Environmental Research 164 (2018) 53–64

Contents lists available at ScienceDirect

Environmental Research

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

A new approach to modeling temperature-related mortality: Non-linear autoregressive models with exogenous input

Cameron C. Lee*, Scott C. Sheridan

Kent State University; Department of Geography, 413 McGilvrey Hall, 325 S. Lincoln St., Kent, OH 44242 USA

ARTICLE INFO

Keywords:
Non-linear auto-regressive models
NARX models
Temperature-related mortality
DLNM

ABSTRACT

Temperature-mortality relationships are nonlinear, time-lagged, and can vary depending on the time of year and geographic location, all of which limits the applicability of simple regression models in describing these associations. This research demonstrates the utility of an alternative method for modeling such complex relationships that has gained recent traction in other environmental fields: nonlinear autoregressive models with exogenous input (NARX models). All-cause mortality data and multiple temperature-based data sets were gathered from 41 different US cities, for the period 1975–2010, and subjected to ensemble NARX modeling. Models generally performed better in larger cities and during the winter season. Across the US, median absolute percentage errors were 10% (ranging from 4% to 15% in various cities), the average improvement in the r-squared over that of a simple persistence model was 17% (6–24%), and the hit rate for modeling spike days in mortality (> 80th percentile) was 54% (34–71%). Mortality responded acutely to hot summer days, peaking at 0–2 days of lag before dropping precipitously, and there was an extended mortality response to cold winter days, peaking at 2–4 days of lag and dropping slowly and continuing for multiple weeks. Spring and autumn showed both of the aforementioned temperature-mortality relationships, but generally to a lesser magnitude than what was seen in summer or winter. When compared to distributed lag nonlinear models, NARX model output was nearly identical. These results highlight the applicability of NARX models for use in modeling complex and time-dependent relationships for various applications in epidemiology and environmental sciences.

1. Introduction

There is a large, robust body of research showing an association between temperature and human mortality across a wide range of climate regions (Sheridan and Allen, 2015). Nearly universally, the highest temperatures experienced at any location are associated with increased mortality rates, as is cold winter weather, yielding an approximate J- or U-shaped curve to the temperature-mortality relationship (Ryti et al., 2016; Donaldson et al., 2003; McMichael et al., 2008). While such generalities persist across most of the literature, the precise nature of this association can differ markedly by climate region, cause of death, the modeling technique used, and other confounding factors (e.g. air quality, demographic differences) that play into this weather-health relationship (Sheridan and Allen, 2015). Along with these differences, the potential impacts of climate change (e.g. Sheridan and Dixon, 2017; Sheridan et al., 2012a, 2012b), and the application of novel modeling techniques (e.g. Gasparrini et al., 2017), have sparked a litany of different studies investigating temperature-related mortality in various locations, especially heat-related mortality (Gosling et al., 2009; Basu and Samet, 2002; Basu, 2009; Gasparrini et al., 2015).

Most often, the impacts of excessive heat events and cold events have been studied separately, either owing to an investigator's specific area of expertise, or due to the differences in terms of how each extreme is modeled. Studies linking extreme heat events and human mortality have generally observed a threshold thermal metric, above which mortality is observed to rise (Sheridan and Allen, 2015). This threshold varies spatially, with higher thresholds typically found in warmer locations. Further, the nature of the threshold may be different from location to location; high humidity plays a substantive role in reducing the human body's ability to cool itself and thus may be as critical of a factor as high temperatures. While there are direct deaths due to hyperthermia, excessive heat has been associated with increased mortality across many causes of death, in particular those related to cardiovascular and respiratory diseases (Gasparrini et al., 2015). Excessive heat events typically result in very immediate health responses, as increases in mortality are generally strongest from the day of the event to 2 days afterwards.

In contrast to heat-related mortality, cold-related mortality is less

* Corresponding author.
E-mail address: cclee@kent.edu (C.C. Lee).

https://doi.org/10.1016/j.envres.2018.02.020
Received 13 September 2017; Received in revised form 18 January 2018; Accepted 14 February 2018
0013-9351/

Fig. 1. The locations of the 41 cities used in this research (see Table 1 for city names for the abbreviations noted here).

often studied, despite the fact that the majority of annual deaths in many mid-latitude climates occur in the winter (Allen and Sheridan, 2014; Gasparrini et al., 2015). It is more difficult to directly link cold-related mortality to temperature effects due to the extended lag-time associated with increased mortality afterwards, usually peaking after 3–4 days, and continuing to be associated with increased mortality for up to 2–3 weeks afterwards (Lee, 2016; Allen and Sheridan, 2015; Anderson and Bell, 2009). The direct effects of cold temperatures can cause constriction of blood vessels, increased blood pressure, thickening of the blood and increases in platelet and red blood cell counts, all of which contribute to an increased risk of negative outcomes such as heart attacks and strokes (Ryti et al., 2016; Näyhä, 2002). The extended delayed response is likely an indirect result of lagged exposure to cold temperatures that can result in acute respiratory issues such as influenza and pneumonia, which may lead to death (Davis et al., 2016b; Kysely et al., 2009). Cold and dry multivariate weather types (Lee et al., 2016; Morabito et al., 2006; Grass and Kane, 2008; McGregor, 1999) and wind (Kim et al., 2016) have also been linked to increased wintertime mortality.

Due to the shape of the relationship between mortality and temperatures, traditional linear regression-based approaches to modeling are not sufficient for capturing both winter and summer temperature-mortality associations. One modeling technique gaining popularity recently is the distributed-lag non-linear model (DLNM) which not only overcomes the issue of non-linearity in the relationship, but also allows the lag-structure of the temperature-mortality relationship to be incorporated into the model (Gasparrini et al., 2010). A recently-developed artificial neural network (ANN)-based time-series modeling framework known as non-linear autoregressive models with exogenous input (or NARX models) shares this ability to approximate nonlinear time-dependent relationships (Guzman et al., 2017). Further, NARX models have the added benefits of: 1) not having to assume any particular shape to the temperature-mortality association; 2) being internally cross-validated on a separate portion of the dataset; and 3) being easily modified to run in a 'closed-loop' (feedback) mode allowing them to take on real-time, multi-step-ahead predictions and hindcasts (e.g. Lee et al., 2016).

The main goal of this paper is to demonstrate the effectiveness of this NARX modeling technique in temperature-related mortality research (and to the wider epidemiology and biometeorology communities in general), as it has gained popularity recently in other environmental research (e.g. Guzman et al., 2017, and references therein), including a recently completed project by the investigators (Lee et al., 2016). While a thorough description of ANNs and NARX models is outside the scope of this paper (c.f., Maier and Dandy, 2000; Diaconescu, 2008; Beale et al., 2014), in describing the NARX modeling framework specific to this study, we attempt to clarify NARX modeling lingo by using regression-specific terms when applicable. The goal of NARX modeling is the same as for a standard multiple regression: to develop a model that uses predictors in order to yield a best estimate of a predictand. To demonstrate this new methodology, we apply a suite of NARX models to examine the year-round apparent-temperature-mortality relationship across 41 different cities in the United States, and discuss the benefits and limitations of this technique in regards to its modeling performance. We also broadly compare the results of the NARX methodology with the output of DLNM models for four cities.

2. Materials and methods

2.1. Data

Mortality data for the United States have been acquired from the

Table 1
The 41 metropolitan areas used in this research. Listed for each area is the airport code for the station whose meteorological data are used, the 2010 population, mean apparent temperature, and the 10th and 90th percentile values for apparent temperature.

Metropolitan area	Airport Code	Population (millions)	Mean Daily AT		
			10th %ile	Mean	90th %ile
Albany	ALB	1.2	−11.0	5.8	22.4
Atlanta	ATL	5.3	0.0	15.0	28.2
Baltimore	BWI	2.7	−4.9	10.8	26.6
Birmingham	BHM	1.1	1.3	16.3	29.5
Boston	BOS	4.5	−8.8	6.4	21.9
Buffalo	BUF	1.1	−11.8	5.0	21.4
Chicago	ORD	9.5	−10.8	6.6	24.0
Cincinnati	CVG	2.1	−7.2	9.7	25.7
Cleveland	CLE	2.1	−9.8	6.8	23.1
Columbus	CMH	1.9	−8.0	8.9	25.0
Dallas	DFW	6.4	1.6	17.2	31.3
Denver	DEN	2.5	−7.3	6.5	20.8
Detroit	DTW	4.3	−10.7	6.3	23.2
Hartford	BDL	1.2	−8.9	7.2	23.4
Houston	IAH	5.9	6.5	20.7	32.3
Indianapolis	IND	1.9	−8.5	8.8	25.6
Kansas City	MCI	2.0	−8.2	9.6	27.0
Los Angeles	LAX	12.8	10.2	15.7	21.0
Memphis	MEM	1.3	−0.9	15.5	30.5
Miami	MIA	5.6	17.1	25.3	31.8
Milwaukee	MKE	1.6	−11.8	4.9	22.1
Minneapolis	MSP	3.3	−15.7	4.2	22.8
Nashville	BNA	1.7	−2.4	13.5	28.1
New York	LGA	19.6	−6.8	9.0	24.8
Orlando	MCO	2.1	11.5	22.3	30.7
Philadelphia	PHL	6.0	−5.9	10.2	26.3
Phoenix	PHX	4.2	9.6	21.8	34.7
Pittsburgh	PIT	2.4	−8.8	7.7	23.3
Portland	PDX	2.2	0.1	9.6	19.5
Providence	PVD	1.6	−8.1	7.1	22.8
Riverside	RIV	4.2	7.7	16.4	26.0
Rochester	ROC	1.1	−11.2	5.5	21.9
Sacramento	SAC	2.1	4.8	14.1	23.6
Saint Louis	STL	2.8	−6.3	11.3	28.3
San Antonio	SAT	2.1	5.9	20.0	31.4
San Diego	SAN	3.1	11.1	16.6	22.5
San Francisco	SFO	4.3	5.9	11.0	16.2
Seattle	SEA	3.4	0.0	8.3	17.2
Tampa	TPA	2.8	11.8	23.1	31.7
Virginia Beach	ORF	1.7	−1.8	13.4	27.8
Washington	DCA	5.6	−3.8	12.0	27.7

National Center for Health Statistics for the period 1975–2010. Daily all-cause mortality totals are aggregated to the metropolitan area level for all major metropolitan areas in the US, using the boundaries defined by the US Census in 2010. Use of these data required that no date have a total mortality of under 5; eliminating all metropolitan areas in which this occurred leaves 41 metropolitan areas for all analyses in this research (Fig. 1).

The universal apparent temperature (AT) developed by Steadman (1984) is one of the most commonly used metrics in assessing the impacts of extreme temperature events on human health outcomes (Anderson et al., 2013). In addition to temperature, this metric also accounts for atmospheric humidity, which can be critical in delineating summer heat events, as well as wind speed, critical for assessing the wind chill in cold events. Beyond the AT itself, many studies have shown an added heat wave effect on mortality, that is, that several consecutive days of very high temperatures will result in augmented impacts on human health beyond what would be expected by those days individually. The Excess Heat Factor (EHF), developed by Nairn and Fawcett (2014), is a statistical way of accounting for longer sequences of hot weather, especially when preceded by relatively cool conditions. It is defined in this work as the number of degrees of exceedance of the three-day mean AT above the 95th percentile AT for a location, multiplied by the difference between that three-day value and the 30-day mean before it. Thus, it highlights periods with extended hot weather that were preceded by cooler conditions.

Table 2
Median predictor variable importance (in NARX models) by season across all 41 cities examined. Darker red (green) coloring indicates increasing (decreasing) importance.

National	AT05	AT17	EHF	ATanom	Season	Trend
WINTER	21.0%	22.9%	0.0%	35.9%	7.6%	12.7%
SPRING	21.2%	25.8%	0.0%	29.2%	6.7%	17.2%
SUMMER	31.2%	25.7%	3.3%	17.1%	7.0%	15.7%
AUTUMN	24.1%	23.3%	0.0%	27.3%	5.1%	20.1%
ANNUAL	22.4%	24.9%	6.3%	24.6%	5.7%	16.0%

Table 3
NARX model monthly performance statistics averaged across all 41 locations examined. MdAPE refers to the median absolute percentage error between actual mortality and the RITE data set; R2IM reflects the improvement of the r2 statistic (of the RITE data set vs. actual mortality) over that of the simple 1-day autocorrelation of the actual observed mortality; and hit rate is the percentage of the time a 'spike day' (> 80th percentile) in actual mortality was also a 'spike day' (> 80th percentile) in the RITE output. Darker red (green) coloring indicates increasingly better (poorer) performance.

MONTH	MdAPE	R2IM	HitRate
JAN	9.9%	14.9%	46.6%
FEB	9.9%	15.4%	47.8%
MAR	9.9%	17.2%	49.5%
APR	10.0%	14.7%	44.0%
MAY	10.2%	14.2%	42.2%
JUN	10.3%	13.7%	41.6%
JUL	10.4%	13.5%	42.7%
AUG	10.4%	13.1%	41.2%
SEP	10.3%	13.5%	41.8%
OCT	10.2%	13.6%	40.7%
NOV	10.1%	14.3%	40.8%
DEC	9.9%	15.6%	46.8%
ANNUAL	10.1%	17.4%	53.7%

For this study, in addition to using the raw value of AT at two specific times (0500 and 1700 EDT), a mean daily AT (calculated from these two values) was transformed into a deseasonalized anomaly. The seasonally-relative baseline for this deseasonalized anomaly was computed by first setting each day in the month equal to its corresponding monthly mean AT (computed across all 36 years), and then subjecting the resulting step-like seasonal curve to a 31-day centered moving average to remove the step-like discontinuities. This smoothed seasonally-relative baseline was then subtracted from the observed mean AT on a daily basis, yielding the deseasonalized apparent temperature anomaly (AT_{anom}) that was included in each city's model. By using these three variables, we account for the fact that the best predictors of mortality – in terms of time of day – vary spatially (e.g. Davis et al., 2016a), and that there is evidence that the impact of heat may vary according to the time of year (e.g. Anderson and Bell, 2010; Ng et al., 2014).

Two additional time-relative variables were also included in every model: a sinusoidal seasonal-signal, and a simple linear (secular trend) variable counting the days in the time series (e.g. 1 January 1975 = 1) to account for population-related changes in raw mortality counts. A day-of-the-week signal was also trialed, but found to have a negative (if any) effect on model performance in many cities, and thus was excluded from final modeling.

The resulting set of six predictor variables (AT05, AT17, AT_{anom}, EHF, seasonal signal, secular trend) was ultimately chosen after multiple rounds of modeling using various datasets (including other meteorological variables) and input variable selection processes. When additional variables were used, the collinearity of these different meteorological variables and the lag-response differences of the

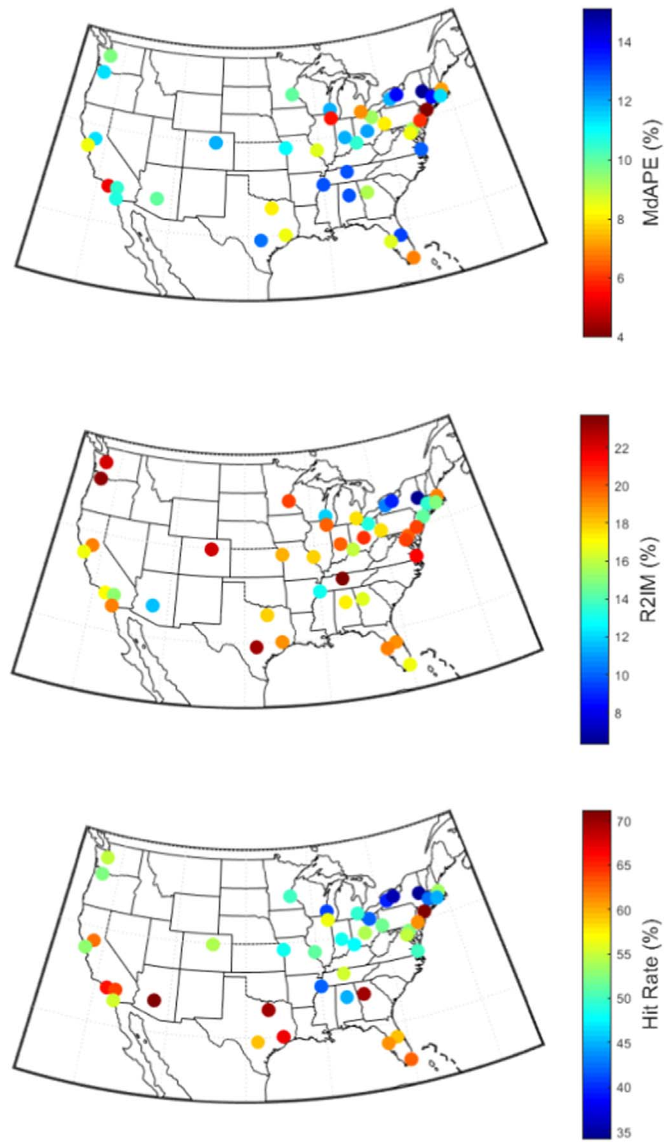

Fig. 2. NARX model performance metrics: MdAPE (top), R2IM (middle) and hit rate (bottom) for each city. Note that the color scales are different for each figure, and red indicates better performance in each figure (i.e. MdAPE is inverse). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

temperature-mortality relationship between cities resulted in highly diverse sets of predictor variables between different cities, which was considered detrimental to the aim of highlighting the modeling technique rather than predictor selection. Thus, we chose to use a standard set of predictors that both account for collinearity (since AT incorporates temperature, humidity and wind) and have been shown to correlate with all-cause mortality in previous research (Anderson et al., 2013).

2.2. NARX modeling methodology

The set of six daily weather variables (X) for each city were input as exogenous (predictor) variables into separate NARX models for each city using customized functions in Matlab's Neural Network Toolbox (Beale et al., 2014). In addition to the set of input variables, NARX models allow for a few different user-defined parameters: the training algorithm, the number of neurons, and the number of delays. Two commonly used NARX training algorithms in environmental research are Levenberg-Marquardt (LM) optimization and Bayesian Regularization (e.g. Guzman et al., 2017). Both of these training algorithms were tested, with neither yielding a consistently better performance across test cities, and the former ultimately was chosen due to its use in previous research by the investigators (Lee et al., 2016). Since the goal of this research is merely to demonstrate the ability of NARX models in describing the temperature-mortality relationship rather than to actually make predictions or hindcasts (as was the focus in Lee et al., 2016), we do not train or run the NARX models in a closed-loop fashion.

The LM-based NARX method begins by dividing the time series into three separate time-blocks of data, the training block (60%), the internal validation block (20%) and the independent testing block (20%). The training block is used by the ANN to 'learn' the relationship between the 6 predictors (X) and the predictand (mortality) by iteratively adjusting the weights (somewhat akin to beta coefficients) of the ANN in a manner that improves the performance of the model from the previous iteration (i.e. reduces the mean squared error (MSE) between the observations and the model output). After each training iteration, the resulting model is then simulated on the internal validation block of data to determine if the MSE on that portion of the data also improves. The NARX model continues iterating through these training and validation steps (known as epochs) until the MSE calculated from this internal validation data set fails to decrease (improve) after 10 successive iterations, indicating that the model trained 10 epochs ago is the most optimized model for both time-blocks of data. This internal validation step helps reduce the likelihood that the model becomes overfit to the training data set. Since the internal validation portion of the data cannot be considered completely independent of the model (because it is used to stop the training), the testing block is held out for independent, external assessment of the model's extrapolability by the user.

The number of neurons (N) in a NARX model corresponds to the amount of complexity the model can incorporate, and the number of delays (D) are equal to the number of time-lags (i.e. days). The number of delays can further be specified separately for the predictors (D_x) and the previous values of the predictand (D_y). Similar to a regression, the more total terms (T, with $T \approx (X * N * D_x) + (N * D_y)$) that are included in the model, the more likely it is to improve performance, however, it is also more likely to become overfit on the training data. This could dramatically increase the time it takes to stop the training (using the internal validation block of the data), and possibly prevent it. The optimum number of neurons and delays to incorporate into the model is not known beforehand, and thus brute-force testing of different combinations of N, D_y, and D_x is necessary (Lee et al., 2016). Herein, it was found that after 14 days of lag in D_y, there was very limited improvement in model performance (MSE) by further increasing D_y, and thus D_y was set to 14 across all locations. However, for N and D_x, each city was tested with different combinations to see which combination yielded the best performance (within the range of $N = 1–10$ and $D_x = 1–4$); the resulting winning NARX model-architectures for N and D_x can be found in supplementary material (Table S.1).

While the LM methodology described above helps prevent over-fitting, a drawback is that it would result in conclusions being drawn based only upon the last 20% of the dataset. To overcome this limitation, Lee et al. (2016) used a block-jackknifing technique, whereby multiple NARX models were fully trained and internally validated using different *settings* for the partitioning of the time-blocks. Herein, we use the same approach, using five separate settings in which a different 20% of the data was used as the testing-block (and by extension, a different 60% for training and a different 20% for internal validation) in its own NARX model. By chronologically stitching together the testing-block of each of these five NARX models' output time-series, this allows us to re-construct an entire 36-year time-series of output data that can be considered independent from training. We acknowledge that, since the weather-mortality relationship is not stationary over time (e.g. Sheridan and Dixon, 2017), there are limitations in terms of how mortality is predicted using block-jackknifing, in particular for the

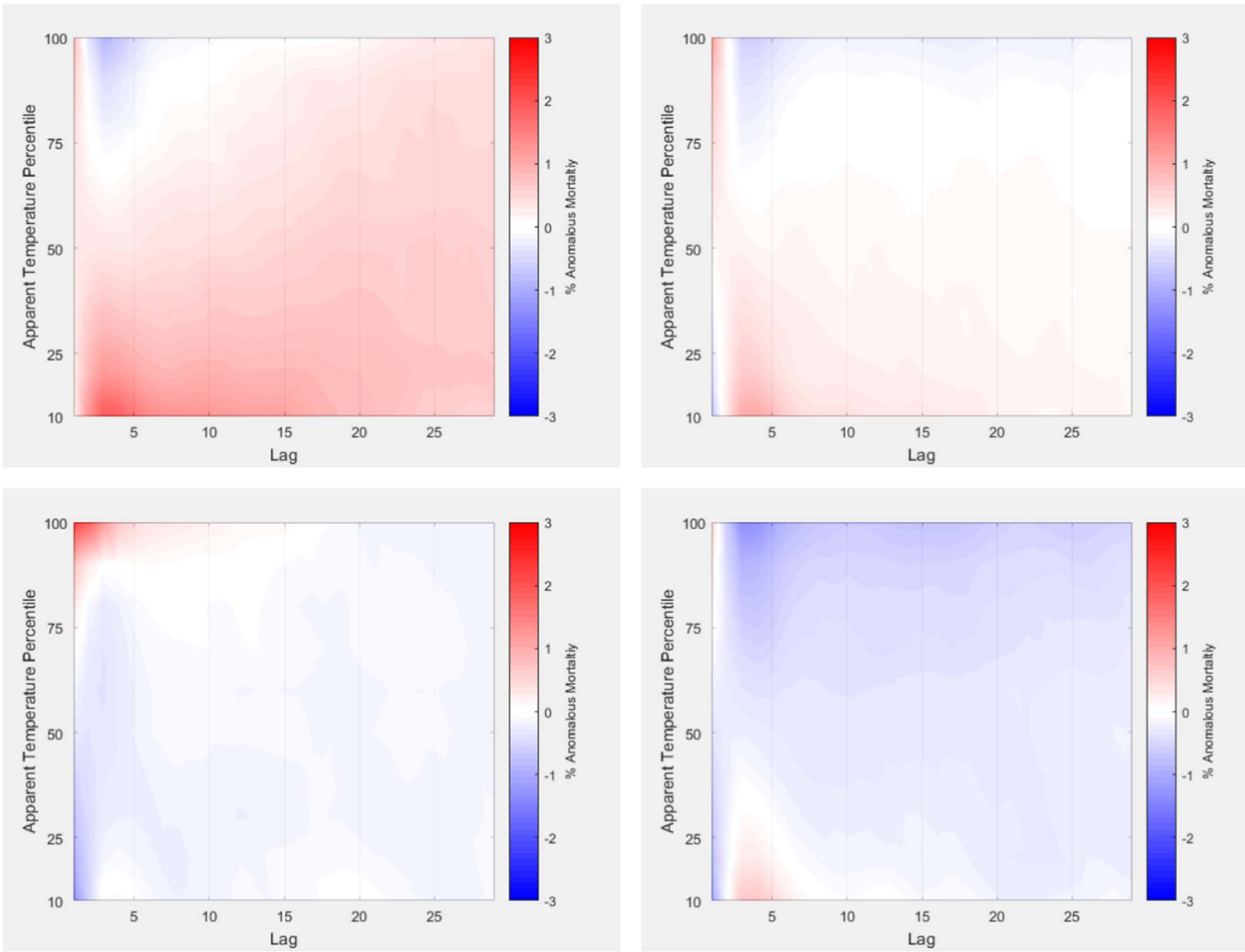

Fig. 3. Nationally-averaged NARX-modeled anomalous mortality by lag day (0–28; x-axis) and AT_{anom} percentile (y-axis) for winter (top left), spring (top right), summer (bottom left) and autumn (bottom right).

earliest and latest parts of the time series.

It is important to note that the training process begins using random values for the input weights, leading to differences in the evolution of the model through successive epochs and eventually to slight differences in each model's output (predicted mortality). Similar to Lee et al. (2016), to account for this, an ensemble-modeling approach was used whereby 10 separate NARX models were trained for ensemble-averaging. When combined with the five different time-block settings described in the paragraph above, these 10 model *permutations* result in the final modeled mortality output for each city consisting of a synthesis of 50 different NARX models. Unless otherwise noted, all performance metrics discussed below are those derived from this reconstructed independent testing ensemble (RITE) output data set.

2.3. Estimating predictor variable importance

One drawback of NARX modeling is the complex interaction between the weights connecting each neuron to each predictor variable; essentially each predictor variable (X) has N x D_x different weights associated with it, which makes determining the 'influence' of each predictor variable on the final model difficult to quantify directly (Olden and Jackson, 2002; Lee et al., 2016). Similar to Lee et al. (2016), in order to estimate the importance of each of the 6 input variables, the final set of 50 NARX models for each city were each re-run (not re-trained) 6 times, each time with a different one of the input variables set to a constant (effectively removing it from the model). The change in the member-model's performance (using median absolute error, MdAE) was calculated and the median MdAE across these 50 models was found. This statistic was then divided by the sum of the six median MdAEs (representing the removal of each of the six variables), yielding a percentage of relative influence of each of the six predictor variables in each city's final NARX model-ensemble. While the national summaries of variable importance are available in Table 2, city-by-city and seasonally-relative statistics were computed as well and are available in supplementary material (Table S.2).

2.4. Comparisons with DLNM

For four cities (Chicago, Los Angeles, Miami, and Houston), distributed-lag nonlinear models (DLNMs) were created using well-established methods (e.g. Sheridan and Dixon, 2017; Gasparrini et al., 2010). The DLNM was run in R using software package *dlnm*, and is set up as:

$$\mathrm{Log}(\mathrm{mortality}) = \mathrm{intercept} + S(\mathrm{AT05}) + \mathrm{S}(\mathrm{AT17}) + \mathrm{S}(\mathrm{AT}_{anom}) + \mathrm{S}(\mathrm{EHF}) + \mathrm{ns}(\mathrm{time}),$$

where mortality is assumed to have a quasi-Poisson distribution, ns (time) is a natural spline fit to the time series with 9 df per year, to account for the season cycle and long-term trends, and S denotes a cubic-spline fit to each of the four meteorological variables, with df =

Table 4
Mean daily NARX-modeled anomalous mortality by location. Statistics computed over the Lag 0–6 day period for heat, and Lag 0–13 day period for cold. Table is sorted from coldest to hottest average annual apparent temperature. Cold days defined as the bottom 10%-ile of AT_{anom} and heat days are defined as top 10%-ile of AT_{anom} in each season. Darker red (green) coloring indicates increasingly positive (negative) mortality.

CITY	Cold Days (Lag 0-13)				Heat Days (Lag 0-6)			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
Minneapolis	1.1%	0.6%	-0.3%	0.5%	-0.3%	-0.5%	0.6%	-1.0%
Milwaukee	1.8%	0.3%	-0.4%	0.3%	-0.8%	-0.6%	1.6%	-0.8%
Buffalo	1.6%	0.4%	-0.1%	0.4%	-0.8%	-0.7%	1.6%	-0.6%
Rochester	1.8%	0.6%	0.0%	0.1%	-1.0%	-0.3%	0.5%	-1.0%
Albany	1.2%	0.3%	0.1%	0.3%	-0.4%	-0.8%	1.1%	-1.0%
Detroit	1.4%	-0.1%	-0.5%	0.2%	-0.4%	-0.2%	2.3%	-0.7%
Boston	1.7%	0.5%	-0.2%	0.7%	-1.0%	-0.4%	2.7%	-1.1%
Denver	0.5%	0.8%	-0.1%	-0.4%	0.6%	0.0%	0.5%	-0.6%
Chicago	2.5%	0.1%	-0.7%	0.7%	-0.8%	-0.5%	2.5%	-0.8%
Cleveland	1.9%	0.5%	0.1%	0.6%	-0.8%	-0.8%	1.2%	-1.1%
Providence	2.6%	0.4%	-0.3%	0.8%	-1.1%	-0.7%	2.9%	-1.6%
Hartford	1.3%	0.5%	-0.6%	0.4%	-0.3%	-0.6%	2.9%	-1.4%
Pittsburgh	2.7%	1.0%	0.0%	0.6%	-1.3%	-1.0%	0.7%	-1.8%
Seattle	2.2%	1.1%	-0.5%	0.1%	-1.1%	-0.2%	1.5%	-0.8%
Indianapolis	1.3%	0.2%	0.0%	-0.2%	0.0%	0.1%	0.4%	-1.3%
Columbus	1.7%	0.7%	-0.1%	0.0%	-1.0%	-0.2%	0.4%	-1.1%
New York	1.7%	-0.1%	-0.7%	1.0%	-0.5%	-0.4%	3.0%	-1.3%
Portland	2.0%	0.9%	0.0%	0.2%	-0.4%	0.1%	2.0%	-0.9%
Kansas City	1.1%	0.2%	-0.7%	0.1%	0.3%	-0.7%	2.0%	-1.1%
Cincinnati	2.4%	0.7%	0.1%	0.6%	-1.2%	-1.1%	1.2%	-1.6%
Philadelphia	2.3%	0.5%	-0.5%	0.6%	-0.8%	-0.4%	1.9%	-1.3%
Baltimore	1.2%	0.4%	-0.5%	0.0%	0.7%	0.2%	1.4%	-0.5%
San Francisco	2.2%	-0.2%	-0.1%	-0.1%	-0.4%	0.6%	0.7%	-0.2%
Saint Louis	1.9%	0.2%	-0.4%	0.2%	-0.7%	-0.8%	2.3%	-0.9%
Washington	1.2%	0.7%	-0.2%	0.3%	0.4%	0.1%	0.9%	-1.2%
Virginia Beach	1.8%	0.6%	-0.1%	-0.2%	-0.7%	-0.1%	0.9%	-1.4%
Nashville	2.4%	0.8%	-0.1%	0.0%	-0.8%	-0.7%	1.4%	-2.3%
Sacramento	0.3%	1.4%	-0.2%	-0.8%	0.6%	0.4%	1.6%	0.2%
Atlanta	4.1%	1.7%	1.3%	1.1%	-0.1%	0.2%	1.7%	-1.5%
Memphis	1.7%	0.3%	-0.7%	0.4%	-0.3%	-0.1%	2.3%	-0.7%
Los Angeles	2.2%	0.2%	-0.1%	0.3%	0.6%	0.3%	1.6%	0.1%
Birmingham	1.2%	0.3%	-0.6%	0.0%	-0.2%	-0.3%	2.6%	-1.3%
Riverside	0.8%	1.3%	0.0%	-1.2%	1.6%	1.5%	2.0%	0.0%
San Diego	1.5%	0.3%	-0.7%	-0.4%	-2.1%	0.2%	0.8%	-0.5%
Dallas	2.2%	0.5%	-0.6%	-0.3%	-0.6%	-0.2%	1.1%	-1.7%
San Antonio	1.1%	0.5%	-0.2%	-0.4%	-0.1%	-0.1%	0.4%	-1.6%
Houston	2.5%	0.8%	0.2%	0.6%	-1.3%	-0.1%	-0.3%	-2.2%
Phoenix	1.9%	0.9%	-1.2%	-1.1%	2.9%	0.8%	1.5%	-1.5%
Orlando	2.4%	0.8%	-0.3%	-0.1%	-0.3%	0.9%	0.4%	-2.0%
Tampa	2.3%	1.3%	-0.9%	-0.3%	-1.5%	-0.4%	0.1%	-1.7%
Miami	2.1%	1.1%	-1.0%	-1.4%	-0.8%	-0.4%	0.4%	-0.7%
AVG	1.8%	0.6%	-0.3%	0.1%	-0.4%	-0.2%	1.4%	-1.1%

5. The lagged influence of these meteorological variables is assessed over a 14-day period, by fitting a spline with 3 equally spaced knots. Similar to Rocklöv et al. (2012), we tested several different permutations of degrees of freedom and model, with very minimal differences in output. Fitted model values were exported for comparison. Since DLNMs are not validated on a separate block of the time series, for a more 'apples-to-apples' comparison with NARX models, the NARX-based model performance results presented in Section 3.4 are based upon the ensemble means of the training-blocks of data rather than the RITE output data set.

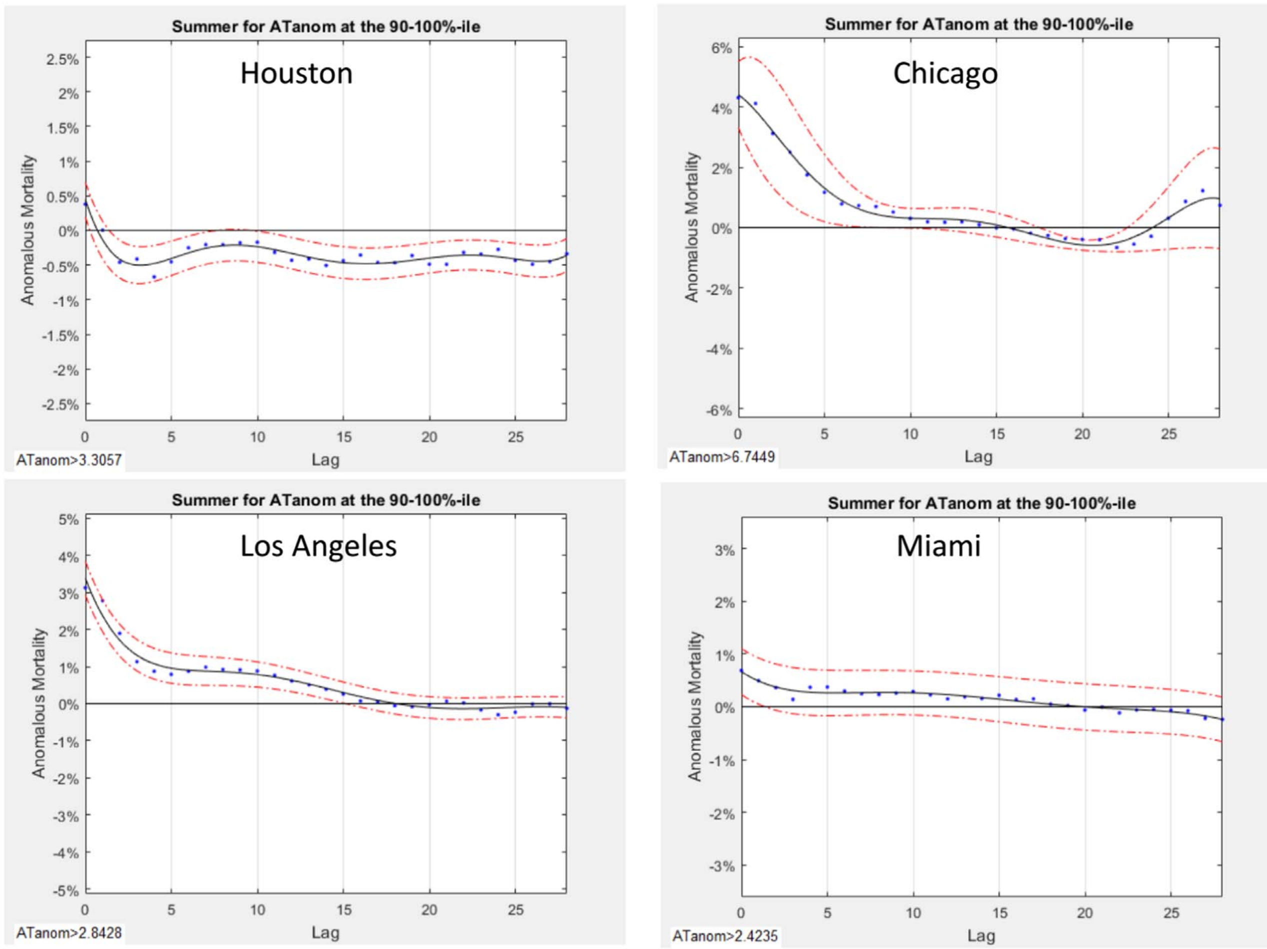

Fig. 4. NARX modeled lagged (0–28 days) mortality response to summer heat days (calculated as days on which AT_{anom} is in the top 10%-ile for the season) for Houston, Chicago, Los Angeles and Miami. Note that the thresholds delineating these events are noted in the bottom left-corner of each image.

3. Results and discussion

The effectiveness of NARX modeling is demonstrated below in four ways: the efficacy of the models using different model performance metrics (Section 3.1), assessing the relative importance of the different predictor variables in the models (Section 3.2), exploring the modeled temperature-mortality relationship in a 'more-traditional' manner using summary statistics and an examination of four different cities as examples (Section 3.3), and directly comparing NARX models with DLNM (Section 3.4).

3.1. NARX model performance

The ability of the NARX models to mimic the actual observed daily mortality in each city is measured using three percentage-based metrics: the median absolute percentage error (MdAPE) between actual mortality and the RITE data set; the percentage-point improvement of the r-squared (of the RITE data set vs. actual mortality) over that of the simple 1-day autocorrelation of the actual observed mortality (R2IM); and the hit rate: the percentage of the actual mortality 'spike day' (> 80th percentile) occurrences that were also 'spike days' (> 80th percentile) in the RITE output. We chose these metrics largely due to fact that these can be interpreted consistently across the cities in this study; since the calculations used in these metrics result in statistics that can be compared across cities with different base populations and rates of demographic change.

On average, year-round MdAPEs across the country were 10.1%, but model performance varies both spatially and temporally (Table 3, Fig. 2). Overall, model performance as measured by MdAPE is best in winter, and worst in summer. Spatially, no clear pattern is observed with relation to different climate regions. Instead, generally the cities with the largest populations performed best in terms of having the lowest MdAPEs, as New York (4%), Los Angeles (5%) and Chicago (6%) were the top three cities in year-round performance, while Albany (15%), Hartford (14%) and Rochester (14%) ranked as the three worst. This relationship between population and performance is unsurprising, as a greater average daily mortality sample size allows a temperature-based signal to emerge from the 'background' noise of all-cause mortality.

In terms of R2IM, NARX models in every city show skill, as they are able to improve upon a persistence-based model. Across the calendar year, once again the cold season is better modeled than the warm season, although the highest model improvement is observed in March. Across all locations, the average improvement in the r-squared is 17%, however, in comparison to MdAPEs, model performance with R2IM is affected by more than population size. Nashville (24%), Portland (23%) and San Antonio (23%) are the top 3, while large cities such as Los Angeles (17%) and Chicago (20%), are near the national average. Further, while many small cities (e.g. Albany, 6%; Rochester, 9%, Buffalo, 11%) perform poorly using this metric, highly-populated New York (14%) and Phoenix (12%) do as well.

All cities showed some accuracy in modeling the top 20% of mortality days ('spikes'), as no hit rate was below what would be expected by random chance (20%). Nationally, mortality spikes were correctly modeled 54% of the time, with NARX models in warmer-weather and populous cities being the top performers for hit rates; Phoenix (71%),

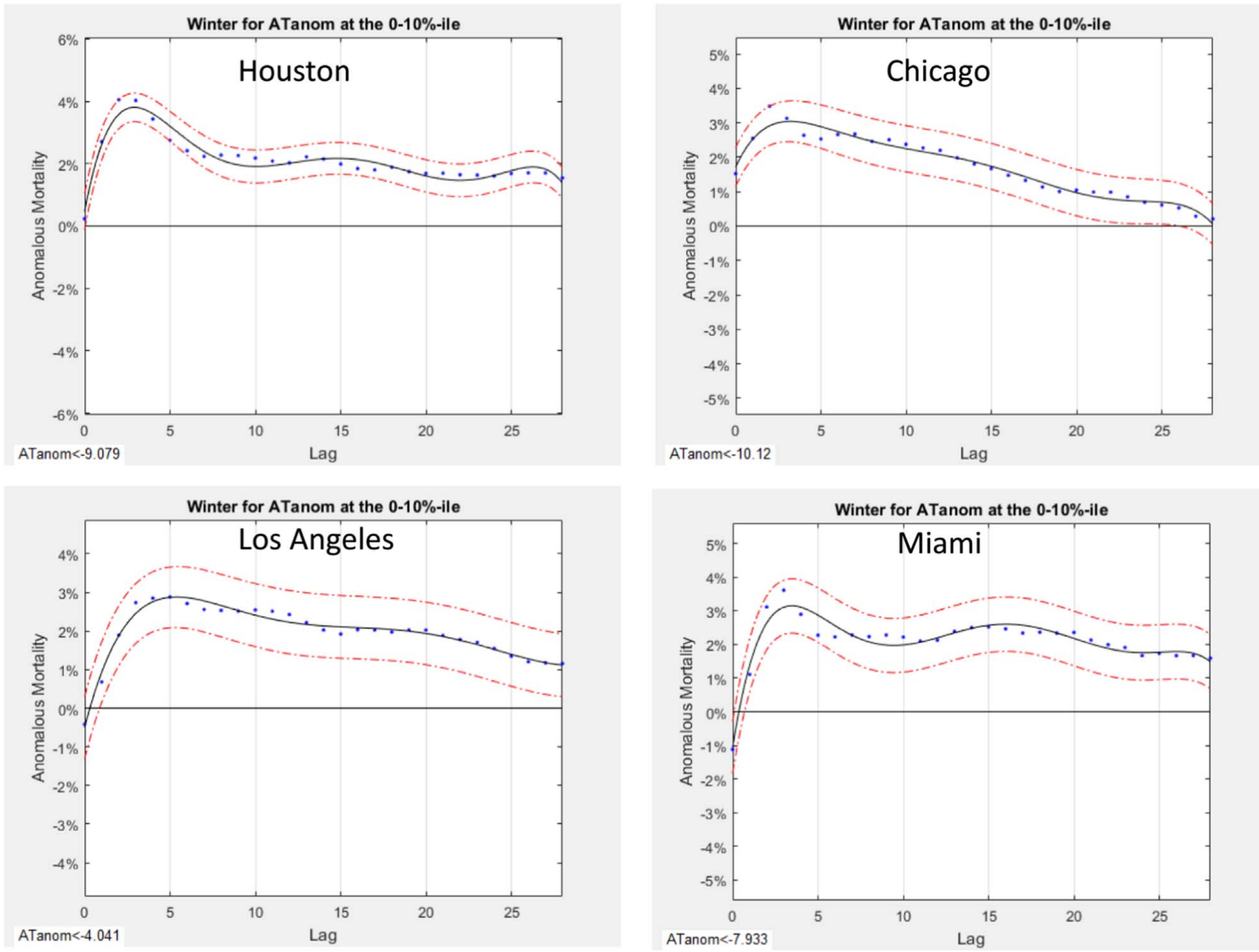

Fig. 5. NARX modeled lagged (0–28 days) mortality response to winter cold days (calculated as days on which AT_{anom} is in the bottom 10%-ile for the season) for Houston, Chicago, Los Angeles and Miami. Note that the thresholds delineating these events are noted in the bottom left-corner of each image.

New York (71%), Dallas (70%) and Atlanta (70%) performed best, and cold-weather and less populated cities like Albany (34%), Rochester (36%), and Buffalo (39%), performed worst. Temporally, spikes were best predicted between December and April, with worse performance in summer and autumn.

Interannual variability was also examined, with all 3 metrics generally improving throughout the 1975–2010 time period, especially MdAPE and R2IM. Again, as population increases over time, so too does the background mortality, likely leading to the improvement in NARX model performance.

3.2. Predictor variable importance

Among the six predictor variables used to train the models, the three temperature-based variables (AT05, AT17, and AT_{anom}) played the most prominent roles, as they collectively accounted for about 72% of the influence across all 41 cities (Table 2). The linear trend variable is most important in cities with the greatest population growth over the period of study, including Phoenix, Orlando, Riverside, Atlanta, Sacramento, and San Antonio (Table S.2, Supplementary material). The seasonal cycle variable is of only marginal importance, largely since AT05 and AT17 inherently mimic the seasonal cycle, and therefore, when it is removed from the model, there is little impact on the final output. The EHF played only a minor role in the outcomes of the NARX models, though this is to be expected because it is only a relevant variable in a small percentage of warm days.

Many cities have either AT05 or AT17 as much more important than the other in terms of their influence on the model output, suggesting that some cities' temperature-mortality relationship is more sensitive to either minimum (AT05) or maximum (AT17) temperature than the other, but both play roles in influencing year-round mortality (Barnett et al., 2010; Davis et al., 2016a). When broken down by season, the AT_{anom} variable is much more important in winter, while it is less important during the summer, when minimum temperature (AT05) is more of a determining factor in the NARX model output. Barnett et al. (2010) explored various temperature metrics in terms of their ability to predict mortality, and found no clear 'winner' either geographically or seasonally. Thus, this finding of minimum temperatures being key in summer and AT_{anom} being key in winter is somewhat surprising, though some previous research has noted the wintertime association with *relative* temperatures as opposed to *absolute* temperatures (e.g. Lee, 2015; Ryti et al., 2016; Anderson and Bell, 2009; Gerber et al., 2006). Despite the high degree of correlation that would otherwise suggest that the relative importance of AT05 and AT17 should both be quite small (by virtue of the way the metric is calculated), this spatial and seasonal variability in the importance of both predictors results in the nation-wide year-round average importance of both being high and largely equal. This also highlights a caveat of using the methods outlined above (in Section 2.3) to estimate predictor variable importance – because the performance of the models is impacted by both the variable that is removed (the real variable of interest) and the variables that remain, gleaning information about any correlated predictors might be difficult.

3.3. NARX-modeled temperature-mortality relationship

The modeled temperature-mortality relationship is discussed below

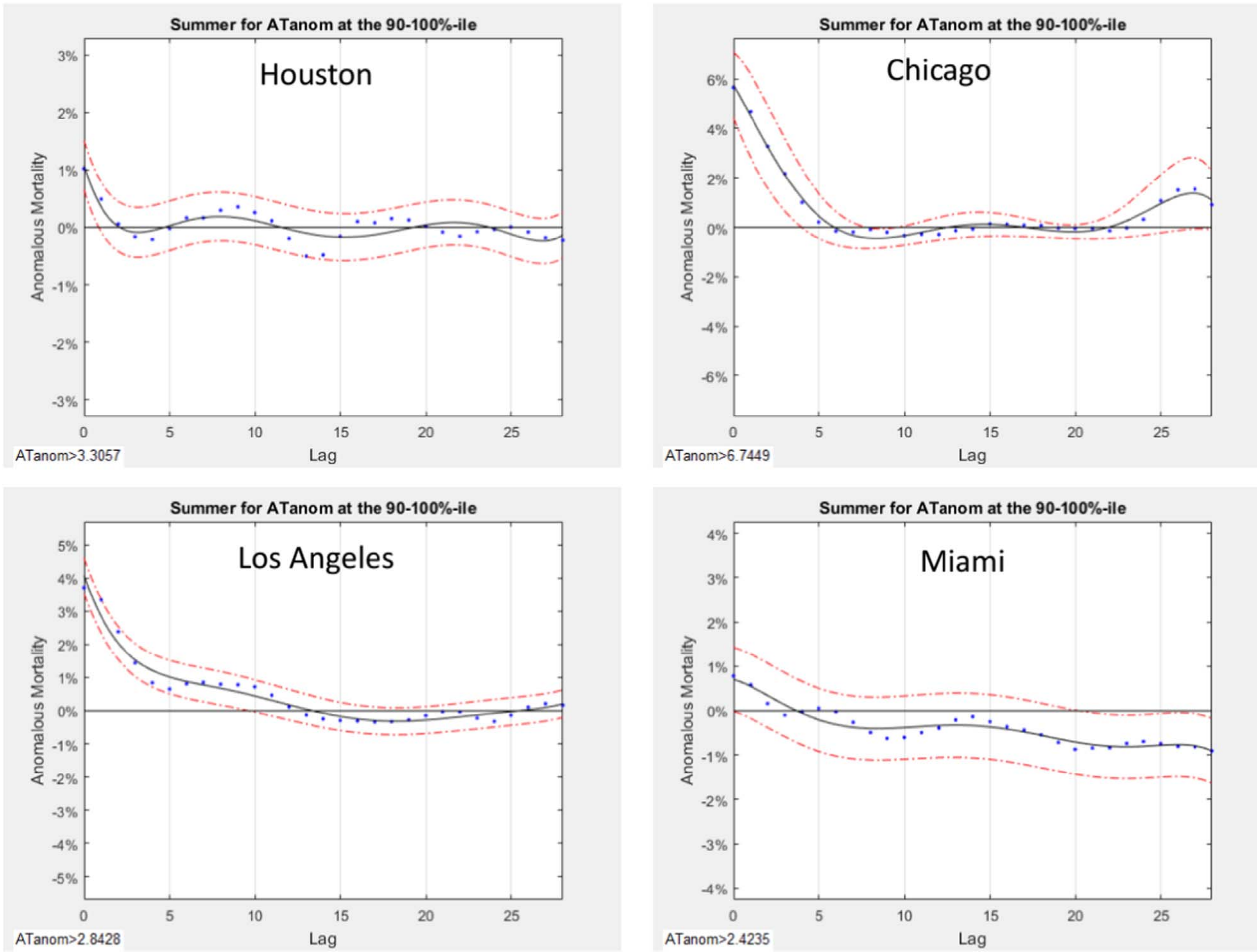

Fig. 6. Same as Fig. 4, except for the DLNM results.

in terms of the relationship between AT_{anom} and percent anomalous mortality relative to season and time in the entire time series. Percent anomalous mortality was calculated by removing seasonal and secular trends from the NARX-modeled mortality data in each city. We used the same deseasonalization process described above (for AT_{anom}) while also removing a simple linear trend from each city's mortality data likely influenced mostly by population changes. Note that this anomalous mortality metric was not used in the NARX models themselves, just for displaying the results in this section.

In winter, when averaged across all 41 locations, the greatest increase in anomalous mortality is associated with the lowest AT_{anom} at 2–4 days of lag (Fig. 3), while the greatest decrease in mortality is associated with the most anomalously warm conditions at the same lags. In summer, mortality increases occur during the warmest AT_{anom} days and quickly declines thereafter (Lag 0–2), with anomalous mortality eventually becoming negative after Lag 15. This phenomenon, sometimes referred to as mortality displacement or a "harvesting" effect (Hajat et al., 2005), is a common finding in heat-related mortality research, but as is also shown here, is not often associated with cold-related mortality. Anomalously cool summer temperatures show large decreases in mortality, especially at Lag 0. Spring and autumn show traits of both summer and winter in terms of the relationship between AT_{anom} and mortality, as both warm and cold temperatures show increases in mortality at the same lags noted above for summer and winter, respectively, although not to the same magnitude. The largest differences between these transitional seasons compared to summer and winter is with heat-related mortality: the heat-effect at Lag 0 extends nearly to the 75th percentile of AT_{anom}, and the Lag 0 rise in mortality is quickly followed by a precipitous decrease in mortality at Lag 2 and onward (whereas in summer, mortality displacement is delayed by over 2 weeks). There is limited research that explores temperature-related mortality specifically in spring or autumn, however, the results herein may be comparable with previous studies that have noted early heat waves (i.e. late-spring to early-summer) and cold-waves (i.e. late-autumn to early-winter) can be the most deadly (Allen and Sheridan, 2015; Lee et al., 2014)

Geographic variability in the temperature-mortality relationship is more subtle than seasonal variability. Table 4 shows the average anomalous mortality for heat days (averaged over Lag 0–6 days) and cold days (averaged over Lag 0–13) by location in each season. The most noticeable geographic pattern is for summer heat days: with only a few exceptions (e.g. Phoenix, Los Angeles, Denver, Minneapolis), the warmest cities have some of the smallest increases in mortality; while the coldest cities have some of the largest increases in mortality. The four locations with the highest summer heat-related mortality are all on the East Coast (New York, Hartford, Providence, and Boston). Spatial variability in the relationship is also apparent in a more detailed examination of four US cities in different climate zones (Los Angeles, Miami, Chicago and Houston; Fig. 4). Chicago shows the typical pattern of large increases in heat-related mortality at Lag0–5, however, hot days are actually related to a significant *decrease* in mortality in Houston for nearly the entire 4 week period. Miami shows almost no significant relationship between heat and mortality, while Los Angeles exhibits significantly increased mortality for nearly 2 weeks after a heat event. Interestingly, two of the warmest cities, Phoenix and Riverside, have the greatest mortality increase in response to anomalously warm

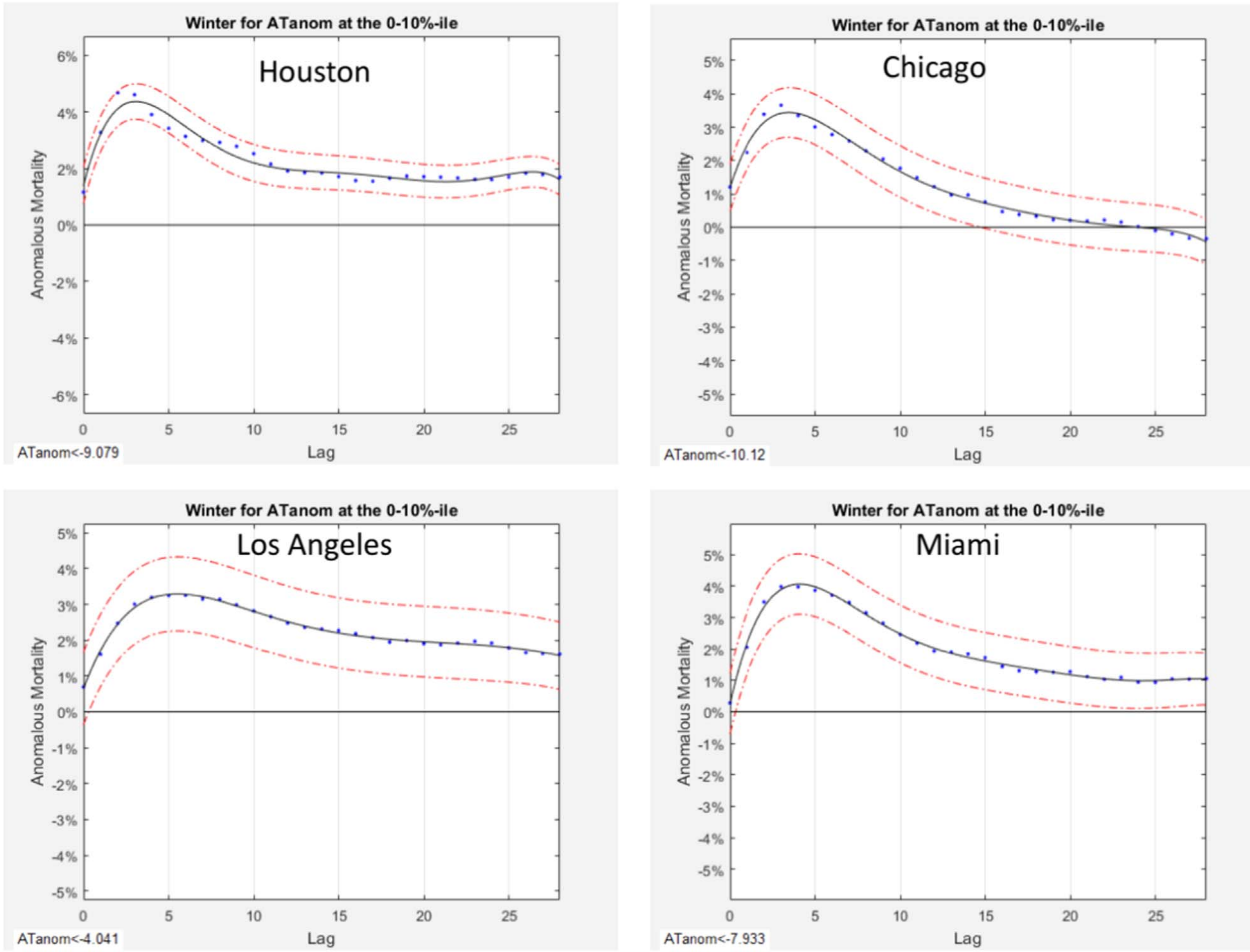

Fig. 7. Same as Fig. 5, except for the DLNM results.

Table 5
Seasonal and annual comparison of NARX and DLNM performance statistics for Chicago (ORD), Houston (IAH), Los Angeles (LAX) and Miami (MIA).

R2IM	DLNM				NARX			
	ORD	IAH	LAX	MIA	ORD	IAH	LAX	MIA
Winter	25%	22%	21%	22%	20%	20%	20%	19%
Spring	25%	21%	25%	23%	22%	21%	22%	21%
Summer	29%	21%	19%	23%	26%	21%	16%	21%
Autumn	21%	21%	23%	25%	17%	20%	20%	23%
Annual	28%	21%	23%	21%	25%	20%	22%	19%
MdAPE	DLNM				NARX			
	ORD	IAH	LAX	MIA	ORD	IAH	LAX	MIA
Winter	5%	8%	5%	6%	5%	8%	5%	6%
Spring	5%	8%	5%	6%	5%	8%	5%	6%
Summer	5%	8%	5%	6%	5%	8%	5%	7%
Autumn	5%	8%	5%	7%	5%	8%	5%	7%
Annual	5%	8%	5%	6%	5%	8%	5%	7%
HitRate (80%)	DLNM				NARX			
	ORD	IAH	LAX	MIA	ORD	IAH	LAX	MIA
Winter	58%	64%	65%	58%	56%	65%	62%	54%
Spring	53%	63%	52%	56%	54%	66%	51%	56%
Summer	41%	61%	39%	52%	39%	63%	38%	47%
Autumn	45%	66%	46%	46%	42%	67%	45%	46%
Annual	59%	65%	66%	64%	57%	67%	66%	63%

weather in *winter*.

Mortality responses to wintertime cold days are more consistent across the country, as 38 of the 41 cities have a mortality increase of at least 1% in the two weeks after cold days, led by Atlanta (+4.1%). This consistency is also exhibited in Fig. 5, with the peak increases occurring around Lag 3 and tapering off slowly enough to remain significantly positive for nearly four weeks in each of the four cities examined. Cities in the Southeastern US appear to have the highest wintertime mortality response to the coldest temperatures, as Atlanta, Houston, Nashville, Tampa, Dallas, Orlando, and Miami all have at least a 2.1% increase in mortality within the Lag 0–13 window (Table 4). This geographic tendency for warmer locations to experience greater cold-related mortality has been noted in previous research (Allen and Sheridan, 2015; Lee, 2015).

3.4. Comparison with DLNM

NARX training-block output and DLNM models exhibit nearly indistinguishable temperature-mortality relationships in the four cities examined (Los Angeles, Miami, Chicago and Houston; Figs. 6 and 7). Their output time series are also nearly identical, with the correlations between the two model outputs ranging from r = 0.95 (Chicago) to r = 0.99 (Houston and Miami). DLNMs perform slightly better in many modeling metrics (Table 5) for these 4 cities, though these differences are not universal and are generally small, especially with MdAPE (all less than 0.2%). DLNMs perform best relative to NARX models when examining R2IM, improving upon persistence by anywhere from 0.9% to 3.1% points annually among these cities. The R2IM difference between the two techniques is most notable in the winter season, and most similar in spring and summer. NARX models do perform well in examining the ability of the models to identify spikes in mortality – especially in Houston, where they outperform DLNM.

It should be noted that comparisons between NARX and DLNM should be viewed in light of the goals of each respective technique. Whereas NARX aims to replicate the time series, DLNM aims primarily to assess the dose-response relationship. Thus, NARX models need to be validated to avoid overfitting and to have confidence in the model's ability. However, DLNM models need not be directly validated on a testing dataset to meet their aim of assessing the dose-response relationship, and therefore, may or may not be able to 'apply' the model to new predictor data presented to it. In other words, NARX models are inherently constrained by the internal validation requirement; if enough neurons were incorporated into a NARX model without consideration of how it replicates the internal validation dataset, or if the NARX model was trained on 100% of the time series (rather than 60%), then NARX models would have a substantially higher performance. Thus, this comparison of the two methods is not straightforward.

4. Conclusions

The aim of this research is to demonstrate the usefulness of a new modeling methodology in biometeorological research that has gained popularity recently in other environmental sciences: NARX models. NARX models are ANN-based time-series models that need no a priori assumption of the predictor(s)-predictand relationship, and are thus well-suited to temperature-related mortality research where the relationship is often lagged, nonlinear, and differs depending upon the time of the year. Due to the flexibility of NARX models, one model can be trained for the entire time series (i.e. all seasons) for a city and can be easily modified to make predictions, prospectively making them useful for weather-based health warning systems. NARX models are not without limitations, however. Due to the fact that the ideal model architecture is not known beforehand, to find the best model, the user must assess many different options, and thus, the process can be computationally expensive. Further, the complexity of the neuron connections in NARX models makes a direct interpretation of variable importance within the model tenuous; however, an indirect attempt was made herein to quantify this variable importance.

Across the 41 cities examined, NARX model performance varied geographically and seasonally, and appeared best in larger cities. While the best monthly hit rates and improvements in r-squared occurred in March, when aggregated by season, models performed best in all metrics during the winter, and worst in summer. NARX models showed skill in modeling mortality spikes, and improved upon using simple persistence of the previous day's mortality, on average by 17% points (with Nashville improving explained variance by 24% points). While using multiple temperature-based variables to model these relationships, it was also noted that on average, anomalous apparent temperature was the most important temperature variable associated with mortality, except in summer, when minimum apparent temperature became the key variable for modeling mortality.

A traditional analysis of the time-lagged temperature-mortality relationships in various locations also helped highlight the utility of the method in biometeorological research. Output from the models displayed temperature-mortality relationships largely in agreement with previous literature; primarily that the warmest summer days (at Lag 0–2) and the coldest winter days (peaking at Lag 2–4) correspond to increases in all-cause mortality. The spring and autumn seasons showed traits of both summer and winter temperature-mortality relationships, but were generally weaker in magnitude. When compared among a subset of 4 cities, NARX and DLNM each output highly correlated mortality time series with nearly identical structures in their temperature-mortality relationships. Quantitatively, DLNM performed slightly better by many performance metrics, though the differences between NARX and DLNM modeling goals makes these comparisons indirect at best.

While different in concept from other time-series modeling techniques used in environmental health research, we believe that NARX models offer advantages that outweigh their limitations and provide additional diversity in modeling techniques that may be most efficacious depending on a project's goals. NARX models would be a suitable choice for modeling various other complex, time-dependent relationships in numerous other environmental applications, especially if the goal of the research is to have a cross-validated replication (or prediction) of a time series more-so than a better understanding of a dose-response relationship.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors declare no conflicts of interest.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.envres.2018.02.020.

References

Allen, M.J., Sheridan, S.C., 2014. High-mortality days during the winter season: comparing meteorological conditions across 5 US cities. Int. J. Biometeorol. 58 (2), 217–225. http://dx.doi.org/10.1007/s00484-013-0640-4.

Allen, M.J., Sheridan, S.C., 2015. Mortality risks during extreme temperature events (ETEs) using a distributed lag non-linear model. Int. J. Biometeorol. 58, 217–225. http://dx.doi.org/10.1007/s00484-015-1117-4.

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.

Anderson, B.G., Bell, M.L., 2010. Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 U.S. communities. Environ. Health Perspect. 119, 210–218.

Anderson, G.B., Bell, M.L., Peng, R.D., 2013. Methods to calculate the heat index as an exposure metric in environmental health research. Environ. Health Perspect. 121 (10), 1111–1119.

Barnett, A.G., Tong, S., Clements, A.C.A., 2010. What measure of temperature is the best predictor of mortality? Environ. Res. 110, 604–611. http://dx.doi.org/10.1016/j.envres.2010.05.006.

Basu, R., 2009. High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. Environ. Health 8 (40), 1–13. http://dx.doi.org/10.1186/1476-069X-8-40.

Basu, R., Samet, J.M., 2002. Relation between elevated ambient temperature and mortaltiy: a review of the epidemiological evidence. Epidemiol. Rev. 24 (2), 190–202. http://dx.doi.org/10.1093/epirev/mxf007.

Beale, M.H., Hagan, M.T., Demuth, H.B., 2014. Neural network toolboxTM user's guide, R2014a. white paper. Mathworks Inc.

Davis, R.E., Hondula, D.M., Patel, A.P., 2016a. Temperature observation time and type influence estimates of heat-related mortality in seven US cities. Environ. Health Perspect. 124 (6), 795–804. http://dx.doi.org/10.1289/ehp.1509946.

Davis, R.E., Dougherty, E., McArthur, C., Huang, Q.S., Baker, M.G., 2016b. Cold, dry air is associated with influenza and pneumonia mortality in Auckland, New Zealand (DOI). Influenza Other Respir. Virus. 10 (4), 210–313. http://dx.doi.org/10.1111/irv.12369.

Diaconescu, E., 2008. The use of NARX neural networks to predict chaotic time series. WSEAS Trans. Comput. Res. 3 (3), 182–191.

Donaldson, G.C., Keatinge, W.R., Näyhä, S., 2003. Changes in summer temperature and heat-related mortality since 1971 in North Carolina, South Finland, and Southeast England. Environ. Res. 91, 1–7.

Gasparrini, A., Armstrong, B., Kenward, M.G., 2010. Distributed lag non-linear models. Stat. Med. 29, 2224–2234. http://dx.doi.org/10.1002/sim.3940.

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklov, J., Forsberg, B., Leone, M., DeSario, M., Bell, M.L., Guo, Y.L.L., Wu, C., Kan, H., Yi, S.M., Coelho, M.S., 2015. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. Lancet 386 (9991), 369–375. http://dx.doi.org/10.1016/S0140-6736(14)62114-0.

Gasparrini, A., Scheipl, F., Armstrong, B., Kenward, M.G., 2017. A peanalized framework for distributed lag non-linear modeling. Biometrics 1–11. http://dx.doi.org/10.1111/biom.12645.

Gerber, Y., Jacobsen, S.J., Killian, J.M., Weston, S.A., Roger, V.L., 2006. Seasonality and daily weather conditions in relation to myocardial infarction and sudden cardiac death in Olmstead County, Minnesota, 1979–2002. J. Am. Coll. Cardiol. 48 (2), 287–292. http://dx.doi.org/10.1016/j.jacc.2006.02.065.

Gosling, S.N., Lowe, J.A., McGregor, G.R., Pelling, M., Malamud, B.D., 2009. Associations between elevated atmospheric temperature and human mortality: a critical review of the literature. Clim. Change 92, 299–341. http://dx.doi.org/10.1007/s10584-008-9441-x.

Grass, D., Kane, M., 2008. The effects of weather on air pollution on cardiovascular and respiratory mortality in Santiago, Chile, during the winters of 1988–1996. Int. J.

Climatol. 28, 1113–1126. http://dx.doi.org/10.1002/joc.1592.
Guzman, S.M., Paz, J.O., Tagert, M.L.M., 2017. The use of NARX neural networks to forecast daily groundwater levels. Water Resour. Manag. 31, 1591–1603. http://dx.doi.org/10.1007/s11269-017-1598-5.
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 (5), 613–620.
Kim, J., Yoon, K., Choi, J.C., Kim, H., Song, J.-K., 2016. The association between wind-related variables and stroke symptom onset: a case-crossover study on Jeju Island. Environ. Res. 150, 97–105. http://dx.doi.org/10.1016/j.envres.2016.05.041.
Kysely, J., Pokorna, L., Kyncl, J., Kriz, B., 2009. Excess cardiovascular mortality associated with cold spells in the Czech Republic. BMC Public Health 9 (1), 1–11. http://dx.doi.org/10.1186/1471-2458-9-19.
Lee, C.C., 2015. A systematic evaluation of the lagged effects of spatiotemporally relative surface weather types on wintertime cardiovascular-related mortality in 19 US cities. Int. J. Biometeorol. 59 (11), 1633–1645. http://dx.doi.org/10.1007/s00484-015-0970-5.
Lee, C.C., Sheridan, S.C., Barnes, B.B., Hu, C., Pirhalla, D.E., Ransibrahmanakul, V., Shein, K., 2016. The development of a non-linear auto-regressive model with exogenous input to model climate-water clarity relationships: reconstructing a historical water clarity index for the coastal waters of the southeastern USA. Theor. Appl. Climatol. http://dx.doi.org/10.1007/s00704-016-1906-7.
Lee, M., Nordio, F., Zanobetti, A., Kinney, P., Vautard, R., Schwartz, J., 2014. Acclimatization across space and time in the effects of temperature on mortality: a time-series analysis. Environ. Health 13 (89), 1–9. http://dx.doi.org/10.1186/1476-069X-13-89.
Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Model. Softw. 15, 101–124.
McGregor, G.R., 1999. Winter ischemic heart disease deaths In Birmingham, United Kingdom: a synoptic climatological analysis. Clim. Res. 13, 17–31.
McMichael, A.J., Wilkinson, P., Kovats, R.S., Pattenden, S., Hajat, S., Armstrong, B., Vajanapoom, N., Niciu, E.M., Mahomed, H., Kingkeow, C., Kosnik, M., O'Neill, M.S., Romieu, I., Ramirez-Aguilar, M., Barreto, M.L., Gouveia, N., Nikiforov, B., 2008. International study of temperature heat and urban mortality: the 'ISOTHURM' project. Int. J. Environ. Epidemiol. 37, 1121–1131.
Morabito, M., Crisci, A., Grifoni, D., Orlandini, S., Cecchi, L., Bacci, L., Modesti, P.A., Gensini, G.F., Maracchi, G., 2006. Winter air-mass-based synoptic climatological approach and hospital admissions for myocardial infarction in Florence, Italy. Environ. Res. 102, 52–60.
Nairn, J.R., Fawcett, R.J., 2014. The excess heat factor: a metric for heatwave intensity and its use in classifying heatwave severity. Int. J. Environ. Res. Public Health 12 (1), 227–253.
Näyhä, S., 2002. Cold and the risk of cardiovascular diseases. A review. Int. J. Circumpolar Health 61 (4), 373–380.
Ng, C.F.S., Ueda, K., Ono, M., Nitta, H., Takami, A., 2014. Characterizing the effect of summer temperature on heatstroke-related emergency ambulance dispatches in the Kanto area of Japan. Int. J. Biometeorol. 58, 941–948.
Olden, JD, Jackson, DA, 2002. Illuminating the Bblack boxˆ: a randomization approach for understanding variable contributions in artificialneural networks. Ecol. Model. 154, 135-15.
Rocklöv, J., Barnett, A.G., Woodward, A., 2012. On the estimation of heat-intensity and heat-duration effects in time series models of temperature-related mortality in Stockholm, Sweden. Environ. Health 11, 23.
Ryti, N.R.I., Guo, Y., Jaakkola, J.J.K., 2016. Global association of cold spells and adverse health effects: a systematic review and meta-analysis. Environ. Health Perspect. 124, 12–22. http://dx.doi.org/10.1289/ehp.1408104.
Sheridan, S.C., Allen, M.A., 2015. Changes in the frequency and intensity of extreme temperature events and human health concerns. Curr. Clim. Change Rep. 1 (3), 155–162. http://dx.doi.org/10.1007/s40641-015-0017-3.
Sheridan, S.C., Dixon, P.G., 2017. Spatiotemporal trends in human vulnerability and adaptation to heat across the United States. Anthropocene 20, 61–73.
Sheridan, S.C., Allen, M.J., Lee, C.C., Kalkstein, L.S., 2012a. Future heat vulnerability in California Part II: projecting future heat related mortality. Clim. Change 115 (2), 291–309.
Sheridan, S.C., Lee, C.C., Allen, M.J., Kalkstein, L.S., 2012b. Future heat vulnerability in California Part I: projecting future weather types and heat events. Clim. Change 115 (2), 311–326.
Steadman, R.G., 1984. A universal scale of apparent temperature. J. Clim. Appl. Meteorol. 23 (12), 1674–1687.