

Models for the Relationship Between Ambient Temperature and Daily Mortality

Ben Armstrong

Background: Ambient temperature is an important determinant of daily mortality that is of interest both in its own right and as a confounder of other determinants investigated using time-series regressions, in particular, air pollution. The temperature–mortality relationship is often found to be substantially nonlinear and to persist (but change shape) with increasing lag. We review and extend models for such nonlinear multilag forms.

Development: Popular models for mortality by temperature at given lag include polynomial and natural cubic spline curves, and the simple but more easily interpreted linear thresholds model, comprising linear relationships for temperatures below and above thresholds and a flat middle section. Most published analyses that have allowed the relationship to persist over multiple lags have done so by assuming that spline or threshold models apply to mean temperature in several lag strata (eg, lags 0–1, 2–6, and 7–13). However, more flexible models are possible, and a modeling framework using products of basis functions (“cross-basis” functions) suggests a wide range, some used previously and some new. These allow for stepped or smooth changes in the model coefficients as lags increase. Applying a range of models to data from London suggest evidence for relationships up to at least 2 weeks’ lag, with smooth models fitting best but lag-stratified threshold models allowing the most direct interpretation.

Conclusions: A wide range of multilag nonlinear temperature–mortality relationships can be modeled. More awareness of options should improve investigation of these relationships and help control for confounding by them.

(*Epidemiology* 2006;17: 624–631)

Ambient temperature is a major cause of fluctuations in mortality over time.^{1,2} To study this association from time-series data, investigators have used a wide variety of statistical methods recently dominated by explicit “time-series regression” approaches. Many of the developments in these methods for air pollution epidemiology, in particular,

for modeling temperature as a potential confounder, have direct relevance for studies in which temperature is the focus.³ Furthermore, several studies of effects of temperature have extended the air pollution methodology for this purpose.^{2,4–16} However, there has been no systematic discussion of methods for studies seeking to elucidate temperature effects. By providing such a discussion, we believe that the selection of models is liable to be less ad hoc and better suited to its purpose.

Specifically, we seek to review, clarify, and extend methods for investigating associations of mortality with temperature focusing on 2 key issues: the nonlinear form of the relationship and the spread of effect over several days or weeks (distributed lags). We begin the article by describing the data we use to illustrate the problems and models. After describing the general time-series regression model, we focus on specific components to address nonlinearity in single-lag models and multiple lags in linear-threshold models. We propose a general modeling framework for multilag nonlinear relationships. This framework places into a single context several methods that have been used previously^{2,4–16} and suggests some new alternatives. Finally, we discuss model choice.

For easier reading of the main text, we have placed more technical content in an appendix, available with the online version of this article.


EXAMPLE DATA SET: LONDON

Table 1 shows univariate descriptive statistics for daily cardiovascular disease (CVD) mortality, weather, and potential confounder data from London for 1993 through 2003. Relative humidity, black smoke (an indicator of particulate air pollution), and counts of laboratory confirmed influenza (A and B) were potentially confounding variables for which data were available.

The left 2 graphs on Figure 1 show the patterns of temperature and of mortality over time showing the expected seasonal patterns. These figures and all results shown and programs used were obtained using Stata software (Stata Corp., College Station, TX). The right graphs show scatter plots of the crude relationship of mortality with temperature; the top right panel displays one point per day and the bottom right panel shows data aggregated to centile bins as a crude transparent smoothing device. The figures on the right suggest a reverse “J” shape with increases in mortality at temperatures lower and higher than approximately 17 degrees. However, the plots on the left indicate that there is potential confounding by other seasonal factors.

Submitted 1 August 2005; accepted 9 May 2006.

From the London School of Hygiene and Tropical Medicine, London, UK.

 Supplemental material for this article is available with the online version of the journal at www.epidem.com

Correspondence: Ben Armstrong, London School of Hygiene and Tropical Medicine, Keppel St., London, UK, WC1E 7HT. E-mail: ben.armstrong@lshtm.ac.uk

Copyright © 2006 by Lippincott Williams & Wilkins

ISSN: 1044-3983/06/1706-0624

DOI: 10.1097/01.ede.0000239732.50999.8f

TABLE 1. Distribution of Variables Over Days, London, 1994–2003 (n = 3652)

Variable	Minimum	Median	Maximum
Temperature (°C); mean	−2.4	11.8	26.5
Relative humidity (%)	42.2	77.3	99.7
Black smoke ($\mu\text{g}/\text{m}^3$)	1.0	10	78.8
Influenza (sample count)	0	2	29
Cardiovascular disease deaths; no.	35	71	114

TIME-SERIES REGRESSION MODEL

To model more formally, and in particular to control for confounding, we follow usual practice³ in assuming an overdispersed Poisson regression model for Y_i , mortality count on day i on observed temperature t_i , and other time-varying factors:

$$E(Y_i) = \exp\{\alpha + f(t_i | \beta) + (\text{specific covariates}) + (\text{smooth function of time})\} \quad (1)$$

Our main focus in this article is the functional form of the temperature term $f(t_i | \beta)$, for which we discuss options in depth subsequently. For the remainder of the model (confounder control), we merely summarize a single variant we used for the London illustration, which is broadly in line with current practice. We acknowledge that this approach places optimal choice of these terms (an important discussion) beyond the scope of this article. For our London analyses, specific potentially confounding time-varying risk factors (eg, pollution) comprised:

- Black smoke (linear, lags 0, 1, and 2);

- Weekly reports of influenza virus isolated by laboratories routinely making assays for the National Health Service (linear, lags 0, 1, and 2 weeks); and
- Relative humidity (natural cubic spline 3 df, lag 0).

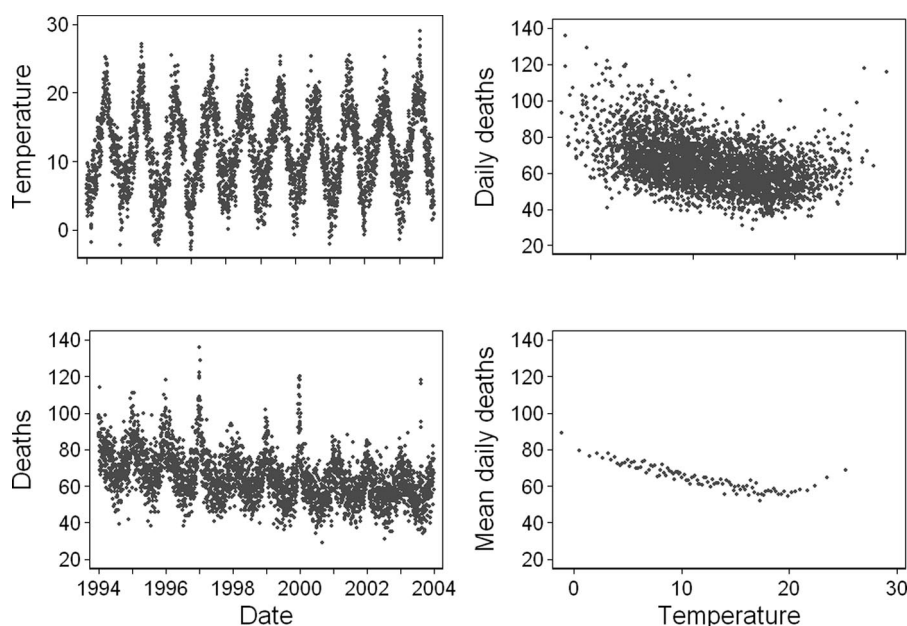
The smooth function of time (date), to allow for changes due to seasonal effects and demographic shift or other slow change not captured in the covariates, comprised a combination of trigonometric terms to 6 harmonics to capture repeated annual (ie, seasonal) patterns and a natural cubic spline of date with 2 degrees of freedom to capture slow change over the 11 years.

In Figure 2, we show the same-day temperature–mortality relationship aggregating temperature to centile bins as before unadjusted and with confounders modeled as described previously.

MODELING THE TEMPERATURE EFFECT

Generally, like in the London data, mortality is found to be lowest at average temperatures and higher at low and high temperatures. Thus, a simple log linear function is not usually adequate. The relationships are often described as U-, V-, or J-shaped.⁹ In air pollution studies, investigators usually model temperature as a general smooth curve, and this also serves well (at least for graphic presentation) if temperature is the focus of interest.

We prefer the natural cubic spline because it is constrained to be linear at the boundaries of the temperature range, in which data are usually too sparse to support more detail. For all smooth curves, investigators must choose the degree of smoothness (usually degrees of freedom) and, for some (such as natural cubic splines), the placing of “knots.” Choice of smoothness (number of knots) is usually more important than precise knot placement, with many knots tending to show many wiggles (some possibly spurious) and

**FIGURE 1.** Daily cardiovascular disease mortality and temperature (°C) in London 1994–2003: patterns over time and crude association.

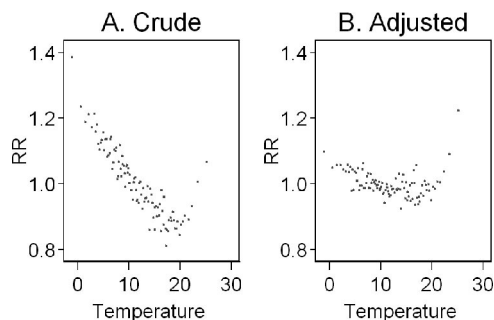


FIGURE 2. Daily cardiovascular disease mortality and temperature (°C) in London: crude and adjusted association.

too few knots losing real wiggles. (We discuss approaches to this and other model choices subsequently.)

General smooth curves are useful for describing relationships, but a simpler quantitative summary is also useful, in particular if relationships are to be compared across places or population subgroups. Classification of temperature into bins is possible (ie, indicator basis variables). Another simple model attractive for interpretability is the 3-piece linear spline (segmented linear) function in which the middle section is constrained to have zero slope. In this model:

$$f(t_i | \beta) = \beta_C t_{C,i} + \beta_H t_{H,i},$$

with

$$t_{C,i} = \max[(\tau_C - t_i), 0] \quad (2)$$

and

$$t_{H,i} = \max[(t_i - \tau_H), 0]$$

The basis variables $t_{C,i}$ and $t_{H,i}$ represent excesses of temperature below the cold-effect threshold (τ_C) and above the heat-effect threshold (τ_H). The attraction of this model is the interpretability of its 4 parameters:

1. The high threshold temperature τ_H above which mortality rises;
2. The coefficient β_H of the high temperature effect;
3. The low threshold temperature τ_C below which mortality rises; and
4. The coefficient β_C of the low temperature effect.

We call this model the linear-thresholds model. The simpler “V” model^{5,17} is a special case, with $\tau_H = \tau_C$; as is the hockey-stick model,² with $\beta_C = 0$ or $\beta_H = 0$. An obvious extension is to allow the “middle” section to have nonzero slope, although this loses some of the interpretational advantages of the simpler model. The algorithm of Muggeo¹⁸ can be used to estimate thresholds; although in smaller and irregular data (especially rounded temperature data), this can converge to a local rather than the global maximum. Evaluation of profile likelihoods (ie, maximized for slopes) over a grid of threshold values is a more robust alternative but takes more processor time. Confidence intervals (CIs) for thresh-

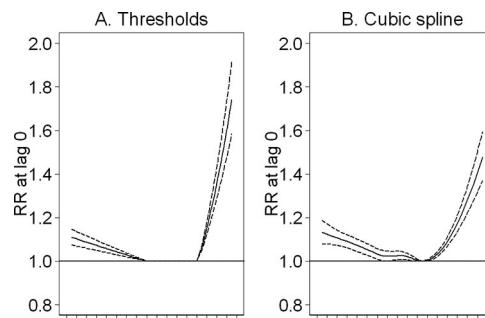


FIGURE 3. Thresholds and natural cubic spline models for lag 0 temperature (London cardiovascular disease deaths 1994–2003).

olds can be estimated from standard errors using Muggeo’s method or from the grid of profile likelihoods described previously.

Figure 3 shows a linear-thresholds model fit to (lag 0) temperature alongside a natural cubic spline with 5 df (default knot placing). Mortality is shown in the graph relative to the minimum fitted mortality over the observed temperature (see the online appendix).

For the thresholds model, the deviance was minimized at $\hat{\tau}_C = 12.1$, $\hat{\tau}_H = 22.3$. The likelihood profile 95% CI for τ_C was 9.6 to 12.9 and for τ_H , it was 21.2 to 24.4 found as the minimum and maximum values of τ_H , τ_C within 3.84 of the minimum deviance. Mortality increased by 0.7% (0.5 to 0.9) per degree below 12.1 degrees and by 8.6% (7.1 to 10.2) per degree above 22.3 degrees. Here and typically, the slopes are highly dependent on the thresholds. For example, the heat slope was correlated at $r = 0.73$ with the heat threshold (given by Muggeo’s method).

A General Framework for Modeling Nonlinear Functions: Basis Variables

To aid the extension to distributed lag models below, we use a framework—“basis variables”—for describing a wide range of nonlinear functions. The linear thresholds and other segmented linear models, cubic and natural cubic spline functions, as well as indicators (dummy variables) can all be represented in the general notation $f(x) = \sum_{b=1,B} \beta_b x_b$, where the x_b are called basis functions.

The form of basis functions depends on the basis used (eg, $x_b = x^b$ for a polynomial). In the regression context, we fit these nonlinear functions by 1) calculating variables (basis variables) from the original variable and then 2) including each of these basis variables in the regression model. The number of basis variables B is called the degrees of freedom for the curve.

Nonparametric curves such as LOESS, smoothing splines, or penalized splines cannot be represented directly as basis variables. Because later development uses the basis variable framework, we restrict our attention to curves that can be represented in this way.

DISTRIBUTED LAG MODELS

Several studies have found that the cold effect is spread over a week or more after a cold day (the delay varying by cause of death), whereas the heat effect is more immediate.^{2,7,14} Some form of distributed lag model seems a natural choice to describe such patterns.

Distributed lag models were introduced to epidemiologists in the air pollution literature,¹⁹ but for linear models. To extend this standard theory most simply to apply in our nonlinear context, we describe first the distributed lag model applied to the linear-thresholds model² with thresholds (T_H , T_C) constrained to be the same at all lags. The distributed lag model extension of expression² is then:

$$f(t_{\sim i} | \beta) = \sum_{l=0}^L \beta_{C,l} t_{C,i-l} + \sum_{l=0}^L \beta_{H,l} t_{H,i-l} \quad (3)$$

The function is now of $t_{\sim i}$ representing history of temperatures on days up to day i ; $\{t_i, t_{i-1}, t_{i-2}, \dots\}$. In this unconstrained distributed lag model, $L+1$ low and $L+1$ high temperature variables, with lags 0 to L , are simultaneously entered into the model. The parameters $\beta_{C,l}$ and $\beta_{H,l}$ represent the effect of low and high temperatures delayed by l days.

This model with lags up to $L = 27$ in London maximum likelihood gave $\hat{\tau}_C = 22.3$ (19.9 to 23.1) and $\hat{\tau}_H = 22.7$ (22.1 to 23.1). (The thresholds, estimated by maximum likelihood, are different from those estimated using lag one only because of the additional information in lags 1–28.) The top graphs on Figure 4 show the 2 X 28 coefficients of this model for cold (left) and heat (right).

The coefficients of linear distributed lag models have 2 equally valid interpretations. First, if temperature increases

by 1 degree over the heat threshold (or decreases 1 degree below the cold threshold), on 1 day, then mortality will change on that day and on each of the next 27 days as estimated by the 28 coefficients (points on the graph). Second, if temperature had increased 1 degree over the heat threshold for one day (say, day i) and each the previous 27 days, then each of those days would have an impact on mortality on day i , also represented by the 28 coefficients.

From either of these viewpoints, one summary of interest is the combined effect of a 1 degree increase over all 28 lags: $\beta_{C,sum} = \sum_{l=0}^L \beta_{C,l}$ (and analogously for heat). In the London data, this total effect of a 1 degree decrease in temperature below the cold threshold over 28 days was 1.9% (95% CI = 1.4 to 2.5), and above the heat threshold was 21.3% (15.3 to 27.7).

Although this unconstrained distributed lag model is useful to obtain overall estimates of effect and to control confounding by temperature when investigating pollution effects,^{13,16} it is problematic for elucidating the pattern of delay in effects because each individual coefficient is imprecisely estimated and is usually highly correlated with estimates of other coefficients. To overcome this problem, investigators have put constraints on the β 's, obtaining a constrained distributed lag model.

Most simply, the coefficients β_l for a range of lags (say, 1–3) can be constrained to be equal ($\beta_1 = \beta_2 = \beta_3 = \beta_{1-3}$). We call this the lag-stratified distributed lag model. It is equivalent to entering the mean of t_C over days 1 to 3, in place of the 3 separate days' temperatures in the regression, except that the regression coefficient in the mean model is 3 times that of the constrained model. The constrained coefficient represents the increment in log mortality per degree increase

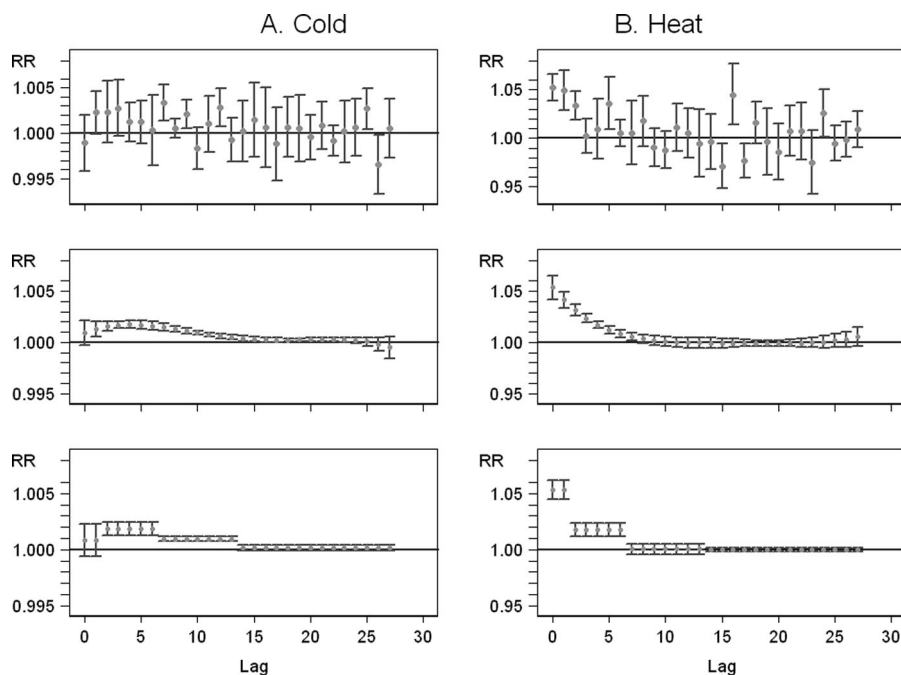


FIGURE 4. Distributed lag linear thresholds model (London cardiovascular disease deaths 1994–2003). Left: cold (below 22.3°C); right: heat (above 22.7°C). Top: unconstrained; middle: quartic; bottom: strata (lags 0–1, 2–6, 7–13, 14–27).

in temperature on one of lags 1 through 3; the coefficient of the mean represents the increment per degree sustained over lags 1 through 3.

The lower graph in Figure 4 shows a lag-stratified distributed lag model applied to the London data with coefficients assumed equal in lag intervals 0–1, 2–6, 7–13, and 14–27. These lag-lengthening intervals were chosen a priori to reflect evidence from previous work of greater change in coefficients over shorter than longer lags.^{12–14,17}

Other possible constraints allow smooth change in the β_l by making the β_l follow a polynomial,¹⁹ cubic spline, or penalized spline²⁰ function in lag l . The quartic constraint is illustrated in Figure 4. When interpreting coefficients from constrained distributed lag models (including lag-stratified models), we should be aware that they reflect information from the constraint as well as the data. For example, the degree of certainty with which the constrained distributed lag models for heat in Figure 4 indicate the persistence of the effect up to lag 7 seems speculative given the individual coefficients (top left). Estimated overall effect is, on the other hand, largely insensitive to constraints.

In comparison with some of the more flexible models we discuss subsequently, the linear-thresholds distributed lag models, in particular the lag-stratified model, have advantages in ease of presentation and interpretation and the ease with which epidemiologic measures such as relative risks and attributable fraction can be attached to them. Table 2 illustrates this. The left column shows the relative risk increment per degree of “cold” or “heat” sustained over each lag interval considered. Thus, the figure of 9.2% for heat in lags 2 through 6 is 5 times the increment per day in this period (seen in Fig. 4 as approximately 2%).

The calculation of percent of deaths attributable to heat and cold (Table 2, right column) complements this usefully, reflecting the number and degree of “cold” and “hot” days as well as the risk on those days. This is obtained from the standard

formula²¹ as the mean over all deaths of $(RR - 1)/RR = 1 - [\exp \sum_{l=0}^L \beta_{C,l} t_{C,l-i}]^{-1}$, and similarly for heat. These estimates are, of course, subject to the validity of the model.

Nonlinear Distributed Lag Models

To easily adapt standard distributed lag methods, we were obliged to simplify the temperature–mortality relationship at each lag to linear-threshold form. This may not be sufficiently flexible for all purposes. We now propose a general framework that includes and systematizes the linear thresholds model and some other curvilinear distributed lag models previously proposed^{7,11,16} and suggest extensions and a programming strategy. This framework requires that:

1. The dependence of mortality on temperature at each lag can be represented as a sum of “basis” functions of temperature $\{t_b, b = 1, B\}$ and coefficients, and
2. The dependence of those coefficients on lag can be similarly represented (basis $\{l_p, P = 1, P\}$).

This framework allows temperature–mortality and distributed lag dependences to be chosen independently from a menu, including natural cubic spline, polynomial, linear thresholds, or strata (indicators).

All such models can be fit using standard software with some preprocessing (the online appendix includes more detail, and Stata ado files are available on request from the author). Specifically, we compute BXP derived temperature variables $\overline{t_{b,p,i}}$ from the data for each day i and use these as linear terms in model 1. The $\overline{t_{b,p,i}} = \sum_{l=0}^L l_p t_{b,i-l}$ comprise sums of products of the basis functions for temperature and lag. We thus call them cross-basis functions. The coefficients of the derived temperature variables are not usually easily interpreted directly, but interpretable graphs, summaries, and statistical inference can be obtained from their estimates and standard errors (see the online appendix).

Table 3 displays the main options suggested by the cross-basis framework with references where they have had published description or use.

We illustrate in Figure 5 an “NCS–NCS” cross-basis model fit to the London data with a 5-df natural cubic spline for the temperature–mortality curves and a 4-df natural cubic spline for the coefficient–lag curves, hence $5 \times 4 = 20$ degrees of freedom (ie, derived temperature variables $\overline{t_{b,p,i}}$) overall. For the lag spline, interior knots were chosen at equal intervals logarithmically over 0 to L so as to allow more flexibility at lower lags. Although the 3-dimensional graph in Figure 5 shows the entire relationship of mortality against l and t , these graphs are not always easy to follow and cannot show confidence intervals, so graphs of “slices” and other derived measures of risk, shown in Figure 6, are often more useful.

Starting interpretation of Figure 6 with the graph of total impact over all 28 lags (bottom left graph), we see the familiar broadly “U”-shaped relationship with minimum at 18.8°C. We chose this multilag “minimum mortality temperature” as baseline for scaling relative risks in all the cross-basis graphs (see the online appendix). Risks would rise

TABLE 2. Simple Summary of Results From Lag-Stratified Distributed Lag-Thresholds Model (London, cardiovascular disease deaths, 1994–2003)

Variable (cold or heat)	Delay (lag) in Days	Relative Risk: % per °C Below or Above Threshold (95% CI)	Percent of Deaths Attributable to Hot or Cold Temperatures (95% CI)
Cold*	0–1	0.2 (–0.0 to 0.4)	1.9 (–0.3 to 3.9)
	2–6	0.9 (0.7 to 1.2)	9.5 (7.1 to 11.8)
	7–13	0.7 (0.4 to 0.9)	7.0 (4.6 to 9.3)
	14–27	0.2 (–0.1 to 0.5)	2.2 (–0.8 to 5.1)
	Total	2.0 (1.6 to 2.4)	19.0 (15.7 to 22.1)
Heat*	0–1	10.8 (8.8 to 12.9)	0.3 (0.3 to 0.4)
	2–6	9.2 (6.6 to 11.9)	0.3 (0.2 to 0.3)
	7–13	0.3 (–2.5 to 3.1)	0.0 (–0.1 to 0.1)
	14–27	0.0 (–3.5 to 3.7)	0.0 (–0.1 to 0.1)
	Total	21.4 (15.6 to 27.4)	0.6 (0.4 to 0.8)

*Cold threshold = 22.3°C; heat threshold = 22.7°C.

TABLE 3. A Selection of Distributed Lag Models Falling Within the Cross-Basis Framework With Specification of Basis Functions and References to Some Published Applications

Temperature–Mortality Model (basis t_b)	Distributed Lag Model (basis l_b)			
	Unconstrained (indicators for each lag 0–L)	Quartic Polynomial ($1, l, l^2, l^3, l^4$)	Strata (indicators for lag intervals)	NCS (as provided by package)
Linear (t)	Schwartz ^{19*}	Schwartz ^{19*}	—*	Zanobetti ^{20†}
Quadratic (t, t^2)	—	Braga ^{7,9}	—	—
Natural cubic spline (as provided by package)	—	—	Curriero, ¹¹ Welty ^{16‡}	This paper
Linear thresholds (t_C, t_H)	Keatinge, ¹³ Hajat ¹²	This paper	Kunst, ¹⁷ Ballester, ⁵ Pattenden ¹⁴	—

*These models have been used extensively for air pollution. The lag-strata model is equivalent to fitting means in lag strata.

†Zanobetti used unrestricted cubic splines.

‡Fitted NCS to mean temperature in lag strata—similar but see the online appendix for slight difference. Also, many other air pollution researchers used this model to control temperature effects.

according to the model to approximately 1.5 times this minimum at extremes of low and high temperature if these temperatures were sustained over 28 days.

Considering now the temperature–mortality relationship at different lags (from top left), we see that heat (temperature above 18.8°C) shows an adverse effect at lags 0 and 3, but by lags 11 and 21, a slight apparent deficit of deaths is observed. This long lag deficit is also seen in the graph of mortality by lag at 25°C (bottom right). It has been noted in other studies using other models^{7,12} and interpreted as evidence for a degree of “harvesting” heat-causing deaths in persons who would have otherwise died anyway 2 or 3 weeks later. The deficits at longer lags are subtracted from the “overall impact” estimate of the bottom graph.

The absence of the long lag heat deficit in the linear thresholds distributed lag model (Fig. 4) is noteworthy and seems likely to be due to the high heat threshold estimated

(by maximum likelihood) in that model. Linear threshold-distributed lag models with lower heat thresholds showed the long lag deficits. It is itself noteworthy that the estimated minimum mortality temperature in the NCS-NCS model (18.8°C) is lower than the 2 estimated thresholds in the linear threshold-distributed lag model (22.3 and 22.7), suggesting sensitivity of this aspect of the relationship to model choice.

The pattern for cold, seen either from the graphs of mortality by temperature (left) or by lag (right), was however similar to that shown by the linear threshold-distributed lag model.

Choosing Among Modeling Options

Our review and proposed framework has indicated a wide variety of possible models, each with choices of how to implement. In particular, analysts must choose:

- Form of temperature–mortality curve at a given lag (polynomial, strata, natural cubic spline, linear thresholds);
- Form of change in coefficients by lag (unconstrained, polynomial, strata, natural cubic spline);
- Maximum lag; and
- Extent and place of wiggles:
 - for lag-stratified models: strata intervals,
 - for NCS models: number and placing of knots, and
 - for polynomial models: order.

Each of these choices will depend on the objectives of the analysis as well as the model fit. In general, simpler models (eg, lag-stratified thresholds) have advantages of easy interpretability, but more complex models (eg, natural cubic spline) may be required to better fit the data or clarify subtle points of the relationship.

The simpler variants seem particularly attractive in multicity studies in which one seeks to compare associations across cities (although parameters from any cross-basis model could in theory be input into a multivariate meta-analysis framework to explore cross-city variation). However, the more flexible models may be useful in exploratory single-city studies and can serve to indicate to what extent

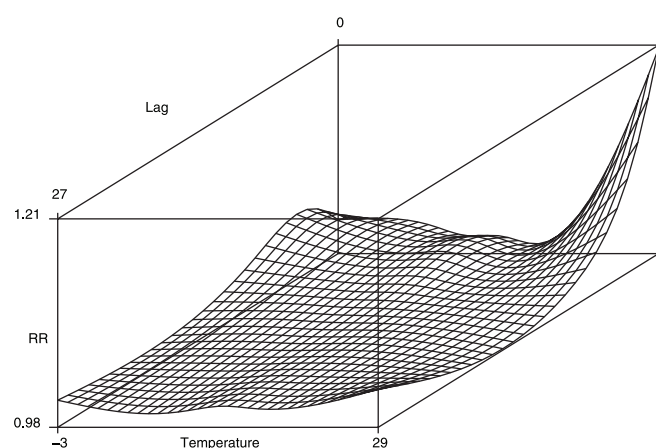


FIGURE 5. Relative risk of mortality by temperature (°C) and lag using cross-basis smoothing by natural cubic splines of 5 df (temperature) and 4 df (lag) (London cardiovascular disease deaths 1994–2003).

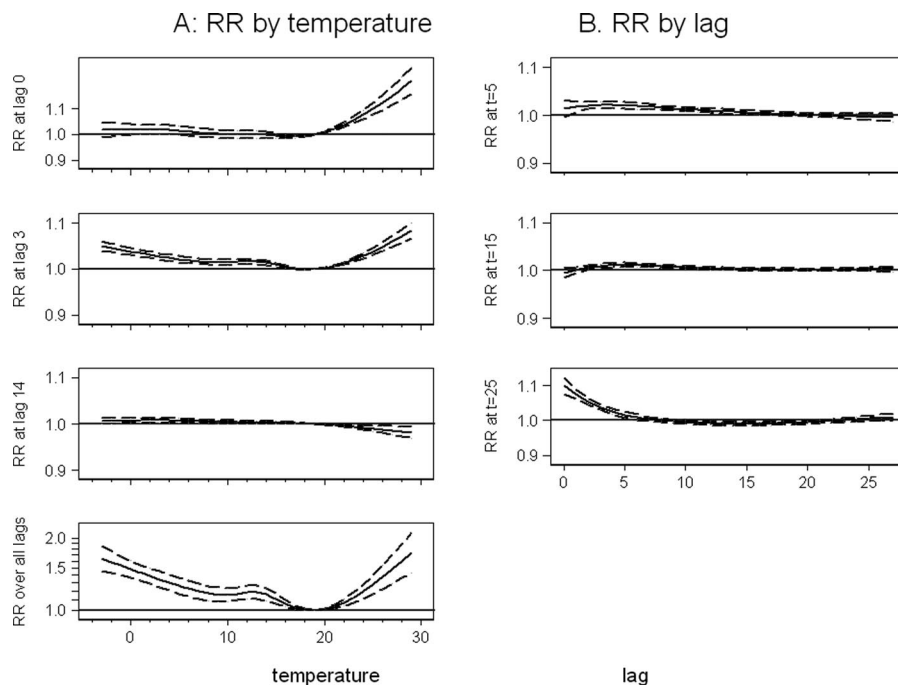


FIGURE 6. Relative risk of mortality by temperature and lag using cross-basis smoothing with natural cubic splines of 5 df (temperature) and 4 df (lag). Left from top: Mortality by temperature at selected lags 0, 3, 11, and 20. Left bottom: mortality by temperature sustained over 28 days. Right from top: Mortality by lag at selected temperatures 5, 15, and 25°C. Mortality is scaled to be relative to that at 18.8°C, the temperature that is sustained so mortality would be lowest.

there are systematic effects of temperature beyond that captured in the simple more easily interpretable model. Furthermore, when the objective is to control confounding by temperature, for example, if air pollution is the focus of interest, a more flexible model is less likely to leave residual confounding.

Choice of specific model may also be informed by model fit criteria such as deviance, Akaike's information criterion (AIC), or tests for white noise in residuals.²² These have the feature of being objective criteria, although they implement criteria (optimizing prediction of outcomes in new data in the case of AIC), which are not obviously relevant in epidemiology.

In the London data set, among a wide variety of models within the cross-basis framework with maximum lag 27, AIC favored the model with 8-df natural cubic spline for temperature and 4-df natural cubic spline for lag, and in larger data sets, AIC favored even more complex models. However, much simpler models (for example, the linear thresholds with 5 lag strata) were not far behind. Although it was useful to be aware that the simpler model had not captured all the relationship, for interpretation, the multiwiggled 8-df natural cubic spline seemed less useful than the 5-df natural cubic spline or linear-thresholds model.

Other specific modeling options may be suggested by the plots of flexible ones. For example, the cold effect appears more linear than the heat effect, suggesting fewer knots for lower temperatures. However, such ad hoc data-led model choices are not reflected in formal inference.

The choice of maximum lag similarly might draw on a priori or model fit criteria. If objectives include investigation of possible negative coefficients at longer lags consistent with harvesting, then we would choose a higher maximum lag. In

lag-stratified models, strength of evidence for association of longer lag intervals with mortality can be assessed directly. For example, in the lag-stratified linear thresholds model in London (Table 2), the 14- to 27-day lag period had relatively small nonsignificant coefficients, suggesting that reducing the maximum lag to 13 would lose little. An AIC comparison of models with maximum lag 13 and 27 supported this.

When using model-fit criteria to inform model choice, we must keep in mind that relative performance of each of the choices in the bulleted list previously depends on the other choices and on choices for the confounder model. For example, although AIC favored a maximum lag of 13 over 27 with the linear thresholds model as noted previously, with the 20-df natural cubic spline model (Figs. 5 and 6), a maximum lag of 27 was preferred over 13.

As well as considering choice of a preferred model, it is also desirable to consider sensitivity of conclusions to model choice not only in relation to the temperature component, but also that for season and other specific factors.

DISCUSSION

We have sought to make a systematic presentation of methods to model effects of temperature on mortality, which are generally simultaneously nonlinear and experienced with a variety of delays (distributed lags). In particular, we have proposed the cross-basis framework, which allows a wide variety of models across the spectrum of simplicity, including most of the models previously proposed and some new possibilities (Table 3).

Modeling temperature effects raises many other issues that we have addressed briefly or not at all: optimal control for confounding by season and other causes of fluctuations in

mortality over time; model diagnostics; “harvesting”¹²; modifiers of the temperature effect by other factors, in particular season^{5,6,16}; interactions of temperature with other weather variables (including the use of apparent temperature or synoptic indices); and cross-lag interactions (eg, sustained periods of heat or changes in temperature). Season is particularly problematic because of its close association (conceptual and empiric) with temperature. We believe, however, that the framework we proposed for multilag nonlinear models can allow these issues to be addressed with greater confidence that at least the main effects of temperature have been adequately modeled.

A limitation to the generality of the cross-basis model family is that the basis functions (and hence the number and placement of knots in splines and of thresholds in the linear-thresholds model) must be the same across lags. Relaxing this generally in a meaningful way appears difficult, but simple special cases are useful and present no problems. In particular, lag-stratified natural cubic spline or linear-threshold models could have different knots (thresholds) at different places for different lag strata without imposing problems, although estimating thresholds by maximum likelihood seems likely to be tricky.

We have used the London data to illustrate the methods discussed but do not advance the results as a definitive description of the effects of temperature in London. An analysis aimed at more robust substantive results would consider the sensitivity issues discussed previously and in the previous section. Furthermore, relationships will differ across cities and calendar time and with outcome. The results we show are however broadly in line with other studies of London mortality.¹⁴

In conclusion, temperature has a substantial impact on mortality in many places. To model this relationship, either for interest in its own right or for control of confounding of effects of other risk factors, we need to capture its multilag and nonlinear nature. A wide range of models are possible for this; the cross-basis family of models provides one framework to guide choice and computing.

REFERENCES

1. Basu R, Samet JM. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiol Rev*. 2002;24:190–202.
2. The Eurowinter G. Cold exposure and winter mortality from ischaemic heart disease, cerebrovascular disease, respiratory disease, and all causes in warm and cold regions of Europe. *Lancet*. 1997;349:1341–1346.
3. Dominici F. Time-series analysis of air pollution and mortality: a statistical review. *Res Rep Health Eff Inst*. 2004;123:3–27; discussion 29–33.
4. Alberdi JC, Diaz J, Montero JC, et al. Daily mortality in Madrid community 1986–1992: relationship with meteorological variables. *Eur J Epidemiol*. 1998;14:571–578.
5. Ballester F, Corella D, Perez-Hoyos S, et al. Mortality as a function of temperature. A study in Valencia, Spain, 1991–1993. *Int J Epidemiol*. 1997;26:551–561.
6. Basu R, Dominici F, Samet JM. Temperature and mortality among the elderly in the United States: a comparison of epidemiologic methods. *Epidemiology*. 2005;16:58–66.
7. Braga AL, Zanobetti A, Schwartz J. The time course of weather-related deaths. *Epidemiology*. 2001;12:662–667.
8. Braga AL, Zanobetti A, Schwartz J. The lag structure between particulate air pollution and respiratory and cardiovascular deaths in 10 US cities. *J Occup Environ Med*. 2001;43:927–933.
9. Braga AL, Zanobetti A, Schwartz J. The effect of weather on respiratory and cardiovascular deaths in 12 US cities. *Environ Health Perspect*. 2002;110:859–863.
10. Carder M, McNamee R, Beverland I, et al. The lagged effect of cold temperature and wind chill on cardiorespiratory mortality in Scotland. *Occup Environ Med*. 2005;62:702–710.
11. Currier FC, Heiner KS, Samet J, et al. Temperature and mortality in eleven cities of the eastern United States. *Am J Epidemiol*. 2002;155:80–87.
12. Hajat S, Armstrong BG, Gouveia N, et al. Mortality displacement of heat-related deaths: a comparison of Delhi, São Paulo, and London. *Epidemiology*. 2005;16:613–620.
13. Keatinge WR, Donaldson GC. Mortality related to cold and air pollution in London after allowance for effects of associated weather patterns. *Environ Res*. 2001;86:209–216.
14. Pattenden S, Nikiforov B, Armstrong BG. Mortality and temperature in Sofia and London. *J Epidemiol Community Health*. 2003;57:628–633.
15. Schwartz J, Samet JM, Patz JA. Hospital admissions for heart disease: the effects of temperature and humidity. *Epidemiology*. 2004;15:755–761.
16. 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. *Am J Epidemiol*. 2005;162:80–88.
17. Kunst AE, Looman CW, Mackenbach JP. Outdoor air temperature and mortality in The Netherlands: a time-series analysis. *Am J Epidemiol*. 1993;137:331–341.
18. Muggeo VM. Estimating regression models with unknown break-points. *Stat Med*. 2003;22:3055–3071.
19. Schwartz J. The distributed lag between air pollution and daily deaths. *Epidemiology*. 2000;11:320–326.
20. Zanobetti A, Wand MP, Schwartz J, et al. Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics*. 2000;1:279–292.
21. Bruzzi P, Green SB, Byar DP, et al. Estimating the population attributable risk for multiple risk factors using case-control data. *Am J Epidemiol*. 1985;122:904–914.
22. Peng R, Dominici F, Louis T. Model choice in multi-city time series studies of air pollution and mortality. *J Royal Stat Soc*. In press.