative effect of a 10-unit increase in PM10 along lags")
```

The function includes the `pred.pm` object with the stored results, and the argument `"slices"` defines that we want to graph the relationship at specific values of the two dimensions (predictor and lag). With `var=10` we specify this relationship along lags for a specific value of PM_{10}, i.e. 10 μgr/m^3. This effect is compared to the reference value of 0 μgr/m^3, giving the lag-specific effects for a 10-unit increase. We also chose a different colour for the first plot. The argument `cumul` indicates if cumulative effect, previously saved in `pred.pm`, must be plotted. The results are shown in Figures 1a-1b. Confidence intervals are set to the default value `"area"` for the argument `ci`. In the left panel, additional arguments are passed to the low-level plotting function `polygon()` through `ci.arg`, to draw instead shading lines as confidence intervals.

The interpretation is twofold: the curve represents the increase in risk in each future day following an increase of 10 μgr/m^3 in PM_{10} in a specific day (*forward interpretation*), or otherwise the contributions

of each past day with the same PM_{10} increase to the risk in a specific day (*backward interpretation*). The plots in Figures 1a-1b suggest that the initial increase in risk of PM_{10} is reversed at longer lags. The overall effect for a 10-unit increase in PM_{10} over 15 days of lag (i.e. summing all the effects up to the maximum lag), together with its 95% confidence intervals can be extracted by the objects `allRRfit`, `allRRhigh` and `allRRlow` included in `pred.pm`, typing:

```
> pred.pm$allRRfit["10"]

       10
0.9997563

> cbind(pred.pm$allRRlow, pred.pm$allRRhigh)["10",]

[1] 0.9916871 1.0078911
```

5.3 Example 2: seasonal analysis

The purpose of the second example is to illustrate an analysis where the data are restricted to a specific season. The main feature of these analysis is that the data are assumed to be composed by multiple equally-spaced and ordered series of the same season for each year, and do not represent a single continuous series. In this case, we assess the effect of ozone and temperature on overall mortality up to 5 and 10 days of lag, respectively, using the same steps already seen in Section 5.2.

First, we create the new data restricting to the summer period (June-September) the dataframe `chicagoNMMAPS`:

```
> chicagoNMMAPSseas <- subset(chicagoNMMAPS, month %in% 6:9)
```

Again, we first create the cross-basis matrices:

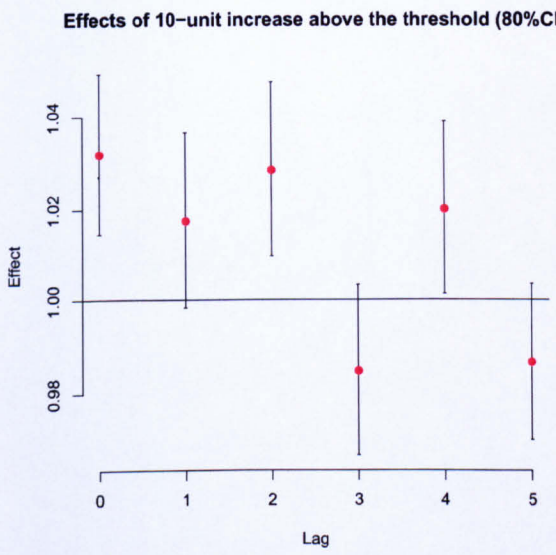
```
> basis.o3 <- crossbasis(chicagoNMMAPSseas$o3, group=chicagoNMMAPSseas$year,
          vartype="hthr", varknots=40.3, lagtype="integer", maxlag=5)
> basis.temp <- crossbasis(chicagoNMMAPSseas$temp, group=chicagoNMMAPSseas$year,
          vartype="dthr", varknots=c(15,25), lagtype="strata", lagknots=c(2,6),
          maxlag=10)
```

The argument `group` indicates the variable which defines multiple series: the function then breaks the series at the end of each group and replaces the first `maxlag` rows of the cross-basis matrix in the following series with `NA`. Each series must be consecutive, complete and ordered. Here we make the assumption that the effect of O_3 is null up to 40.3 $\mu gr/m^3$ and then linear, applying an high threshold parameterization. For temperature, we use a double threshold with the assumption that the effect is linear below 10°C and above 25°C, and null in between. Regarding the lag dimension, we specify an unconstrained function for O_3, applying one parameter for each lag (`lagtype="integer"`) up to a 5 days. For temperature, we define 3 strata intervals at lag 0-1, 2-5, 6-10. A summary of the choices made for the cross-bases can be shown by the function `summary.crossbasis()`.

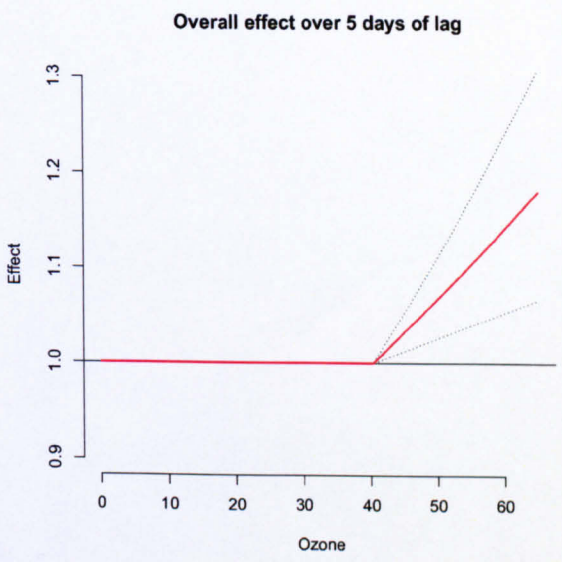
The regression model includes a natural spline for day of the year (with 4 df) in order to describe the seasonal effect within each year. Apart from that, the estimates and predictions are carried out in the same way as in Section 5.2. The code is:

Figure 2

(a)



(b)




```
> model <- glm(death ~  basis.o3 + basis.temp + ns(doy, 4) + dow,
        family=quasipoisson(), chicagoNMMAPSseas)
> pred.o3 <- crosspred(basis.o3, model, at=c(0:65,40.3,50.3))
```

The values for which the prediction must be computed are specified in `at`: here we define the integers from 0 to 65 μgr/m^3 (approximately the range of ozone distribution), plus the threshold and the value 50.3 μgr/m^3 corresponding to a 10-unit increase above the threshold, which is automatically set as the reference point for `type="hthr"` (see Section 4.2). The vector is automatically ordered. We can plot the lag-specific effects, similarly to Section 5.2, and also the overall effect of a 10-unit increase in O_3 with 95% confidence intervals. The related code is (results in Figures 2a-2b):

```
> plot(pred.o3, "slices", var=50.3, ci="bars", type="p", pch=19, ci.level=0.80,
        main="Effects of 10-unit increase above the threshold (80%CI)")
> plot(pred.o3,"overall",xlab="Ozone", ci="lines", ylim=c(0.9,1.3), lwd=2,
    ci.arg=list(col=1,lty=3), main="Overall effect over 5 days of lag")
```

In the first statement, the argument `ci="bars"` dictates that, differently from the default `"area"` seen in Figures 1a-1b, the confidence intervals are represented by bars. In addition, the argument `ci.level=0.80` states that 80% confidence intervals must be plotted. Finally, we chose points, instead of the default line, with specific symbol, by the arguments `type` and `pch`. In the second statement, the argument `type="overall"` indicates that the overall effects (summed upon lags) must be plotted, with confidence intervals as lines, `ylim` defining the range of the y-axis, `lwd` the thickness of the line. Similarly to the previous example, the display of confidence intervals are refined through the list of arguments specified by `ci.arg`, passed in this case to the low-level function `lines()`.

Similarly to the previous example, we can extract from `pred.o3` the estimated overall effect for a 10-unit increase in ozone above the threshold (50.3 − 40.3 μgr/m^3), together with its 95% confidence intervals:

```
> pred.o3$allRRfit["50.3"]

    50.3
1.069768

> cbind(pred.o3$allRRlow, pred.o3$allRRhigh)["50.3",]

[1] 1.026563 1.114791
```

The same plots and computation can be applied to the cold and heat effects of temperatures. For example, we can describe the increase in risk for 1°C beyond the low or high thresholds. The user can perform this analysis repeating the steps above.

5.4 Example 3: a complex DLNM

In the previous examples, the effects of air pollution (PM_{10} and O_3, respectively) were assumed completely linear or linear above a threshold. This assumption facilitates both the interpretation and the representation of the association: the dimension of the predictor is never considered, and the lag-specific or overall effects for a 10-unit increase are easily plotted. In contrast, when considering the non-linear effects of temperature, we need to adopt a bi-dimensional perspective in order to represent effects which vary non-linearly along the space of the predictor and lags.

In this last example we specify a more complex DLNM, where the effects are estimated using smooth non-linear functions for both dimensions. Despite the higher complexity of the relationship, we will see how the steps required to specify and fit the model and predict the results are exactly the same as for the simpler models see before in Sections 5.2-5.3, only requiring different plotting choices. The user can apply the same steps to investigate the effects of temperature in previous examples, and extend the plots for PM_{10} and O_3. In this case we run a DLNM to investigate the effects of temperature and PM_{10} on overall mortality up to lag 30 and 1, respectively.

These are the cross-basis matrices:

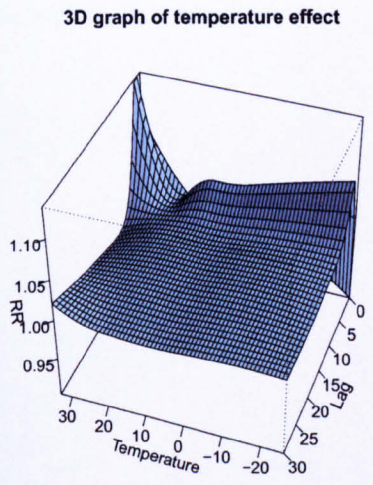
```
> basis.pm <- crossbasis(chicagoNMMAPS$pm10,vartype="lin", lagtype="strata",
          cen=F, maxlag=1)
> basis.temp <- crossbasis(chicagoNMMAPS$temp, vartype="bs", vardf=5, vardegree=2,
          lagdf=5, cenvalue=21, maxlag=30)
```

The chosen basis functions for the space of the predictor are a linear function for the effect of PM_{10} and a quadratic B-spline (`vartype="bs"`) with 5 degrees of freedom for temperature (with `varknots` placed by default at equally spaced quantiles in the space of the predictor). The basis for temperature is centered at 21°C, which will represent the reference point for the predicted effects. Regarding the space of lags, we assume a simple lag 0-1 parameterization for PM_{10} (i.e. a single strata up to lag 1, keeping the default values of `lagdf=1`), while we define another cubic spline, this time with the natural constraint (`lagtype="ns"` by default) for the lag dimension of temperature. For this space, `lagknots` are located by default at equally spaced values in the log scale of lags, while the boundary knots are set to 0 and `maxlag`. The estimation, prediction and plotting of the effects of temperature are performed by:

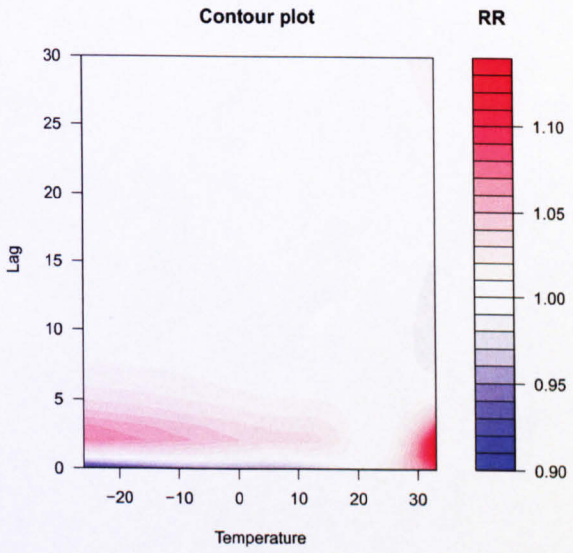
```
> model <- glm(death ~  basis.pm + basis.temp + ns(time, 7*14) + dow,
          family=quasipoisson(), chicagoNMMAPS)
> pred.temp <- crosspred(basis.temp, model, by=1)
```

Figure 3

(a)



(b)




```
> plot(pred.temp, xlab="Temperature",zlab="RR", theta=200, phi=40, lphi=30,
        main="3D graph of temperature effect")
> plot(pred.temp, "contour", xlab="Temperature", key.title=title("RR"),
        plot.title=title("Contour plot",xlab="Temperature",ylab="Lag"))
```

Note that prediction values are chosen only with the argument `by=1` in `crosspred()`, defining all the integer values within the predictor range. The first plotting expression produces a 3-D plot illustrated in Figure 3a, with non-default choices for perspective and lightning obtained through the arguments `theta-phi-lphi`. The second plotting expression specifies the contour plot in Figure 3b with titles and axis labels chosen by arguments `plot.title` and `key.title`. The user can find additional information and a complete list of arguments in the help pages of the original high-level plotting functions (typing `?persp` and `?filled.contour`). The plot of the overall effects can be obtained by (result not shown):

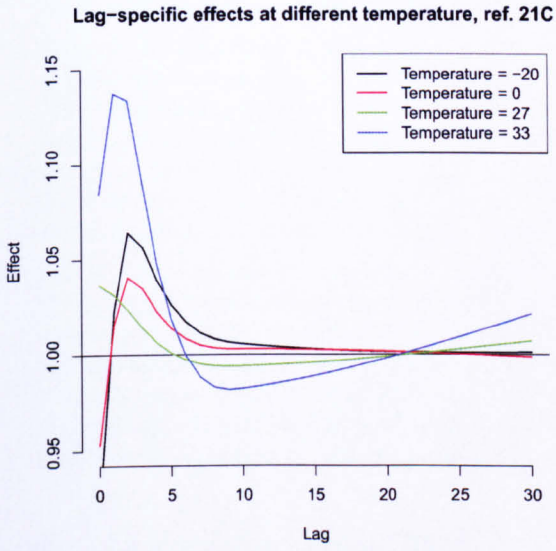
```
> plot(pred.temp, "overall", xlab="Temperature", ylim=c(0.8,1.7),
        main="Overall effect of temperature over 30 days of lag")
```

Plots in Figures 3a - 3b offer a comprehensive summary of the bi-dimensional relationship, but are limited in their ability to inform on effects at specific values of predictor or lags. In addition, they are also limited for inferential purposes, as the uncertainty of the estimated effects is not reported in 3-D and contour plots. A more detailed analysis is provided by plotting "slices" of the effect surface for specific predictor and lag values. The code is:

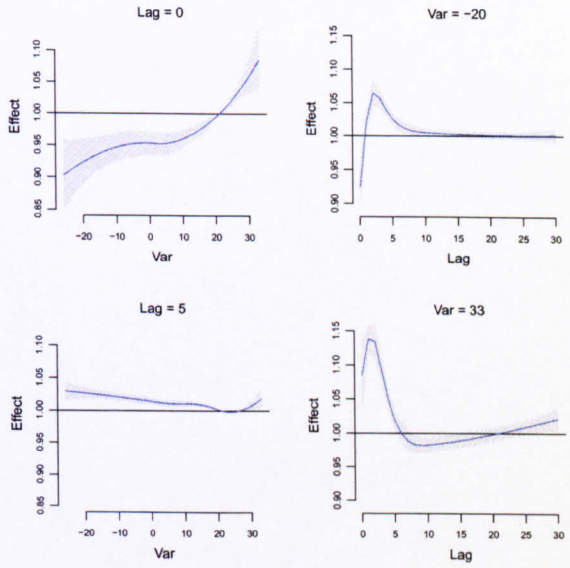
```
> plot(pred.temp, "slices", var=-20, ci="n", col=1, ylim=c(0.95,1.15), lwd=1.5,
        main="Lag-specific effects at different temperature, ref. 21C")
> for(i in 1:3) lines(pred.temp, "slices", var=c(0,27,33)[i], col=i+1, lwd=1.5)
> legend("topright",paste("Temperature =",c(-20,0,27,33)), col=1:4, lwd=1.5)
```

Figure 4

(a)



(b)




```
> plot(pred.temp, "slices", var=c(-20,33), lag=c(0,5), col=4,
    ci.arg=list(density=40,col=grey(0.7)))
```

The results are reported in Figures 4a - 4b. Figure 4a illustrates lag-specific effects for mild and extreme cold and hot temperatures of -20°C, 0°C, 27°C, and 33°C (with reference at 21°C). Figures 4b depicts both effects along the predictor range at lag 0 and 5 (left column), and effects along lags at temperatures -20°C and 33°C (right column). The arguments `var` and `lag` define the "slices" to be cut in the effect surface in Figure 3a - 3b. The argument `ci="n"` in the first expression states that confidence intervals must not be plotted. In the multi-panel Figure 4b, the list argument `ci.arg` is used to plot confidence intervals as shading lines with increased grey contrast, more visible here.

The preliminary interpretation suggests that cold temperatures are associated with longer mortality risk than heat, but not immediate, showing a "protective" effect at lag 0. This analytical proficiency would be hardly achieved with simpler models, probably losing important details of the association.

6 Conclusions

This document illustrates the functionalities of the **dlnm** package, providing a detailed overview of the process to specify and run a DLNM and then to predict and plot its results. The main advantage of this family of models is to unify many of the previous methods to deal with delayed effects in a unique framework, also providing more flexible alternatives regarding the shape of the relationships. Section 3 provides a brief summary of the theory underpinning DLNM's: a more detailed overview has been published elsewhere (Armstrong, 2006; Gasparrini, 2011; Gasparrini et al., 2010), together with a complete specification of the algebra (Gasparrini et al., 2010).

The flexibility is kept when this framework is implemented in the **dlnm** package: several different models with an increasing level of complexity can be performed using a simple and general procedure,

as showed in the examples in Section 5. As already explained, this method is not limited to the examples on the effect of air pollution and temperature on mortality, but can be applied to investigate the relationship between any predictor and outcomes in time-series data.

The choice of keeping separated the two steps of cross-basis specification and parameters estimation offers several advantages. First, as illustrated in the example, more than one variable showing delayed effects can be transformed through cross-basis functions and included in the model. Second, standard regression commands can be used for estimation, with the default set of diagnostic tools and related functions. More importantly, this implementation provides an open platform where additional models specified with different regression commands can be included as well, aiding the development of these methodology in other contexts or study designs.

The DLNM's framework introduced here is developed for time series design. The general expression of the model in allows this methodology to be applied for any family distribution and link function within generalized linear models (GLM), with extensions to GAM or models based on generalized estimating equations (GEE). Anyway, the current implementation of of DLNM's requires single series of equally-spaced and ordered data. Preliminary tests on the application of the functions included in the package dlnm in case-control, cohort and longitudinal data are promising. Further development may lead to a general framework to describe delayed effects, which spans different study designs.

7 Acknowledgements

This work was supported by the Medical Research Council (UK), Research Grant RES-G0707030.

We gratefully acknowledge the valuable suggestions of Fabio Frascati regarding the procedures to build and document this package. Other package vignettes were used as examples (in particular gnm by Heather Turner and David Firth). We are thankful to the colleagues who tested the beta-version of this package and suggested some important improvements included in the current version (in particular Marie-France Valois and Adrian Barnett).

Finally, we express our gratitude to all the people working to develop and maintain the R Project.

References

S. Almon. The distributed lag between capital appropriations and expenditures. *Econometrica*, 33: 178–196, 1965.

B. Armstrong. Models for the relationship between ambient temperature and daily mortality. *Epidemiology*, 17(6):624–31, 2006.

M. Baccini, A. Biggeri, G. Accetta, T. Kosatsky, K. Katsouyanni, A. Analitis, H. R. Anderson, L. Bisanti, D. D'Ippoliti, J. Danova, B. Forsberg, S. Medina, A. Paldy, D. Rabczenko, C. Schindler, and P. Michelozzi. Heat effects on mortality in 15 European cities. *Epidemiology*, 19(5):711–9, 2008.

A. L. Braga, A. Zanobetti, and J. Schwartz. The time course of weather-related deaths. *Epidemiology*, 12(6):662–7, 2001.

J. Cao, M. F. Valois, and M. S. Goldberg. An S-Plus function to calculate relative risks and adjusted means for regression models using natural splines. *Computer Methods and Programs in Biomedicine*, 84(1):58–62, 2006.

A. Gasparrini. Distributed lag linear and non-linear models in R: the package dlnm. *Journal of Statistical Software*, 43(8):1–20, 2011. URL `http://www.jstatsoft.org/v43/i08/`.

A. Gasparrini, B. Armstrong, and M. G. Kenward. Distributed lag non-linear models. *Statistics in Medicine*, 29(21):2224–2234, 2010.

S. Hajat, B. G. Armstrong, N. Gouveia, and P. Wilkinson. Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. *Epidemiology*, 16(5):613–20, 2005.

S. Pattenden, B. Nikiforov, and B. G. Armstrong. Mortality and temperature in Sofia and London. *Journal of Epidemiology and Community Health*, 57(8):628–33, 2003.

J. M. Samet, S. L. Zeger, F. Dominici, F. Curriero, I. Coursac, and D. W. Dockery. The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 2. Morbidity and mortality from air pollution in the United States. Technical report, Health Effects Institute, 2000a.

J. M. Samet, S. L. Zeger, F. Dominici, D. Dockery, and J. Schwartz. The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 1. Methods and methodological issues. Technical report, Health Effects Institute, 2000b.

J. Schwartz. The distributed lag between air pollution and daily deaths. *Epidemiology*, 11(3):320–6, 2000.

J. Schwartz. Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? *Epidemiology*, 12(1):55–61, 2001.

J. Schwartz, J. M. Samet, and J. A. Patz. Hospital admissions for heart disease: the effects of temperature and humidity. *Epidemiology*, 15(6):755–61, 2004.

L. J. Welty and S. L. Zeger. Are the acute effects of particulate matter on mortality in the National Morbidity, Mortality, and Air Pollution Study the result of inadequate control for weather and season? A sensitivity analysis using flexible distributed lag models. *American Journal of Epidemiology*, 162(1):80–8, 2005.

S.N. Wood. *Generalized additive models: an introduction with R*. Chapman & Hall/CRC, 2006.

A. Zanobetti and J. Schwartz. Mortality displacement in the association of ozone with mortality: an analysis of 48 cities in the United States. *American Journal of Respiratory and Critical Care Medicine*, 177(2):184–9, 2008.

A. Zanobetti, M. P. Wand, J. Schwartz, and L. M. Ryan. Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics*, 1(3):279–92, 2000.

Chapter 8

Commentary II

Title: Multivariate meta-analysis: a method to summarize non-linear associations.

Author(s): Antonio Gasparrini, Ben Armstrong.

Journal/Publisher: Statistics in Medicine.

Type of publication: Commentary.

Stage of publication: Accepted for publication. To appear in Volume 30, Issue 20, September 2011.

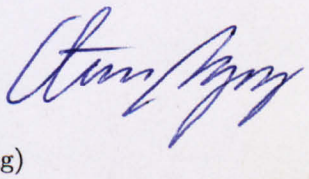
URL: Not available.

Academic peer-reviewed: No.

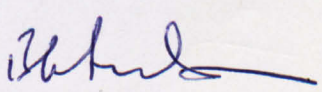
Copyright: Not required.

Candidate's role: See Section 2.3.

Candidate's signature:



Senior author: (Prof. Ben Armstrong)



Multivariate meta-analysis: a method to summarize non-linear associations

Antonio Gasparrini*,† and Ben Armstrong

Department of Social and Environmental Health Research
London School of Hygiene and Tropical Medicine, UK

Abstract

Multivariate meta-analysis represents a promising statistical tool in several research areas. Here we provide a brief overview of the application of this methodology to combining complex multi-parameterized relationships, such as non-linear or delayed associations, in multi-site studies. The discussion focuses on the advantages over simpler univariate methods, estimation and computational issues and directions for further research.

In this issue of *Statistics in Medicine*, Jackson and collaborators offer a comprehensive overview of the recent methodological advancements on multivariate meta-analysis, also highlighting limitations and research directions. Among the potential areas of application illustrated in their examples, we find particularly valuable the use of this methodology to combine multi-parameterized effects in multi-site observational studies, such as time series studies to assess the short term effects of environmental stressors. These studies usually adopt a two-stage approach, where a common first-stage model is applied to different cities or regions to derive site-specific estimates, and a second-stage meta-analysis is performed to combine these effects [1]. The presence of complex regression models with a high number of nuisance parameters to account for confounding factors makes the two-stage analysis attractive, circumventing the specification of a very highly parameterized hierarchical structure in a single multilevel development.

The usual approach proposed so far is based on first-stage models which simplify or summarize the city-specific effect in a single parameter, allowing the application of standard univariate meta-analytic techniques in the second stage. However, in the presence of complex associations, this choice could provide biased results with wrong assumptions about the simplified exposure-response shape (e.g. linear), or offer only a partial description of the phenomenon if the relationship is reduced to simple summaries. Multivariate meta-analysis has been proposed to combine non-linear dependencies [2, 3] and distributed lag structures [4], but there is no overview of methodological options. As a motivating example we illustrate the association between mean daily temperature and all-cause mortality in 108 USA cities [5], estimated through a quadratic B-spline with 5 degrees of freedom (with 3 equally-spaced knots) on lag 0-3. The associations in 4 cities are depicted in Figure 1.

The two-stage approach described above may be applied to model these relationships across cities, assuming that the k estimated parameters $\hat{\boldsymbol{\theta}}_i$ of the B-spline, defining the association in each of the $i = 1, \ldots, m$ cities, follow a multivariate normal distribution with

$$\hat{\boldsymbol{\theta}}_i \sim \mathcal{N}(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{S}_i + \boldsymbol{\Sigma}) \tag{1}$$

where $\mathbf{S}_i$ and $\boldsymbol{\Sigma}$ are the within and between-city (co)variance matrices, respectively. The term $\mathbf{X}_i$ represents a $k \times kp$ block-diagonal matrix, with each $1 \times p$ block containing city-specific meta- variables $\mathbf{x}_i$ (usually with intercept). The kp-dimensional vector $\boldsymbol{\beta}$ contains the coefficients specifying the change (effect modification) in each of the k true parameters $\boldsymbol{\theta}$ for a unit increase in each of the p meta-variables $\mathbf{x}_i$. When no modifier is included,

*Correspondance: Antonio Gasparrini, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK

†E-mail: antonio.gasparrini@lshtm.ac.uk

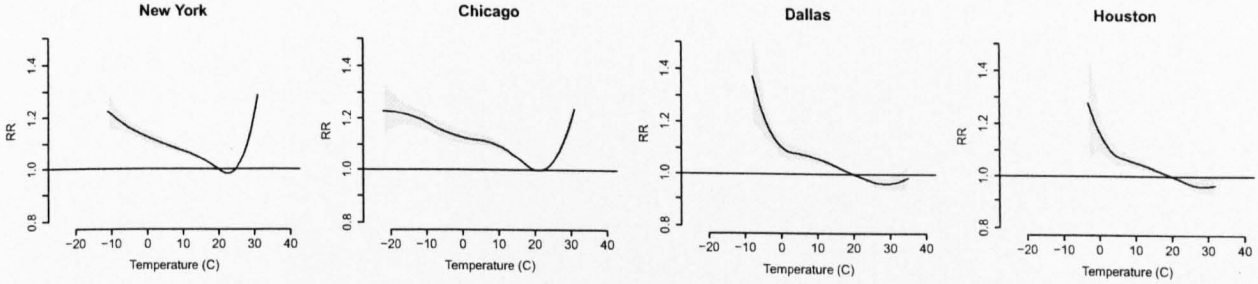


Figure 1: Temperature-mortality relationship (relative risk) in 4 USA cities, with reference at 20°C.

$\mathbf{X}_i\boldsymbol{\beta} \equiv \boldsymbol{\theta}$, the vector of true overall (population) parameters, and the model in (1) reduces to Eq. 3 in the paper by Jackson and colleagues.

The need for the more complex meta-regression model in (1), more elaborated than the framework described by the authors for their examples, is motivated by the different focus of the analysis: the main interest here is not to obtain a pooled estimate of the association, but to characterize the heterogeneity of the effects through city-specific meta-variables, while accounting for a random residual component in $\boldsymbol{\Sigma}$. In the specific example illustrated in Figure 1, our aim is to model a temperature-mortality relationship reflecting patterns such as shapes relatively similar within pairs of northern (New York and Chicago) and southern cities (Dallas and Houston), but different between them. This pattern may be explained by meta-variables $\mathbf{x}_1, \ldots, \mathbf{x}_p$, representing geographical, climatological, demographical or socio-economic determinants. Such analytical proficiency is not obviously achieved with simpler univariate methods.

There are issues of estimation and computation specific to this area of application. Usually, the study design allows complete control of the first-stage model, thus making the within-study covariances in $\mathbf{S}_i$ available. However, dimensionality needs to be taken into account: as the association is described by a growing number of parameters $\boldsymbol{\theta}$, estimation of the $k(k+1)/2$ (co)variance parameters in $\boldsymbol{\Sigma}$ could be problematic. Potential solutions may involve the simplification of $\boldsymbol{\Sigma}$, imposing for example an autoregressive, diagonal or compound-symmetry structure. The problem is worsened by the inclusion of a high number p of meta-variables, involving the estimation of kp coefficients. A simpler alternative is offered by meta-smoothing [6], a method based on a series of univariate meta-analysis of the effects estimated on a grid of exposure values, in order to recover the combined underlying relationship. While this method offers flexibility, an overall estimate of residual heterogeneity and significance tests are not easily provided. Finally, the model in (1) implies that exactly the same function is applied in every city, in order for the parameters $\hat{\boldsymbol{\theta}}_i$ to be meaningfully combined. In the example in Figure 1, the knots of the spline must be placed at the same values and this might represent a problem given the different temperature ranges between cities.

In conclusion, multivariate meta-analysis represents a promising methodology to combine multi-parameterized associations across studies. Compared to other examples described by Jackson and colleagues, the problem here is inherently multivariate, as each parameter is not interpretable on its own, and simplifications or approximations to re-express it in univariate terms are often limited or biased. However, the current framework could be infeasible for complex associations such as distributed lag non-linear relationships, involving a high number of parameters [7]. Further research is needed to address this problem of dimensionality, also providing some guidance on the limitations and comparative performances of different estimation methods in relation to number of studies m, parameters k, modifiers p and complexity of the structure of $\boldsymbol{\Sigma}$. This framework applies to other multi-parameter functions summarizing non-linear associations, such as strata or polynomials, and may be extended to other multi-unit studies such as multi-centre randomized controlled trials or multi-country cohort studies.

References

[1] Dominici F, Samet JM, Zeger SL. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. *Journal of the Royal Statistical Society: Series A* 2000; **163**(3):263–302.

[2] Baccini M, Biggeri A, Accetta G, Kosatsky T, Katsouyanni K, Analitis A, Anderson HR, Bisanti L, D'Ippoliti D, Danova J, *et al.*. Heat effects on mortality in 15 European cities. *Epidemiology* 2008; **19**(5):711–9.

[3] Dominici F, Daniels MJ, Zeger SL, Samet JM. Air pollution and mortality: estimating regional and national dose-response relationships. *Journal of the American Statistical Association* 2002; **97**:100–111.

[4] Analitis A, Katsouyanni K, Biggeri A, Baccini M, Forsberg B, Bisanti L, Kirchmayer U, Ballester F, Cadum E, Goodman PG, *et al.*. Effects of cold weather on mortality: results from 15 European cities within the PHEWE Project. *American Journal of Epidemiology* 2008; **168**(12):1397.

[5] iHAPSS. Internet-based Health and Air Pollution Surveillance System. Mortality, air pollution, and meteorological data for 108 US cities 19872000 ; :http://www.ihapss.jhsph.edu.

[6] Schwartz J, Zanobetti A. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. *Epidemiology* 2000; **11**(6):666–72.

[7] Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Statistics in Medicine* 2010; **29**(21):2224–2234.

Chapter 9

Research paper IV

Title: Multivariate meta-analysis for non-linear and other multi-parameter associations.

Author(s): Antonio Gasparrini, Ben Armstrong, Michael G. Kenward.

Journal/Publisher: Statistics in Medicine.

Type of publication: Research paper.

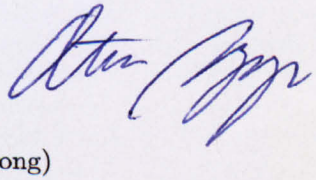
Stage of publication: Submitted (with online appendix).

URL: Not available.

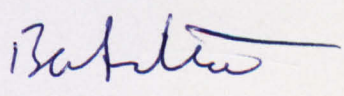
Academic peer-reviewed: Not yet.

Copyright: Not required.

Candidate's role: See Section 2.3.

Candidate's signature: 

Senior author: (Prof. Ben Armstrong)



Multivariate meta-analysis for non-linear and other multi-parameter associations

Antonio Gasparrini[1*,†] Ben Armstrong[1] and Michael G. Kenward[2]

[1]*Department of Social and Environmental Health Research*
[2]*Department of Medical Statistics*
London School of Hygiene and Tropical Medicine, UK

Abstract

In this paper we formalize the application of multivariate meta-analysis and meta-regression to synthesize estimates of multi-parameter associations obtained in different studies. This modelling approach extends the standard two-stage analysis used to combine results across different sub-groups or populations. The most straightforward application is for non-linear relationships, described for example by regression coefficients of splines or other functions, but the methodology easily generalizes to settings where complex associations are described by multiple correlated parameters. The modelling framework is implemented in the package mvmeta within the statistical environment R. As an illustrative example, we propose a two-stage analysis for investigating the non-linear exposure-response relationship of temperature and the distributed lag curve of ozone for all-cause mortality, using a real multi-city dataset including 98 cities in the USA. Multivariate meta-analysis represents a useful analytical tool for studying complex associations through a two-stage procedure.

1 Introduction

Meta-analysis is a standard, well-grounded statistical procedure for combining the evidence from independent studies that address the same research hypothesis [1]. This methodology was developed originally for pooling the results from published observational or experimental studies, for which individual data were not available. Recently, meta-analysis has been described more broadly as a research synthesis method, with the aim of estimating an average association and to explore the degree and sources of heterogeneity over multiple sub-groups or populations [2]. The analytical approach adopted in this context may be described as a two-stage hierarchical model: in the first stage, group-specific estimates of the association of interest are calculated, controlling for individual-level covariates; in the second stage, meta-analytical procedures are applied to combine these estimates, optionally exploring the association with group-level predictors. The two-stage approach has been proven to be a flexible and efficient method [3], and has been adopted in different contexts: to pool estimates from multiple randomized controlled trials [4]; to combine results from survival models on time-to-event data in multi-centre cohorts [5]; and to synthesize associations from Poisson time series models in multi-city analyses [6]. Without loss of generality, we will retain the meta-analytic terminology and refer to the sub-group analysis as the first-stage model, and to the first-stage units as studies.

The majority of applications of two-stage analyses has been characterized by fairly complex first-stage models, compared to relatively simple second-stage meta-analytic procedures. Although finely controlled for individual-level confounders, the association is usually summarized in a single parameter in the first step. This restriction reduces the amount of information in the data carried forward to the second stage, possibly producing inadequate or biased results in the presence of complex dependencies. A more flexible and general approach should potentially retain greater complexity, and so provide a method to synthesize multi-parameter associations, for example non-linear

*Correspondance: Antonio Gasparrini, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK
†E-mail: antonio.gasparrini@lshtm.ac.uk

relationships defined by splines or other functions. However, such extension requires a more elaborate meta-analytic model, capable of handling the multivariate nature of the summary estimates. Multivariate meta-analysis, a method originally developed to pool multiple correlated outcomes in randomized controlled trials [7, 8, 9], can be used to extend the standard two-stage analytical approach.

The aim of this article is to formalize the application of multivariate meta-analytic techniques to synthesize multi-parameter associations in two-stage hierarchical analyses, describing the statistical framework, methodological issues, limitations and research directions. This contribution originates from a commentary, to be published in this Journal [10], to the seminal paper by Jackson and collaborators on multivariate meta-analysis [11]. The article also offers the opportunity to describe the implementation in the package mvmeta within the R software [12], designed to perform multivariate meta-analysis and meta-regression in this and other contexts. The document is structured as follows. In Section 2, we describe the modelling framework of multivariate meta-analysis, with a specific focus on the setting of multi-parameter associations. An application is provided in Section 3, illustrating a two-stage analysis for the estimation of the non-linear exposure-response relationship of temperature and the distributed lag curve of ozone for all-cause mortality, using a real multi-city dataset including 98 cities in the USA. Specific methodological issues are discussed in Section 4. Finally, a general discussion is provided in Section 5. The Supplementary Web Appendix contains additional information on the software and the complete R code to replicate the results of the analysis illustrated in Section 3.

2 Modelling framework

The theoretical arguments that underpin the definition of the modelling framework of multivariate meta-analysis closely follow the simple univariate model, recently re-evaluated in detail [13]. The multivariate extension has been previously presented [8, 9, 14, 15], and a thorough overview has been also provided [11]. However, in contrast to the multiple outcomes scenario in which the method has been originally developed, in the context of multi-parameter associations the parameters may not be individually interpretable, and the association is instead characterized through their joint distribution. This specific feature constitutes the object of our re-assessment. In this Section, as an illustrative example, we will often refer to the application for estimating non-linear exposure-response relationship through spline functions in regression models, although the framework generalizes easily to other multi-parameter dependencies. A random-effects multivariate meta-regression model will be presented throughout, with fixed-effects models or simple meta-analysis as special cases.

2.1 The model

The framework we use is nested within that of the multivariate normal linear mixed model, and so follows well-developed lines [16]. Here the modelling development will be presented in the specific context of multi-parameter associations. We assume that a first-stage model has been fitted to the data from each of the $i = 1, \ldots, m$ studies, obtaining a k-dimensional set of regression coefficients $\hat{\boldsymbol{\theta}}_i$, and accompanying $k \times k$ estimated (co)variance matrix $\mathbf{S}_i$. Following our example, $\boldsymbol{\theta}$ may represent the parameters of the spline function, applied in the first-stage to model a non-linear dependency. The regression coefficients estimated in the first stage are used as outcomes for the second stage. In meta-regression models, these outcomes are modelled in terms of a set of p meta-predictors $\mathbf{x}_i = [x_{1i}, x_{2i}, \ldots, x_{pi}]^{\mathsf{T}}$ associated with the i^{th} study, where usually $x_1 = 1$ specifies the intercept. Regression coefficients from the first stage are termed from here on as outcome parameters, in order to distinguish them from the coefficients of the meta-analytic model.

Following Jackson and colleagues [11], we can write the marginal model for $\hat{\boldsymbol{\theta}}_i$, assuming a multivariate normal distribution of dimension k, as:

$$\hat{\boldsymbol{\theta}}_i \sim \mathrm{N}_k\left(\mathbf{X}_i\boldsymbol{\beta}\,,\,\boldsymbol{\Sigma}_i\right)\,. \tag{1}$$

Here $\boldsymbol{\Sigma}_i = \mathbf{S}_i + \boldsymbol{\Psi}$, where $\boldsymbol{\Psi}$ is the unknown between-study covariance matrix. The $k \times kp$ block-diagonal matrix $\mathbf{X}_i$, of rank kp, is derived by the Kronecker product between an identity matrix of dimension k and the vector $\mathbf{x}_i$,

following:

$$\mathbf{X}_i = \mathbf{I}_{(k)} \otimes \mathbf{x}_i^\top = \begin{Bmatrix} x_{1i} & x_{2i} & \dots & x_{pi} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & x_{1i} & x_{2i} & \dots & x_{pi} \end{Bmatrix} . \tag{2}$$

The kp-dimensional vector $\boldsymbol{\beta}$ defines how the p meta-predictors are associated with each of the k outcome parameters, for example defining intercepts and linear terms. The problem can also be re-expressed in the form of a conventional linear mixed model, defining random effects $\boldsymbol{u}_i \sim \mathrm{N}_k(\mathbf{0}, \boldsymbol{\Psi})$ which represent study-specific deviations. The model in (1) is then written as:

$$\hat{\boldsymbol{\theta}}_i \mid \boldsymbol{u}_i \sim \mathrm{N}_k\left(\mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{u}_i, \mathbf{S}_i\right) . \tag{3}$$

The matrix $\boldsymbol{\Psi}$ is completely defined by a set of parameters $\boldsymbol{\xi}$, dependent on the chosen structure and parameterization. If no a-priori structure is assumed, $k(k+1)/2$ terms are needed. Optionally, under the assumption that each outcome parameter is explained only in terms of a subset of the p variables, the related columns of $\mathbf{X}$ and entries of $\boldsymbol{\beta}$ can be excluded, defining different linear predictors. When no study-level variable is included, $\mathbf{X} \equiv \mathbf{I}_{(k)}$ and $\boldsymbol{\beta} \equiv \boldsymbol{\theta}$, the vector of average parameters, and the model in (1) reduces to conventional multivariate meta-analysis. Fixed-effects meta-analytic models presuppose that no heterogeneity exists in the outcome parameters distribution, and that the random variability is explained only by sampling error, assuming $\boldsymbol{\Sigma}_i \equiv \mathbf{S}_i$. As for the univariate case, estimation procedures treat $\mathbf{S}$ as known. The unknown parameters are therefore $\boldsymbol{\beta}$ and, for random-effects meta-analytic models, $\boldsymbol{\xi}$.

2.2 Estimation

Different estimation methods have been proposed for random-effects multivariate meta-analysis: likelihood-based methods [9, 15], estimating equations [17], variants of iterative generalized least squares [8, 18], Bayesian approaches [14] and multivariate extensions of the method of moments [19]. Here we will concentrate on maximum likelihood (ML) and restricted maximum likelihood (REML), following an extensive literature within the framework of linear mixed models [16, 20, 21]. These methods are implemented in the R package mvmeta and applied to perform the analysis in Section 3.

The marginal log-likelihood function $\ell(\boldsymbol{\beta}, \boldsymbol{\xi})$ for model (3) may be written as [16]:

$$\ell(\boldsymbol{\beta}, \boldsymbol{\xi}) = -\tfrac{1}{2} n \log \pi - \tfrac{1}{2} \sum_{i=1}^{m} \log |\boldsymbol{\Sigma}_i| - \tfrac{1}{2} \sum_{i=1}^{m} \left[\left(\hat{\boldsymbol{\theta}}_i - \mathbf{X}_i\boldsymbol{\beta}\right)^\top \boldsymbol{\Sigma}_i^{-1} \left(\hat{\boldsymbol{\theta}}_i - \mathbf{X}_i\boldsymbol{\beta}\right) \right] , \tag{4}$$

with n as the total number of observations (usually equal to km where there are no missing values). Assuming that $\boldsymbol{\xi}$, and therefore $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$, are known, the maximum likelihood (ML) estimates for $\boldsymbol{\beta}$ and its (co)variance matrix $V(\hat{\boldsymbol{\beta}})$ conditional on $\boldsymbol{\xi}$, are obtained by maximizing (4). In this case, closed-form equations are given by generalized least squares estimators:

$$\begin{aligned} \hat{\boldsymbol{\beta}}(\boldsymbol{\xi}) &= \left(\sum_{i=1}^{m} \mathbf{X}_i^\top \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{m} \mathbf{X}_i^\top \boldsymbol{\Sigma}_i^{-1} \hat{\boldsymbol{\theta}}_i , \\ V(\hat{\boldsymbol{\beta}}) &= \left(\sum_{i=1}^{m} \mathbf{X}_i^\top \boldsymbol{\Sigma}_i^{-1} \mathbf{X}_i \right)^{-1} . \end{aligned} \tag{5}$$

When $\boldsymbol{\Psi}$ is not known, the joint likelihood function in (4) needs to be maximized with respect to both $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$, and iterative methods are required. However, the ML estimator of the (co)variance parameters $\boldsymbol{\xi}$ is usually biased downward, as it does not account for the loss of degrees of freedom from the estimation of $\boldsymbol{\beta}$. An alternative estimator can be obtained from the maximization of the log-likelihood function based on a set of $n - q$ linearly

independent error contrasts, with q as the number of fixed-effects coefficients in $\boldsymbol{\beta}$. This restricted log-likelihood (REML) function $\ell_R(\boldsymbol{\xi})$, not dependent on $\boldsymbol{\beta}$, may be conveniently expressed as [16, 20]:

$$\ell_R(\boldsymbol{\xi}) = -\tfrac{1}{2}(n-q)\log\pi - \tfrac{1}{2}\sum_{i=1}^{m}\log|\boldsymbol{\Sigma}_i| - \tfrac{1}{2}\log\left|\sum_{i=1}^{m}\mathbf{X}_i^{\mathsf{T}}\boldsymbol{\Sigma}_i^{-1}\mathbf{X}_i\right| - \tfrac{1}{2}\sum_{i=1}^{m}\left[\left(\hat{\boldsymbol{\theta}}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}\right)^{\mathsf{T}}\boldsymbol{\Sigma}_i^{-1}\left(\hat{\boldsymbol{\theta}}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}\right)\right], \quad (6)$$

where $\hat{\boldsymbol{\beta}}$ is defined in (5).

The ML estimates of $\boldsymbol{\beta}$ in fixed-effects meta-analysis are simply obtained by (5), given that, as discussed in Section 2.1, $\boldsymbol{\Sigma}_i$ equals $\mathbf{S}_i$, and is therefore completely known. The ML and REML estimates in random-effects models can be instead obtained through Newton-type iterative algorithms. For computational purpose, the objective functions in (4) and (6) are both expressed with respect to $\boldsymbol{\xi}$ only, and maximization of $\ell(\boldsymbol{\xi})$ and $\ell_R(\boldsymbol{\xi})$ can be achieved by plugging-in at each iteration the conditional estimate of $\hat{\boldsymbol{\beta}}(\boldsymbol{\xi})$ in (5) using the current estimate of $\boldsymbol{\xi}$, until convergence. Additional information on the estimation algorithms used here are provided in the Supplementary Web Appendix.

2.3 Hypothesis testing and model comparison

We can separate inferences about the parameters in model (3) into those about fixed effects $\boldsymbol{\beta}$, which will typically be of prime interest, and between-study (co)variance matrix $\boldsymbol{\Psi}$. Inferential procedures, again, follow the theory of linear mixed models [16, Chap. 6].

Regarding fixed effects, under the marginal model in (3) and conditionally on $\boldsymbol{\xi}$, $\hat{\boldsymbol{\beta}}$ follows a multivariate normal distribution with mean and (co)variance matrix given in (5). As already mentioned, these coefficients represent the average outcome parameters in multivariate-meta-analysis, or their intercepts and linear dependencies on meta-predictors in meta-regression models. The correspondent entries of $\hat{\boldsymbol{\beta}}$ and $V(\hat{\boldsymbol{\beta}})$ may be used for obtaining significance tests or confidence intervals. However, in the context of multi-parameter associations, inferential procedures about single coefficients are of limited use: in the example proposed above, tests and confidence intervals referring to single parameters of a spline function offer little information on the association of interest. A more pertinent approach is to evaluate the relationship in a multivariate context, for example by testing if the non-linear spline curve changes depending on study-level meta-variables. This may be achieved by a multivariable Wald test for the null hypothesis $\mathbf{L}\boldsymbol{\beta} = 0$, with $\mathbf{L}$ as a contrast matrix selecting only the k coefficients which define the linear relationship of a specific meta-predictor with the outcome parameters. An alternative is to compare nested models differing by the same set of coefficients as before, through conventional likelihood ratio (LR) test. Note, however, that this test is appropriate only for ML models, as the general likelihood theory does not hold when comparing REML models with different fixed-effects structures [20]. The extension to testing more complex outcomes-predictors dependencies is straightforward. A common issue of the inferential procedures discussed above is that no account is taken of the uncertainty in the estimate of $\boldsymbol{\Psi}$ when calculating the precision of the fixed effects estimates. A suitable adjustment for this bias has been provided, also with an application to bivariate meta analysis [22, 23], although not as yet been implemented in mvmeta.

For random effects, the focus is on comparing models involving different choices about the structure of the between-study (co)variance matrix. In this setting, an interesting hypothesis to test is $\boldsymbol{\Psi} = \mathbf{0}$, namely that no heterogeneity between studies exists, beyond that explained by sampling variability. Similarly, a likelihood ratio test between nested models may be performed, which is appropriate in REML models as well given the identical fixed-effects structures. Note, however, that for alternative hypotheses which constrain (co)variance matrices to be positive-definite (see the Supplementary Web Appendix), the null value lies on a boundary of the parameter space. Under these conditions, the conventional null asymptotic $\chi^2(n-q)$ distribution does not hold, and some adjustment has been proposed [24]. A score test for the same null hypothesis and distribution has also been developed as the multivariate extension of the Cochran Q test for (residual) heterogeneity [7, 17]. The test is based on the statistic:

$$Q = \sum_{i=1}^{m}\left[\left(\hat{\boldsymbol{\theta}}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}\right)^{\mathsf{T}}\mathbf{S}_i^{-1}\left(\hat{\boldsymbol{\theta}}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}\right)\right], \quad (7)$$

where $\hat{\boldsymbol{\beta}}$ are estimated by the correspondent fixed-effects model. An extension of this heterogeneity test for a subset of $\boldsymbol{\beta}$ has also been proposed [17]. What seems less well known is that this test suffers exactly the same boundary value problems as the corresponding likelihood ratio test under the constrained, one sided-alternative, being based on the identical null asymptotic distribution [25].

In addition, in this meta-analytical setting, the quantification of the heterogeneity among studies, or the residual amount beyond that explained by specific covariates, is also of interest. Indices of heterogeneity analogous to the univariate case may easily been derived from the Q statistic in (7), such as the $H^2 = \max\{1\,,\, Q/(n-q)\}$ and $I^2 = (H^2-1)/H^2$ [26]. These measures are interpreted as the relative excess in heterogeneity above those explained by sampling error, and the proportion of total variation attributable to heterogeneity, respectively. Although recently criticized for being dependent on precision of the estimates from the first-stage model [27], these statistics provide simple summaries on the extent of heterogeneity.

More broadly, non-nested models may be compared using fit statistics, in particular Akaike information criterion AIC $= -2\,\ell(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Psi}}) + 2q$ and Bayesian information criterion BIC $= -2\,\ell(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Psi}}) + q\,\log(n)$, where $\ell(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Psi}})$ is the maximum log-likelihood. These statistics may also be used with REML models, with the additional requirement that fixed-effects structure be held constant.

2.4 Prediction

In the context of multi-parameter associations, the general tests and fit criteria described above, although important, are usually insufficient for interpretation. Coefficients in $\boldsymbol{\beta}$ refer to single outcome parameters which are rarely interpretable on their own, and the tests only offer a statistical belief on *whether* the multivariate distribution of outcome parameters depends on study-level covariates. However, these procedures fail to inform on *how* the latter modifies the former.

In the current setting, prediction represents an important tool to extend the inference from multivariate meta-regression models, offering a method to link specific values of study-level meta-variables with outcome parameters expectations. Given a set of meta-predictor values $\mathbf{x}_0$, the model predicted mean $\hat{\boldsymbol{\theta}}_0$ and (co)variance matrix $V(\hat{\boldsymbol{\theta}}_0)$ are obtained by:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_0 &= \mathbf{X}_0\hat{\boldsymbol{\beta}}\,, \\ V(\hat{\boldsymbol{\theta}}_0) &= \mathbf{X}_0 V(\hat{\boldsymbol{\beta}})\mathbf{X}_0^{\intercal}\,, \end{aligned} \tag{8}$$

with $\mathbf{X}_0$ computed from $\mathbf{x}_0$ following (2). The equations in (8) may be used to recover the predicted multi-parameterized association over a set of values observed in the individual data used in the first stage, together with confidence intervals. In the illustrative example, different exposure-response curves may be predicted for specific meta-predictor values $\mathbf{x}_0$, or simply the average curve for models with no predictors. The same equations may be used to predict the association in a new study characterized by a specific set of study-level variables, simply increasing the uncertainty in the estimates by adding $\boldsymbol{\Psi}$ to $V(\hat{\boldsymbol{\theta}}_0)$ in (8).

In addition, the assumptions outlined in Section 2.1 regarding the random-effects multivariate distribution may be exploited to extend the inference regarding study-specific outcome parameters $\boldsymbol{\theta}_i$ estimated in the first-stage model, computing the (asymptotic) best linear unbiased prediction (BLUP) [16, Section 7.4]. The predicted $\hat{\boldsymbol{\theta}}_{b(i)}$ and associated (co)variance matrix $V(\hat{\boldsymbol{\theta}}_{b(i)})$ are:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{b(i)} &= \mathbf{X}_i\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Sigma}}_i^{-1}\left(\hat{\boldsymbol{\theta}}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}\right)\,, \\ V(\hat{\boldsymbol{\theta}}_{b(i)}) &= \mathbf{X}_i V(\hat{\boldsymbol{\beta}})\mathbf{X}_i^{\intercal} + \hat{\boldsymbol{\Psi}} - \hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Sigma}}_i^{-1}\hat{\boldsymbol{\Psi}}\,, \end{aligned} \tag{9}$$

for $\hat{\boldsymbol{\Sigma}}_i = \mathbf{S}_i + \hat{\boldsymbol{\Psi}}$. The BLUP equations in (9) merely represent the sum of two components: the predicted averaged outcome parameters in (8) and study-specific deviations, predicted as random effects $\boldsymbol{u}_i$ in (3). Associations predicted with BLUP represent a trade-off between city-specific and average estimates, with weights inversely proportional to the two components $\boldsymbol{\Psi}$ and $\mathbf{S}$ of the total variability $\boldsymbol{\Sigma}$, respectively. The BLUP estimates borrow strength from the assumption of an underlying distribution of outcome parameters, with city-specific predictions

being shrunk toward the average: this shrinkage effect is stronger when the first-stage model provides imprecise estimates. It is noteworthy, in this multivariate setting, that the BLUP estimates of missing parameters from the first stage exploit the information about the other study-specific parameters and the between-study (co)variance matrix $\hat{\boldsymbol{\Psi}}$, and may be therefore different from predicted values from (8).

3 An application

As an illustration, we propose a two-stage analysis using time series data from multiple cities. The aim of the analysis is to investigate the risk of all-cause mortality with two environmental stressors, temperature and ozone, during summertime. The non-linear exposure-response relationship of temperature and the delayed effect of ozone are described in the first stage with functions specified through multiple parameters, which are then combined using the multivariate meta-analytic techniques illustrated in Section 2. Our intention is to illustrate the application of the methodology with real data, more than to provide substantive evidence on the associations under study. Therefore, several analytical steps, such as model selection and checking, are intentionally omitted. Moreover, we will also skip details on the interpretation of results in favour of methodological matters.

3.1 Data

The multi-city time series data used in the analysis were collected as part of the National Morbidity, Mortality and Air Pollution Study (NMMAPS) (http://www.ihapss.jhsph.edu). This publicly available database contains, among other information, daily series of mortality counts and weather and pollution measurements for the period 1987-2000 in 108 cities in USA. The analysis here is restricted to summer months (May-September) in the 98 cities reporting ozone measurements. In addition, the database includes city-level measures of several variables on geographical, climatological, demographic, and socio-economic characteristics. Given the illustrative purpose of this example, we limit our assessment to meta-regression models for only 3 city-level meta-predictors: latitude, population size and population percentage living in poverty.

3.2 Model specification

In the first stage, we adopt a standard analytical approach for time series environmental data [28, 29]. In each city, we fit a common generalized linear model for the quasi-Poisson family, and obtain estimates of the associations of temperature and ozone with all-cause mortality. The model also includes: a natural cubic spline of day of the year with 3 equally-spaced knots to model the within-summer seasonal variation; a natural cubic spline of year with 2 equally-spaced knots to allow long-term trends; indicator variables for day of the week.

The exposure-response relationship for temperature is modelled through a quadratic spline. The index is chosen as the lag 0-3 moving average of mean daily temperature. Given the adaptation of populations to their own climate [30], we define the association on a relative scale, reporting the exposure-response in terms of percentiles. In order to derive estimated parameters comparable across studies, we place the two knots of the quadratic spline at the 25^{th} and 75^{th} percentiles of the city-specific distribution. The spline basis function is centered at the 50^{th} percentile: the estimates are then reported as relative risk (RR) at each percentile versus the reference centering point. Interpretation of results must conform to the relative scale chosen here.

The effect of ozone is assumed linear with delay, and described through a distributed lag model [31]. Briefly, the linear effect of a specific exposure event is assumed to be distributed over a specific time period, measured in terms of *lags*: in this specific context, a given increase in ozone in a given day is expected to cause an increase in mortality on the same and the following days, up to a maximum lag period. The overall effect is represented by the sum of lag-specific contributions. These contributions are modelled through a function expressed in the new lag dimension, estimating a distributed lag curve. In this analysis we use a natural cubic spline including an intercept, with 3 knots at equally-spaced values on the logarithmic scale of lag, to allow more flexibility in the first part of the curve, where more variability is expected. The results are reported as the RR at each lag for a 10 μgr/m^3 increase in ozone.

Table 1: Mean, range and specific percentiles for city-level variables in in 98 USA cities, summers 1987-2000.

	Mean	Min	25%	75%	Max
Average daily mortality counts	18.7	2.0	7.3	18.8	178.1
Average temperature (°C)	22.7	16.7	20.0	25.6	31.9
Average ozone level (μgr/m^3)	31.3	13.2	27.7	34.5	51.6
Latitude (degree North)	37.2	21.3	33.8	41.1	47.7
Population size (x 100,000)	10.7	1.5	4.1	10.0	95.2
Percentage living in poverty (%)	13.5	6.5	10.5	15.7	27.9

In the second stage, multivariate meta-analyses and meta-regressions are applied independently to model the exposure-response of temperature and the distributed lag curve of ozone, given the estimates obtained by the common first-stage model. The models for temperature are based on the 4 outcome parameters of the quadratic spline, while models for ozone on the 5 outcome parameters of the distributed lag function. Multivariate meta-analyses are defined by intercepts-only models, while multivariate meta-regression models, specified for each of the 3 city-level meta-predictors in turn, include an intercept and linear term for each outcome parameters. An unconstrained form for the between-study (co)variance matrix $\mathbf{\Psi}$ is always chosen. The models are estimated through maximum likelihood.

3.3 Software

The analysis is performed in R (version 2.13). The package mvmeta (version 0.2.3) is used to run multivariate meta-analysis and meta-regression. The package dlnm (version 1.4.1) [32] is used to specify the basis matrices for the quadratic spline for temperature and the distributed lag spline for ozone, and to predict and plot the effects. The data are accessed through functions in the package NMMAPSlite. The code of the analysis to replicate all the results of Section 3.4 is available in the Supplementary Web Appendix.

3.4 Results

A descriptive analysis of the city-level variables included in the analysis is provided in Table 1. The cities are quite heterogeneous, particularly in respect of population size and related daily mortality count. The distribution of population size is also highly skewed, while latitude and poverty proportion are more symmetrically distributed.

The results on the exposure-response relationship between all-cause mortality and temperature are illustrated in Figure 1. The top left panel shows the estimated average curve computed from multivariate meta-analysis of the 4 coefficients of the quadratic spline, which represent the outcome parameters in (1) and (3). These coefficients are used to identify the exposure-response curve over a set of percentiles. In order to aid interpretation, the x axis is scaled in such a way that percentiles match those of the average temperature distribution of all the cities included in the analysis. As expected, the curve is very flat for low percentiles, representing mild summer temperatures, and rises for high relative temperatures. Fit statistics, the test of heterogeneity in (7) and I^2 are reported on top of the plot. Results indicate a high degree of heterogeneity between city-specific estimates. The Cochran Q test is highly significant and the I^2 indicates that 61.0% of the variation in the first-stage estimates is due to heterogeneity of the true city-specific associations.

The other 3 panels in Figure 1 summarize the results from multivariate meta-regression models, each including one city-level meta-predictor. These models comprise 4 additional fixed-effects coefficients, representing the linear change in outcome parameters accordingly to meta-predictor levels. The plots show the predicted curve for specific percentiles of the meta-predictor distribution, computed through (8). The analysis suggests that the effect of a high relative summer temperature is stronger in population living at higher latitudes, in more populated conurbations and in cities characterized by an higher percentage of people living in poverty, with an markedly increased steep in the hot tail. For example, the average percentage increase in risk for the 99$^{\text{th}}$ percentile versus the median temperature for cities at the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentiles of latitude distribution, are 8.9% (95%CI: 6.9 to 11.0%) and 5.3% (95%CI: 3.4 to 7.2%), respectively. This evidence is confirmed by tests and fit statistics. In particular,

both the LR and Wald tests suggest that each meta-predictor is significantly associated with the multivariate outcome. These results are confirmed by the lower AIC of models including the meta-predictors, while the BIC, highly penalized by the number of observations in the analysis, indicates instead a slight preference for the more parsimonious model. Although significant, the city-level meta-predictors seem to explain a limited amount of the heterogeneity, as showed by the small decrease of the Q and I^2 statistics, also reported on top of each panel.

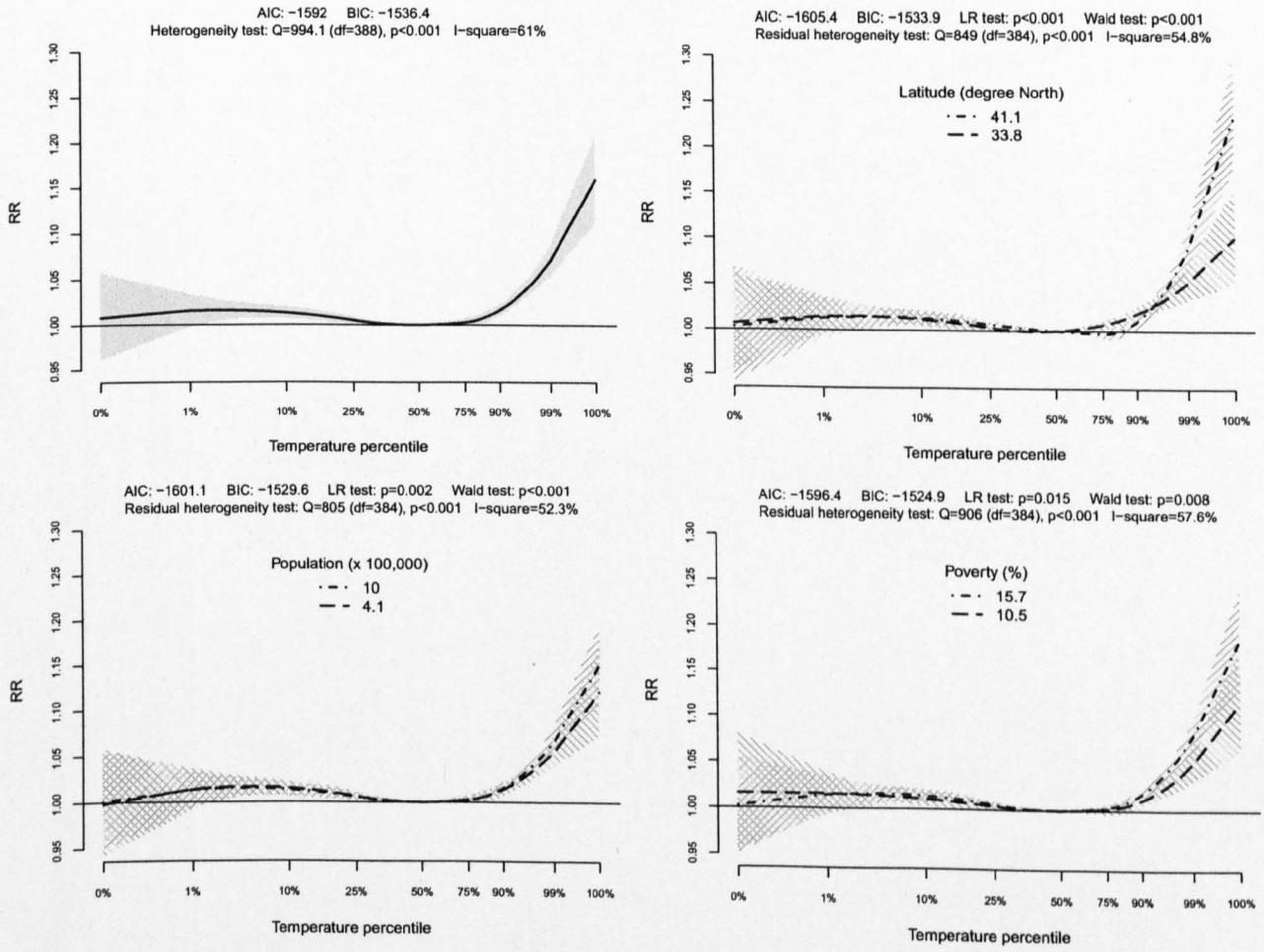


Figure 1: Pooled and predicted exposure-response relationships in relative risk (RR) between relative temperature (percentiles) and all-cause mortality in 98 USA cities, summers 1987-2000. The x axis is scaled so that percentiles represent the average temperature distribution of all the cities. The figure illustrates the population-average curve from meta-analysis (top left) and the predicted curves from meta-regression for the 25^{th} (dash line) and 75^{th} (dash-dot line) percentiles of latitude (top right), population size (bottom left) and population percentage in poverty (bottom right). Fit statistics, test for heterogeneity and I-square are reported on top of each panel. For meta-regression models, the likelihood ratio (LR) and Wald tests versus the model with no meta-predictor are also included.

Figure 2 illustrates the distributed lag curves estimated from models for ozone. The average relationship depicted in the top left panel suggests that a 10 μgr/m^3 increase in ozone is associated to a steep increase in risk in the same and following day, with a subsequent protective effect at longer lags, consistent with an harvesting effect

net overall effect is a percent increases of 0.2% (95%CI: 0.0 to 0.5%) (result not reported). Differently than in the temperature example, only latitude significantly modifies the association, with a higher initial effects followed by a stronger decrease in northern cities. However, the overall net effect predicted for the 25^{th} and 75^{th} percentiles of latitude distribution is almost identical, with estimated percent increases of 0.2% (95%CI: -0.1 to 0.5%) and 0.2% (95%CI: -0.2 to 0.6%), respectively (result not reported). Tests and information criteria for the other models clearly indicate no evidence that the effect of ozone varies by population size or percentage of poor people. The degree of heterogeneity is lower than for temperature. In particular, latitude seems to explain a large part of the variability between the true city-specific associations. However, the statistic for the related Q test lies near the boundaries of the parameter space, and under this condition the test suffers the problems described in Section 2.3. The results should be therefore interpreted with caution.

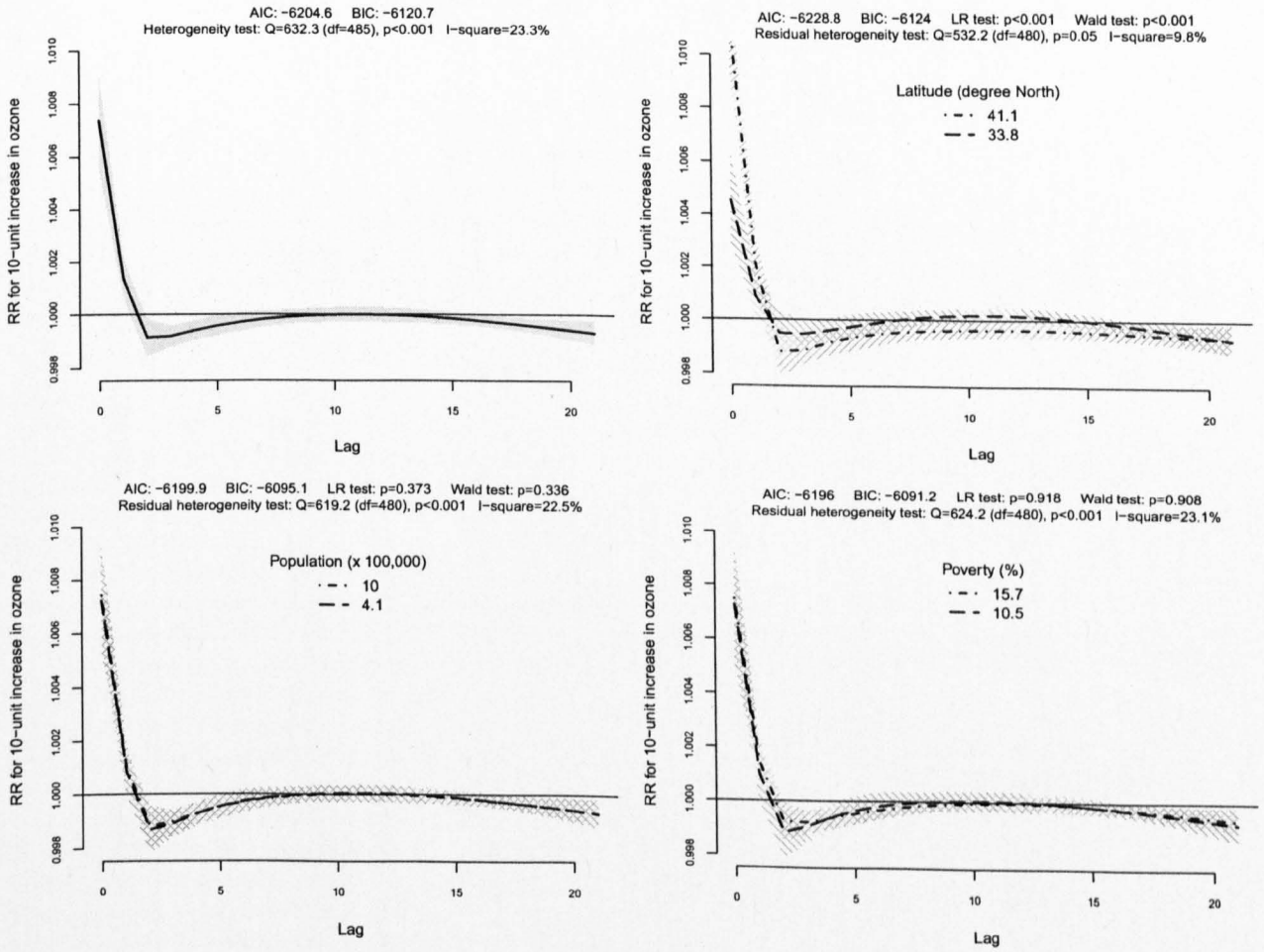


Figure 2: Pooled and predicted distributed lag curves in relative risk of all-cause mortality (RR) for a 10 μgr/m^3 increase in ozone in 98 USA cities, summers 1987-2000. The figure illustrates the population-average curve from meta-analysis (top left) and the predicted curves from meta-regression for the 25^{th} (dash line) and 75^{th} (dash-dot line) percentiles of latitude (top right), population size (bottom left) and population percentage in poverty (bottom right). Fit statistics, test for heterogeneity and I-square are reported on top of each panel. For meta-regression models, the likelihood ratio (LR) and Wald tests versus the model with no meta-predictor are also included.

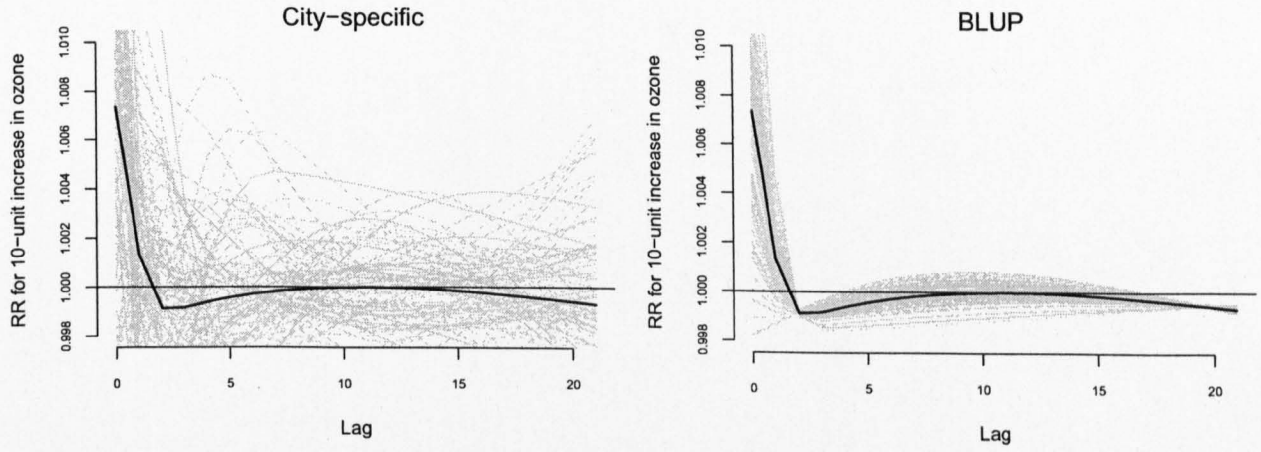


Figure 3: City-specific (left) and best linear unbiased predicted (right) estimates of the distributed lag curve in relative risk of all-cause mortality (RR) for a 10 $\mu gr/m^3$ increase in ozone in 98 USA cities, summers 1987-2000. The bold black line represents the population-average curve, while the grey lines the estimates for each city.

As discussed in Section 2.4, the assumptions about between-city variability, namely the distribution of the random effects in (3), may be used to extend the inference regarding city-specific estimates. The left panel of Figure 3 illustrates the distributed lag curve of ozone as estimated separately by each city-specific first-stage model, and the same population-average curve as depicted in Figure 2, top left plot. Variability around the average is due to both heterogeneity between cities and uncertainty in the first-stage model. The BLUP estimates, computed from (9), account for the latter and shrink city-specific curves toward the average, as shown in the right panel of Figure 3.

Given the high difference in population size showed in Table 1, the shrinkage effect is expected to vary considerably among cities. Figure 4 shows the predicted city-specific and BLUP exposure-response relationships for temperature in two cities, together with the population-average as depicted in Figure 1, top left plot. As expected, the BLUP estimate is closer to the original estimate from the city-specific first-stage model in Chicago, a large city characterized by a high number of daily deaths, while the BLUP curve for the small city of Kingston is heavily shrunk toward the population average. Interestingly, the shrinkage is higher in the left tail, corresponding to relative mild summer temperatures, if compared to the effect of heat: this is probably due to the degree of precision of the related part of the curve in the original city-specific estimate.

The choice of maximum likelihood estimators allows the comparison of models with different fixed-effect structures through likelihood ratio test and information criteria, as described in Sections 2.2 - 2.3. Given the relative high number of cities included in the analysis, we do not anticipate important differences with REML models, which indeed provide almost identical estimates (results not shown). However, this may not apply in analyses of smaller datasets. The extension to multivariable multivariate meta-regression is straightforward: tests and statistics are defined exactly in the same way, and predicted effects showed in Figures 1 - 2, controlled for the effect of other meta-variables, may be similarly computed. An example is included in the code provided in the Supplementary Web Appendix, together with other results described in this section.

4 Further considerations

As mentioned earlier, the methodology of multivariate meta-analysis has been largely developed in the context of randomized controlled trials to pool estimates on multiple outcomes. Moreover, as showed in Section 2, the statistical framework may be placed within linear mixed models, although with particular characteristics. Nevertheless, specific

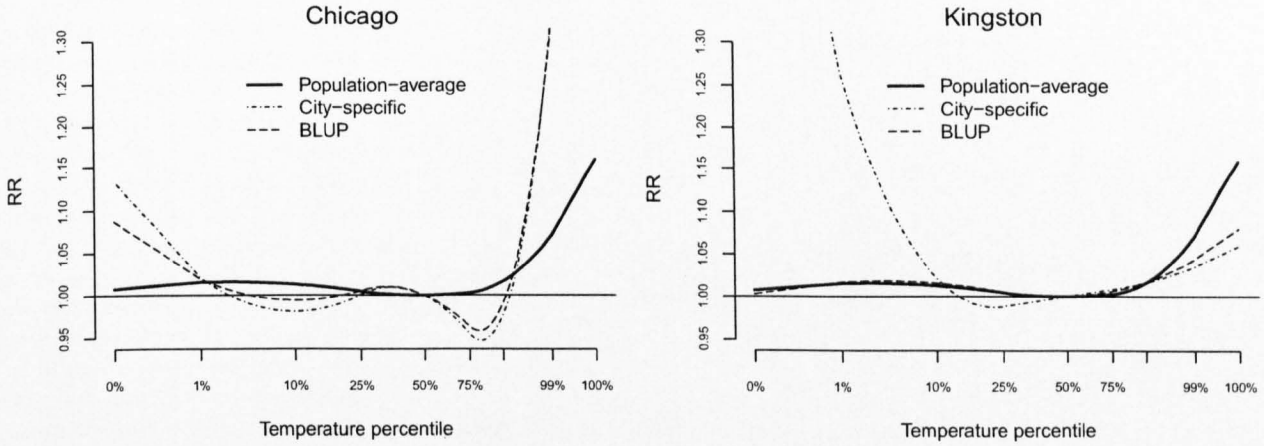


Figure 4: Population-average (continuous bold line), city-specific (dash-dot line) and best linear unbiased predicted (dash line) exposure-response relationship in relative risk (RR) between relative temperature (percentiles) and all-cause mortality in 2 USA cities, summers 1987-2000. The figure illustrates a large (Chicago, left) and small (Kingston, right) city included in the analysis.

issues arise when this methodology is applied to multi-parameter associations in two-stage analyses. Here we provide some comments on these aspects, highlighting advantages and limitations and directions for future research.

Advantages of multi-parameter synthesis. As anticipated in Section 1, the application of multivariate meta-analysis extends the standard two-stage design, where the data on the associations of interest are usually summarized in the estimate of a single parameter. For complex associations this choice may be too limited to characterize the phenomenon under study. Referring to the examples illustrated in Section 3, a standard analysis can be based on the pooling of single estimated effects at specific percentiles for temperature [30], or just on the overall net effect of ozone [34]. The estimate of the whole exposure-response relationship or distributed lag curve offers a more comprehensive picture, revealing additional important features. This approach may be more broadly described in the context of multi-parameter evidence synthesis, [35, 36].

Dealing with complexity. The two-stage approach discussed above provides tools to analyze complex associations. In the first step, the estimate is controlled for potential confounders, while reducing the relationship to a limited number of parameters of a chosen function, corresponding to the outcomes for the second stage meta-analytic model. The amount of complexity retained in the first stage represents a trade-off between synthesis and detail. Ideally, this balance should be fine-tuned only to the purpose of the analysis. However, in practice, concrete problems such as mathematical and statistical properties of the function or the maximum number of parameters needs to be taken into account, as discussed below.

Dimensionality. The number of parameters k which is possible to combine in multivariate-meta-analysis is of course limited. The models presented in Section 3 are defined by kp fixed-effects coefficients and $k(k+1)/2$ variance parameters: in the ozone example, the statistical problem requires the estimation of 25 total parameters using 490 correlated observations available in 98 cities. For instance, in a previous assessment based on distributed lag non-linear models, 25 parameters were used to model the non-linear and delayed effect of temperature in a single-city analysis [31]. A simple multivariate meta-regression would then require the estimate of 375 among coefficients and variance parameters, in order to combine such a complex multivariate dependency across multiple cities. A possible solution to reduce the number of parameters is to structure the between-study (co)variance matrix $\mathbf{\Psi}$, for example imposing a compound-symmetry, diagonal or autoregressive forms. Robust estimation to account for wrong correlation structures has been proposed for meta-regression of correlated outcomes [37], although further research is needed for this approach.

The two-stage design. In the setting of randomized controlled trials, the two-stage approach is often compared with the so-called individual patient data analysis, usually performed through a single multilevel model. Although the latter has been advocated as more efficient and less prone to bias if compared to the meta-analysis of published studies [38, 39], the two-stage alternative has been proved as competitive when applied to individual data [40, 41]. In addition, a single multilevel development is not always feasible or advisable, especially in the presence of many individual-level covariates, which would require the definition of intricate study-specific dependencies. For examples, in the models illustrated in Section 3, excluding the untenable assumption of a common seasonal variation in all the cities, $4 \times 97 = 388$ additional fixed-effects coefficients would be required to model city-specific seasonal trends. In the two-stage framework, parameters related to the association of interest are treated as nuisance terms in the first-stage model, offering computational efficiency and flexibility regarding model specification and assumptions.

Analysis of published studies. Although the modelling framework proposed here is focused on two-stage analysis on complete study-specific datasets, most of the original development of meta-analysis is based on the combination of estimates from published results. This also applies to the multivariate extension, as described later in Section 5. The meta-analysis of published studies poses additional problems. First, the outcome parameters defining the association in each study may not be comparable, for instance if estimated from different functions. Referring to the application in which non-linear dependencies are modelled, exposure categories may be defined with different cut-offs or spline functions with different knots. Solutions have been previously proposed to retrieve estimates of comparable outcome parameters from available study-specific information, as discussed in Section 5. Another issue is that correlation between estimated outcome parameters are rarely reported. Methods to deal with missing correlations have been developed [42, 43], although mainly limited to the bivariate case. This issue needs to be explored further, especially if dimensionality increases.

Exposure ranges. In the examples we provide in Section 3, the functions applied in the first stage for estimating the associations are defined exactly in the same range: a predetermined lag period for ozone, and the scale of percentile for temperature. In the latter case, this choice is motivated by existing evidence on modelling the effect of temperature, but in some instances, an absolute scale is preferred, and the analytical approach needs to cope with exposures defined in different ranges across studies. In such an analysis, some parameters of the function may be inestimable or meaningless. If not estimated, the parameters can be considered missing and the analysis performed under specific assumptions on the missing value mechanism. In any case, careful consideration needs to be given to the interpretation of the city-specific outcome parameters and their meta-analysis under this scenario.

Interpretational issues. The point discussed above is closely linked to the more general problem of interpretation of estimates of complex associations. Although the results illustrated in Section 3.4 are described on the original scale of the first-stage model, estimation is carried out in the multivariate dimension of the spline parameters. In practice, we read the association in the usual exposure-response frame, but we model it through coefficients of a function. We presuppose that these coefficients, in this multi-study assessment, still preserve their interpretation, and that the way we model the relationship between meta-predictors and their multivariate distribution reflects the association of interest. If, for example, different combinations of parameters define exactly the same association, this link vanishes, and interpretation of the results would be less straightforward. This issue requires further consideration.

Model selection. The choices of the models in the application we show in Section 3 are motivated only for illustrative purposes, and many alternative specifications may be suggested, possibly producing different results and conclusions. Model selection criteria have been proposed for multi-site studies adopting a two-stage approach [44], but the results are not conclusive and this issue needs further research, in particular in the multivariate setting proposed here.

5 Discussion

In this contribution we have provided a methodological overview of the application of multivariate meta-analysis and meta-regression analysis for the investigation of complex associations which are described by multiple parameters. This final section offers a review of previous research in this area, focusing first on the most straightforward application for non-linear relationships, in two-stage analyses of complete data and then in meta-analyses of published results. We then describe previous applications for modelling other multi-parameter associations. Finally,

we consider the advantages of two-stage procedures based on multivariate meta-analysis, also considering future research directions.

Two-stage analyses for non-linear exposure-response relationships based on complete data have previously been presented, although mostly focused on applied aspects. A recent paper has discussed the statistical methods for two-stage analysis of multi-site cohorts, also illustrating the use of multivariate meta-analysis for pooling dose-response associations that have been estimated using multiple categories [5]. Other examples include applications in multi-city time series studies to assess potential non-linear effects of air pollution [45, 46], using approaches similar to the example for temperature in Section 3. Methods based on Bayesian hierarchical models have also been presented [47]. An alternative approach already proposed is the so-called meta-smoothing [48]. This method is based on a series of univariate meta-analyses performed on estimated effects, for a grid of exposure values, in order to re-construct the pooled non-linear relationship. Although very flexible, as it provides complete freedom on the choice of the first-stage model in each study, it ignores the dependence among the analyses, which must be introduced subsequently for making valid inferences.

Methods for obtaining pooled dose-response dependencies from published epidemiological studies have been investigated in previous research. Pioneering works [49, 50] describes an analysis based on log-RR estimates for different exposure categories compared with a common reference, in which the whole within-study (co)variance matrix is reconstructed using ad-hoc approximations. The estimates of linear and (optionally) quadratic terms were then combined using fixed meta-analytic methods, and then the random counterpart based on method of moments. This approach has also been applied with splines or fractional polynomials to model non-linearity [51, 52, 53]. More recently, a general methodological treatment of the meta-analysis of published estimates for non-linear associations has been provided [54, 55, 56].

However, the framework illustrated here is not limited to model multi-parameterized non-linear exposure-response dependencies: investigators have also applied the methodology to synthesize survival curves [18, 57], longitudinal profiles [58], ROC curves [59] and heat wave effects [60]. Other studies have adopted multivariate meta-analysis to explore the distributed lag effects of air pollution [61] and temperature [62]. In particular, two studies have assessed the lagged effects of ozone [63, 64], with results comparable to those produced in the second example in Section 3. Finally, the same methods have also been applied to pool main and interactions terms across studies [50, 65].

The main limitation of the traditional approach based on univariate meta-analysis rests in the mismatch between the process of data synthesis applied in the first stage, and the details of the description offered by the second-stage meta-analytic model. This choice is limited by the requirement to summarize the association into a single outcome parameter. Multivariate meta-analysis relaxes this limitation, allowing a flexible specification of the two-stage development. In the application illustrated in Section 3, we propose a common first-stage model to study the relationship of all-cause mortality with two environmental stressors, then performing independent multivariate meta-analyses and meta-regressions for combining the study-specific estimates. The two sets of outcome parameters define different features of the association of each stressor with mortality, namely an exposure-response curve and a distributed lag pattern.

In its traditional setting for pooling multiple health endpoints in randomized controlled trials, multivariate meta-analysis offers parameter estimates with better statistical properties, in particular potentially increased precision from accommodating the estimated between-study covariance structure [11]. Nonetheless, the analysis could be carried out with multiple univariate meta-analysis, although often less efficiently. In the application we have described, instead, estimates of complex associations, such as those illustrated in Figures 1 - 4, cannot be provided by simple univariate models, without important limitations or additional assumptions. In this context, multivariate meta-analysis offers clear advantages.

As discussed in Section 2, this modelling framework can be seen as an example of a multivariate linear mixed model. The extensive body of research defining this statistical framework may therefore be exploited for this context, for example in the definition of tests discussed in Section 2.3. There are, of course, specific issues which deserve further research, for example statistics for heterogeneity, handling missing correlations or critical comparison of estimation methods. Other important issues specific to multivariate meta-analysis have been illustrated and discussed by Jackson and colleagues [11]. In the specific context of multi-parameter associations, a main limitation is related to dimensionality, as the number of outcome parameters which can be accommodated is currently limited. This and other issues will be hopefully addressed in future research on the development of this methodology.

References

[1] Normand SLT. Meta-analysis: formulating, evaluating, combining, and reporting. *Statistics in Medicine* 1999; **18**(3):321–359.

[2] Sutton AJ, Higgins JP. Recent developments in meta-analysis. *Statistics in Medicine* 2008; **27**(5):625–50.

[3] Stukel TA, Demidenko E, Dykes J, Karagas MR. Two-stage methods for the analysis of pooled data. *Statistics in Medicine* 2001; **20**(14):2115–2130.

[4] Simmonds MC, Higgins J, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clinical Trials* 2005; **2**(3):209–217.

[5] Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J, et al. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. *International Journal of Epidemiology* 2010; **39**:1345–1359.

[6] Dominici F, Samet JM, Zeger SL. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. *Journal of the Royal Statistical Society: Series A* 2000; **163**(3):263–302.

[7] Berkey CS, Anderson JJ, Hoaglin DC. Multiple-outcome meta-analysis of clinical trials. *Statistics in Medicine* 1996; **15**(5):537–557.

[8] Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine* 1998; **17**(22):2537–2550.

[9] Arends LR, Vokó Z, Stijnen T. Combining multiple outcome measures in a meta-analysis: an application. *Statistics in Medicine* 2003; **22**(8):1335–1353.

[10] Gasparrini A, Armstrong B, Kenward MG. Multivariate meta-analysis: a method to summarize non-linear associations. *Statistics in Medicine* 2011; **30**(20):To appear.

[11] Jackson D, Riley R, White IR. Multivariate meta-analysis: Potential and promise. *Statistics in Medicine* 2010; **30**(20):To appear. DOI: 10.1002/sim.4172.

[12] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria 2011. URL `http://www.R-project.org/`, ISBN 3-900051-07-0.

[13] Higgins J, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A* 2009; **172**(1):137–159.

[14] Nam IS, Mengersen K, Garthwaite P. Multivariate meta-analysis. *Statistics in Medicine* 2003; **22**(14):2309–2333.

[15] White IR. Multivariate random-effects meta-analysis. *Stata Journal* 2009; **9**(1):40–56.

[16] Verbeke G, Molenberghs G. *Linear mixed models for longitudinal data*. Springer Verlag: New York, 2000.

[17] Ritz J, Demidenkob E, Spiegelman D. Multivariate meta-analysis for data consortia, individual patient meta-analysis, and pooling projects. *Journal of Statistical Planning and Inference* 2008; **139**(7):1919–1933.

[18] Dear KBG. Iterative generalized least squares for meta-analysis of survival data at multiple times. *Biometrics* 1994; **50**:989–1002.

[19] Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine* 2010; **29**(12):1282–1297.

[20] Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* 1977; **72**(358):320–338.

[21] Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS*. Springer Verlag: New York, 2000.

[22] Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* 1997; **53**(3):983–997.

[23] Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics and Data Analysis* 2009; **53**(7):2583–2595.

[24] Stram DO, Lee JW. Variance components testing in the longitudinal mixed effects model. *Biometrics* 1994; **50**(4):1171–1177.

[25] Verbeke G, Molenberghs G. The use of score tests for inference on variance components. *Biometrics* 2003; **59**(2):254–262.

[26] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; **21**(11):1539–1558.

[27] Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on i2 in assessing heterogeneity may mislead. *BMC Medical Research Methodology* 2008; **8**(1):79.

[28] Dominici F. Time-series analysis of air pollution and mortality: a statistical review. *Research report - Health Effects Institute* 2004; **123**:3–27; discussion 29–33.

[29] Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology* 2006; **17**(6):624–31.

[30] Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology* 2009; **20**(2):205–213.

[31] Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Statistics in Medicine* 2010; **29**(21):2224–2234.

[32] Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. *Journal of Statistical Software* 2011; **43**(8):1–20. URL `http://www.jstatsoft.org/v43/i08/`.

[33] Schwartz J. Harvesting and long term exposure effects in the relation between air pollution and mortality. *American Journal of Epidemiology* 2000; **151**(5):440–8.

[34] Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F. Ozone and short-term mortality in 95 U.S. urban communities, 1987-2000. *Journal of the American Statistical Association* 2004; **292**(19):2372–8.

[35] Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Statistics in Medicine* 2003; **22**(19):2995–3016.

[36] Ades AE, Sutton AJ. Multiparameter evidence synthesis in epidemiology and medical decision-making: current approaches. *Journal of the Royal Statistical Society: Series A* 2006; **169**(1):5–35.

[37] Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods* 2010; **1**(1):39–65.

[38] Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Statistics in Medicine* 2002; **21**(3):371–87.

[39] Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of Clinical Epidemiology* 2002; **55**(1):86–94.

[40] Olkin I, Sampson A. Comparison of meta-analysis versus analysis of variance of individual patient data. *Biometrics* 1998; **54**(1):317–322.

[41] Mathew T, Nordstrom K. Comparison of one-step and two-step meta-analysis models using individual patient data. *Biometrical Journal* 2010; **52**(2):271–287.

[42] Riley RD, Thompson JR, Abrams KR. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics* 2008; **9**(1):172.

[43] Riley RD. Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A* 2009; **172**(4):789–811.

[44] Peng RD, Dominici F, Louis TA. Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society: Series A* 2006; **169**(2):179–203.

[45] Samoli E, Analitis A, Touloumi G, Schwartz J, Anderson HR, Sunyer J, Bisanti L, Zmirou D, Vonk JM, Pekkanen J, *et al.*. Estimating the exposure-response relationships between particulate matter and mortality within the APHEA multicity project. *Environmental Health Perspectives* 2005; **113**(1):88–97.

[46] Baccini M, Biggeri A, Accetta G, Kosatsky T, Katsouyanni K, Analitis A, Anderson HR, Bisanti L, D'Ippoliti D, Danova J, *et al.*. Heat effects on mortality in 15 European cities. *Epidemiology* 2008; **19**(5):711–9.

[47] Dominici F, Daniels MJ, Zeger SL, Samet JM. Air pollution and mortality: estimating regional and national dose-response relationships. *Journal of the American Statistical Association* 2002; **97**:100–111.

[48] Schwartz J, Zanobetti A. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. *Epidemiology* 2000; **11**(6):666–72.

[49] Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *American Journal of Epidemiology* 1992; **135**(11):1301–1309.

[50] Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology* 1993; **4**(3):218–28.

[51] Bagnardi V, Zambon A, Quatto P, Corrao G. Flexible meta-regression functions for modeling aggregate dose-response data, with an application to alcohol and mortality. *American Journal of Epidemiology* 2004; **159**(11):1077–1086.

[52] Hartemink N, Boshuizen HC, Nagelkerke NJ, Jacobs MA, van Houwelingen HC. Combining risk estimates from observational studies with different exposure cutpoints: a meta-analysis on body mass index and diabetes type 2. *American Journal of Epidemiology* 2006; **163**(11):1042–52.

[53] Vlaanderen J, Portengen L, Rothman N, Lan Q, Kromhout H, Vermeulen R. Flexible meta-regression to assess the shape of the benzene-leukemia exposure-response curve. *Environmental Health Perspectives* 2010; **118**(4):526–532.

[54] Shi JQ, Copas JB. Meta-analysis for trend estimation. *Statistics in Medicine* 2004; **23**(1):3–19.

[55] Liu Q, Cook NR, Bergstrom A, Hsieh CC. A two-stage hierarchical regression model for meta-analysis of epidemiologic nonlinear dose-response data. *Computational Statistics and Data Analysis* 2009; **53**(12):4157–4167.

[56] Rota M, Bellocco R, Scotti L, Tramacere I, Jenab M, Corrao G, La Vecchia C, Boffetta P, Bagnardi V. Random-effects meta-regression models for studying nonlinear dose-response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Statistics in Medicine* 2010; **29**:2679–2687.

[57] Arends LR, Hunink MG, Stijnen T. Meta-analysis of summary survival curve data. *Statistics in Medicine* 2008; **27**(22):4381–4396.

[58] Ishak KJ, Platt RW, Joseph L, Hanley JA, Caro JJ. Meta-analysis of longitudinal studies. *Clinical Trials* 2007; **4**(5):525–39.

[59] Arends LR, Hamza TH, Van Houwelingen JC, Heijenbrok-Kal MH, Hunink MGM, Stijnen T. Bivariate random effects meta-analysis of ROC curves. *Medical Decision Making* 2008; **28**(5):621–638.

[60] Gasparrini A, Armstrong B. The impact of heat waves on mortality. *Epidemiology* 2011; **22**(1):68–73.

[61] Zanobetti A, Schwartz J, Samoli E, Gryparis A, Touloumi G, Atkinson R, Le Tertre A, Bobros J, Celko M, Goren A, *et al.*. The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement. *Epidemiology* 2002; **13**(1):87–93.

[62] Analitis A, Katsouyanni K, Biggeri A, Baccini M, Forsberg B, Bisanti L, Kirchmayer U, Ballester F, Cadum E, Goodman PG, *et al.*. Effects of cold weather on mortality: results from 15 European cities within the PHEWE Project. *American Journal of Epidemiology* 2008; **168**(12):1397.

[63] Zanobetti A, Schwartz J. Mortality displacement in the association of ozone with mortality: an analysis of 48 cities in the United States. *American Journal of Respiratory and Critical Care Medicine* 2008; **177**(2):184–9.

[64] Samoli E, Zanobetti A, Schwartz J, Atkinson R, Le Tertre A, Schindler C, Perez L, Cadum E, Pekkanen J, Paldy A, *et al.*. The temporal pattern of mortality responses to ambient ozone in the APHEA project. *Journal of Epidemiology and Community Health* 2009; **63**:960–966.

[65] Zanobetti A, Schwartz J. Temperature and mortality in nine US cities. *Epidemiology* 2008; **19**(4):563–70.

PAGE

NUMBERING

AS ORIGINAL

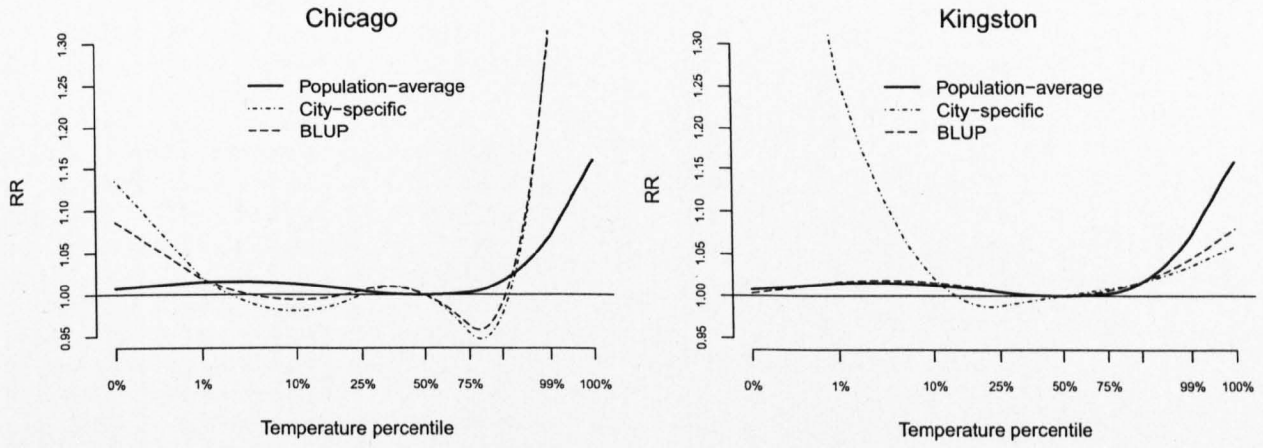


Figure 4: Population-average (continuous bold line), city-specific (dash-dot line) and best linear unbiased predicted (dash line) exposure-response relationship in relative risk (RR) between relative temperature (percentiles) and all-cause mortality in 2 USA cities, summers 1987-2000. The figure illustrates a large (Chicago, left) and small (Kingston, right) city included in the analysis.

issues arise when this methodology is applied to multi-parameter associations in two-stage analyses. Here we provide some comments on these aspects, highlighting advantages and limitations and directions for future research.

Advantages of multi-parameter synthesis. As anticipated in Section 1, the application of multivariate meta-analysis extends the standard two-stage design, where the data on the associations of interest are usually summarized in the estimate of a single parameter. For complex associations this choice may be too limited to characterize the phenomenon under study. Referring to the examples illustrated in Section 3, a standard analysis can be based on the pooling of single estimated effects at specific percentiles for temperature [30], or just on the overall net effect of ozone [34]. The estimate of the whole exposure-response relationship or distributed lag curve offers a more comprehensive picture, revealing additional important features. This approach may be more broadly described in the context of multi-parameter evidence synthesis, [35, 36].

Dealing with complexity. The two-stage approach discussed above provides tools to analyze complex associations. In the first step, the estimate is controlled for potential confounders, while reducing the relationship to a limited number of parameters of a chosen function, corresponding to the outcomes for the second stage meta-analytic model. The amount of complexity retained in the first stage represents a trade-off between synthesis and detail. Ideally, this balance should be fine-tuned only to the purpose of the analysis. However, in practice, concrete problems such as mathematical and statistical properties of the function or the maximum number of parameters needs to be taken into account, as discussed below.

Dimensionality. The number of parameters k which is possible to combine in multivariate-meta-analysis is of course limited. The models presented in Section 3 are defined by kp fixed-effects coefficients and $k(k+1)/2$ variance parameters: in the ozone example, the statistical problem requires the estimation of 25 total parameters using 490 correlated observations available in 98 cities. For instance, in a previous assessment based on distributed lag non-linear models, 25 parameters were used to model the non-linear and delayed effect of temperature in a single-city analysis [31]. A simple multivariate meta-regression would then require the estimate of 375 among coefficients and variance parameters, in order to combine such a complex multivariate dependency across multiple cities. A possible solution to reduce the number of parameters is to structure the between-study (co)variance matrix $\mathbf{\Psi}$, for example imposing a compound-symmetry, diagonal or autoregressive forms. Robust estimation to account for wrong correlation structures has been proposed for meta-regression of correlated outcomes [37], although further research is needed for this approach.

The two-stage design. In the setting of randomized controlled trials, the two-stage approach is often compared with the so-called individual patient data analysis, usually performed through a single multilevel model. Although the latter has been advocated as more efficient and less prone to bias if compared to the meta-analysis of published studies [38, 39], the two-stage alternative has been proved as competitive when applied to individual data [40, 41]. In addition, a single multilevel development is not always feasible or advisable, especially in the presence of many individual-level covariates, which would require the definition of intricate study-specific dependencies. For examples, in the models illustrated in Section 3, excluding the untenable assumption of a common seasonal variation in all the cities, $4 \times 97 = 388$ additional fixed-effects coefficients would be required to model city-specific seasonal trends. In the two-stage framework, parameters related to the association of interest are treated as nuisance terms in the first-stage model, offering computational efficiency and flexibility regarding model specification and assumptions.

Analysis of published studies. Although the modelling framework proposed here is focused on two-stage analysis on complete study-specific datasets, most of the original development of meta-analysis is based on the combination of estimates from published results. This also applies to the multivariate extension, as described later in Section 5. The meta-analysis of published studies poses additional problems. First, the outcome parameters defining the association in each study may not be comparable, for instance if estimated from different functions. Referring to the application in which non-linear dependencies are modelled, exposure categories may be defined with different cut-offs or spline functions with different knots. Solutions have been previously proposed to retrieve estimates of comparable outcome parameters from available study-specific information, as discussed in Section 5. Another issue is that correlation between estimated outcome parameters are rarely reported. Methods to deal with missing correlations have been developed [42, 43], although mainly limited to the bivariate case. This issue needs to be explored further, especially if dimensionality increases.

Exposure ranges. In the examples we provide in Section 3, the functions applied in the first stage for estimating the associations are defined exactly in the same range: a predetermined lag period for ozone, and the scale of percentile for temperature. In the latter case, this choice is motivated by existing evidence on modelling the effect of temperature, but in some instances, an absolute scale is preferred, and the analytical approach needs to cope with exposures defined in different ranges across studies. In such an analysis, some parameters of the function may be inestimable or meaningless. If not estimated, the parameters can be considered missing and the analysis performed under specific assumptions on the missing value mechanism. In any case, careful consideration needs to be given to the interpretation of the city-specific outcome parameters and their meta-analysis under this scenario.

Interpretational issues. The point discussed above is closely linked to the more general problem of interpretation of estimates of complex associations. Although the results illustrated in Section 3.4 are described on the original scale of the first-stage model, estimation is carried out in the multivariate dimension of the spline parameters. In practice, we read the association in the usual exposure-response frame, but we model it through coefficients of a function. We presuppose that these coefficients, in this multi-study assessment, still preserve their interpretation, and that the way we model the relationship between meta-predictors and their multivariate distribution reflects the association of interest. If, for example, different combinations of parameters define exactly the same association, this link vanishes, and interpretation of the results would be less straightforward. This issue requires further consideration.

Model selection. The choices of the models in the application we show in Section 3 are motivated only for illustrative purposes, and many alternative specifications may be suggested, possibly producing different results and conclusions. Model selection criteria have been proposed for multi-site studies adopting a two-stage approach [44], but the results are not conclusive and this issue needs further research, in particular in the multivariate setting proposed here.

5 Discussion

In this contribution we have provided a methodological overview of the application of multivariate meta-analysis and meta-regression analysis for the investigation of complex associations which are described by multiple parameters. This final section offers a review of previous research in this area, focusing first on the most straightforward application for non-linear relationships, in two-stage analyses of complete data and then in meta-analyses of published results. We then describe previous applications for modelling other multi-parameter associations. Finally,

we consider the advantages of two-stage procedures based on multivariate meta-analysis, also considering future research directions.

Two-stage analyses for non-linear exposure-response relationships based on complete data have previously been presented, although mostly focused on applied aspects. A recent paper has discussed the statistical methods for two-stage analysis of multi-site cohorts, also illustrating the use of multivariate meta-analysis for pooling dose-response associations that have been estimated using multiple categories [5]. Other examples include applications in multi-city time series studies to assess potential non-linear effects of air pollution [45, 46], using approaches similar to the example for temperature in Section 3. Methods based on Bayesian hierarchical models have also been presented [47]. An alternative approach already proposed is the so-called meta-smoothing [48]. This method is based on a series of univariate meta-analyses performed on estimated effects, for a grid of exposure values, in order to re-construct the pooled non-linear relationship. Although very flexible, as it provides complete freedom on the choice of the first-stage model in each study, it ignores the dependence among the analyses, which must be introduced subsequently for making valid inferences.

Methods for obtaining pooled dose-response dependencies from published epidemiological studies have been investigated in previous research. Pioneering works [49, 50] describes an analysis based on log-RR estimates for different exposure categories compared with a common reference, in which the whole within-study (co)variance matrix is reconstructed using ad-hoc approximations. The estimates of linear and (optionally) quadratic terms were then combined using fixed meta-analytic methods, and then the random counterpart based on method of moments. This approach has also been applied with splines or fractional polynomials to model non-linearity [51, 52, 53]. More recently, a general methodological treatment of the meta-analysis of published estimates for non-linear associations has been provided [54, 55, 56].

However, the framework illustrated here is not limited to model multi-parameterized non-linear exposure-response dependencies: investigators have also applied the methodology to synthesize survival curves [18, 57], longitudinal profiles [58], ROC curves [59] and heat wave effects [60]. Other studies have adopted multivariate meta-analysis to explore the distributed lag effects of air pollution [61] and temperature [62]. In particular, two studies have assessed the lagged effects of ozone [63, 64], with results comparable to those produced in the second example in Section 3. Finally, the same methods have also been applied to pool main and interactions terms across studies [50, 65].

The main limitation of the traditional approach based on univariate meta-analysis rests in the mismatch between the process of data synthesis applied in the first stage, and the details of the description offered by the second-stage meta-analytic model. This choice is limited by the requirement to summarize the association into a single outcome parameter. Multivariate meta-analysis relaxes this limitation, allowing a flexible specification of the two-stage development. In the application illustrated in Section 3, we propose a common first-stage model to study the relationship of all-cause mortality with two environmental stressors, then performing independent multivariate meta-analyses and meta-regressions for combining the study-specific estimates. The two sets of outcome parameters define different features of the association of each stressor with mortality, namely an exposure-response curve and a distributed lag pattern.

In its traditional setting for pooling multiple health endpoints in randomized controlled trials, multivariate meta-analysis offers parameter estimates with better statistical properties, in particular potentially increased precision from accommodating the estimated between-study covariance structure [11]. Nonetheless, the analysis could be carried out with multiple univariate meta-analysis, although often less efficiently. In the application we have described, instead, estimates of complex associations, such as those illustrated in Figures 1 - 4, cannot be provided by simple univariate models, without important limitations or additional assumptions. In this context, multivariate meta-analysis offers clear advantages.

As discussed in Section 2, this modelling framework can be seen as an example of a multivariate linear mixed model. The extensive body of research defining this statistical framework may therefore be exploited for this context, for example in the definition of tests discussed in Section 2.3. There are, of course, specific issues which deserve further research, for example statistics for heterogeneity, handling missing correlations or critical comparison of estimation methods. Other important issues specific to multivariate meta-analysis have been illustrated and discussed by Jackson and colleagues [11]. In the specific context of multi-parameter associations, a main limitation is related to dimensionality, as the number of outcome parameters which can be accommodated is currently limited. This and other issues will be hopefully addressed in future research on the development of this methodology.

References

[1] Normand SLT. Meta-analysis: formulating, evaluating, combining, and reporting. *Statistics in Medicine* 1999; **18**(3):321–359.

[2] Sutton AJ, Higgins JP. Recent developments in meta-analysis. *Statistics in Medicine* 2008; **27**(5):625–50.

[3] Stukel TA, Demidenko E, Dykes J, Karagas MR. Two-stage methods for the analysis of pooled data. *Statistics in Medicine* 2001; **20**(14):2115–2130.

[4] Simmonds MC, Higgins J, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clinical Trials* 2005; **2**(3):209–217.

[5] Thompson S, Kaptoge S, White I, Wood A, Perry P, Danesh J, et al. Statistical methods for the time-to-event analysis of individual participant data from multiple epidemiological studies. *International Journal of Epidemiology* 2010; **39**:1345–1359.

[6] Dominici F, Samet JM, Zeger SL. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. *Journal of the Royal Statistical Society: Series A* 2000; **163**(3):263–302.

[7] Berkey CS, Anderson JJ, Hoaglin DC. Multiple-outcome meta-analysis of clinical trials. *Statistics in Medicine* 1996; **15**(5):537–557.

[8] Berkey CS, Hoaglin DC, Antczak-Bouckoms A, Mosteller F, Colditz GA. Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine* 1998; **17**(22):2537–2550.

[9] Arends LR, Vokó Z, Stijnen T. Combining multiple outcome measures in a meta-analysis: an application. *Statistics in Medicine* 2003; **22**(8):1335–1353.

[10] Gasparrini A, Armstrong B, Kenward MG. Multivariate meta-analysis: a method to summarize non-linear associations. *Statistics in Medicine* 2011; **30**(20):To appear.

[11] Jackson D, Riley R, White IR. Multivariate meta-analysis: Potential and promise. *Statistics in Medicine* 2010; **30**(20):To appear. DOI: 10.1002/sim.4172.

[12] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria 2011. URL `http://www.R-project.org/`, ISBN 3-900051-07-0.

[13] Higgins J, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A* 2009; **172**(1):137–159.

[14] Nam IS, Mengersen K, Garthwaite P. Multivariate meta-analysis. *Statistics in Medicine* 2003; **22**(14):2309–2333.

[15] White IR. Multivariate random-effects meta-analysis. *Stata Journal* 2009; **9**(1):40–56.

[16] Verbeke G, Molenberghs G. *Linear mixed models for longitudinal data*. Springer Verlag: New York, 2000.

[17] Ritz J, Demidenkob E, Spiegelman D. Multivariate meta-analysis for data consortia, individual patient meta-analysis, and pooling projects. *Journal of Statistical Planning and Inference* 2008; **139**(7):1919–1933.

[18] Dear KBG. Iterative generalized least squares for meta-analysis of survival data at multiple times. *Biometrics* 1994; **50**:989–1002.

[19] Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine* 2010; **29**(12):1282–1297.

[20] Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* 1977; **72**(358):320–338.

[21] Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS*. Springer Verlag: New York, 2000.

[22] Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* 1997; **53**(3):983–997.

[23] Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics and Data Analysis* 2009; **53**(7):2583–2595.

[24] Stram DO, Lee JW. Variance components testing in the longitudinal mixed effects model. *Biometrics* 1994; **50**(4):1171–1177.

[25] Verbeke G, Molenberghs G. The use of score tests for inference on variance components. *Biometrics* 2003; **59**(2):254–262.

[26] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; **21**(11):1539–1558.

[27] Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on i2 in assessing heterogeneity may mislead. *BMC Medical Research Methodology* 2008; **8**(1):79.

[28] Dominici F. Time-series analysis of air pollution and mortality: a statistical review. *Research report - Health Effects Institute* 2004; **123**:3–27; discussion 29–33.

[29] Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology* 2006; **17**(6):624–31.

[30] Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology* 2009; **20**(2):205–213.

[31] Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Statistics in Medicine* 2010; **29**(21):2224–2234.

[32] Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. *Journal of Statistical Software* 2011; **43**(8):1–20. URL `http://www.jstatsoft.org/v43/i08/`.

[33] Schwartz J. Harvesting and long term exposure effects in the relation between air pollution and mortality. *American Journal of Epidemiology* 2000; **151**(5):440–8.

[34] Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F. Ozone and short-term mortality in 95 U.S. urban communities, 1987-2000. *Journal of the American Statistical Association* 2004; **292**(19):2372–8.

[35] Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Statistics in Medicine* 2003; **22**(19):2995–3016.

[36] Ades AE, Sutton AJ. Multiparameter evidence synthesis in epidemiology and medical decision-making: current approaches. *Journal of the Royal Statistical Society: Series A* 2006; **169**(1):5–35.

[37] Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods* 2010; **1**(1):39–65.

[38] Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Statistics in Medicine* 2002; **21**(3):371–87.

[39] Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *Journal of Clinical Epidemiology* 2002; **55**(1):86–94.

[40] Olkin I, Sampson A. Comparison of meta-analysis versus analysis of variance of individual patient data. *Biometrics* 1998; **54**(1):317–322.

[41] Mathew T, Nordstrom K. Comparison of one-step and two-step meta-analysis models using individual patient data. *Biometrical Journal* 2010; **52**(2):271–287.

[42] Riley RD, Thompson JR, Abrams KR. An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics* 2008; **9**(1):172.

[43] Riley RD. Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A* 2009; **172**(4):789–811.

[44] Peng RD, Dominici F, Louis TA. Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society: Series A* 2006; **169**(2):179–203.

[45] Samoli E, Analitis A, Touloumi G, Schwartz J, Anderson HR, Sunyer J, Bisanti L, Zmirou D, Vonk JM, Pekkanen J, *et al.*. Estimating the exposure-response relationships between particulate matter and mortality within the APHEA multicity project. *Environmental Health Perspectives* 2005; **113**(1):88–97.

[46] Baccini M, Biggeri A, Accetta G, Kosatsky T, Katsouyanni K, Analitis A, Anderson HR, Bisanti L, D'Ippoliti D, Danova J, *et al.*. Heat effects on mortality in 15 European cities. *Epidemiology* 2008; **19**(5):711–9.

[47] Dominici F, Daniels MJ, Zeger SL, Samet JM. Air pollution and mortality: estimating regional and national dose-response relationships. *Journal of the American Statistical Association* 2002; **97**:100–111.

[48] Schwartz J, Zanobetti A. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. *Epidemiology* 2000; **11**(6):666–72.

[49] Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *American Journal of Epidemiology* 1992; **135**(11):1301–1309.

[50] Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology* 1993; **4**(3):218–28.

[51] Bagnardi V, Zambon A, Quatto P, Corrao G. Flexible meta-regression functions for modeling aggregate dose-response data, with an application to alcohol and mortality. *American Journal of Epidemiology* 2004; **159**(11):1077–1086.

[52] Hartemink N, Boshuizen HC, Nagelkerke NJ, Jacobs MA, van Houwelingen HC. Combining risk estimates from observational studies with different exposure cutpoints: a meta-analysis on body mass index and diabetes type 2. *American Journal of Epidemiology* 2006; **163**(11):1042–52.

[53] Vlaanderen J, Portengen L, Rothman N, Lan Q, Kromhout H, Vermeulen R. Flexible meta-regression to assess the shape of the benzene-leukemia exposure-response curve. *Environmental Health Perspectives* 2010; **118**(4):526–532.

[54] Shi JQ, Copas JB. Meta-analysis for trend estimation. *Statistics in Medicine* 2004; **23**(1):3–19.

[55] Liu Q, Cook NR, Bergstrom A, Hsieh CC. A two-stage hierarchical regression model for meta-analysis of epidemiologic nonlinear dose-response data. *Computational Statistics and Data Analysis* 2009; **53**(12):4157–4167.

[56] Rota M, Bellocco R, Scotti L, Tramacere I, Jenab M, Corrao G, La Vecchia C, Boffetta P, Bagnardi V. Random-effects meta-regression models for studying nonlinear dose-response relationship, with an application to alcohol and esophageal squamous cell carcinoma. *Statistics in Medicine* 2010; **29**:2679–2687.

[57] Arends LR, Hunink MG, Stijnen T. Meta-analysis of summary survival curve data. *Statistics in Medicine* 2008; **27**(22):4381–4396.

[58] Ishak KJ, Platt RW, Joseph L, Hanley JA, Caro JJ. Meta-analysis of longitudinal studies. *Clinical Trials* 2007; **4**(5):525–39.

[59] Arends LR, Hamza TH, Van Houwelingen JC, Heijenbrok-Kal MH, Hunink MGM, Stijnen T. Bivariate random effects meta-analysis of ROC curves. *Medical Decision Making* 2008; **28**(5):621–638.

[60] Gasparrini A, Armstrong B. The impact of heat waves on mortality. *Epidemiology* 2011; **22**(1):68–73.

[61] Zanobetti A, Schwartz J, Samoli E, Gryparis A, Touloumi G, Atkinson R, Le Tertre A, Bobros J, Celko M, Goren A, *et al.*. The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement. *Epidemiology* 2002; **13**(1):87–93.

[62] Analitis A, Katsouyanni K, Biggeri A, Baccini M, Forsberg B, Bisanti L, Kirchmayer U, Ballester F, Cadum E, Goodman PG, *et al.*. Effects of cold weather on mortality: results from 15 European cities within the PHEWE Project. *American Journal of Epidemiology* 2008; **168**(12):1397.

[63] Zanobetti A, Schwartz J. Mortality displacement in the association of ozone with mortality: an analysis of 48 cities in the United States. *American Journal of Respiratory and Critical Care Medicine* 2008; **177**(2):184–9.

[64] Samoli E, Zanobetti A, Schwartz J, Atkinson R, Le Tertre A, Schindler C, Perez L, Cadum E, Pekkanen J, Paldy A, *et al.*. The temporal pattern of mortality responses to ambient ozone in the APHEA project. *Journal of Epidemiology and Community Health* 2009; **63**:960–966.

[65] Zanobetti A, Schwartz J. Temperature and mortality in nine US cities. *Epidemiology* 2008; **19**(4):563–70.

Supplementary Web Appendix for: "Multivariate meta-analysis for non-linear and other multi-parameter associations"

Antonio Gasparrini[1*,†], Ben Armstrong[1] and Michael G. Kenward[2]

[1]*Department of Social and Environmental Health Research*
[2]*Department of Medical Statistics*
London School of Hygiene and Tropical Medicine, UK

This web appendix contains some information on the estimation procedures adopted in the R package mvmeta (version 0.2.3), used to perform the analysis illustrated in the manuscript, and details on the related R code also provided as supplementary material. The package is under constant development, and some changes are likely to occur in future releases. Moreover, the usage of existing functions may also change, although portability of the existing code in future versions will be preserved whenever possible. For further information, type `help('mvmeta-package')` in R.

A Details about estimation procedures

In this section we provide some additional details on the estimation algorithms used in the current version of the package mvmeta, already discussed in Section 2.2 of the manuscript. As already mentioned, the unknown parameters to be estimated are $\boldsymbol{\beta}$ and, for random-effects meta-analytic models, $\boldsymbol{\xi}$, a set of components which uniquely define the between-study (co)variance matrix $\boldsymbol{\Psi}$.

The current implementation of mvmeta only supports an unstructured form for $\boldsymbol{\Psi}$, although options for additional structures will be added in the future versions. Actually, here $\boldsymbol{\Psi}$ is expressed in term of its Cholesky decomposition, with $\boldsymbol{\Psi} = \mathbf{R}^{\mathsf{T}}\mathbf{R}$, in order to assure positive-definiteness, and $\boldsymbol{\xi}$ corresponds to the $k(k+1)/2$ upper-triangular terms of $\mathbf{R}$. For computational convenience, the problem is re-arranged taking a second Cholesky decomposition of the marginal (co)variance matrix $\boldsymbol{\Sigma}_i = \mathbf{U}_i^{\mathsf{T}}\mathbf{U}_i$. The generalized least square problem in Eq. 5 of the manuscript, applied to obtain the conditional estimate of the fixed-effects coefficients $\boldsymbol{\beta}$, is then re-arranged as a simple least square fit procedure, carried out by minimizing the modified objective $\lambda = \sum_i |\mathbf{U}_i^{-\mathsf{T}}\hat{\boldsymbol{\theta}}_i - \mathbf{U}_i^{-\mathsf{T}}\tilde{\mathbf{X}}_i\boldsymbol{\beta}|$. An appropriate QR decomposition of the transformed objects $\mathbf{U}_i^{-\mathsf{T}}\hat{\boldsymbol{\theta}}$ and $\mathbf{U}_i^{-\mathsf{T}}\tilde{\mathbf{X}}$ is performed to guarantee stability. The related (co)variance matrix $V(\hat{\boldsymbol{\beta}})$ is also derived. See [1, pag. 13 and 49] for details.

The procedure above is used for fitting fixed-effects meta-analytic models. The random-effects counterparts are also specified in terms of the (co)variance components $\boldsymbol{\xi}$, and estimation is performed using iterative algorithms. As mentioned in the manuscript, ML models are fitted through a profiled (concentrated) likelihood approach, specifying the objective function $\ell(\boldsymbol{\xi})$ in Eq. 4 of the manuscript in terms of $\boldsymbol{\xi}$ only, while the conditional estimate $\boldsymbol{\beta}(\boldsymbol{\xi})$ is computed as above, and plugged in at each iteration [2, Chapter 2]. Estimation of models fitted through REML follows the same lines, using the objective function $\ell_R(\boldsymbol{\xi})$ in Eq. 6 of the manuscript. By excluding the parameters for the fixed part of the model, this method reduces the dimensionality of the optimization problem, in particular for meta-regression models, with obvious computational advantages.

*Correspondance: Antonio Gasparrini, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK
†E-mail: antonio.gasparrini@lshtm.ac.uk

In the current implementation of mvmeta, the maximization of the objective functions is obtained through a quasi-Newton iterative algorithm, a variation of the Newton-Raphson method [3], exploiting the built-in R function `optim()`. Briefly, in the quasi-approach, the computation of the updated guess only requires the vector of first partial derivative of $\ell(\boldsymbol{\xi})$ or $\ell_R(\boldsymbol{\xi})$ with respect to $\boldsymbol{\xi}$, while using an approximation of the inverse of the Hessian, the matrix of second partial derivative, obtained from previous iterations. The equations for the vectors of first partial derivatives are provided in [4]. Convergence of Newton methods is heavily dependent on optimal starting values $\boldsymbol{\xi}^{[0]}$: these are provided performing few runs of an iterative generalized least square algorithm [5, 6].

Missing values in the estimated outcome parameters or (co)variance matrix for study i are naturally handled in the optimization algorithms by excluding the corresponding entries of $\hat{\boldsymbol{\theta}}_i$ and rows of $\tilde{\mathbf{X}}_i$, although no missing parameter or element of the (co)variance matrices occur in the analysis proposed in this paper.

B The R code

The R scripts provided as supplementary material reproduce all the results illustrated in the manuscript, figures included. Although the code could have been written in a more concise and faster version, we have privileged clarity here. The R packages mvmeta, dlnm and NMMAPSlite, available on the R CRAN, need to be installed.

The first script is used to generate the data, producing a list of databases for the 98 NMMAPS cities included in the analysis and a database with the city-level meta-predictors used in multivariate meta-regression models. Additional meta-variables have been included, so the reader may extend the investigation. Note that the script takes several minutes to complete, as the data are downloaded by an external repository.

The second script performs the first-stage time-series Poisson model. It first produces the basis matrix for temperature and ozone using the function `crossbasis()` in the package dlnm. Although this package is expressly meant to be used for distributed lag (non-linear) models, it is applied also to produce the basis for the temperature spline, which is automatically centered and conveniently lagged. In addition, as described below, other functions in the package dlnm help extracting the parameters from the fitted model, and facilitates the prediction and plotting of the estimated exposure-response relationships. After the Poisson models are fitted, the estimated parameters for temperature and ozone are extracted and stored, together with associated (co)variance matrices.

The third scripts runs the second-stage models, namely multivariate meta-analyses and meta-regressions, and computes the predicted effects. This step is carried out through the functions in mvmeta and dlnm. After the models have been fitted through function `mvmeta()`, basis matrices are created to obtain the predictions for a set of values in the range of the original scale for temperature and ozone, built by `crossbasis()` using the same specifications as in the original bases used for estimation. Prediction is computed through the function `crosspred()`. Given the relative scale for temperature, the prediction is computed for values corresponding to percentiles of an average distribution. The script then provides the code to derive vectors of city-specific and BLUP estimated associations, the latter through the function `blup()`, and to store the results in matrices. Finally, prediction from meta-analytic models are produced. First, for models with no predictor, then for meta-regression models. For the latter, outcomes parameters and associated (co)variance matrices for specific percentiles of the meta-variables are predicted through the function `predict()`, then the association is predicted on the original scale of temperature and ozone through `crosspred()`.

The fourth script define some functions to perform likelihood ratio and Wald tests, and to carry out the multivariate Cochran heterogeneity test and the I^2 statistic. The functions are used to print the results in the figures. The function `qtest()` is used here. Proper functions to perform the other tests and statistics will be added in future releases of the mvmeta package. In particular, specific `anova()` methods will be provided for hypothesis tests and model comparison.

The fifth script produces the main results included in the manuscript, namely Table 1 and the Figures, conveniently saved as pdf files. The code exploits the `plot()` and `lines()` method functions for objects `crosspred` where the predictions have been saved, which facilitates the graphical representation.

The sixth script provides additional results presented in the manuscript, in particular the comparison with REML models and results from multivariable multivariate meta-regression.

References

[1] Wood SN. *Generalized Additive Models: an Introduction with R*. Chapman & Hall/CRC, 2006.

[2] Pinheiro JC, Bates DM. *Mixed-Effects Models in S and S-PLUS*. Springer Verlag: New York, 2000.

[3] Nocedal J, Wright SJ. *Numerical Optimization*. Springer: New York, 2006.

[4] Lindstrom MJ, Bates DM. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association* 1988; **83**(404):1014–1022.

[5] Goldstein H. Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika* 1986; **73**(1):43.

[6] Goldstein H. Restricted unbiased iterative generalized least-squares estimation. *Biometrika* 1989; **76**(3):622.

Chapter 10

Research paper V

Title: The short-term influence of temperature on daily mortality in the temperate climate of Montreal,Canada.

Author(s): Mark S.Goldberg, Antonio Gasparrini, Ben Armstrong, Marie-France Valois.

Journal/Publisher: Environmental Research.

Type of publication: Research paper.

Stage of publication: Published online on June 17, 2011 as doi:10.1016/j.envres.2011.05.022.

URL: `http://www.sciencedirect.com/science/article/pii/S0013935111001514`.

Academic peer-reviewed: Yes.

Copyright: Permission obtained from the publisher.

Candidate's role: See Section 2.3.

Candidate's signature:

Senior author: (Prof. Ben Armstrong)

Environmental Research 111 (2011) 853–860

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

The short-term influence of temperature on daily mortality in the temperate climate of Montreal, Canada ☆

Mark S. Goldberg [a,b,*], Antonio Gasparrini [c], Ben Armstrong [c], Marie-France Valois [a,b]

[a] Department of Medicine, McGill University, Montreal, Canada
[b] Division of Clinical Epidemiology, Department of Medicine, McGill University Health Center—RVH, 687 Pine Avenue West, R4.29, Montreal, Quebec, Canada H3A 1A1
[c] Social & Environmental Health Research Department, London School of Hygiene and Tropical Medicine, University of London, London, England

ARTICLE INFO

Article history:
Received 8 January 2011
Received in revised form
18 May 2011
Accepted 31 May 2011

Keywords:
Temperature
Mortality
Time series study
Distributed lag non-linear models
Air pollution

ABSTRACT

The purpose of this study was to determine whether short-term changes in ambient temperature were associated with daily mortality among persons who lived in Montreal, Canada, and who died in the urban area between 1984 and 2007. We made use of newly developed distributed lag non-linear Poisson models, constrained to a 30 day lag period, and we adjusted for temporal trends and nitrogen dioxide and ozone. We found a strong non-linear association with high daily maximum temperatures showing an apparent threshold at about 27 °C; this association persisted until about lag 5 days. For example, we found across all lag periods that daily non-accidental mortality increased by 28.4% (95% confidence interval: 13.8–44.9%) when temperatures increased from 22.5 to 31.8 °C (75–99th percentiles). This association was essentially invariant to different smoothers for time. Cold temperatures were not found to be associated with daily mortality over 30 days, although there was some evidence of a modest increased risk from 2 to 5 days. The adverse association with colder temperatures was sensitive to the smoother for time. For cardio-respiratory mortality we found increased risks for higher temperatures of a similar magnitude to that of non-accidental mortality but no effects at cold temperatures.

1. Introduction

Extreme weather conditions have been shown to increase daily mortality (Basu and Samet, 2002b; Gasparrini and Armstrong, 2010; Gosling et al., 2009). For example, the heat wave in Europe during the summer of 2003 may have caused an additional 22,000 deaths in France, Italy, Great Britain, and Spain (Conti et al., 2005; Kosatsky, 2005; le Tertre et al., 2006). These figures may be underestimates as not all heat-related deaths are recognized as such (Donoghue et al., 1997). The response to increasing temperature does not occur just at the upper range; there is a steep gradient in daily mortality and daily hospitalizations, usually above a location-specific "threshold" (Basu and Samet, 2002a; Gosling et al., 2009; Gouveia et al., 2003; Kovats et al., 2005, 1998; Martens, 1998; McMichael et al., 2006). The effects of increased temperatures are primarily found within a few days of the hot day, although longer lag effects are sometimes found, and the increased risks attenuate with increasing lag time (Braga et al., 2002; Conti et al., 2005; Curriero et al., 2002; Davis et al., 2003a, 2003b; Dessai, 2002; Gosling et al., 2009; Hajat et al., 2006; Hajat et al., 2005).

In addition, there are data suggesting that colder than normal temperatures can increase risk (Anderson and Bell, 2009; Curriero et al., 2002). These effects may be delayed for as many as two weeks into the future (Pattenden et al., 2003).

The heat events in Europe and elsewhere suggested that the elderly may be at higher risk, and this may have been due to a lack of support structures to ensure sufficient hydration and other measures to alleviate effects from extreme heat. As well, certain other sub-groups may be at higher risk; for example, a recent study from our group suggested that persons with congestive heart failure may be susceptible to increasing temperatures (Kolb et al., 2007). Other investigations of the effects of weather in this subpopulation have shown that the risk of hospitalization for congestive heart failure among elderly persons living in Denver, Colorado, during the summer increased monotonically with increasing maximum temperature (13% increase for a 5.3 °C increase) (Koken et al., 2003).

Abbreviations: df, degrees of freedom; dlnm, distributed lag non-linear models; ICD, International Classification of Diseases; NO_2, nitrogen dioxide; O_3, ozone; $PM_{2.5}$, particulate matter having an aerodynamic diameter of 2.5 μm or less

☆ **Funding sources**: None.

Ethics approval: Ethics approval was obtained from the Institutional Review Board of McGill University and from the Commission de l'accès à l'information du Québec. This research was completed in accordance with the Helsinki Declaration.

* Corresponding author at: Division of Clinical Epidemiology, Department of Medicine, McGill University Health Center—RVH, 687 Pine Avenue West, R4.29, Montreal, Quebec, Canada H3A 1A1. Fax: +514 843 1493.
E-mail address: mark.goldberg@mcgill.ca (M.S. Goldberg).

0013-9351/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.envres.2011.05.022

An analysis of the acute effects of weather needs to account for the fact that the response pattern for temperature is non-linear, the effects can be delayed in time, and different response functions may apply at different lag periods. In most of the previous studies the methods did not allow estimates of the complexities of these joint effects. Recently, Armstrong and colleagues (Armstrong, 2006; Gasparrini et al., 2010) have developed a formal theory and software for distributed lag non-linear models that is a generalization of the usual (linear) distributed lag linear model (Almon, 1965; Moshammer et al., 2006; Schwartz, 2000; Wyzga, 1978; Zanobetti et al., 2000). Using these new methods, our primary objective was to assess in Montreal, Canada, a city with a large range in daily temperatures, the duration and response function of the effect of higher temperature and to determine whether there were effects from colder temperatures.

2. Material and methods

2.1. The study population

The study population comprised residents of Montreal who died in the city between 1984 and 2007 of any non-accidental cause. Montreal has about two million inhabitants (in 2001) and they live in an area of about 500 km^2. The city is in a temperate zone and experiences both very cold and quite hot temperatures. A feature of the city and the province is that there is little air conditioning in homes (25.6% in the Province of Quebec in 2008 (Statistics Canada, 2010)) but that the buildings are well-heated during cold periods (usually mid-October until mid-April, with December to February being the coldest periods of the year).

Deceased subjects were identified from the computerized provincial database of death certificates and they were provided to us without personal identifiers. Approval to have access to the mortality data was granted by the provincial agency responsible for allowing access (Commission de l'accès à l'information du Québec) and ethical approval was granted by the Institutional Review Board of the Faculty of Medicine, McGill University.

2.2. Weather and air pollution data

Previous papers described in detail the environmental data (Goldberg et al., 2003, 2009). Daily weather data, comprising hourly measurements of temperature and other parameters, were provided by Environment Canada from their monitoring station located at the Pierre-Elliott-Trudeau International Airport (latitude: 45°28′05″N; longitude: 73°44′29″W) situated approximately 30 km west of downtown Montreal. We computed daily averages of temperature, humidex, humidity, and maximum temperature.

The air pollution data comprised bi-hourly or hourly measurements in Montreal of a number of criteria gaseous pollutants (sulfur dioxide, carbon monoxide, nitrogen dioxide (NO_2), and ozone (O_3)) at 12 fixed-site monitoring stations. We chose to include two of these as covariates in the substantive analysis: NO_2 was measured at eight stations and O_3 was measured at nine stations, both pollutants were measured using chemiluminescence (Thermo electron 14 V). Mean daily concentrations of NO_2 and O_3 were derived by taking a simple daily average for each monitor and then averaging these across monitors to obtain a final daily mean value. Respirable and fine particles were measured using high-volume samplers approximately every six days during 1984–2004 period and in 1996 these were replaced by tapered element oscillating microbalances. Because of the large number of missing days in the early part of the study period and the difficulty of combining high-volume samples with the measurements from the tapered element oscillating microbalances, we excluded fine particles from all analyses.

2.3. Statistical methods

We selected maximum temperature as the exposure metric and we assessed the association with non-accidental mortality using a time series approach (Goldberg et al., 2004) that has been generalized to handle the distributed lag non-linear models (Armstrong, 2006; Gasparrini et al., 2010). Specifically, we used quasi-likelihood Poisson regression in a generalized linear model to model the natural logarithm of daily counts of cause-specific deaths as functions of predictor variables. We accounted for the over-dispersed Poisson data by assuming that the total variance was proportional to the number of counts, with the over-dispersion constant estimated through quasi-likelihood. To remove seasonal and sub-seasonal cycles in the mortality time series, we included a natural cubic spline function on day of study and we included a factor for day-of-the-week. Following the analyses of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) (Dominici et al., 2004; Samet et al., 2000), we specified a "primary" model using a smoother for time of 7 degrees of freedom (df) per annum and we investigated the sensitivity of the results using temporal smoothers having 5, 9, and 13 df.

2.3.1. Other potential confounding variables

We accounted for the effects of air pollution by including mean daily concentrations of nitrogen dioxide and ozone. Our previous work showed that the effects of these two air pollutants were linear and that their effects persisted over the concurrent day (lag 0 days) and the two previous days (lags 1 and 2 days) (Brook et al., 2007; Goldberg et al., 2001d) We could not account for influenza epidemics, as monitoring and recording of these epidemics is not carried out routinely.

2.3.2. Distributed lag non-linear models of the effects of temperature on mortality

We made use of the distributed lag non-linear models developed by two of us (B.A. and A.G.; referred to as dlnm) (Armstrong, 2006; Gasparrini et al., 2010) to describe simultaneously non-linear and delayed dependencies in the association between mortality and temperature. Briefly, these models are a generalization of the traditional distributed lag models (Almon, 1965; Moshammer et al., 2006; Schwartz, 2000; Wyzga, 1978; Zanobetti et al., 2000) to allow the model to contain a flexible representation of the time-course of the exposure–response relationship, which also provides an estimate of the overall effect in the presence of delayed contributions or "harvesting". The dlnm allow for the simultaneous estimation of different non-linear functions of the associations with temperature at each lag period and also allows for the estimation of non-linear effects across lags. The methodology is based on the definition of a "cross-basis" function, a bi-dimensional space of functions specifying the possible non-linear association between temperature and mortality across lag periods. The cross-basis functions are combined from the basis functions for the two dimensions (temperature and lag), chosen among a set of possible bases. We used the dlnm package in the R project for statistical computing (version 2.10.1; http://www.r-project.org/) that was written by one of us (A.G.).

The amount of smoothing chosen for the temperature and lag spaces is independent, as they are modeled by two different functions. Having equally spaced knots over the temperature space does indeed imply similar degree of flexibility across the range, and one could use alternative positioning of the knots, a priori or data-based, although there are issues associated with selecting the knots. Our approach to investigating curvature assumptions has been through sensitivity analyses. The knots in the lag space were, however, placed unequally across the lag space, following the default in the dlnm package (equally on a logarithmic scale), to reflect greater expected smoothness as lags increase (e.g., smoother over lags 29–30 days versus lags 0–1 days).

Among the possible non-linear functions, including linear thresholds, polynomials, and spline transformations, we selected cubic b-splines to model the temperature effect, as they are flexible at the endpoints where some degree of non-linearity is expected. Using a dlnm model that was constrained to assess effects for a lag period of 30 days, we placed knots evenly across the range of maximum temperature and selected a priori, following the work of Gasparrini et al. (2010), the "primary" model having a natural cubic spline with 5 degrees of freedom (df) in the lag space (knots placed at logarithmically equal intervals) and a cubic b-spline with 6 df (three equally spaced knots) in the temperature space.

As sensitivity analyses, we also investigated b-splines having three knots chosen from the quantiles of the temperature distribution (6 df total) and we investigated threshold models. Analyses were also conducted in the smaller group of individuals who were under age 65 years at time of death. In addition, we conducted analyses that were not adjusted for air pollution to determine the extent of confounding on temperature.

3. Results

Table 1 shows that during the study period, 1984–2007 (comprising 8766 days), the average maximum daily temperature was 11.5 °C (average mean daily temperature was 6.8 °C), varying from −23.9 to 36.2 °C (interquartile range of 20.6 °C). Air pollution was relatively low in Montreal as compared to most North American cities, with mean daily concentrations of NO_2 and O_3 of 38 and 33 $\mu g/m^3$, respectively. Table 2 shows that the different metrics for temperature were highly correlated (Pearson correlation coefficients of 0.99) and that, as expected, NO_2 and O_3 were negatively and positively correlated with temperature, respectively. In what follows, we will only show results for maximum temperature.

Table 3 shows the distributions of mortality from non-accidental causes and from cardiovascular diseases (International Classification of Diseases (ICD), revision-9 390–459; ICD-10 I00–I99) and

Table 1
Distribution of selected weather and air pollution variables, Montreal, 1984–2007.

	Units	Number of days of measurements	Mean	Standard deviation	Minimum	Percentiles				Interquartile range
						25th	50th	75th	100th	
Maximum temperature	°C	8703	11.55	12.36	−23.9	1.9	12.4	22.5	36.2	20.6
Mean temperature	°C	8700	6.78	11.86	−27.3	−2.1	7.7	17.1	29.3	19.2
Minimum temperature	°C	8720	1.99	11.61	−31.8	−6.1	2.9	11.7	24.6	17.8
Maximum humidex[a]	°C	8766	11.14	15.55	−29.4	−1.1	11.0	24.6	46.1	25.7
Mean relative humidity	%	8766	69.57	12.41	28	61	70	78	100	17
Change in pressure in 24 h ending at 08:00	kPa	8756	0.00	0.92	−4.22	−0.54	0.00	0.54	5.03	1.08
NO_2	$\mu g/m^3$	8764	37.99	14.95	7.34	27.36	35.88	45.91	165.67	18.55
O_3	$\mu g/m^3$	8764	32.77	18.00	1.86	19.55	30.28	42.96	163.93	23.41

[a] Humidex is calculated as mean temperature (°C) + 0.5555 (6.11E − 10) where E = exp(5417.753 × (1/273.16)) − (1/Dew Point Temperature (°K)) (see for the definition http://www.weatheroffice.gc.ca/mainmenu/faq_e.html; accessed June 2011).

Table 2
Pearson correlation coefficients between selected weather and air pollution variables, Montreal, 1984–2007.

	Daily maximum temperature (°C)	Daily mean temperature (°C)	Daily minimum temperature (°C)	Daily humidex (°C)	NO_2 ($\mu g/m^3$)	O_3 ($\mu g/m^3$)
Daily maximum temperature	1	0.99	0.95	0.99	−0.21	0.41
Daily mean temperature		1	0.99	0.99	−0.25	0.39
Daily minimum temperature			1	0.96	−0.26	0.35
Daily humidex				1	−0.23	0.40
NO_2					1	−0.20

Table 3
Distribution of mortality from non-accidental causes, cardiovascular diseases, and respiratory diseases, by age and sex, Montreal, 1984–2007.

	Number of days of measurements	Mean	Standard deviation	Minimum	Percentiles				Interquartile range
					25th	50th	75th	100th	
Non-accidental mortality									
All	8766	38.08	7.60	10	33	38	43	95	10
< 65 years of age	8766	7.80	3.01	0	6	8	10	23	4
≥ 65 years	8766	30.29	6.92	6	26	30	35	79	9
Men	8766	18.59	4.75	4	15	18	22	43	7
Women	8766	19.50	5.17	2	16	19	23	52	7
Respiratorymortality									
All	8766	3.38	2.12	0	2	3	5	15	3
< 65 years of age	8766	0.34	0.59	0	0	0	1	4	1
≥ 65 years	8766	3.04	2.00	0	2	3	4	15	2
Men	8766	1.74	1.40	0	1	2	3	9	2
Women	8766	1.64	1.41	0	1	1	2	11	1
Cardiovascular mortality									
All	8766	14.34	4.58	1	11	14	17	48	6
< 65 years of age	8766	2.11	1.57	0	1	2	3	11	2
≥ 65 years	8766	12.23	4.08	1	9	12	15	40	6
Men	8766	6.83	2.89	0	5	7	9	22	4
Women	8766	7.51	3.06	0	5	7	9	33	4
Cardio-respiratory mortality									
All	8766	17.73	5.34	2	14	17	21	52	7
< 65 years of age	8766	2.45	1.70	0	1	2	3	13	2
≥ 65 years	8766	15.27	4.82	2	12	15	18	50	6
Men	8766	8.57	3.34	0	6	8	11	24	5
Women	8766	9.15	3.47	1	7	9	11	34	4

respiratory diseases (ICD-9 460–519; ICD-10 J00–J99). The mean number of daily non-accidental deaths was 38.1 and the variance was 57.8. The daily mean number of deaths from respiratory diseases was 3.4 and from cardiovascular diseases it was 14.3. (Time series plots for the endpoints and for temperature are shown in Supplementary Annex Figs. 1 and 2.)

The model that accounted only for seasonal and secular trends (a time smoother of 7 df per annum and a term for day-of-the-week) had an over-dispersion parameter of 1.11 and a serial autocorrelation coefficient that was close to zero by lag 5 days. (See Supplementary Annex Table 1 for these parameters across all of the time smoothers used: 5, 7, 9, and 13 df.)

Fig. 1 shows a three-dimensional plot of non-accidental mortality and temperature that was modeled as a cubic b-spline having three equally spaced knots (total of 6 df), constrained to a lag period of 30 days. This model included natural cubic splines for the temporal

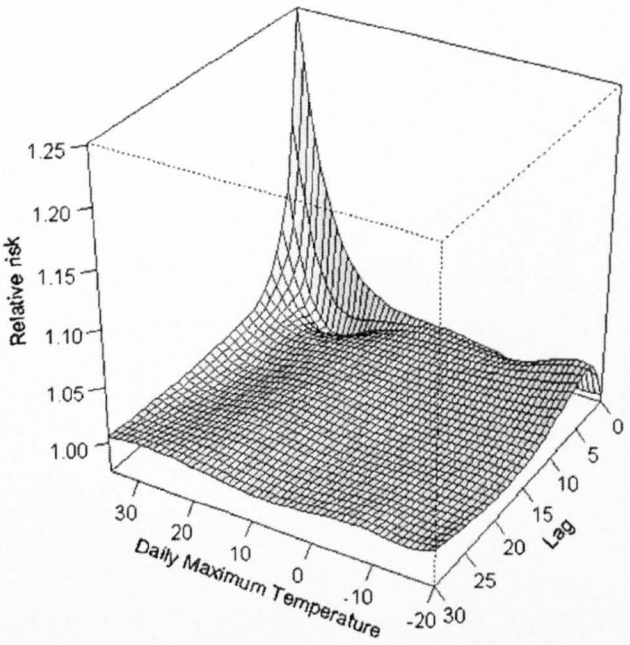


Fig. 1. The relative risk of daily non-accidental mortality and maximum temperature by lag period, from a distributed lag non-linear model, adjusted for nitrogen dioxide and ozone, Montreal, 1984–2007. The model comprised a smoother for time of 7 df per annum, three equally spaced knots for the effect of temperature (6 df total) and 5 df for the lag space. The *z*-axis represents the relative increase in daily counts of mortality with respect to the mean value (11.55 °C) and the other axes represent maximum temperature and lag period.

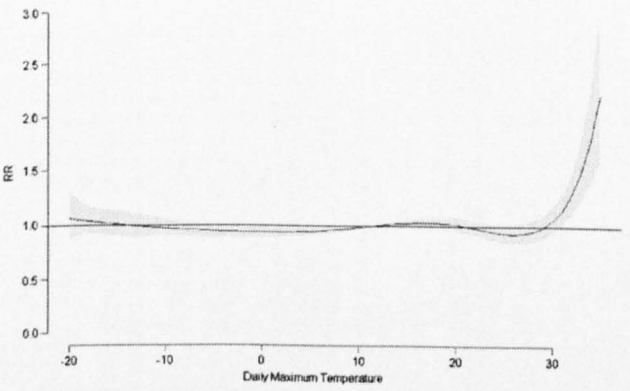


Fig. 2. Cumulative effects between daily non-accidental mortality and maximum temperature, from a distributed lag non-linear model, adjusted for nitrogen dioxide and ozone, Montreal, 1984–2007. The model comprised a smoother for time of 7 df per annum, three equally spaced knots for the effect of temperature (6 df total) and 5 df for the lag space. The *y*-axis represents the relative increase in daily counts of mortality with respect to the mean value (11.55 °C). The maximum likelihood estimate is shown as a smooth line and the pointwise 95% confidence intervals are shown in the shaded area.

smoother (7 df) and for the lag space (5 df), a term for day-of-the-week, as well as terms for the two air pollutants. The figure shows the strong effect at high temperatures that persisted up to lag 5 days and a cold effect starting at about −15 °C between lags 2 and 5 days. This graph must be interpreted cautiously as it is not possible to provide estimates of variability, but it does show the general pattern of risk by lag and by temperature. The following results help explain this overall pattern.

Based on the same model, Fig. 2 shows the fitted cumulative distributed non-linear lag function. This function is interpreted as the total effect on mortality on the concurrent day from the effects of temperature accumulated over the concurrent day and out to lag 30 days, inclusive, in the hypothetical case in which temperature is constant over that period. The relative increase in the number of daily deaths (referred to as "relative risk") for temperature is compared to the average maximum temperature of 11.55 °C. At high maximum temperatures, there is a strong monotonic increase in the number of deaths starting at about 27 °C. (Similar effects were observed for other temporal smoothers and other smooth functions for temperature; Supplementary Annex Fig. 4.) We also found a small non-significant cold effect at about −18 °C. Most of the other models for which we used different temporal smoothers showed no or protective overall effects at colder temperatures. The protective effects for colder temperatures were more pronounced as the number of df on the smoother for time increased (Supplementary Annex Fig. 4).

To illustrate the delayed effects of maximum temperature on mortality, we show the response function at lag 4 days (Fig. 3).

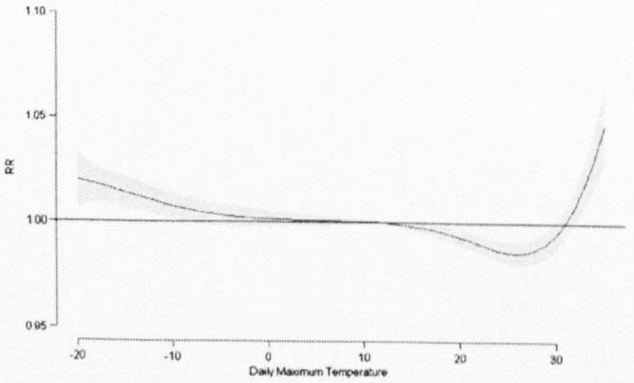


Fig. 3. Effects on daily non-accidental mortality evaluated at lag 4 days and maximum temperature, from a distributed lag non-linear model, adjusted for nitrogen dioxide and ozone, Montreal, 1984–2007. The model comprised a smoother for time of 7 df per annum and three equally spaced knots for the effect of temperature (6 df total) and 5 df for the lag space. The *y*-axis represents the excess mortality with respect to the mean value (11.55 °C). The maximum likelihood estimate is shown as a smooth line and the pointwise 95% confidence intervals are shown in the shaded area.

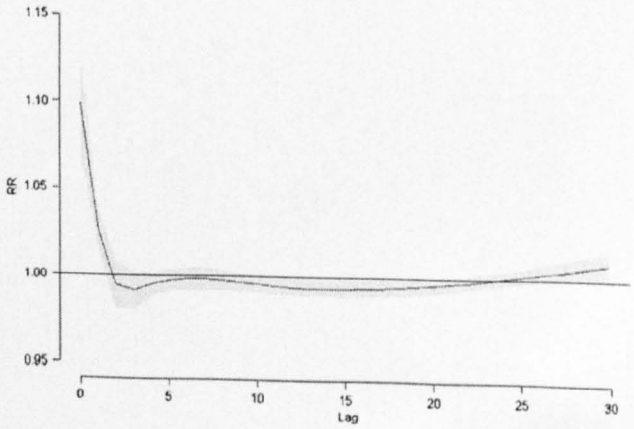


Fig. 4. Effects on daily non-accidental mortality evaluated at warm maximum temperatures (30 °C), from a distributed lag non-linear model, adjusted for nitrogen dioxide and ozone, Montreal, 1984–2007. The model comprised a smoother for time of 7 df per annum and three equally spaced knots for the effect of temperature (6 df total) and 5 df for the lag space. The *y*-axis represents the relative increase in daily counts of mortality with respect to the mean value (11.55 °C). The maximum likelihood estimate is shown as a smooth line and the pointwise 95% confidence intervals are shown in the shaded area.

A pronounced effect at high temperatures was found (and in all models) as well as a cold effect starting about −10 °C. The effect at cold temperatures vanished when 13 df per year was used as the temporal smoother (Supplementary Annex Fig. 5).

Fig. 4 shows the increased risk of high maximum daily temperatures (30 °C) relative to the average maximum temperature (11.55 °C). Effects were found for lags 0 and 1 days independent of the model (Supplementary Annex Table 4 and Fig. 6), with protective effects seen from lags 2–23 days, and a suggestion of an increase starting at lag 25 days.

Fig. 5 shows the results comparing a temperature of −15 °C to the average maximum. The maximum cold effect occurred at lag 3 days, with the models having smoothers for time of 5 and 7 df per year showing significant effects (Supplementary Annex Fig. 7 and Table 3). The figure also shows a slight apparent protective effect from lags 11–27 days, and those models with 9 and 13 df showed much stronger protective effects (in the Supplementary Annex).

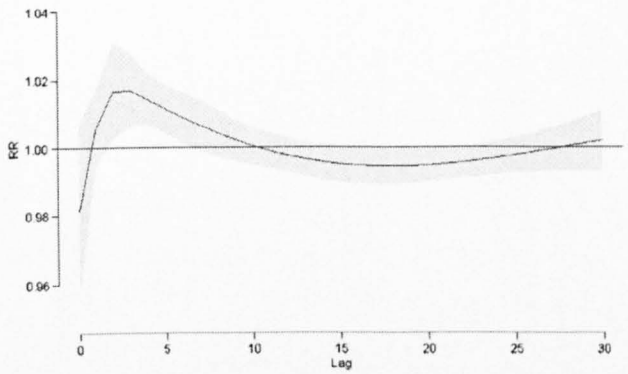


Fig. 5. Effects on daily non-accidental mortality evaluated at cold temperatures (−15 °C), from a distributed lag non-linear model, adjusted for nitrogen dioxide and ozone, Montreal, 1984–2007. The model comprised using a smoother for time of 7 df per annum and three equally spaced knots for the effect of temperature (6 df total) and 5 df for the lag space. The *y*-axis represents the relative increase in daily counts of mortality with respect to the mean value (11.55 °C). The maximum likelihood estimate is shown as a smooth line and the pointwise 95% confidence intervals are shown in the shaded area.

Table 4 summarizes the results of these figures by tabulating the percentage change in daily non-accidental mortality for the cumulative distributive lag model as well as for lagged effects from 0 days to 14 days. The first column compares the 1st percentile to the tenth percentile (risks at −16.3 °C relative to −5.3 °C) and it shows essentially the cold effects, with small increases in risk found at lags 2–5 days. (The positive sign on the percentage change is interpreted as an increase in risk as temperatures decrease.) The cumulative effect for cold temperatures was large but had considerable statistical variability; significant effects were found, however, from around lag 2 days until lag 5 days.

The second and third columns show the effects of heat on daily mortality, comparing the 99th percentile (31.8 °C) to the 75th (22.5 °C) and to the 90th (26.9 °C) percentiles, respectively. For hot temperatures, we found strong positive increases in daily mortality for the cumulative model (28.4% and 34.3% increases, respectively) and the effects declined in magnitude but persisted until lag 6 days. (These results are slightly different than shown in the figures as we are now comparing different temperature ranges, but they are derived from the same statistical model.)

3.1. *Sensitivity analyses for non-accidental mortality*

We found that the effects of heat were fairly insensitive to the smoother for time used but that there was some variation in the cold effects (Supplementary Annex Tables 3–7). In particular, use of smoothers for time having more than 7 df removed the deleterious cold effect and showed protective effects at higher lags. Although it is difficult to assess which models are preferable, it is possible that the smoothers for time using 9 and 13 df are over-fitting the data as the serial autocorrelation coefficients within seven day lags are mostly negative (Supplementary Annex Table 1).

We also made use of other smoothers for temperature (cubic b-splines with three knots based on quantiles of the distribution of temperature with 5 or 6 df) and threshold models, but we did not find any important differences in the response functions

Table 4
Percentage change in daily non-accidental mortality, and associated 95% confidence intervals (CI), for changes in maximum temperature between selected cut-points in the distribution, adjusted for nitrogen dioxide and ozone[a], Montreal, 1984–2007.

Lagged effect (days)	1st percentile relative to the 10th percentile[b]		99th percentile relative to the 75th percentile[c]		99th percentile relative to the 90th percentile[d]	
	% Change	95% CI	% Change	95% CI	% Change	95% CI
Cumulative	7.80	−3.87–20.90	28.40	13.76–44.93	34.34	19.18–51.43
0 days	−0.61	−2.96–1.81	11.93	8.95–15.00	8.48	6.08–10.93
1	1.13	−0.02–2.30	6.88	5.67–8.10	6.14	5.10–7.19
2	1.79	0.40–3.20	3.90	2.48–5.35	4.34	3.11–5.57
3	1.60	0.56–2.65	2.54	1.48–3.60	3.05	2.13–3.97
4	1.24	0.50–1.99	1.76	1.06–2.48	2.11	1.45–2.78
5	0.96	0.22–1.70	1.18	0.51–1.85	1.44	0.78–2.09
6	0.74	−0.03–1.53	0.75	0.04–1.46	0.97	0.28–1.66
7	0.59	−0.18–1.36	0.44	−0.27–1.16	0.67	−0.02–1.36
8	0.47	−0.23–1.18	0.23	−0.43–0.90	0.49	−0.14–1.13
9	0.39	−0.23–1.01	0.09	−0.50–0.68	0.40	−0.17–0.97
10	0.32	−0.22–0.87	−0.01	−0.54–0.52	0.35	−0.16–0.87
11	0.27	−0.25–0.78	−0.09	−0.59–0.41	0.31	−0.18–0.81
12	0.21	−0.29–0.72	−0.16	−0.66–0.33	0.28	−0.21–0.77
13	0.16	−0.35–0.68	−0.22	−0.72–0.28	0.25	−0.26–0.75
14	0.12	−0.41–0.65	−0.26	−0.77–0.25	0.21	−0.30–0.73

[a] Model included a cubic b-spline using three equally spaced knots (total of 6 df) for maximum temperature, a natural cubic spline with 5 df for the lag space, a natural cubic spline with 7 df per year for the time filter, day of the week, and NO_2 and O_3.

[b] 10th percentile = −5.3 °C, 1st percentile = −16.3 °C.

[c] 75th percentile = 22.5 °C, 99th percentile = 31.8 °C.

[d] 90th percentile = 26.9 °C, 99th percentile = 31.8 °C.

(selected results shown in Supplementary Annex Tables 3–7, sensitivity model 4, and under additional sensitivity analyses). Results of analyses amongst persons who died under the age of 65 years of age had large variability and were not informative and we did not find major differences between men and women (data not shown). The findings did not change when we did not adjust for the two air pollution variables, NO_2 and O_3 (Supplementary Annex Table 8). We also found similar response functions for different metrics of temperature (see additional sensitivity analyses).

3.2. Analyses of cardio-respiratory mortality

Because of uncertainties regarding the actual underlying cause of death, we combined deaths from cardiovascular and respiratory diseases. Fig. 6 shows the cumulative effects for cardio-respiratory mortality for a model using the same parameters as used in the analyses of non-accidental mortality. We did not find any effects at colder temperatures but we observed effects from hot weather that persisted to lag 5 days (Table 5).

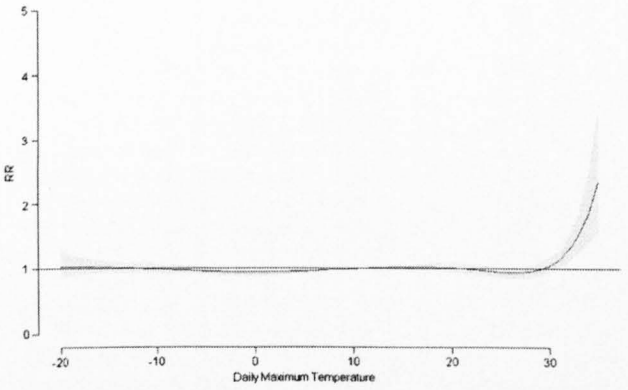


Fig. 6. Cumulative effects between daily mortality from cardio-respiratory diseases and maximum temperature, from a distributed lag non-linear model, adjusted for nitrogen dioxide and ozone, Montreal, 1984–2007. The model comprised a smoother for time of 7 df per annum, three equally spaced knots for the effect of temperature (6 df total) and 5 df for the lag space. The y-axis represents the relative increase in daily counts of mortality with respect to the mean value (11.55 °C). The maximum likelihood estimate is shown as a smooth line and the pointwise 95% confidence intervals are shown in the shaded area.

4. Discussion

As in most other studies, we have found that higher temperatures conferred excess risks of daily deaths from non-accidental causes and from cardio-respiratory disease, with the models being fairly robust to various specifications. The approximate "threshold" of maximum temperature for cumulative effects for non-accidental deaths over a 30 day period was about 27 °C (the 91st percentile of the temperature distribution). We did not find any cumulative effects at the colder end of the spectrum, although increased risks were apparent between lags 2 and 5 days. These effects were not, however, robust to the type of smoother for time used, with smoothers of more than 7 df per annum causing these effects to disappear. The analyses of cardio-respiratory mortality showed similar effects to that of non-accidental mortality, but there were no apparent effects at colder temperatures.

In the analyses of the cold effects, the occurrence of lagged effects with no cumulative effect is consistent with a "harvesting" effect, whereby the dates of death are moved up just a few days among a subpopulation at higher risk.

Association of elevated mortality with cold temperatures has been reported widely (e.g., Curriero et al., 2002; Keatinge et al., 1997; Analitis et al., 2008; Anderson and Bell, 2009; Barnett et al., 2005; Curriero et al., 2002; Keatinge et al., 1997). Studies from Europe (Pattenden et al., 2003; Analitis et al., 2008) have often found effects delayed by two weeks or more. In North America there seems to be less evidence of such a long delay, though Anderson and Bell (2009) did find effects up to two weeks. The absence of a strong association of cold temperatures with elevated mortality in Montreal is thus unusual. However, in Montreal, in contrast to many cities in milder climates, all homes are well-heated, so that this adaptation to continuously uncomfortable climatic conditions may explain the lack of a pronounced cold effect. The very small effects of

Table 5
Comparison of the estimated percentage change in daily mortality by cause of death, and associated 95% confidence intervals (CI), for changes in maximum temperature between the 75th and 99th percentiles and between the 10th and 1st percentiles, adjusted for nitrogen dioxide and ozone,[a] Montreal, 1984–2007.

Lagged effect (days)[a]	99th percentile relative to the 75th percentile				1st percentile relative to the 10th percentile			
	Non-accidental[b]		Cardio-respiratory[b]		Non-accidental[b]		Cardio-respiratory[b]	
	% Change	95% CI	% Change	95% CI	% Change	95% CI	% Change	95% CI
Cumulative	28.40	13.76–44.93	24.01	3.40–48.74	7.80	−3.87–20.90	6.39	−9.55–25.14
0	11.93	8.95–15.00	10.22	5.91–14.70	−0.61	−2.96–1.81	−0.08	−3.38–3.34
1	6.88	5.67–8.10	9.22	7.42–11.05	1.13	−0.02–2.30	0.86	−0.75–2.51
2	3.90	2.48–5.35	7.32	5.17–9.50	1.79	0.40–3.20	1.24	−0.70–3.22
3	2.54	1.48–3.60	4.96	3.38–6.56	1.60	0.56–2.65	1.16	−0.29–2.63
4	1.76	1.06–2.48	2.96	1.91–4.02	1.24	0.50–1.99	0.95	−0.09–2.01
5	1.18	0.51–1.85	1.55	0.55–2.55	0.96	0.22–1.70	0.75	−0.29–1.79
6	0.75	0.04–1.46	0.61	−0.45–1.67	0.74	−0.03–1.53	0.55	−0.54–1.65
7	0.44	−0.27–1.16	0.04	−1.01–1.10	0.59	−0.18–1.36	0.36	−0.72–1.46
8	0.23	−0.43–0.90	−0.26	−1.24–0.73	0.47	−0.23–1.18	0.19	−0.80–1.20
9	0.09	−0.50–0.68	−0.38	−1.25–0.51	0.39	−0.23–1.01	0.04	−0.83–0.92
10	−0.01	−0.54–0.52	−0.41	−1.19–0.39	0.32	−0.22–0.87	−0.09	−0.86–0.69
11	−0.09	−0.59–0.41	−0.42	−1.17–0.33	0.27	−0.25–0.78	−0.19	−0.91–0.53
12	−0.16	−0.66–0.33	−0.44	−1.18–0.30	0.21	−0.29–0.72	−0.27	−0.99–0.44
13	−0.22	−0.72–0.28	−0.46	−1.21–0.30	0.16	−0.35–0.68	−0.33	−1.06–0.40
14	−0.26	−0.77–0.25	−0.48	−1.25–0.30	0.12	−0.41–0.65	−0.37	−1.12–0.38

[a] 1st percentile = −16.3 °C, 10th percentile = −5.3 °C, 75th percentile = 22.5 °C, and 99th percentile = 31.8 °C.

[b] Main model using a cubic spline (BS) three equally spaced knots with a total of 6 df for maximum temperature, a cubic spline (NS) with 5 df for the lag space, cubic spline (NS) with 7 df per year for the time filter and is also adjusted for day of the week, NO_2 and O_3.

cold on mortality has also been observed in some other very cold climates, for example in Finland (Keatinge et al., 1997), but the only publication we are aware of reporting no association is in Yakutsk, Siberia (Donaldson et al., 1998). In an analysis of the MONICA project, Barnett et al. (2005) found that rates of morbidity and mortality from coronary problems, mostly myocardial infarctions, were higher amongst persons living in warmer climates as compared to those living in colder climates.

The "adaptation" to cold is mirrored by a lack of adaptation to heat. As we noted in Section 1, air conditioning of homes is rather limited in Quebec, and thus many individuals will be exposed to higher temperatures and thus the strong effect at high temperatures is indeed plausible physiologically and is certainly consistent with the literature.

We note that the selection of the smoothing functions is critical regarding the shape of the curve near the tails. Natural cubic splines are constrained to define a linear relationship beyond the boundaries and this often affects the shape near the ends. Although it is frequently reported that natural cubic splines have an "optimal behavior" in the tails, meaning that they are less prone to the effect of outliers and more able to capture the true curve, there is actually very little written on this. One of us (Gasparrini) has conducted some simulations (unpublished) and the natural cubic splines performed worse (by an Aikaike Information Criterion) than unconstrained cubic b-splines of the same df. In particular, it is true that the linearity constraint on the natural cubic splines could produce some underestimate of the width of the confidence intervals near the tails. This is why we have preferred simple cubic b-splines for modeling the relationship in the space of the predictor within dlnms.

There are some limitations that need to be considered in interpreting these results. We discussed the issue of misclassification of causes of death previously (Goldberg et al., 2001b), where we indicated that respiratory and cardiovascular diseases are often confused because the conditions can occur concurrently and both can contribute to death, so that there may be some uncertainty about which cause should be selected as the primary underlying cause. As well, we suggested that there may be errors in selecting one underlying cause in a complex chain of health events (e.g., cancer leading to pneumonia and then to respiratory failure).

In our analysis of air pollution in Montreal (Goldberg et al., 2000, 2001a, 2001c, 2003), we have found much higher risks in some sub-populations, such as those with diabetes and cardiovascular disease and those with congestive heart failure. Indeed, in a case-crossover analysis of the sub-group of persons who died between 1984 and 1993 from non-accidental causes but who had congestive heart failure one year before death, the adjusted odds ratios comparing temperatures between 30 and 35 °C were 1.08, 1.22, and 1.13 for the concurrent day and lags 1 and 2 days, respectively (Kolb et al., 2007). We also found in these analyses a delayed cold effect in the colder seasons of the year.

In the present analyses, we made use of distributed lag regression models to identify possible associations. Although these analyses are complex, and have many tunable parameters, our extensive sensitivity analyses indicate that the findings are robust. Although the cold effect did disappear with the use of temporal smoother that explain more of the short-term variation (i.e., 9 df or more per annum), it is possible that these smoothers are obscuring important signals from the data and, thus, may lead to biased findings.

The effects of air pollution were included in our analyses as possible confounding variables. We could not account for daily variations of concentrations of fine particles during the study period because these were measured every six days, so there was a considerable amount of missing data. We adjusted for NO_2 and O_3 because NO_2 is somewhat higher in the colder months, as are particles, and ozone is high in the warmer months. The Pearson correlation coefficients between the high-volume sampling of $PM_{2.5}$ and NO_2 was 0.61 and for O_3 it was −0.01; the correlation between $PM_{2.5}$ measured by the tapered element oscillating microbalances and NO_2 was 0.54, for ozone it was 0.13, and for high-volume samples for $PM_{2.5}$ it was 0.89. It is also possible that the effect of pollution is on a causal pathway between weather, indexed by temperature, and mortality (weather causes fluctuations in concentrations), in which case controlling for these variables may not be warranted. However, we found that adjustments for them did not greatly change the unadjusted estimates of effect, although we cannot exclude the possibility of sensitivity to control for fine particles. Our findings that air pollution did not confound the association are consistent with that of other studies, notably of the 107 American cities included in the analysis of Anderson and Bell (2009).

We could not control explicitly for the effects of infectious disease epidemics (e.g., influenza, which occurs mostly in the fall and winter) because there are no databases that could be used for this purpose. However, the smooth function of time should have eliminated most such residual secular effects, and there is no reason to expect an association of influenza with cold temperatures (after accounting for season).

We have shown that there are indeed heat islands in the city and that slightly stronger response functions for mortality were found in areas where temperatures were generally higher (Smargiassi et al., 2009). The analysis presented herein ignored these local effects and, given that the datasets used in the two papers overlapped, we may be underestimating effects in heat islands.

In summary, we have found that in Montreal hot weather was clearly associated with increases in short-term risk of mortality, but cold weather was associated with at most a small association with increased risk and at an intermediate lag with subsequent compensating decreased risk.

Acknowledgments

We thank the Montreal Urban Community, and Environment Canada for providing the air pollution data and the meteorological data and we are grateful to the Ministère de la santé et des services sociaux de Québec for providing the health data. We thank Chia-Yi Tseng for assistance with the statistical analyses.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.envres.2011.05.022.

References

Almon, S., 1965. The distributed lag between capital appropriations and expenditures. Econometrica 33, 178–196.

Analitis, A., et al., 2008. Effects of cold weather on mortality: results from 15 European cities within the PHEWE project. American Journal of Epidemiology 168, 1397.

Anderson, B., Bell, M., 2009. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. Epidemiology 20, 205.

Armstrong, B., 2006. Models for the relationship between ambient temperature and daily mortality. Epidemiology 17, 624–631.

Barnett, A.G., et al., 2005. Cold periods and coronary events: an analysis of populations worldwide. Journal of Epidemiology and Community Health 59, 551–557.

Basu, R., Samet, J.M., 2002a. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. Epidemiologic Reviews 24, 190–202.

Basu, R., Samet, J.M., 2002b. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. Epidemiologic Reviews 24, 190–202.

Braga, A.L., et al., 2002. The effect of weather on respiratory and cardiovascular deaths in 12 U.S. cities. Environmental Health Perspectives 110, 859–863.

Brook, J.R., et al., 2007. Further interpretation of the acute effect of nitrogen dioxide observed in Canadian time-series studies. Journal of Exposure Science and Environmental Epidemiology 17, S36–S44.

Conti, S., et al., 2005. Epidemiologic study of mortality during the summer 2003 heat wave in Italy. Environmental Research 98, 390–399.

Curriero, F.C., et al., 2002. Temperature and mortality in 11 cities of the eastern United States. American Journal of Epidemiology 155, 80–87.

Davis, R.E., et al., 2003a. Changing heat-related mortality in the United States. Environmental Health Perspectives 111, 1712–1718.

Davis, R.E., et al., 2003b. Decadal changes in summer mortality in U.S. cities. International Journal of Biometeorology 47, 166–175.

Dessai, S., 2002. Heat stress and mortality in Lisbon Part I. model construction and validation. International Journal of Biometeorology 47, 6–12.

Dominici, F., et al., 2004. Hierarchical bivariate time series models: a combined analysis of the effects of particulate matter on morbidity and mortality. Biostatistics 5, 341–360.

Donaldson, G.C., et al., 1998. Cold related mortalities and protection against cold in Yakutsk, eastern Siberia: observation and interview study. British Medical Journal 317, 978–982.

Donoghue, E.R., et al., 1997. Criteria for the diagnosis of heat-related deaths: National Association of Medical Examiners. Position paper. National Association of Medical Examiners Ad Hoc Committee on the Definition of Heat-Related Fatalities. American Journal of Forensic Medicine & Pathology 18, 11–14.

Gasparrini, A., Armstrong, B., 2010. Time series analysis on the health effects of temperature: advancements and limitations. Environal Research 110, 633–638.

Gasparrini, A., et al., 2010. Distributed lag non-linear models. Statistics in Medicine 29, 2224–2234.

Goldberg, M.S., et al., 2000. Identifying subgroups of the general population that may be susceptible to short-term increases in particulate air pollution: a time-series study in Montreal. Quebec Reseach Report. Health Effects Institute, 7–113.

Goldberg, M.S., et al., 2001a. Identification of persons with cardiorespiratory conditions who are at risk of dying from the acute effects of ambient air particles. Enviromental Health Perspectives 109 (Suppl. 4), 487–494.

Goldberg, M.S., et al., 2003. Associations between ambient air pollution and daily mortality among persons with congestive heart failure. Environmental Research 91, 8–20.

Goldberg, M.S., et al., 2001b. The association between daily mortality and ambient air particle pollution in Montreal, Quebec. 2. Cause-specific mortality. Environmental Research 86, 26–36.

Goldberg, M.S., et al., 2001c. Identification of persons with cardiorespiratory conditions who are at risk of dying from the acute effects of ambient air particles. Environmental Health Perspectives 109 (Suppl-94).

Goldberg, M.S., et al., 2001d. Associations between daily cause-specific mortality and concentrations of ground-level ozone in Montreal, Quebec. American Journal of Epidemiology 154, 817–826.

Goldberg, M.S., et al., 2004. A review of time series studies used to evaluate the short-term effects of air pollution on human health. Reviews on Environmental Health 18, 269–303.

Goldberg, M.S., et al., 2009. Shortness of breath at night and health status in congestive heart failure: effects of environmental conditions and health-related and dietary factors. Environtal Research 109, 166–174.

Gosling, S.N., et al., 2009. Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. Climatic Change 92, 299–341.

Gouveia, N., et al., 2003. Socioeconomic differentials in the temperature–mortality relationship in Sao Paulo, Brazil. International Journal of Epidemiology 32, 390–397.

Hajat, S., et al., 2006. Impact of high temperatures on mortality: is there an added heat wave effect? Epidemiology 17, 632–638.

Hajat, S., et al., 2005. Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. Epidemiology 16, 613–620.

Keatinge, W.R., et al., 1997. Cold exposure and winter mortality from ishaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe. The Lancet 349, 1341–1346.

Koken, P.J., et al., 2003. Temperature, air pollution, and hospitalization for cardiovascular diseases among elderly people in Denver. Environmental Health Perspectectives 111, 1312–1317.

Kolb, S., et al., 2007. The short-term influence of weather on daily mortality in congestive heart failure. Archives of Environmental and Occupational Health 62, 169–176.

Kosatsky, T., 2005. The 2003 European heat waves. European Surveillance 10, 148–149.

Kovats, R.S., et al., 2005. Climate change and human health: estimating avoidable deaths and disease. Risk Analalysis 25, 1409–1418.

Kovats, S., et al., 1998. Global climate change and environmental health: proceedings of the 1997 annual conference of the Society for Occupational and Environmental Health. International Journal of Occupational and Environmental Health 4, 41–52.

le Tertre, A., et al., 2006. Impact of the 2003 heatwave on all-cause mortality in 9 French cities. Epidemiology 17, 75–79.

Pattenden, S., et al., 2003. Mortality and temperature in Sofia and London. Journal of Epidemiology and Community Health 57, 628–633.

Martens, W.J., 1998. Health impacts of climate change and ozone depletion: an ecoepidemiologic modeling approach. [Review] [56 refs]. Environmental Health Perspectives 106 (Suppl. 51).

McMichael, A.J., et al., 2006. Climate change and human health: present and future risks. Lancet 367, 859–869.

Moshammer, H., et al., 2006. Polynomial distributed lag models of short term air pollution effects on lung function. Epidemiology 17, S261.

Samet, J.M., et al., 2000. The national morbidity, mortality, and air pollution study. Part I: methods and methodologic issues. Research Reports. Health Effects Institutes, 5–14.

Schwartz, J., 2000. The distributed lag between air pollution and daily deaths. Epidemiology 11, 320–326.

Smargiassi, A., et al., 2009. Variation of daily warm season mortality as a function of micro-urban heat islands. Journal of Epidemiological & Community Health 63, 659–664.

Statistics Canada, 2010. Survey of household spending (SHS), household equipment at the time of interview, by province, territory and selected metropolitan areas, annual, CANSIM (database). In: Statistics, C., (Ed.), Government of Canada, Ottawa.

Wyzga, R.E., 1978. The effect of air pollution upon mortality: a consideration of distributed lag models. Journal of the American Statistical Association 73, 463–472.

Zanobetti, A., et al., 2000. Generalized additive distributed lag models: quantifying mortality displacement. Biostatistics 1, 279–292.

Chapter 11

Research paper VI

Title: The impact of heat waves on mortality.

Author(s): Antonio Gasparrini, Ben Armstrong.

Journal/Publisher: Epidemiology.

Type of publication: Research paper.

Stage of publication: Published in Volume 22, Issue 1, January 2011, Pages 68-73 (with online appendix).

URL: `http://www.ncbi.nlm.nih.gov/pubmed/21150355`.

Academic peer-reviewed: Yes.

Copyright: Permission obtained from the publisher.

Candidate's role: See Section 2.3.

Candidate's signature:

Senior author: (Prof. Ben Armstrong)

The Impact of Heat Waves on Mortality

Antonio Gasparrini and Ben Armstrong

Background: Heat waves have been linked with an increase in mortality, but the associated risk has been only partly characterized.
Methods: We examined this association by decomposing the risk for temperature into a "main effect" due to independent effects of daily high temperatures, and an "added" effect due to sustained duration of heat during waves, using data from 108 communities in the United States during 1987–2000. We adopted different definitions of heat-wave days on the basis of combinations of temperature thresholds and days of duration. The main effect was estimated through distributed lag nonlinear functions of temperature, which account for nonlinear delayed effects and short-time harvesting. We defined the main effect as the relative risk between the median city-specific temperature during heat-wave days and the 75th percentile of the year-round distribution. The added effect was defined first using a simple indicator, and then a function of consecutive heat-wave days. City-specific main and added effects were pooled through univariate and multivariate meta-analytic techniques.
Results: The added wave effect was small (0.2%–2.8% excess relative risk, depending on wave definition) compared with the main effect (4.9%–8.0%), and was apparent only after 4 consecutive heat-wave days.
Conclusions: Most of the excess risk with heat waves in the United States can be simply summarized as the independent effects of individual days' temperatures. A smaller added effect arises in heat waves lasting more than 4 days.

(*Epidemiology* 2011;22: 68–73)

Heat is a well-known public health hazard.[1] The relationship between high temperatures and a number of health outcomes, in particular mortality, has been documented in many epidemiologic studies.[2–5] Extended periods of extreme heat, usually defined as heat waves, have been linked with a substantial increase in mortality,[6] and specific events have been reported as public health disasters—such as in Chicago during July 1995[7] and in France during August 2003.[8,9] The characterization of the relationship of heat and heat waves with health assumes a particular importance, given the predicted increase in their frequency and intensity based on climate change scenarios.[10,11]

Past approaches to investigate the health effects of heat are of 2 types—episode analysis and continuous-temperature time-series analyses.[12] In episode analysis, a heat wave is considered as a distinct event (episode), and excess risk associated with it is estimated by comparison with non-heat-wave periods.[13–15] A time-series analysis usually considers temperature as a continuous risk factor, using linear threshold parameterization,[16,17] or smooth functions[18,19] to specify its exposure-response relationship, sometimes allowing for lagged effects.

A few studies have recently brought these 2 approaches together, investigating the increase in risk during heat waves in a time-series regression model that also includes daily temperature as a numeric explanatory variable, possibly allowing for lagged effects. This method has been used to quantify harvesting during single events, as in August 2003 in Europe[8,20] and July 1995 in Chicago,[21] and also extended in studies with multiple heat-wave periods.[2,22] The rationale under this methodology assumes that the effect of heat may be described as the sum of 2 contributions: an increased risk because of the independent effects of daily temperature levels, and an additional risk due to duration of heat sustained for several consecutive days. The former is predicted by the usual exposure-response function for the temperature-health relationship, characterizing both heat-wave and non-heat-wave days, whereas the latter is commonly estimated by an indicator, usually defined as 2 or more consecutive days above a specified temperature. In this study, we refer to these contributions as main and added effect of heat, respectively.

This approach entails a more developed definition of heat-wave effects, identified as that not merely due to a series of days with extremely hot temperature, but because of periods when sustained heat produces an excess mortality beyond that predicted by independent contributions of daily temperature occurrences. In consequence, this method allows a more accurate prediction of the effect of heat on health by distinguishing between impacts from isolated days of heat and from sustained days of heat in waves. A substantial added effect implies the presence of additional pathophysiologic mechanisms that arise when the exposure to hot temperatures is protracted for several days, not occurring in single sporadic

Submitted 4 May 2010; accepted 17 August 2010.
From the Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, United Kingdom.
Supported by Medical Research Council (United Kingdom), contract/grant number G0707030.
SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).
Correspondence: Antonio Gasparrini, Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, 15–17 Tavistock Place, London WC1H 9SH, United Kingdom. E-mail: antonio.gasparrini@lshtm.ac.uk.

ISSN: 1044-3983/11/2201-0068
DOI: 10.1097/EDE.0b013e3181fdcd99

days of extreme heat. In contrast, a weak (if any) added effect would suggest that the increased risk during waves may be explained by the sole main effect, estimated by simpler models based on temperature-mortality exposure-response functions. Such evidence has a clear implication to plan public health interventions or to estimate the future burden of heat-related deaths under predicted climate change scenarios.

Studies on multiple heat-wave periods have indeed shown a substantial added effect.[2,22] However, the extent of the wave effect appears to be sensitive to model features, in particular, the specific function used to model the main exposure-response relationship.[22] In this paper, we seek to characterize more clearly the relationship between heat and mortality, analyzing the excess risks in heat-wave periods, by comparing the contributions of main and added effects, as defined previously, under different wave definitions. In addition, we propose a new, more flexible model to describe the added effect in terms of duration, allowing the risk to vary smoothly by the number of consecutive heat-wave days.

METHODS

Data

The analysis includes the data for 108 urban communities in the United States during the period 1987–2000. The series for mortality, weather, and pollution data were assembled from publicly available data sources as part of the National Morbidity, Mortality, and Air Pollution Study.[23,24] Daily overall mortality consists of death counts among residents, excluding injuries and external causes (International Classification of Diseases, 9th revision (ICD-9) codes 800 and above, ICD-10 codes S and above). Maximum and minimum temperatures are computed as the highest and lowest hourly measurements registered within each day, with mean temperature as the average between them. General information about how the data were collected and assembled has been previously reported, together with a detailed summary of descriptive statistics for each community (http://www.ihapss.jhsph.edu). For the current analyses, we restrict the period to summer months (June–September), to avoid the complexities of having to model cold as well as heat effects.[3,22,25]

Statistical Analysis

The analytic strategy follows a scheme already proposed for multicity studies, with a common model applied to each community and then the use of meta-analytic procedures to derive the pooled estimates.[26] At this time, the effect of heat is decomposed into the main and added effects introduced previously, by including 2 terms for mean temperature in the city-specific model. An algebraic representation is given by:

$$\log[E(Y_i)] = \alpha + \sum_{j=1}^{p} g_j(x_{ij}) + m(t_i) + w(t_i)$$

where Y_i is the mortality count, assumed to follow an overdispersed Poisson distribution for each day i. The covariates x_j, with effects expressed by the functions g_j, include an indicator for day of the week and spline functions for dew point temperature, day of the year, and time. These last 2 terms describe a regular seasonal trend, forced to be identical each year, and a smooth long-time trend, using 5 and 3 degrees of freedom (*df*), respectively, following a parsimonious approach previously applied for analyses restricted to summer months.[3,22,25]

The main effect of heat on day i is described by the function m of the series of lagged temperatures t_{i-l}, with $l = 0,\ldots, L_m$, and L_m as maximum lag. To allow flexibility, m is specified as a 2-dimensional spline function, defining a distributed lag nonlinear model that allows the main effect to vary smoothly along both dimensions of temperature and lags.[18,27] The relationship in the temperature space is modeled by a cubic spline with 6 degrees of freedom (*df*). Changes in the shape along lags are modeled by a natural cubic spline with 5 *df*, up to a maximum lag $L_m = 10$. This flexible model accounts simultaneously for nonlinear and lagged effects and short-time harvesting. Despite this flexibility, the relationship specified by the term m still assumes that effects of temperature at each lag are independent. We summarize the main effect from each city-specific model from the term $m(t)$, predicting the relative risk between the median temperature among heat-wave days versus the 75th percentile of annual temperature distribution. This reference was chosen as a temperature at which little, if any, adverse effect of temperature on mortality is expected.[19]

The pooled main effect across cities is computed through a random effect meta-analysis based on restricted maximum likelihood.[28]

The additional risk of sustained heat is left to the added effect described by the function w. The choices for this function are introduced in the proceeding discussion.

Heat-wave Indicator

In a first analysis, we specify $w(t_i)$ with:

$$w(t_i) = \prod_{l=0}^{L_w} I(t_{i-l} \geq \tau)$$

where I is an indicator which assumes value 1 if t_{i-l} is greater than or equal to a threshold level τ. In practice, in this first analysis $w(t)$ is the usual indicator defining heat-wave days as those with temperature greater than or equal to an intensity criterion τ for at least $L_w + 1$ days of duration. Following the definitions already proposed in literature,[2,22] we set τ equal to the 97th, 98th, or 99th percentiles of the year-round city-specific distribution, and L_w equal to 1 or 3 (2 or 4 days of duration). The city-specific added effect is estimated as the exponential of the coefficient for the indicator variable.

The same meta-analytic techniques used for the main effect were applied to estimate the pooled added effect across cities.

Numeric Measure of Heat-wave Duration

The second approach to characterize the added effect retains the temperature dichotomy (at the 97th percentile), but replaces the duration dichotomy by allowing risk to depend on how many consecutive days there have been greater than or equal to the threshold. In this case, $w(t) = f(d)$, where:

$$d_i = \sum_{l=1}^{L_w} [I(t_{i-l} \geq \tau) \prod_{j=0}^{l} I(t_{i-j} \geq \tau)]$$

Here, d_i is defined as the consecutive day the temperature has by date i reached the threshold τ. The product term in the equation above ensures that all the preceding days show a temperature greater than or equal to τ. Note that, d is 0 for non-heat-wave days and for the first day greater than or equal to the threshold τ, then 1 for the second day, and so on, up to the day the temperature comes back below the limit, with a maximum of L_w days. Here we set τ equal to the 97th city-specific percentile and a maximum duration L_w of 10 days. The function f describing the added effects in terms of consecutive heat-wave days d is specified in the following 2 ways: through a step function (strata: 1, 2–3, 4–5, 6–7, 8–10), or through quadratic splines with 5 df (without natural constraints, 3 knots at 2, 5, 8 days).

The estimates and variance-covariance matrix for the 5 parameters of the function $f(d)$ are then included in a multivariate meta-analysis,[29] to obtain the pooled added effect along consecutive heat-wave days. The maximum heat-wave length is different in each city, and those with maximum duration less than 10 days might contribute only to a subset of parameters. This is handled by the meta-analytic procedure allocating very large variances to the missing parameters, so that they will receive very small weight and not contribute to the average estimate.[29,30] The limit of 10 consecutive days was set to retain enough cities in the analysis actually contributing to the estimates.

Sensitivity Analysis

Given the complex statistical approaches adopted in the aforementioned analyses, involving several assumptions and a priori choices, a sensitivity analysis was carried out on the parameters for the city-specific model for functions g_j and m. Specifically, we modified the degree of smoothing for seasonality and the complexity of the distributed lag functions, varying the df and type for the splines for day of the year, temperature, and lag dimensions in the models with the mildest (97th percentile, 2 days of duration) and strictest (99th percentile, 4 days of duration) wave definitions.

We also carried out some analyses to elucidate whether the main and added terms are too correlated for their effects to be disentangled. First, we computed the simple correlations between mean temperature and both indicators and continuous measure of consecutive heat-wave days. Then, we more generally assessed the multicollinearity between the full set of main effect terms and the added wave term. Specifically, we computed the R^2 of a model regressing each heat-wave term on the cross-basis variables: a high R^2 indicates that the heat-wave term is almost perfectly predicted by the variables for the main effect, potentially inducing problems of multicollinearity in our regression model.

Further information on modeling choices and residual, correlation and additional sensitivity analyses are provided in the eAppendix (http://links.lww.com/EDE/A437), Sections S1–S3.

Software

The main analyses and graphical representation are performed in the statistical environment R version 2.11.1.[31] Distributed lag nonlinear models are specified through the package dlnm (version 1.2.3), whereas univariate meta-analyses are carried out through the package metafor (version 1.1–0). Multivariate meta-analytical estimates are obtained by Stata 11,[32] using the command mvmeta.

The main results included in the paper are entirely reproducible.[33] The data are freely available using the R package NMMAPSlite (version 0.3–2). The R code to run the main analysis and the Stata code for multivariate meta-analysis are available in the eAppendix (http://links.lww.com/EDE/A437), Section S4.

RESULTS

Mean summer temperature shows a high variability in the 108 communities, ranging from 12.8°C in Anchorage to 33.0°C in Phoenix, with an average of 23.5°C. The number of heat-wave days during the 14-year period, defined by the indicator variable used in the first analysis, varies depending on wave definitions. The average number of heat-wave days in each community is 90.0 (range: 38–129) when using the 97th percentile and 2 days' duration, and 7.2 (range: 0–21) using the 99th percentiles and 4 days' duration.

Table 1 shows the estimated increase in risk due to main and added effects in those days matching the 6 definitions. The average main effect is similar between definitions based on 2 or 4 days of duration, and increases proportionally to the intensity criterion (97th, 98th, and 99th percentiles), being computed on the median temperature among heat-wave days, which increases accordingly. In contrast, the duration criterion plays an important role for the added effect: the models using 2 days' minimum duration show very small increases in risk; if the minimum duration period is extended to 4 days, the average added effect increases proportionally to the selected percentile. Only the strictest definition of days showing a temperature greater than or equal to the 99th percentile for at least 4 past days provides an increase of 2.8% (95% confidence interval [CI] = 0.4%–5.3%) in mortality. The contribution of the main effect substantially ex-

TABLE 1. Pooled Main and Added Effects With Tests for Heterogeneity (*P*) Across Cities Under Different Heat-wave Definitions

			Main Effect		Added Wave Effect	
No. Days	Percentile	No. Cities	% Increase (95% CI)	Test for Heterogeneity	% Increase (95% CI)	Test for Heterogeneity
≥2	≥97th	108	4.9 (3.3 to 6.5)	$P \leq 0.001$	0.3 (−0.5 to 1.1)	$P = 0.536$
	≥98th	108	6.3 (4.7 to 8.0)	$P \leq 0.001$	0.4 (−0.5 to 1.4)	$P = 0.892$
	≥99th	108	8.0 (5.7 to 10.4)	$P \leq 0.001$	0.2 (−1.3 to 1.7)	$P = 0.005$
≥4	≥97th	108	5.4 (3.9 to 6.9)	$P \leq 0.001$	0.7 (−0.5 to 1.9)	$P = 0.186$
	≥98th	108	6.3 (4.5 to 8.1)	$P \leq 0.001$	1.3 (−0.3 to 2.9)	$P = 0.008$
	≥99th	105[a]	7.7 (5.4 to 10)	$P \leq 0.001$	2.8 (0.4 to 5.3)	$P = 0.033$

[a]Three communities do not show any days matching this heat-wave definition and do not contribute to the estimates.

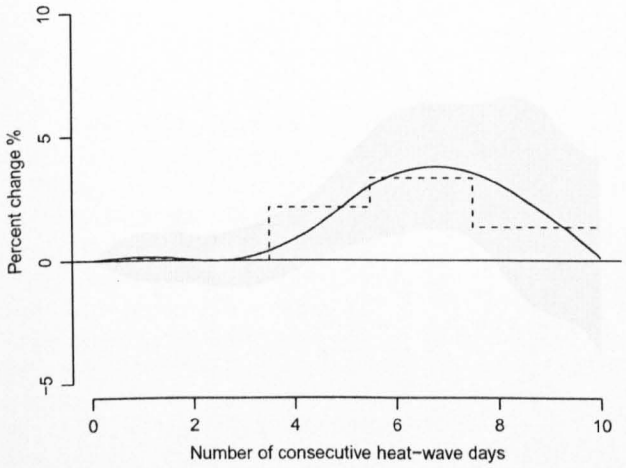


FIGURE. Average wave effect of consecutive heat-wave days (greater than or equal to 97th percentile), as estimated by quadratic spline (continuous line) with 95% CI (gray area), and by a step function (dashed line).

ceeds the added effect during heat-wave days in all the 6 definitions.

Communities show some variability in the length of wave periods, when specified as 2 consecutive days with temperature greater than or equal to the 97th city-specific percentile, with an average maximum length of 9.5 days (range: 4–20 days). Heat waves of at least 10 and 7 days long, were experienced respectively, by 45.4% and 81.5% of communities. Heat-wave periods are usually short, with 76.3% of days within the first 3 days of heat wave. The average added effect, specified by increase in risk for consecutive heat-wave days and modeled alternatively by both quadratic spline and a step functions, is depicted in the Figure. The analysis shows no effect during the beginning of a wave period, then an increase when the heat is sustained for longer than 4 uninterrupted days. The plot also displays a decrease after a peak at around 7 consecutive days, although wide confidence intervals.

The results of sensitivity analysis are illustrated in Table 2. The estimated added effect (0.3% and 2.8% in the original models, respectively) was robust to most of the changes. The most notable exceptions are the results reported in the last 3 rows of Table 2, which showed considerably higher wave effects (up to 3.7% and 7.0%). These models were characterized by either relatively inflexible splines for temperature, inflexible lag structure, or both. The 2 *df* spline with "natural" constraints is forced to be linear beyond the boundaries, further limiting its flexibility to model nonlinear effects for extremely hot days. Because extremely hot days are also likely to be labeled as heat-wave days, this would produce an inflated added effect. The same happens when applying a very simple model with 1 *df* to describe lagged effect, corresponding approximately to a simple moving average of the temperatures in the lag period of 10 days.

The correlation between mean temperature and heat-wave terms is not high: the average correlation coefficient r across cities is 0.40 (range: 0.29–0.51) for the indicator variable based on 97th percentile and 2 days of duration and 0.32 (range: 0.24–0.43) for continuous measure of consecutive heat-wave days. The R^2 of the regression of wave terms on the cross-basis variables for the main effect, an index of multicollinearity, shows an average of 0.63 (range: 0.43–0.76) for the same indicator and of 0.56 (range: 0.38–0.72) for the continuous variable. These results demonstrate that, although the main and added effects are correlated, the model and data still have power to separate these 2 effects.

DISCUSSION

Our approach seeks to characterize the excess risk during heat-wave periods, quantifying how much of this additional burden is simply explained by the increase in temperature and how much is attributable to the heat continuing over several consecutive days. Furthermore, this additional risk during waves is described in terms of duration, proposing a new definition based on consecutive heat-wave days.

TABLE 2. Sensitivity Analysis on the Degrees of Freedom (*df*) and Spline Type for Seasonality and Temperature-Lag Functions on the Pooled Added Effect Across Cities, Under 2 Different Heat Wave Definitions

df for Specific Functions			≥2 Days; ≥97th Percentile	≥4 Days; ≥99th Percentile
Seasonality	Temperature	Lag	% Increase (95% CI)	% Increase (95% CI)
4	6	5	0.3 (−0.5 to 1.1)	2.8 (0.4 to 5.3)
2	6	5	0.3 (−0.4 to 1.1)	2.8 (0.4 to 5.2)
6	6	5	0.3 (−0.4 to 1.1)	2.8 (0.4 to 5.3)
4	4	5	−0.1 (−0.9 to 0.7)	3.0 (0.3 to 5.8)
4	7	5	0.1 (−0.7 to 0.9)	2.5 (0.1 to 5.0)
4	6	3	0.8 (0.0 to 1.6)	3.0 (0.2 to 5.8)
4	6	6	0.3 (−0.5 to 1.1)	2.8 (0.4 to 5.2)
4	2[a]	5	1.0 (0.3 to 1.8)	4.8 (1.4 to 8.3)
4	6	1	3.6 (2.8 to 4.4)	6.7 (3.5 to 10.1)
4	2[a]	1	3.7 (2.8 to 4.5)	7.0 (3.3 to 10.8)

[a]A natural cubic spline is used here instead than a simple B-spline.

This analysis addresses important epidemiologic and public health questions: the implementation of adequate preventive measures such as heat-wave plans (in the short-to-medium term) and the prediction of the burden of future events under the suggested climate change scenarios (in the long term) require a detailed characterization of the association between heat, heat waves, and mortality. The results suggest that most of the excess risk during waves is attributable to (and predictable by) the increase in daily temperatures whether isolated or occurring with other hot days, the effect of which is larger than any added effect. The latter is negligible for short heat-wave periods, although it does bring some additional risks after 4 days of uninterrupted heat.

Our analytic design offers several advantages. First, the choice of flexible distributed lag nonlinear functions gives greater assurance than simpler models that the main effect is adequately accounted for, reducing the risk of confounding of the added effect by a residual main effect of heat. In addition, the analysis takes into account the adaptation of each population to its own climate,[3,19] by allowing community-specific exposure-response functions for the main effect, and wave definitions based on community-specific percentiles. Finally, by modeling the heat-wave effect as a continuous function of duration, we avoid arbitrary duration criteria and allow direct estimation of the duration at which such effect become apparent.

Our findings from the first analysis using an indicator for heat-wave days, as described in Table 1, are rather different from some others previously reported in the literature. An analysis of London, Milan, and Budapest by Hajat et al,[22] with a wave definition based on the 99th percentile for at least 2 days and a natural cubic spline with 3 *df* to specify the unlagged main exposure-response relationship, showed a percentage increase in mortality from 4.3% to 8.3%. Anderson and Bell,[2] analyzing the whole year data on the same dataset considered here and a natural cubic spline with 3 *df* for lag 0–1, found an average increase of 6.5% for a definition based on 99th percentile and 4 days of duration. These results are comparable in magnitude to our estimates for similar models reported in the last 3 rows of Table 2, and can be probably explained by the limited flexibility of the functions used to account for the main effect, a pattern also reported by Hajat et al.[22] The results on the effect of wave duration are consistent with some findings already reported in the literature.[34,35]

We estimated the proposed association between heat, heat waves, and mortality by averaging the effects across different cities and different wave periods, and this average relationship might not accurately represent every specific heat-wave event. The approach we propose showed quite good performance when applied to predict mortality during the extreme heat wave in Chicago in 1995 (eAppendix [http://links.lww.com/EDE/A437], Section S3), but might be biased in describing some waves in some cities if these heat waves are unusual with respect to variables not included in the analytic model and acting as modifiers of the temperature-health association. For instance, a potential synergistic effect between air pollution and heat has been suggested, although specific analyses have reported conflicting results.[2,36,37] The evidence is also unclear for an effect modification by socioeconomic characteristics,[16,38–40] whereas more robust for the prevalence of air conditioning.[2,41,42] These issues may be addressed in further research.

In this paper, we provide a novel analysis of the impact of heat waves on mortality. Our results suggest that the excess risk during heat-wave periods is largely explained by the immediate and lagged effect of daily temperatures, with just a small added impact because of sustained heat limited to waves lasting more than 4 days.

ACKNOWLEDGMENTS

We thank Shakoor Hajat for his valuable suggestions and comments and the researchers of the National Morbidity Mortality and Air Pollution Study (NMMAPS), who made publicly available the data used in our analysis.

REFERENCES

1. Kovats RS, Hajat S. Heat stress and public health: a critical review. *Ann Rev Public Health*. 2008;29:41–55.
2. Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology*. 2009; 20:205–213.
3. Baccini M, Biggeri A, Accetta G, et al. Heat effects on mortality in 15 European cities. *Epidemiology*. 2008;19:711–719.
4. Basu R. High ambient temperature and mortality: a review of epidemiological studies from 2001 to 2008. *Environ Health*. 2009;8:40.
5. Basu R, Samet JM. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiol Rev*. 2002;24:190–202.
6. Kovats RS, Kristie LE. Heatwaves and public health in Europe. *Eur J Public Health*. 2006;16:592–599.
7. Semenza JC, Rubin CH, Falter KH, et al. Heat-related deaths during the July 1995 heat wave in Chicago. *N Engl J Med*. 1996;335:84–90.
8. Le Tertre A, Lefranc A, Eilstein D, et al. Impact of the 2003 heatwave on all-cause mortality in 9 French cities. *Epidemiology*. 2006;17:75–79.
9. Poumadere M, Mays C, Le Mer S, Blong R. The 2003 heat wave in France: dangerous climate change here and now. *Risk Anal*. 2005;25: 1483–1494.
10. Luber G, McGeehin M. Climate change and extreme heat events. *Am J Prev Med*. 2008;35:429–435.
11. O'Neill MS, Ebi KL. Temperature extremes and health: impacts of climate variability and change in the United States. *J Occup Environ Med*. 2009;51:13–25.
12. Gosling SN, McGregor GR, Lowe JA. Climate change and heat-related mortality in six cities Part 2: climate model evaluation and projected impacts from changes in the mean and variability of temperature with climate change. *Int J Biometeorol*. 2009;53:31–51.
13. Conti S, Meli P, Minelli G, et al. Epidemiologic study of mortality during the summer 2003 heat wave in Italy. *Environ Res*. 2005;98:390–399.
14. Huynen MM, Martens P, Schram D, Weijenberg MP, Kunst AE. The impact of heat waves and cold spells on mortality rates in the Dutch population. *Environ Health Perspect*. 2001;109:463–470.
15. Rey G, Jougla E, Fouillet A, et al. The impact of major heat waves on all-cause and cause-specific mortality in France from 1971 to 2003. *Int Arch Occup Environ Health*. 2007;80:615–626.
16. Hajat S, Kovats RS, Lachowycz K. Heat-related and cold-related deaths in England and Wales: who is at risk? *Occup Environ Med*. 2007;64: 93–100.
17. Pattenden S, Nikiforov B, Armstrong BG. Mortality and temperature in Sofia and London. *J Epidemiol Community Health*. 2003;57:628–633.
18. Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology*. 2006;17:624–631.
19. Curriero FC, Heiner KS, Samet JM, Zeger SL, Strug L, Patz JA. Temperature and mortality in 11 cities of the eastern United States. *Am J Epidemiol*. 2002;155:80–87.
20. Hertel S, Le Tertre A, Jockel KH, Hoffmann B. Quantification of the heat wave effect on cause-specific mortality in Essen, Germany. *Eur J Epidemiol*. 2009;24:407–414.
21. Kaiser R, Le Tertre A, Schwartz J, Gotway CA, Daley WR, Rubin CH. The effect of the 1995 heat wave in Chicago on all-cause and cause-specific mortality. *Am J Public Health*. 2007;97(suppl 1):S158–S162.
22. Hajat S, Armstrong B, Baccini M, et al. Impact of high temperatures on mortality: is there an added heat wave effect? *Epidemiology*. 2006;17: 632–638.
23. Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW. *The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 2. Morbidity and Mortality From Air Pollution in the United States*. Cambridge, MA):Health Effects Institute;2000.
24. Samet JM, Zeger SL, Dominici F, Dockery D, Schwartz J. *The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 1. Methods and Methodological Issues*. Cambridge, MA):Health Effects Institute;2000.
25. Michelozzi P, Accetta G, De Sario M, et al. High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities. *Am J Respir Crit Care Med*. 2009;179:383–389.
26. Dominici F. Time-series analysis of air pollution and mortality: a statistical review. *Res Rep Health Eff Inst*. 2004;3–27; discussion 29–33.
27. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Stat Med*. 2010;29:2224–2234.
28. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med*. 2002;21:589–624.
29. Jackson D, White IR, Thompson SG. Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat Med*. 2010;29:1282–1297.
30. White IR. Multivariate random-effects meta-analysis. *Stata J*. 2009;9: 40–56.
31. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria:R Foundation for Statistical Computing;2010.
32. Stata11 [computer program]. College Station, TX: StataCorp LP; 2009.
33. Peng RD, Dominici F, Zeger SL. Reproducible epidemiologic research. *Am J Epidemiol*. 2006;163:783–789.
34. Diaz J, Jordan A, Garcia R, et al. Heat waves in Madrid 1986–1997: effects on the health of the elderly. *Int Arch Occup Environ Health*. 2002;75:163–170.
35. Hajat S, Kovats RS, Atkinson RW, Haines A. Impact of hot temperatures on death in London: a time series approach. *J Epidemiol Community Health*. 2002;56:367–372.
36. Medina-Ramon M, Schwartz J. Temperature, temperature extremes, and mortality: a study of acclimatization and effect modification in 50 United States cities. *Occup Environ Med*. 2007;64:827–833.
37. Ren C, Williams GM, Morawska L, Mengersen K, Tong S. Ozone modifies associations between temperature and cardiovascular mortality: analysis of the NMMAPS data. *Occup Environ Med*. 2008;65:255–260.
38. Stafoggia M, Forastiere F, Agostini D, et al. Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. *Epidemiology*. 2006;17:315–323.
39. Basu R, Ostro BD. A multicounty analysis identifying the populations vulnerable to mortality associated with high ambient temperature in California. *Am J Epidemiol*. 2008;168:632–637.
40. O'Neill MS, Zanobetti A, Schwartz J. Modifiers of the temperature and mortality association in seven US cities. *Am J Epidemiol*. 2003;157: 1074–1082.
41. Barnett AG. Temperature and cardiovascular deaths in the US elderly: changes over time. *Epidemiology*. 2007;18:369.
42. Davis RE, Knappenberger PC, Michaels PJ, Novicoff WM. Changing heat-related mortality in the United States. *Environ Health Perspect*. 2003;111:1712–1719.

Online Supplemental Material for the article "The impact of heat waves on mortality"

Antonio Gasparrini*,† and Ben Armstrong

Department of Social and Environmental Health Research
London School of Hygiene and Tropical Medicine, UK

Contents

*Correspondance: Antonio Gasparrini, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK

†E-mail: antonio.gasparrini@lshtm.ac.uk

S1 Modelling choices

The city-specific model was defined in the manuscript as:

$$log[E(Y_i)] = \alpha + \sum_{j=1}^{P} g_j(x_{ij}) + m(t_i) + w(t_i) \tag{S1.1}$$

The following sections provide some further justifications about the choices on the functions to describe the effects of covariates $g_j(x_j)$, the main $m(t)$ and the added $w(t)$ effects of temperature .

S1.1 Covariates

As explained in the text, the covariates x_j included in the model in (S1.1) are day of the week, dew point temperature, long time trend and seasonality. Their inclusion and specification is decided independently from statistical significance and actual confounding effect in the city-specific estimates, following the rationale of the NMMAPS analysis (Dominici et al., 2005, 2003).
Day of the week is specified as 6 indicator variables, while dew point temperature is characterized through a natural cubic spline with 3 df, 2 knots at equally-spaced percentiles. The effect of seasonality is modelled through a natural cubic spline with 4 df (3 equally-spaced knots), in order to describe the variation within the summer period considered here (June-September). This effect is supposed to remain constant across different years, following the assumptions of other analyses published earlier (Analitis et al., 2008; Baccini et al., 2008; Michelozzi et al., 2009). These studies used an indicator variable for month in order to model the seasonal effect. We use a similar number of df (1 per month), but describing the effect through a smooth function. Long time trend is included as a natural cubic spine with 3 df (2 equally-spaced knots), to capture the residual temporal variability.

S1.2 Main effect of temperature

The main effect of temperature $m(t)$ is specified by a *cross-basis*, a specific set of functions which can describe simultaneously the relationship both in the space of the predictor (temperature) and in the lags (Armstrong, 2006; Gasparrini et al., 2010). This choice allows a strong control of potentially non-linear and lagged effect, also accounting for short-time harvesting (if present), and is motivated by the need to accurately control for the effect of daily temperature occurrences. Given the strong correlation between the parameters used to describe the main and added effect, a weak control for the former might produce biased estimates for the latter, due to residual confounding effect.
The cross-basis functions can be described as tensor-products between the basis functions used to define the relationship in each dimension. Specifically, we use here a cubic spline with 6 df (without natural constraints, 3 knots at equally spaced values) to specify the dependency along the dimension of temperature, and a natural cubic spline with 5 df (3 knots at equally spaced values in the log scale, plus intercept) for the distributed lag effects, with 30 df overall. The maximum lag is fixed at 10, a period of time long enough to include delayed effects and short time harvesting.
We found that the fit of the model improves when relaxing the linearity constraints of the spline at

the boundaries of temperature distribution, using the same amount of df. This may be attributed to a strong non-linear effect of heat at very high temperatures, which is better described by the spline without natural constraints. The days showing high temperatures are likely to be defined as HW days: an underestimation of the main effect in this range can therefore result in a overestimation of the added effect.

We keep a natural cubic spline for the dimension of the lag in order to specify more knots with the same df (for the natural cubic splines $df = k + 1$, while for a simple cubic spline $df = k + 3$, with k number of knots). The knots are placed at equally-spaced values in the log scale (0.8, 1.9, 4.4 lags), assuring enough flexibility in the first lags, where more variability is expected (Muggeo, 2008; Peng and Dominici, 2009).

S1.3 Added effect during HWs

The different HW definitions used in the first analysis with the simple indicator variables follow from choices already proposed in the literature (Anderson and Bell, 2009; Hajat et al., 2006). Regarding the second analysis on the effect of consecutive HW days, we fixed the threshold to the 97^{th} city-specific percentile in order to obtain a suitable amount of HW days, and we pooled the results using a meta-analytical technique based on the multivariate extension of the method of moment estimator of Der Simonian and Laird (Jackson et al., 2010; White, 2009).

Given that many cities show only short HW periods, the maximum length is set to 10 days, coherently with the time frame used to specify the cross-basis functions for the main effect. HW days beyond that point will keep the value of 10. As explained in the manuscript, cities with maximum duration less than 10 days may contribute only to a subset of parameters of the two functions, strata and quadratic B-spline.

S2 Sensitivity analysis

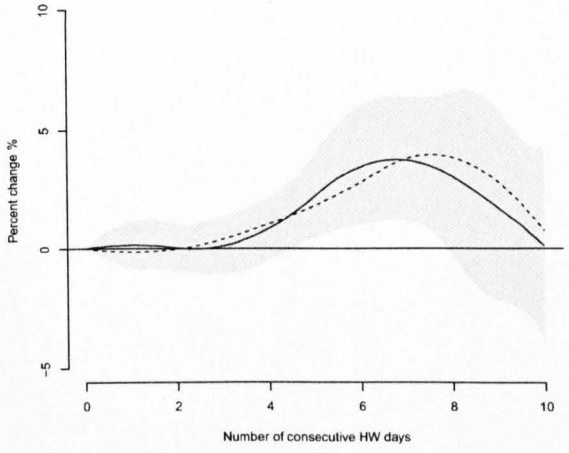
The robustness of the results to the various choices adopted in our modelling approach was tested through a sensitivity analysis. The main results obtained by varying the parameter of the functions g_j, s and w in model (S1.1) are reported in the paper. Here we provide additional sensitivity analyses on the choices regarding the function f of consecutive HW days, evaluating graphically the differences for Figure 1 in the main text.
In particular:

- *13 days*: extending the maximum HW consecutive days to 13.
- *only 10 days*: restricting the analysis to the subsample of cities showing HW periods of at least 10 days (49 cities).
- *98th*: using the 98$^{\text{th}}$ percentile as a cut-off to define consecutive HW days.
- *REML*: using restricted maximum likelihood as estimation procedure for multivariate meta-analysis.

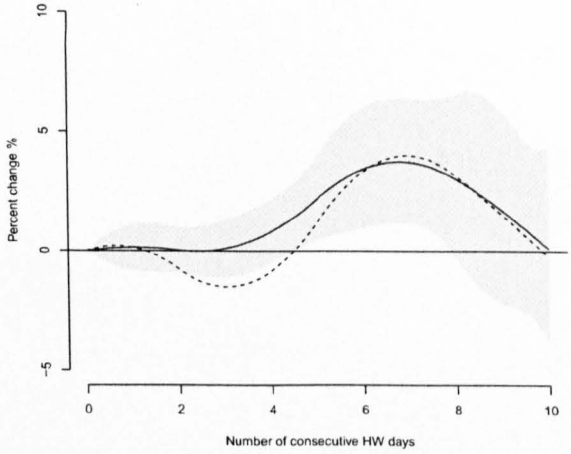
The results are summarized in Figure S1.
The shape of the curve obtained by the original model in the main text does not seem to be strongly influenced by the changes listed above. Increasing the maximum number of consecutive HW days to 13 only slightly postpones the peak in risk. This result suggests that the risk is not confined to the first 10 HW days, but that additional effects can be associated to longer HW periods. Furthermore, this might be compensated by some harvesting effect at longer lags, as previously pointed out (Hertel et al., 2009; Kaiser et al., 2007; Le Tertre et al., 2006). The subsample of cities with maximum HW length of at least 10 days shows approximately the same relationship, indicating that the results are robust to city selection up to this point. Anyway, only a limited number of cities actually shows very long HW's, and this selection precludes the generalizability of the results beyond this HW length. Applying a more stringent definition for consecutive HW days based on the 98$^{\text{th}}$ percentile reveals a similar effect, but starting earlier within the HW periods. The results are robust to the estimation method selected for the multivariate meta-analysis, as expected given the large sample of cities.
The R and Stata code of the main analysis is included in Section S4. The reader is free to perform further sensitivity checks changing the code directly.

Figure S1: Sensitivity analysis for the added effect (consecutive HW days)

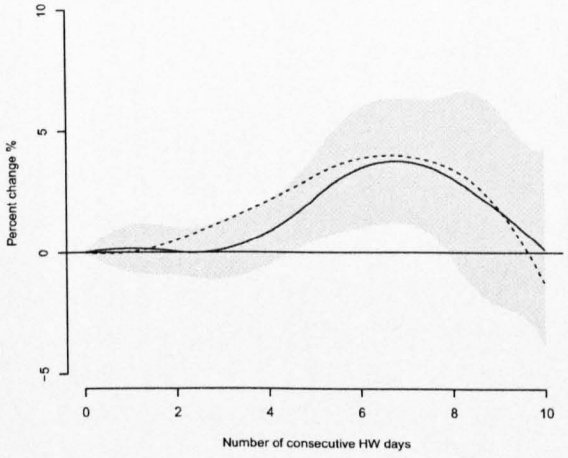
(a) Extending to 13 consecutive HW days



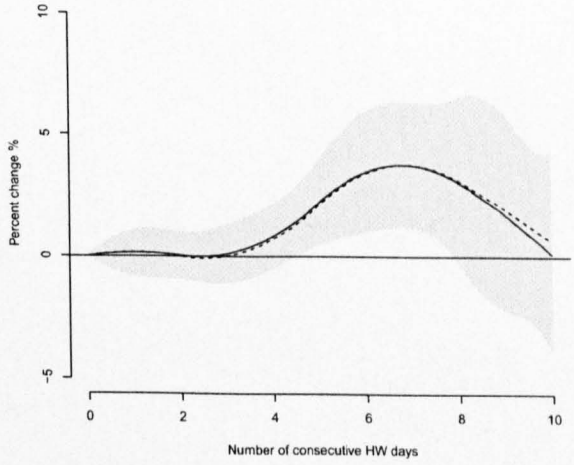
(b) Subsample of cities with 10-days HW's



(c) Choosing $98th$ percentile as threshold



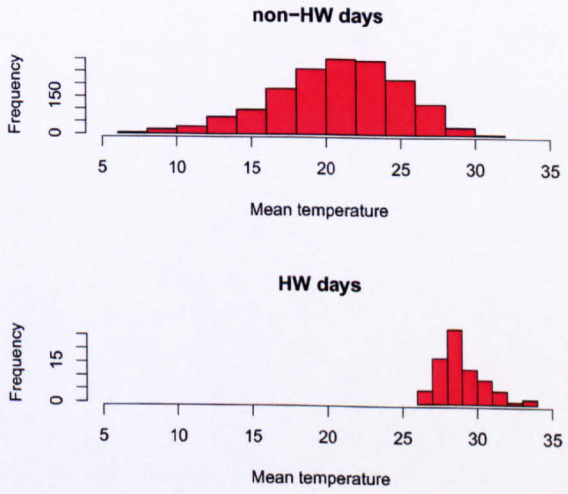
(d) REML



S3 Residual and correlation analysis

In this Section we provide an analysis restricted to the city of Chicago, where two important HWs occurred in August 1988 and, particularly infamous, in July 1995. The results showed here are computed from the model where the added effect is specified with a continuous measure of consecutive HW days, defined using the 97^{ht} percentile and 2 days of minimum duration.
The correlation between mean temperature and HW terms is not very high, as in the rest of the NMMAPS cities. The coefficient r is 0.39 using the simple HW indicator and 0.33 for consecutive HW days. Figure S2 illustrates the temperature distribution in HW and non-HW days. The plot shows a substantial overlap between the two distributions, due to the fact that HW days are defined not just in terms of temperature but also of duration, thus explaining the low correlation with the HW indicator.

Figure S2: Temperature distribution in HW and non-HW days



The analysis of standardized residuals suggests a good fit in general of the model, as illustrated in Figure S3. However, it is possible to detect 2 outliers, corresponding to 2 days in July 1995 (under predicted) and August 1988 (over predicted).
More specifically, as depicted in Figure S4, the model predicts the mortality quite well: in periods identified as HW days, the average observed-predicted number of deaths are 122.4-122.6 (12^{ht}-18^{ht} of August 1988) and 261.3-242.2 (13^{ht}-16^{ht} of July 1995).

Figure S3: Distribution, Q-Q plot and series of standardized residuals

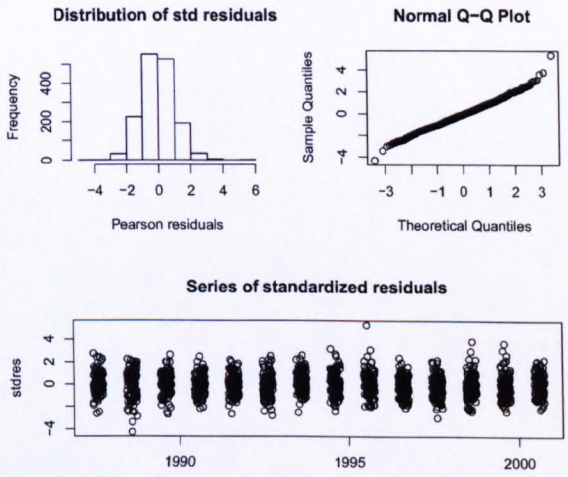
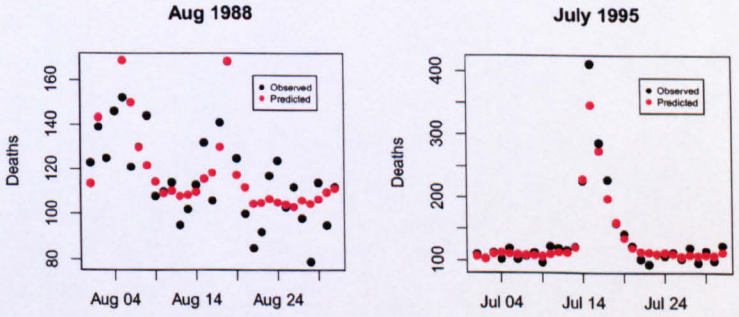


Figure S4: Observed and predicted mortality during August 1988 and July 1995



S4 R and Stata code

R and Stata code to reproduce the main results of the analysis are included below. The first part of the R code (Section S4.1) performs the first-stage (city-level) model and store the results in a file readable from Stata, saved in the current directory. The Stata code (Section S4.2) then runs the multivariate meta-analysis and store the results in other Stata files. Finally, the second part of the R code (Section S4.3) imports the estimates back to R and produces the results for the first and second analysis reported in the paper.

Additional information on the specific analytical steps are provided as comments within the code. The reader should pay attention to run the code in the order explained above.

S4.1 R code (first part)

```
require(dlnm);require(Epi);require(tsModel)
require(NMMAPSlite);require(metafor);require(foreign)

# FUNCTION TO CREATE AN HEAT WAVE INDICATOR FOR A TEMPERATURE SERIES
#  BASED ON THE THRESHOLD AND THE DURATION, BY GROUPS
fun.hw.thr <- function(x,thr,dur,group=NULL) {
  as.numeric(apply(Lag(x>=thr,0:(dur-1),group=group),
  1,sum,na.rm=T)>(dur-1))
}

# INITIALIZE THE DATASET
initDB()
cities <- listCities()

# CREATE THE MATRICES TO STORE THE RESULTS
#  DESCRIPTIVE STATS
descr.tmean <- matrix(NA,length(cities),7, dimnames=list(cities,
  names(summary(c(1:10,NA)))))
hw.N <- matrix(NA,length(cities),6, dimnames=list(cities,
  paste("hw",rep(c(2,4),each=3),rep(c(97,98,99),2),sep=".")))
hw.cons <- matrix(NA,length(cities),4, dimnames=list(cities,
  c("N","Max",">3",">7")))
# REGRESSION MODELS
main.eff <- added.eff <- matrix(NA,length(cities),12,
  dimnames=list(cities,paste("hw",rep(c(2,4),each=6),rep(c(97,98,99),
  each=2),c("est","sd"),sep=".")))
strata.eff <- matrix(NA,length(cities),5,dimnames=list(cities,1:5))
strata.vcov <- vector("list",length(cities)) ; names(strata.vcov) <- cities
quad.eff <- strata.eff
quad.vcov <- strata.vcov
# MEAN SUMMER TEMPERATURE
meantemp <- 0

##############################################################################

# START THE LOOP FOR CITIES
time <- proc.time()
```

```
for(i in seq(length(cities))) {

  # LOAD AND PREPARE DATASET
  datatot <- readCity(cities[i], collapseAge = T)
  datatot$tmean <- (datatot$tmpd-32)*5/9
  datatot$time <- 1:nrow(datatot)
  datatot$year <- as.numeric(substr(datatot$date,1,4))
  datatot$month <- as.numeric(substr(datatot$date,6,7))
  datatot$doy <- sequence(tapply(datatot$year,datatot$year,length))
  datatot$dp01 <- filter(datatot$dptp,c(1,1)/2,side=1)
  percentiles <- quantile(datatot$tmean,c(75,97:99)/100,na.rm=T)
  data <- datatot[datatot$month%in%6:9,]

  # SAVE DESCRIPTIVE STATISTICS FOR TEMPERATURE
  descr.tmean[i,1:6] <- summary(data$tmean)[1:6]
  descr.tmean[i,7] <- sum(is.na(data$tmean))
  meantemp[i] <- mean(data$tmean,na.rm=T)

  # CREATE THE CROSSBASIS FOR THE MAIN TEMPERATURE-MORTALITY RELATIONSHIP
  # CENTERED ON 75TH PERCENTILE, REFERENCE VALUE FOR PREDICTED EFFECTS
  range <- round(range(data$tmean,na.rm=T),0)
  ktemp <- range[1] + (range[2]-range[1])/4*1:3
  basis <- crossbasis(data$tmean,group=data$year,vartype="bs",vardegree=3,
    varknots=ktemp,lagdf=5,maxlag=10,cenvalue=percentiles[1])

  ################################################################
  # FIRST ANALYSIS: INDICATOR FOR DIFFERENT HW DEFINITIONS
  ################################################################

  # HW DEFINITIONS
  hw.def <- cbind(rep(percentiles[2:4],2),rep(c(2,4),c(3,3)))

  # RUN THE MODEL FOR EACH DEFINITION
  for(k in 1:nrow(hw.def)) {

    # CREATE HEATWAVE INDICATOR FOR THE SPECIFIC HW DEFINITION
    hw <- fun.hw.thr(data$tmean,hw.def[k,1],hw.def[k,2],data$year)
    hw.N[i,k] <- sum(hw)

    # RUN THE MODEL
```

```
  model.first <- glm(death ~  hw + basis + dow + ns(year,3) +
    ns(doy,df=4) + ns(dp01,df=3), family=quasipoisson(), data)
  # SAVE MAIN EFFECT
  if(sum(hw)>0) {
  tmedian <- median(data$tmean[hw==1],na.rm=T)
  pred <- crosspred(basis,model.first,
    at=c((range[1]+1):(range[2]-1),tmedian))
  main.eff[i,c(k*2-1,k*2)] <- cbind(pred$allfit,
    pred$allse)[as.character(tmedian),]
  } else main.eff[i,c(k*2-1,k*2)] <- c(NA,NA)
  # SAVE ADDED EFFECT
  added.eff[i,c(k*2-1,k*2)] <- ci.lin(model.first)["hw",1:2]
}

################################################################
# SECOND ANALYSIS: STRATA AND QUAD SPLINE OF CONSECUTIVE HW DAYS
################################################################

# CREATE HEATWAVE INDICATOR AND CONSECUTIVE TERM (97TH PERCENTILE)
hw <- fun.hw.thr(data$tmean,percentiles[2],2,data$year)
# CREATE HW CONSECUTIVE DAYS (UP TO 10 DAYS)
hw.lin <- hw
for(j in 2:10) {
  hw.lin[apply(Lag(hw,0:(j-1),group=data$year),
    1,sum,na.rm=T)==j] <- j
}
# SAVE STATS ON CONSECUTIVE HW DAYS
hw.cons[i,] <- c(sum(hw),max(hw.lin),sum(hw.lin>3),sum(hw.lin>7))

# CREATE THE STRATA OF CONSECUTIVE HW DAYS
strata <- mkbasis(c(1:10,hw.lin),type="strata",
  knots=c(1,2,4,6,8))$basis[-(1:10),]
# RUN THE MODEL
model.strata <- glm(death ~  basis + strata + dow +
  ns(dp01,df=3) + ns(year,3) + ns(doy,df=4),
  family=quasipoisson(), data)
# SAVE THE RELATED COEF AND VCOV (INCLUDING MISSING)
index1 <- grep("strata",names(coef(model.strata)))
index2 <- (1:length(coef(model.strata)))[is.na(coef(model.strata))]
index <- index1[!index1%in%index2]
```

```
  strata.eff[i,!index1%in%index2] <- ci.lin(model.strata)[index,1]
  strata.vcov[[i]] <- matrix(NA,length(index1),length(index1))
  strata.vcov[[i]][!index1%in%index2,!index1%in%index2] <-
    vcov(model.strata)[index,index]

  # CREATE THE SPLINE OF CONSECUTIVE HW DAYS
  quad <- bs(hw.lin,knots=c(2,5,8),Bound=c(0,10),degree=2)
  # RUN THE MODEL
  model.quad <- glm(death ~  basis + quad + dow + ns(dp01,df=3) +
    ns(year,3) + ns(doy,df=4),family=quasipoisson(), data)
  # SAVE THE RELATED COEF AND VCOV (INCLUDING MISSING)
  index1 <- grep("quad",names(coef(model.quad)))
  index2 <- (1:length(coef(model.quad)))[is.na(coef(model.quad))]
  index <- index1[!index1%in%index2]
  quad.eff[i,!index1%in%index2] <- ci.lin(model.quad)[index,1]
  quad.vcov[[i]] <- matrix(NA,length(index1),length(index1))
  quad.vcov[[i]][!index1%in%index2,!index1%in%index2] <-
    vcov(model.quad)[index,index]
}
proc.time()-time
# TAKES APPROXIMATELY 5-6 MIN IN A 2GHz LAPTOP

###############
# TO STATA
###############

index <- cbind(rep(1:5,5),rep(1:5,each=5))
names <- c(paste("b",1:5,sep="_"),
  paste("V",rep(1:5,5),rep(1:5,each=5),sep="_"))
temp1 <- temp2 <- matrix(0,length(cities),length(names))
for(i in 1:length(cities)) {
  temp1[i,] <- c(strata.eff[i,],strata.vcov[[i]][index])
  temp2[i,] <- c(quad.eff[i,],quad.vcov[[i]][index])
}

colnames(temp1) <- colnames(temp2) <- names

library(foreign)
write.dta(as.data.frame(temp1),"strata.dta")
write.dta(as.data.frame(temp2),"quad.dta")
```

S4.2 Stata code

```
*cd "..."
set more off

* QUAD MM
use quad, clear
mvmeta b V, mm bscov
matrix b = e(b)
matrix V = e(V)
clear
svmat b
svmat V
save quad_mm, replace

* STRATA MM
use strata, clear
mvmeta b V, mm bscov
matrix b = e(b)
matrix V = e(V)
clear
svmat b
svmat V
save strata_mm, replace
```

S4.3 R code (second part)

```
###############
# FROM STATA (STATA CODE SHOULD HAVE BEEN RUN)
###############

quad.pool.est <- as.matrix(read.dta("quad_mm.dta")[1,1:5])
quad.pool.vcov <- as.matrix(read.dta("quad_mm.dta")[1:5,6:10])
strata.pool.est <- as.matrix(read.dta("strata_mm.dta")[1,1:5])
strata.pool.vcov <- as.matrix(read.dta("strata_mm.dta")[1:5,6:10])

################################
# RESULTS: DESCRIPTIVE STATISTICS
################################
```

```
# SUMMARY FOR TMEAN
summary(descr.tmean[,c("Mean","NA's")])

# TOTAL NUMBER OF HW DAYS UNDER DIFFERENT HW DEFINITIONS
summary(hw.N)

# CONSECUTIVE HW DAYS (WITH 97TH PERCENTILE)
# % OF CITIES WITH MAX LENGTH >7 AND >9
sum(hw.cons[,"Max"]>6)/nrow(hw.cons)*100
sum(hw.cons[,"Max"]>9)/nrow(hw.cons)*100
# % OF CONSECUTIVE HW DAYS ABOVE 3 AND 7
colSums(hw.cons[,c(">3",">7")])/sum(hw.cons[,"N"])*100

################################
# RESULTS: FIRST ANALYSIS
################################

label <- paste("hw",rep(c(2,4),each=3),rep(c(97,98,99),2),sep=".")
table1 <- matrix(NA,6,7,dimnames=list(label,
  c("N comm","Est.main","95%CI.main","P-het.added","Est.added",
    "95%CI.added","P-het.added")))

for(i in 1:6) {

  # SET TO MISSING IF NO ESTIMATE FOR ADDED EFFECT
  added.eff[added.eff[,2*i]==0,c(2*i-1,2*i)] <- NA
  main.eff[is.na(added.eff[,2*i]),c(2*i-1,2*i)] <- NA

  # RUN THE META-ANALYSIS
  pool.main <- rma.uni(yi=main.eff[,2*i-1],sei=main.eff[,2*i])
  pool.added <- rma.uni(yi=added.eff[,2*i-1],sei=added.eff[,2*i])
  # FILL TABLE1
  table1[i,] <- c(sum(!is.na(added.eff[,2*i-1])),
    round(exp(pool.main$b)*100-100,1),
    paste(round(exp(pool.main$b-1.96*pool.main$se)*100-100,1),"to",
    round(exp(pool.main$b+1.96*pool.main$se)*100-100,1)),
    round(pool.main$QEp,3),
    round(exp(pool.added$b)*100-100,1),
    paste(round(exp(pool.added$b-1.96*pool.added$se)*100-100,1),"to",
```

```
    round(exp(pool.added$b+1.96*pool.added$se)*100-100,1)),
    round(pool.added$QEp,3))
}

# TABLE 1 IN THE MANUSCRIPT
table1

################################
# RESULTS: SECOND ANALYSIS
################################

# CREATE THE BASIS VARIABLES FOR PREDICTION
x <- 0:100/10
x.quad <- bs(x,knots=c(2,5,8),degree=2,Bound=c(0,10))
x.strata <- mkbasis(0:20/2,type="strata",knots=c(1,2,4,6,8))$basis

# PLOT
quad.plot <- cbind(x.quad%*%t(quad.pool.est),
  sqrt(diag(x.quad%*%quad.pool.vcov%*%t(x.quad))))

plot(x,exp(quad.plot[,1]),type="n",ylim=c(0.95,1.10),yaxt="n",
  ylab="Percent change %",
  xlab="Number of consecutive HW days",frame.plot=F)
axis(2,labels=-1:2*5,at=0.95+0:3*0.05)
polygon(c(x,rev(x)),c(exp(quad.plot[,1]+1.96*quad.plot[,2]),
  rev(exp(quad.plot[,1]-1.96*quad.plot[,2]))),border=NA,col=grey(0.9))
abline(h=1)
lines(x,exp(quad.plot[,1]))

strata.plot <- cbind(x.strata%*%t(strata.pool.est),
  sqrt(diag(x.strata%*%strata.pool.vcov%*%t(x.strata))))
lines(0:20/2,exp(strata.plot[,1]),type="S",lty=2)
```

References

Analitis, A., Katsouyanni, K., Biggeri, A., Baccini, M., Forsberg, B., Bisanti, L., Kirchmayer, U., Ballester, F., Cadum, E., Goodman, P., et al. (2008). Effects of cold weather on mortality: results From 15 European cities within the PHEWE Project. *American Journal of Epidemiology*, 168(12):1397.

Anderson, B. G. and Bell, M. L. (2009). Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology*, 20(2):205–213.

Armstrong, B. (2006). Models for the relationship between ambient temperature and daily mortality. *Epidemiology*, 17(6):624–31.

Baccini, M., Biggeri, A., Accetta, G., Kosatsky, T., Katsouyanni, K., Analitis, A., Anderson, H. R., Bisanti, L., D'Ippoliti, D., Danova, J., Forsberg, B., Medina, S., Paldy, A., Rabczenko, D., Schindler, C., and Michelozzi, P. (2008). Heat effects on mortality in 15 European cities. *Epidemiology*, 19(5):711–9.

Dominici, F., McDermott, A., Daniels, M., Zeger, S. L., and Samet, J. M. (2005). Revised analyses of the National Morbidity, Mortality, and Air Pollution Study: mortality among residents of 90 cities. *Journal of Toxicology and Environmental Health: Part A*, 68(13-14):1071–92.

Dominici, F., Sheppard, L., and Clyde, M. (2003). Health effects of air pollution: a statistical review. *International Statistical Review*, 71(2):243–276.

Gasparrini, A., Armstrong, B., and Kenward, M. G. (2010). Distributed lag non-linear models. *Statistics in Medicine*, page Epub ahead of print May 7. DOI: 10.1002/sim.3940.

Hajat, S., Armstrong, B., Baccini, M., Biggeri, A., Bisanti, L., Russo, A., Paldy, A., Menne, B., and Kosatsky, T. (2006). Impact of high temperatures on mortality: is there an added heat wave effect? *Epidemiology*, 17(6):632–8.

Hertel, S., Le Tertre, A., Jockel, K. H., and Hoffmann, B. (2009). Quantification of the heat wave effect on cause-specific mortality in Essen, Germany. *European Journal of Epidemiology*, 24(8):407–14.

Jackson, D., White, I. R., and Thompson, S. G. (2010). Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine*, 29(12):1282–1297.

Kaiser, R., Le Tertre, A., Schwartz, J., Gotway, C. A., Daley, W. R., and Rubin, C. H. (2007). The effect of the 1995 heat wave in Chicago on all-cause and cause-specific mortality. *American Journal of Public Health*, 97 Suppl 1:S158–62.

Le Tertre, A., Lefranc, A., Eilstein, D., Declercq, C., Medina, S., Blanchard, M., Chardon, B., Fabre, P., Filleul, L., Jusot, J. F., Pascal, L., Prouvost, H., Cassadou, S., and Ledrans, M. (2006). Impact of the 2003 heatwave on all-cause mortality in 9 French cities. *Epidemiology*, 17(1):75–9.

Michelozzi, P., Accetta, G., De Sario, M., D'Ippoliti, D., Marino, C., Baccini, M., Biggeri, A., Anderson, H., Katsouyanni, K., Ballester, F., et al. (2009). High temperature and hospitalizations for cardiovascular and respiratory causes in 12 European cities. *American Journal of Respiratory and Critical Care Medicine*, 179(5):383–389.

Muggeo, V. M. (2008). Modeling temperature effects on mortality: multiple segmented relationships with common break points. *Biostatistics*, 9(4):613–620.

Peng, R. D. and Dominici, F. (2009). A Bayesian hierarchical distributed lag model for estimating the time course of risk of hospitalization associated with particulate matter air pollution. *Journal of the Royal Statistical Society: Series A*, 58(1):3–24.

White, I. R. (2009). Multivariate random-effects meta-analysis. *Stata Journal*, 9(1):40–56.

Part III

Discussion

Chapter 12

Final comments

In this final chapter I provide some conclusive comments about my research on statistical methods for investigating the association between temperature and human health. In the first Section 12.1, I anticipate the potential future development of the two statistical frameworks of distributed lag non-linear models and multivariate meta-analysis, within and beyond the field of temperature-health studies. A final discussion is then provided in Section 12.2.

12.1 Future developments

During my PhD project and in drafting the publications included in Part II, I have attempted to provide a comprehensive methodological description, and software implementation, of the two statistical frameworks. However, the research on the two techniques is far from being concluded, and several potential extensions, already been planned, will be hopefully carried out soon.

These future developments are firstly stimulated by the need to improve further the analytical approaches to study the health effects of temperature. However, in the related research papers in Chapters 5-6 and 9, I deliberately provided a very general and wide-ranging definition of distributed lag non-linear models and multivariate meta-analysis for multi-parameter associations, respectively. De-

spite the specific example applications, the two methodologies are illustrated as general statistical tools, and are potentially applicable in different study designs and research fields.

Extension of distributed lag non-linear models

Within future research on DLNMs, relevant problems which need to be addressed are model selection and the related issue of optimal degree of smoothing. As thoroughly discussed in the related publications, several alternative modelling choices are available in order to describe non-linear and delayed effects. These choices refer to the use of different functions in the two dimensions of predictor and lags and, for continuous functions such as splines, to the degree of smoothing. Given that DLNMs may be simply considered as standard regression models involving a complex lag parameterization, traditional selection criteria, like the Akaike and Bayesian-type information criteria suggested in Gasparrini *et al.* (2010), are already available. The three-dimensional structure of the models, however, implies additional complexities which require further research.

As already discussed (Gasparrini, 2011; Gasparrini *et al.*, 2010), the DLNM framework has been developed so far using a completely parametric approach. In models using spline functions, the flexibility and smoothness of the effect surface is only determined by the number and position of knots. A straightforward extension would involve the use of penalized regression with low-rank smoothers, exploiting the ongoing research on semi-parametric approaches (Ruppert *et al.*, 2003; Wood, 2006) and tensor product smoothing (Eilers *et al.*, 2006).

Beside these purely statistical advancements, the main development of DLNMs is focused on extending the method beyond the specific applications in temperature-health studies or more generally in environmental time series. Actually, the basic definition of these models is easily generalizable, both conceptually and algebraically, and also applies to different data structures and study designs. I have already exploited the framework to model delayed effects and latency in case-control, cohort and longitudinal data. The extension of the framework sounds feasible and promising.

I also plan to implement all the extensions of the methodology illustrated above in the dlnm package. Some efforts is being made to promote the use of the software among applied researches in different fields, and to propose it as a general tool for investigating associations with delayed effects.

Research on multivariate meta-analysis

Multivariate meta-analytic techniques have been the object of an intense research in the last few years (Jackson *et al.*, 2011). The main interest still lies in the original application for multiple outcomes in randomized controlled trials, although its use for describing multi-parameter associations is closely related. Among current research topics, the further development of tests and statistics for heterogeneity, and the critical comparison of the relative performance of different estimation methods are of particular importance. It is worth mentioning that multivariate meta-analysis and meta-regression may be specified as linear mixed models, and that the wide research on this framework may be exploited in this particular setting.

Regarding the application in multi-parameter associations within a two-stage design, methodological problems and research directions have already been discussed in the research paper included in Chapter 9. Among other issues, a topic which deserves further research is the problem of dimensionality. As the number of outcomes increases, the specification of the model becomes computationally problematic, in particular regarding the complexity of the between-study (co)variance matrix. A possible solution is to specify simpler structures for the matrix, defined on a limited number of (co)variance components, together with robust estimation of the standard errors for the fixed effects in the model. However, this approach requires further research.

I plan to extend the R package mvmeta accordingly. Although in the research paper in Chapter 9 this software has been applied for describing multi-parameter associations in a two-stage analysis, my aim is to provide a general tool for multivariate meta-analysis and meta-regression, applicable for different research purposes. The availability of the software, together with implementations in alternative statistical programs (White, 2009, 2011) will hopefully boost the application of the methodology among applied researchers.

12.2 Conclusions

In Chapter 1, Section 1.1, I discussed how the health effects of temperature has been a matter of growing concern in the last decade, particularly in relation with extreme weather events and with predicted climate change scenarios. Several epidemiological studies have been performed in order to define

the association between temperature and human health. The results provided by these investigation are important to characterize the physiological mechanisms involved, to assess the exposure-response relationship, to identify vulnerable sub-groups and, in general, to deepen our knowledge on the epidemiology of temperature. This evidence is paramount to set up public health interventions and policies, in order to prevent or mitigate the effects of current and future exposures.

The appropriateness of this research process is dependent on the availability of suitable statistical methods, capable of providing reliable results on the association under study. However, as described in Section 1.4, the analysis of temperature-health dependencies shows peculiar and additional complexities, and traditional statistical tools for environmental time series, largely developed for assessing the effect of air pollution, may turn out to be inadequate in this new context. The development of the two methodologies of distributed lag non-linear models and multivariate meta-analysis for multi-parameter associations provides some tools to improve the analytical approaches in this fields. In addition, the implementation in a freely available statistical software facilitates the application of these methods among applied researchers.

Although recently proposed and published, these statistical methods and related software seems to represent a valid and useful tool for the research community. Even if an accurate literature review has not been performed, I can name at least six publications by other research teams which applied the DLNM methodology and used the dlnm package for investigating temperature-health associations (Barnett *et al.*, 2010; Guo *et al.*, 2011; Lin *et al.*, 2011; Yu *et al.*, 2011a,b,c). In addition, not surprisingly, the framework has also been used for assessing the effects of air pollution (Barnett *et al.*, 2011; Guo *et al.*, 2010a,b,c; Zhou *et al.*, 2011) and for methodological research (Strand *et al.*, 2011). This is reassuring about the importance of the research I carried out within my PhD project, and a strong motivation to develop it further in the future.

Bibliography

Almon, S. (1965) The distributed lag between capital appropriations and expenditures. *Econometrica*, **33**, 178–196.

Analitis, A., Katsouyanni, K., Biggeri, A., Baccini, M., Forsberg, B., Bisanti, L., Kirchmayer, U., Ballester, F., Cadum, E., Goodman, P. G. *et al.* (2008) Effects of cold weather on mortality: results from 15 European cities within the PHEWE Project. *American Journal of Epidemiology*, **168**(12), 1397.

Anderson, B. G. and Bell, M. L. (2009) Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology*, **20**(2), 205–213.

Armstrong, B. (2006) Models for the relationship between ambient temperature and daily mortality. *Epidemiology*, **17**(6), 624–31.

Armstrong, B. G. (2003) Fixed factors that modify the effects of time-varying factors: applying the case-only approach. *Epidemiology*, **14**(4), 467–472.

Armstrong, B. G., Chalabi, Z., Fenn, B., Hajat, S., Kovats, S., Milojevic, A. and Wilkinson, P. (2010) The association of mortality with high temperatures in a temperate climate: England and Wales. *Journal of Epidemiology and Community Health*, **65**, 340–345.

Aylin, P., Morris, S., Wakefield, J., Grossinho, A., Jarup, L. and Elliott, P. (2001) Temperature, housing, deprivation and their relationship to excess winter mortality in Great Britain, 1986-1996. *International Journal of Epidemiology*, **30**(5), 1100–1108.

Baccini, M., Biggeri, A., Accetta, G., Kosatsky, T., Katsouyanni, K., Analitis, A., Anderson, H. R., Bisanti, L., D'Ippoliti, D., Danova, J., Forsberg, B., Medina, S., Paldy, A., Rabczenko, D., Schindler, C. and Michelozzi, P. (2008) Heat effects on mortality in 15 European cities. *Epidemiology*, **19**(5), 711–9.

Ballester, F., Corella, D., Perez-Hoyos, S., Saez, M. and Hervas, A. (1997) Mortality as a function of temperature. A study in Valencia, Spain, 1991-1993. *International Journal of Epidemiology*, **26**(3), 551–61.

Barnett, A. G. (2007) Temperature and cardiovascular deaths in the US elderly: changes over time. *Epidemiology*, **18**(3), 369.

Barnett, A. G., Tong, S. and Clements, A. C. A. (2010) What measure of temperature is the best predictor of mortality? *Environmental Research*, **110**(6), 604–611.

Barnett, A. G., Fraser, J. F. and Munck, L. (2011) The effects of the 2009 dust storm on emergency admissions to a hospital in Brisbane, Australia. *International Journal of Biometeorology*, (Published onlineahead of print. DOI:10.1007/s00484-011-0473-y).

Basu, R. (2009) High ambient temperature and mortality: a review of epidemiological studies from 2001 to 2008. *Environmental Health*, **8**(1), 40.

Basu, R. and Ostro, B. D. (2008) A multicounty analysis identifying the populations vulnerable to mortality associated with high ambient temperature in California. *American Journal of Epidemiology*, **168**(6), 632–7.

Basu, R. and Samet, J. M. (2002a) Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiologic Reviews*, **24**(2), 190–202.

Basu, R. and Samet, J. M. (2002b) An exposure assessment study of ambient heat exposure in an elderly population in Baltimore, Maryland. *Environmental Health Perspectives*, **110**(12), 1219–1224.

Bell, M. L., Samet, J. M. and Dominici, F. (2004) Time-series studies of particulate matter. *Annual Review of Public Health*, **25**, 247–80.

Bell, M. L., Peng, R. D. and Dominici, F. (2006) The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. *Environmental Health Perspectives*, **114**(4), 532–6.

Bell, M. L., O'Neill, M. S., Ranjit, N., Borja-Aburto, V. H., Cifuentes, L. A. and Gouveia, N. C. (2008) Vulnerability to heat-related mortality in Latin America: a case-crossover study in Sao Paulo, Brazil, Santiago, Chile and Mexico City, Mexico. *International Journal of Epidemiology*, **37**(4), 796–804.

Berkey, C. S., Anderson, J. J. and Hoaglin, D. C. (1996) Multiple-outcome meta-analysis of clinical trials. *Statistics in Medicine*, **15**(5), 537–557.

Berkey, C. S., Hoaglin, D. C., Antczak-Bouckoms, A., Mosteller, F. and Colditz, G. A. (1998) Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine*, **17**(22), 2537–2550.

Borrell, C., Mari-Dell'Olmo, M., Rodriguez-Sanz, M., Garcia-Olalla, P., Cayla, J. A., Benach, J. and Muntaner, C. (2006) Socioeconomic position and excess mortality during the heat wave of 2003 in Barcelona. *European Journal of Epidemiology*, **21**(9), 633–40.

Braga, A. L., Zanobetti, A. and Schwartz, J. (2001a) The time course of weather-related deaths. *Epidemiology*, **12**(6), 662–7.

Braga, A. L., Zanobetti, A. and Schwartz, J. (2001b) The lag structure between particulate air pollution and respiratory and cardiovascular deaths in 10 US cities. *Journal of Occupational and Environmental Medicine*, **43**(11), 927–33.

Braga, A. L., Zanobetti, A. and Schwartz, J. (2002) The effect of weather on respiratory and cardiovascular deaths in 12 U.S. cities. *Environmental Health Perspectives*, **110**(9), 859–63.

Carder, M., McNamee, R., Beverland, I., Elton, R., Cohen, G. R., Boyd, J. and Agius, R. M. (2005) The lagged effect of cold temperature and wind chill on cardiorespiratory mortality in Scotland. *Occupational and Environmental Medicine*, **62**(10), 702–10.

Carder, M., McNamee, R., Beverland, I., Elton, R., Van Tongeren, M., Cohen, G. R., Boyd, J., Macnee, W. and Agius, R. M. (2008) Interacting effects of particulate pollution and cold temperature on cardiorespiratory mortality in Scotland. *Occupational and Environmental Medicine*, **65**(3), 197–204.

Carson, C., Hajat, S., Armstrong, B. and Wilkinson, P. (2006) Declining vulnerability to temperature-related mortality in London over the 20th century. *American Journal of Epidemiology*, **164**(1), 77–84.

Conti, S., Meli, P., Minelli, G., Solimini, R., Toccaceli, V., Vichi, M., Beltrano, C. and Perini, L. (2005) Epidemiologic study of mortality during the Summer 2003 heat wave in Italy. *Environmental Research*, **98**(3), 390–9.

Conti, S., Masocco, M., Meli, P., Minelli, G., Palummeri, E., Solimini, R., Toccaceli, V. and Vichi, M. (2007) General and specific mortality among the elderly during the 2003 heat wave in Genoa (Italy). *Environmental Research*, **103**(2), 267–74.

Curriero, F. C., Heiner, K. S., Samet, J. M., Zeger, S. L., Strug, L. and Patz, J. A. (2002) Temperature and mortality in 11 cities of the eastern United States. *American Journal of Epidemiology*, **155**(1), 80–7.

Daniels, M. J., Dominici, F., Samet, J. M. and Zeger, S. L. (2000) Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. *American Journal of Epidemiology*, **152**(5), 397.

Davis, R. E., Knappenberger, P. C., Michaels, P. J. and Novicoff, W. M. (2003) Changing heat-related mortality in the United States. *Environmental Health Perspectives*, **111**(14), 1712–1719.

Diggle, P. (1990) *Time series: a biostatistical introduction.* Oxford University Press.

Dominici, F. (2004) Time-series analysis of air pollution and mortality: a statistical review. *Research report - Health Effects Institute*, **123**, 3–27; discussion 29–33.

Dominici, F., Samet, J. M. and Zeger, S. L. (2000) Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. *Journal of the Royal Statistical Society: Series A*, **163**(3), 263–302.

Dominici, F., Daniels, M. J., Zeger, S. L. and Samet, J. M. (2002) Air pollution and mortality: estimating regional and national dose-response relationships. *Journal of the American Statistical Association*, **97**, 100–111.

Dominici, F., McDermott, A., Zeger, S. L. and Samet, J. M. (2003a) Airborne particulate matter and mortality: timescale effects in four US cities. *American Journal of Epidemiology*, **157**(12), 1055–65.

Dominici, F., Sheppard, L. and Clyde, M. (2003b) Health effects of air pollution: a statistical review. *International Statistical Review*, **71**(2), 243–276.

Eilers, P. H. C., Currie, I. D. and Durban, M. (2006) Fast and compact smoothing on large multidimensional grids. *Computational Statistics and Data Analysis*, **50**(1), 61–76.

Eurowinter Group (1997) Cold exposure and winter mortality from ischaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe. *The Lancet*, **349**(9062), 1341–1346.

Filleul, L., Le Tertre, A., Baldi, I. and Tessier, J. F. (2004) Difference in the relation between daily mortality and air pollution among elderly and all-ages populations in southwestern France. *Environmental Research*, **94**(3), 249–53.

Gasparrini, A. (2011) Distributed lag linear and non-linear models in R: the package dlnm. *Journal of Statistical Software*, **43**(8), 1–20.

Gasparrini, A. and Armstrong, B. (2010) Time series analysis on the health effects of temperature: Advancements and limitations. *Environmental Research*, **110**(6), 633–638.

Gasparrini, A. and Armstrong, B. (2011) The impact of heat waves on mortality. *Epidemiology*, **22**(1), 68–73.

Gasparrini, A., Armstrong, B. and Kenward, M. G. (2010) Distributed lag non-linear models. *Statistics in Medicine*, **29**(21), 2224–2234.

Gasparrini, A., Armstrong, B., Kovats, S. and Wilkinson, P. (2011) The effect of high temperatures on cause-specific mortality in England and Wales. *Occupational and Environmental Medicine*, (Published on-line ahead of print. DOI:10.1136/oem.2010.059782).

Goldberg, M. S., Gasparrini, A., Armstrong, B. and Valois, M. F. (2011) The short-term influence of temperature on daily mortality in the temperate climate of Montreal, Canada. *Environmental Research*, **111**(6), 853–60.

Goodman, P. G., Dockery, D. W. and Clancy, L. (2004) Cause-specific mortality and the extended effects of particulate pollution and temperature exposure. *Environmental Health Perspectives*, **112**(2), 179–85.

Gorjanc, M. L., Flanders, W. D., VanDerslice, J., Hersh, J. and Malilay, J. (1999) Effects of temperature and snowfall on mortality in Pennsylvania. *American Journal of Epidemiology*, **149**(12), 1152–1160.

Gosling, S. N., Lowe, J. A., McGregor, G. R., Pelling, M. and Malamud, B. D. (2009) Association between elevated athmospheric temperatures and human mortality: a critical review of the literature. *Climate Change*, **92**, 299–341.

Gouveia, N., Hajat, S. and Armstrong, B. (2003) Socioeconomic differentials in the temperature-mortality relationship in Sao Paulo, Brazil. *International Journal of Epidemiology*, **32**(3), 390–7.

Guo, Y., Barnett, A., Zhang, Y., Tong, S., Yu, W. and X, P. (2010a) The short-term effect of air pollution on cardiovascular mortality in Tianjin, China: comparison of time series and case-crossover analyses. *Science of the Total Environment*, **409**(2), 300–306.

Guo, Y., Tong, S., Li, S., Barnett, A., , Yu, Zhang, Y. W. and X, P. (2010b) Gaseous air pollution and emergency hospital visits for hypertension in Beijing, China: a time-stratified case-crossover study. *Environmental Health*, **9**, 57.

Guo, Y., Tong, S., Zhang, Y., G., B. A., Jia, Y. and X, P. (2010c) The relationship between particulate air pollution and emergency hospital visits for hypertension in Beijing, China. *Science of the Total Environment*, **408**(20), 4446–4450.

Guo, Y., Barnett, A. G., Yu, W., Pan, X., Ye, X., Huang, C. and Tong, S. (2011) A large change in temperature between neighbouring days increases the risk of mortality. *PloS One*, **6**(2), e16511.

Hajat, S. and Haines, A. (2002) Associations of cold temperatures with GP consultations for respiratory and cardiovascular disease amongst the elderly in London. *International Journal of Epidemiology*, **31**(4), 825–30.

Hajat, S., Kovats, R. S., Atkinson, R. W. and Haines, A. (2002) Impact of hot temperatures on death in London: a time series approach. *Journal of Epidemiology and Community Health*, **56**(5), 367–72.

Hajat, S., Armstrong, B. G., Gouveia, N. and Wilkinson, P. (2005) Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. *Epidemiology*, **16**(5), 613–20.

Hajat, S., Armstrong, B., Baccini, M., Biggeri, A., Bisanti, L., Russo, A., Paldy, A., Menne, B. and Kosatsky, T. (2006) Impact of high temperatures on mortality: is there an added heat wave effect? *Epidemiology*, **17**(6), 632–8.

Hajat, S., Kovats, R. S. and Lachowycz, K. (2007) Heat-related and cold-related deaths in England and Wales: who is at risk? *Occupational and Environmental Medicine*, **64**(2), 93–100.

Hoffmann, B., Hertel, S., Boes, T., Weiland, D. and Jockel, K. H. (2008) Increased cause-specific mortality associated with 2003 heat wave in Essen, Germany. *Journal of Toxicology and Environmental Health: Part A*, **71**(11-12), 759–65.

Huynen, M. M., Martens, P., Schram, D., Weijenberg, M. P. and Kunst, A. E. (2001) The impact of heat waves and cold spells on mortality rates in the Dutch population. *Environmental Health Perspectives*, **109**(5), 463–70.

Ihaka, R. and Gentleman, R. (1996) R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**(3), 299–314.

Jackson, D., Riley, R. and White, I. R. (2011) Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*, **30**(20), To appear. DOI: 10.1002/sim.4172.

Kalkstein, L. S. and Greene, J. S. (1997) An evaluation of climate/mortality relationships in large US cities and the possible impacts of a climate change. *Environmental Health Perspectives*, **105**(1), 84.

Katsouyanni, K., Zmirou, D., Spix, C., Sunyer, J., Schouten, J. P., Ponka, A., Anderson, H. R., Le Moullec, Y., Wojtyniak, B., Vigotti, M. A., Bacharova, L. and Schwartz, J. (1997) Short-term effects of air pollution on health: a European approach using epidemiologic time series data. The APHEA Project. Air Pollution Health Effects–A European Approach. *Public Health Reviews*, **25**(1), 7–18; discussion 19–28.

Keatinge, W. R., Donaldson, G. C., Cordioli, E., Martinelli, M., Kunst, A. E., Mackenbach, J. P., Nayha, S. and Vuori, I. (2000) Heat related mortality in warm and cold regions of Europe: observational study. *British Medical Journal*, **321**(7262), 670–3.

Knowlton, K., Rotkin-Ellman, M., King, G., Margolis, H. G., Smith, D., Solomon, G., Trent, R. and English, P. (2009) The 2006 California heat wave: impacts on hospitalizations and emergency department visits. *Environmental Health Perspectives*, **117**(1), 61–67.

Kovats, R. S. and Hajat, S. (2008) Heat stress and public health: a critical review. *Annual Review of Public Health*, **29**, 41–55.

Le Tertre, A., Lefranc, A., Eilstein, D., Declercq, C., Medina, S., Blanchard, M., Chardon, B., Fabre, P., Filleul, L., Jusot, J. F., Pascal, L., Prouvost, H., Cassadou, S. and Ledrans, M. (2006) Impact of the 2003 heatwave on all-cause mortality in 9 French cities. *Epidemiology*, **17**(1), 75–9.

Lin, Y. K., Ho, T. J. and Wang, Y. C. (2011) Mortality risk associated with temperature and prolonged temperature extremes in elderly populations in Taiwan. *Environmental Research*, (Published on-line ahead of print. DOI:10.1016/j.envres.2011.06.008).

Luber, G. and McGeehin, M. (2008) Climate change and extreme heat events. *American Journal of Preventive Medicine*, **35**(5), 429–435.

McGeehin, M. A. and Mirabelli, M. (2001) The potential impacts of climate variability and change on temperature-related morbidity and mortality in the United States. *Environmental Health Perspectives*, **109(Suppl 2)**, 185–9.

McMichael, A. J., Anderson, H. R., Brunekreef, B. and Cohen, A. J. (1998) Inappropriate use of daily mortality analyses to estimate longer-term mortality effects of air pollution. *International Journal of Epidemiology*, **27**(3), 450–3.

Medina-Ramn, M., Zanobetti, A., Cavanagh, D. P. and Schwartz, J. (2006) Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis. *Environmental Health Perspectives*, **114**(9), 1331.

Medina-Ramon, M. and Schwartz, J. (2007) Temperature, temperature extremes, and mortality: a study of acclimatization and effect modification in 50 United States cities. *Occupational and Environmental Medicine*, **64**, 827–833.

Meehl, G., Zwiers, F., Evans, J., Knutson, T., Mearns, L. and Whetton, P. (2000) Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change. *Bulletin of the American Meteorological Society*, **81**(3), 427–436.

Michelozzi, P., De Sario, M., Accetta, G., de'Donato, F., Kirchmayer, U., D'Ovidio, M. and Perucci, C. A. (2006) Temperature and summer mortality: geographical and temporal variations in four Italian cities. *Journal of Epidemiology and Community Health*, **60**(5), 417–23. HHWWS Collaborative Group.

Michelozzi, P., Kirchmayer, U., Katsouyanni, K., Biggeri, A., McGregor, G., Menne, B., Kassomenos, P., Anderson, H. R., Baccini, M., Accetta, G., Analytis, A. and Kosatsky, T. (2007) Assessment and prevention of acute health effects of weather conditions in Europe, the PHEWE project: background, objectives, design. *Environmental Health*, **6**, 12.

Muggeo, V. M. (2008) Modeling temperature effects on mortality: multiple segmented relationships with common break points. *Biostatistics*, **9**(4), 613–620.

Muggeo, V. M. R. (2010) Analyzing temperature effects on mortality within the R environment: the constrained segmented distributed lag parameterization. *Journal of Statistical Software*, **32**(12), 1–17.

Nafstad, P., Skrondal, A. and Bjertness, E. (2001) Mortality and temperature in Oslo, Norway, 1990-1995. *European Journal of Epidemiology*, **17**(7), 621–7.

National Reasearch Council (NRC) (2000) Reconciling observations of global temperature change. *Technical Report 86*, National Research Council.

Naughton, M. P., Henderson, A., Mirabelli, M. C., Kaiser, R., Wilhelm, J. L., Kieszak, S. M., Rubin, C. H. and McGeehin, M. A. (2002) Heat-related mortality during a 1999 heat wave in Chicago. *American Journal of Preventive Medicine*, **22**(4), 221–7.

O'Neill, M. S. and Ebi, K. L. (2009) Temperature extremes and health: impacts of climate variability and change in the United States. *Journal of Occupational and Environmental Medicine*, **51**(1), 13–25.

ONeill, M. S., Zanobetti, A. and Schwartz, J. (2005) Disparities by race in heat-related mortality in four US cities: the role of air conditioning prevalence. *Journal of Urban Health*, **82**(2), 191–197.

Pattenden, S., Nikiforov, B. and Armstrong, B. G. (2003) Mortality and temperature in Sofia and London. *Journal of Epidemiology and Community Health*, **57**(8), 628–33.

Patz, J. A., Campbell-Lendrum, D., Holloway, T. and Foley, J. A. (2005) Impact of regional climate change on human health. *Nature*, **438**(7066), 310–317.

Peng, R. D. and Dominici, F. (2008) *Statistical Methods for Environmental Epidemiology with R - A Case Study in Air Pollution and Health*. Use R! New York: Springer.

Peng, R. D., Dominici, F. and Louis, T. A. (2006) Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society: Series A*, **169**(2), 179–203.

Poumadere, M., Mays, C., Le Mer, S. and Blong, R. (2005) The 2003 heat wave in France: dangerous climate change here and now. *Risk Analysis*, **25**(6), 1483–94.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rabl, A. (2005) Air pollution mortality: harvesting and loss of life expectancy. *Journal of Toxicology and Environmental Health: Part A*, **68**(13-14), 1175–80.

Rainham, D. G. and Smoyer-Tomic, K. E. (2003) The role of air pollution in the relationship between a heat stress index and human mortality in Toronto. *Environmental Research*, **93**(1), 9–19.

Rey, G., Jougla, E., Fouillet, A., Pavillon, G., Bessemoulin, P., Frayssinet, P., Clavel, J. and Hemon, D. (2007) The impact of major heat waves on all-cause and cause-specific mortality in France from 1971 to 2003. *International Archives of Occupational and Environmental Health*, **80**(7), 615–26.

Rey, G., Fouillet, A., Bessemoulin, P., Frayssinet, P., Dufour, A., Jougla, E. and Hemon, D. (2009) Heat exposure and socio-economic vulnerability as synergistic factors in heat-wave-related mortality. *European Journal of Epidemiology*, **24**(9), 495–502.

Robinson, P. J. (2001) On the definition of a heat wave. *Journal of Applied Meteorology*, **40**, 762–775.

Rocklov, J. and Forsberg, B. (2008) The effect of temperature on mortality in Stockholm 1998–2003: a study of lag structures and heatwave effects. *Scandinavian Journal of Public Health*, **36**(5), 516–23.

Ruppert, D., Wand, M. P. and Carroll, R. J. (2003) *Semiparametric Regression*. Cambridge University Press.

Samet, J. M., Zeger, S. L., Dominici, F., Curriero, F., Coursac, I. and Dockery, D. W. (2000a) The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 2. Morbidity and Mortality from Air Pollution in the United States. *Technical report*, Health Effects Institute.

Samet, J. M., Zeger, S. L., Dominici, F., Dockery, D. and Schwartz, J. (2000b) The National Morbidity, Mortality, and Air Pollution Study (NMMAPS). Part 1. Methods and Methodological Issues. *Technical report*, Health Effects Institute.

Samoli, E., Touloumi, G., Zanobetti, A., Le Tertre, A., Schindler, C., Atkinson, R., Vonk, J., Rossi, G., Saez, M., Rabczenko, D. *et al.* (2003) Investigating the dose-response relation between air pollution and total mortality in the APHEA-2 multicity project. *Occupational and Environmental Medicine*, **60**(12), 977–982.

Samoli, E., Analitis, A., Touloumi, G., Schwartz, J., Anderson, H. R., Sunyer, J., Bisanti, L., Zmirou, D., Vonk, J. M., Pekkanen, J. *et al.* (2005) Estimating the exposure-response relationships between particulate matter and mortality within the APHEA multicity project. *Environmental Health Perspectives*, **113**(1), 88–97.

Samoli, E., Peng, R., Ramsay, T., Pipikou, M., Touloumi, G., Dominici, F., Burnett, R., Cohen, A., Krewski, D., Samet, J. *et al.* (2008) Acute effects of ambient particulate matter on mortality in Europe and North America: results from the APHENA study. *Environmental Health Perspectives*, **116**(11), 1480–1486.

Samoli, E., Zanobetti, A., Schwartz, J., Atkinson, R., Le Tertre, A., Schindler, C., Perez, L., Cadum, E., Pekkanen, J., Paldy, A., Touloumi, G. and Katsouyanni, K. (2009) The temporal pattern of mortality responses to ambient ozone in the APHEA project. *Journal of Epidemiology and Community Health*, **63**, 960–966.

Schwartz, J. (2000) The distributed lag between air pollution and daily deaths. *Epidemiology*, **11**(3), 320–6.

Schwartz, J. (2001) Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? *Epidemiology*, **12**(1), 55–61.

Schwartz, J. (2005) Who is sensitive to extremes of temperature?: A case-only analysis. *Epidemiology*, **16**(1), 67–72.

Schwartz, J., Spix, C., Touloumi, G., Bacharova, L., Barumamdzadeh, T., le Tertre, A., Piekarksi, T., Ponce de Leon, A., Ponka, A., Rossi, G., Saez, M. and Schouten, J. P. (1996) Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *Journal of Epidemiology and Community Health*, **50 (Suppl 1)**, S3–S11.

Schwartz, J., Laden, F. and Zanobetti, A. (2002) The concentration-response relation between PM(2.5) and daily deaths. *Environmental Health Perspectives*, **110**(10), 1025–9.

Semenza, J. C., Rubin, C. H., Falter, K. H., Selanikio, J. D., Flanders, W. D., Howe, H. L. and Wilhelm, J. L. (1996) Heat-related deaths during the July 1995 heat wave in Chicago. *New England Journal of Medicine*, **335**(2), 84–90.

Semenza, J. C., McCullough, J. E., Flanders, W. D., McGeehin, M. A. and Lumpkin, J. R. (1999) Excess hospital admissions during the July 1995 heat wave in Chicago. *American Journal of Preventive Medicine*, **16**(4), 269–77.

Smoyer, K. E., Rainham, D. G. C. and Hewko, J. N. (2000) Heat-stress-related mortality in five cities in Southern Ontario: 1980-1996. *International Journal of Biometeorology*, **44**(4), 190–197.

Stafoggia, M., Forastiere, F., Agostini, D., Biggeri, A., Bisanti, L., Cadum, E., Caranci, N., de' Donato, F., De Lisio, S., De Maria, M., Michelozzi, P., Miglio, R., Pandolfi, P., Picciotto, S., Rognoni, M., Russo, A., Scarnato, C. and Perucci, C. A. (2006) Vulnerability to heat-related mortality: a multicity, population-based, case-crossover analysis. *Epidemiology*, **17**(3), 315–23.

Stafoggia, M., Forastiere, F., Agostini, D., Caranci, N., de'Donato, F., Demaria, M., Michelozzi, P., Miglio, R., Rognoni, M., Russo, A. and Perucci, C. A. (2008a) Factors affecting in-hospital heat-related mortality: a multi-city case-crossover analysis. *Journal of Epidemiology and Community Health*, **62**(3), 209–15.

Stafoggia, M., Schwartz, J., Forastiere, F. and Perucci, C. A. (2008b) Does temperature modify the association between air pollution and mortality? A multicity case-crossover analysis in Italy. *American Journal of Epidemiology*, **167**(12), 1476–85. SISTI Group.

Strand, L. B., Barnett, A. G. and Tong, S. (2011) Methodological challenges when estimating the effects of season and seasonal exposures on birth outcomes. *BMC Medical Research Methodology*, **11**(1), 49.

Touloumi, G., Atkinson, R., Le Tertre, A., Samoli, E., Schwartz, J., Schindler, C., Vonk, J. M., Rossi, G., Saez, M. and Rabszenko, D. (2004) Analysis of health outcome time series data in epidemiological studies. *EnvironMetrics*, **15**(2), 101–117.

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

Verbeke, G. and Molenberghs, G. (2000) *Linear mixed models for longitudinal data*. New York: Springer Verlag.

White, I. R. (2009) Multivariate random-effects meta-analysis. *Stata Journal*, **9**(1), 40–56.

White, I. R. (2011) Multivariate random-effects meta-regression: updates to mvmeta. *Stata Journal*, **11**(2), 255–270.

Wilkinson, P., Pattenden, S., Armstrong, B., Fletcher, A., Kovats, R. S., Mangtani, P. and McMichael, A. J. (2004) Vulnerability to winter mortality in elderly people in Britain: population based study. *British Medical Journal*, **329**(7467), 647.

Wood, S. N. (2006) *Generalized Additive Models: an Introduction with R*. Chapman & Hall/CRC.

Yu, W., Guo, Y., Ye, X., Wang, X., Huang, C., Pan, X. and Tong, S. (2011a) The effect of various temperature indicators on different mortality categories in a subtropical city of Brisbane, Australia. *Science of the Total Environment*, **409**(18), 3431–3437.

Yu, W., Hu, W., Mengersen, K., Guo, Y., Pan, X., Connell, D. and Tong, S. (2011b) Time course of temperature effects on cardiovascular mortality in Brisbane, Australia. *Heart*, **97**(13), 1089–1093.

Yu, W., Mengersen, K., Hu, W., Guo, Y., Pan, X. and Tong, S. (2011c) Assessing the relationship between global warming and mortality: lag effects of temperature fluctuations by age and mortality categories. *Environmental Pollution*, **159**(7), 1789–1793.

Zanobetti, A. and Schwartz, J. (2008) Temperature and mortality in nine US cities. *Epidemiology*, **19**(4), 563–70.

Zeger, S. L., Irizarry, R. and Peng, R. D. (2006) On time series analysis of public health and biomedical data. *Annual Review of Public Health*, **27**, 57–79.

Zhou, J., Ito, K., Lall, R., Lippmann, M. and Thurston, G. (2011) Time-series analysis of mortality effects of fine particulate matter components in Detroit and Seattle. *Environmental Health Perspectives*, **119**(4), 461466.