INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **28**: 545–562 (2008)
Published online 25 June 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1546

Models for mortality associated with heatwaves: update of the Portuguese heat health warning system

Paulo Nogueira* and Eleonora Paixão
Instituto Nacional de Saúde Dr. Ricardo Jorge, Observatório Nacional de Saúde, Portugal

ABSTRACT: In Portugal a Heat Health Watch Warning exists since 1999 – the íCARO Surveillance System. It is a system for monitoring heatwaves with potential impacts on population's morbidity and mortality, which is set in motion every year between May and September. This system was based on a model for the relation between heat and mortality calibrated with the district of Lisbon data for the big heatwaves of June 1981 and July 1991.

The occurrence of 2003's big heatwave brought opportunity for updating existing models.

The fact that this heatwave has been particularly long and had characteristics that were not described in the previously known heatwave episodes also allowed the chance to investigate the mechanism of the relation between occurrence of extreme heat and mortality.

The aim of this work was to update the existing íCARO model and contribute to increase the knowledge on the phenomena of the heatwave's impact on mortality.

Thus, four models were assayed, that represent four distinct proposals for reference temperature's thresholds and in the generalization of the main used variable accumulated thermal overcharge (ATO).

All the assayed models showed a good adaptation to the observed mortality data for the district of Lisbon for the three known big heatwaves.

It is concluded that the rational generalized accumulated thermal overload (GATO) adapts well to the relation between heat and mortality. The best model was chosen as the one that considered a dynamic threshold that follows the ascending phase of the temperatures of summer until reaching its maximum level in the end of the month of August remaining thereafter constant until the end of the summer, thus recurring to a rational of population's adaptation to heat, contrarily to what happens to air temperatures that decrease at end of the summer.

KEY WORDS heatwave; statistical models; surveillance system; heat health; watch warning system

Received 14 November 2005; Revised 16 March 2007; Accepted 18 March 2007

1. Introduction

The occurrence of heatwaves is currently recognized as a public health problem (Clarke, 1972; Kunst *et al.*, 1993; MMWR, 2002; Smoyer-Tomic *et al.*, 2003; Conti *et al.*, 2005), given that it is a phenomenon causally associated to excess of potentially avoidable mortality. Still very present in our minds is the 2003 summer European heatwave, which impact is estimated in about 50 000 excess deaths beyond expected (Brucker, 2005).

Heatwaves' impacts on human mortality are widely and well established and some impacts in the morbidity and in health care services are also recognized when these occur.

There is evidence that the occurrence of these episodes jointly with the conditions in which they occur may have broader impacts in the different sectors of human activity beyond health, in agriculture and animal production, in construction, transports, and the robustness of infrastructures (Smoyer-Tomic *et al.*, 2003).

In scientific literature the episodes of heatwaves are described as occurring all over the globe. Portugal is no exception and there are some indications that in the city of Lisbon episodes of heat excess have sporadically occurred, with probable impacts on the health of the respective population, at least throughout all the twentieth century. A heatwave impact on Lisbon's district population mortality (for June 1981 heatwave) was published for the first time in 1988 (Falcão *et al.*, 1988).

In fact, multiple episodes of heatwaves of different time ranges have been identified in Portugal, for the period of 1980 to 2004, namely:

1. Heat episodes with great impact on mortality – Big heatwaves – in June 1981, July 1991 and July/August 2003. The respective excess mortality at the national level (Portugal mainland) was estimated in about 1900 (Garcia *et al.*, 1999), 1000 (Paixão and Nogueira, 2002) and 2000 (Calado *et al.*, 2004a,b) deaths, respectively;
2. Episodes with lower impacts on mortality – small and moderate heatwaves – of which are examples at

* Correspondence to: Paulo Nogueira, Instituto Nacional de Saúde Dr. Ricardo Jorge, Observatório Nacional de Saúde, Portugal.
E-mail: paulo.nogueira@insa.min-saude.pt

national level (Portugal mainland), 14/07 to 25/07/1990 and 19/05 to 28/05/1991 with an excess mortality estimate of 690 and 475 deaths respectively (Paixão and Nogueira, 2002); 27/05 to 2/06/2001 with estimate of excess of 441 deaths (Paixão and Nogueira, 2003); and at the local level, in the region of the Algarve, 27/07 to 4/08/2004 with an estimate of excess of 80 deaths (Falcão *et al.*, 2004).

In terms of health, the advanced age; cognitive limitations; suffering from illnesses; consumption of some medicines; individuals' hydration level, isolation and housing conditions are pointed as vulnerability conditions (Kilbourne, 1999; Semenza *et al.*, 1999; WHO 2004).

Younger individuals seem to suffer less from heatwave's impacts, but there are cases of severe heat episodes in which all age groups were statistically shown as affected in terms of impact on the respective mortality, as in the case of the heatwave of June 1981 in Portugal (Garcia *et al.*, 1999).

Studies pointed out that the risk to suffer the effect of heatwaves is similar in men and women (WHO 2004). But the heatwave of 2003, showed to have a bigger impact on female mortality, even after controlling for age (Calado *et al.*, 2004a,b; Garcia-Herrera *et al.*, 2005; Nogueira, 2005) as pointed out by Kosatsky (2005).

The impact of summer extreme temperatures in mortality occurs almost immediately, the maximum peak of mortality often occurs 1 day after the maximum peak of temperature, usually exists a delay between the heat occurrence and the respective mortality which is of 1 or 2 days (Diaz *et al.*, 2002; Garcia-Herrera *et al.*, 2005).

The investigation of the heatwave occurrence phenomena presents some problems from its most basic definition of what is a heatwave with effects on human health. It is clear from the literature that a consensual definition does not yet exist (Trigo *et al.*, 2005).

The more commonly used definitions are based on temperature thresholds. Some references to impacts in mortality with air temperature above 32 °C (90 °F) during some consecutive days exist in literature for some time now (Schuman *et al.*, 1964; Henschel *et al.*, 1969), but other thresholds values are currently adopted in different regions. Definitions from official organizations also exist in Canada – temperatures above of 32 °C for more than three consecutive days (Smoyer-Tomic *et al.*, 2003); and in the United States of America – temperatures above of 32 °C for more than two consecutive days (Centers for Disease Control and Prevention (CDC), 1995, 1996, 1997).

The relation between temperature and human mortality is in general described as having a U, V, or J Shape (Saez *et al.*, 1995; Nunes *et al.*, 1999; Dessai, 2002). This description fits several characterizations of the heat-mortality relation as being non-linear (Smoyer *et al.*, 2000). This notion of non-linearity results of evidence gathered in models established for some American and European cities that indicate air temperatures thresholds beyond which mortality sharply increases. These thresholds vary from city to city. In Europe north–south and west–east gradients seem to exist for the definition of these thresholds temperatures (WHO, 2004).

In Portugal, for Lisbon district mortality, in the big heatwaves 1981 and 1991, the 32 °C threshold revealed itself as a good indicator for excess mortality occurrence. In fact Lisbon's district mortality always increased after two or more consecutive days above 32 °C occurrence.

Posterior evaluations showed that this temperature threshold corresponds approximately to percentile 97.5 of Lisbon's observed maximum temperatures distribution for the months of May to September 1980–2000. When the same percentile for Portugal's mainland of 18 districts was calculated, these varied among themselves. However, 97.5 percentile of the daily average maximum temperature of the 18 districts, in the same period was approximately 32 °C (see 3.2).

From the observed summer mortality for some districts of Portugal it was possible to verify that the number of consecutive days of maximum temperatures of air is above 32 °C (HLen32) and the observed excess of temperature above 32 °C (Exc32) relate directly with the observed mortality. Although the relation of these two variables with daily mortality was not particularly high, it was verified that the variable that synthesized these two, which was named accumulated thermal overload (ATO), defined as $ATO_t(32) = Hlen_t(32) \times Exc_t(32)$, related proportionally with daily mortality.

On the basis of this variable, a model was created for the Lisbon district mortality data from 1980 to 1982, which related to the mortality of 1 day with the accumulated thermal overload of the previous day (Nunes *et al.*, 1999).

This model revealed equally adjusted and robust when the same mortality data for the years 1990 to 1992 was added (Nunes *et al.*, 1999). It was seen that when using the data that corresponded only to the summer period, the time series model simplifies substantially, disappearing the auto-correlated errors of the model and the model can be approached using linear regression models (Nunes *et al.*, 2001).

The knowledge of these models, with ability to estimate mortality increase associated with the occurrence of heat waves, allowed the construction of a Heat Health Watch Warning System for heatwaves in Portugal that is issued regularly since 1999 (Nogueira *et al.*, 1999; Nogueira, 2005).

The occurrence of the big heatwave of 2003 and its correct prediction demonstrated that the existing model – the ÍCARO model – and the Heat Health Watch Warning System for heatwaves with probable impact on the population's health – the ÍCARO surveillance system – were plainly justifiable (see 3.1) (Nogueira, 1999, 2005). This unusual episode of heat, with its extension in duration (about 15 days), brought the opportunity to update the existing models as well as of acquiring eventually new perspectives (Calado *et al.*, 2004a,b; Nogueira, 2005). The existing model revealed capability to estimate observed impacts on mortality, whether on relatively small heat episodes that occurred in June and those that occurred in July and August of the same year, with greater amplitude. It also estimated well the initial impact of this big heatwave, but it seemed to underestimate the later observed impacts (Nogueira, 2005). These facts are not negative for a surveillance system of heatwaves. However they place us before the necessity to improve the current models and supply the chance to increase the knowledge of the heatwave mechanism.

This article is intended to bring up to date the existing ÍCARO model for the relation between the occurrence of heatwaves and its impacts on the mortality of Lisbon's district population. This update had as additional objectives to contribute for the improvement of the existing surveillance system and to give new perspectives for a better understanding of the heatwave phenomenon.

This article is organized as follows: in Section 2 the data description and used methodology, the latter establishes the considered construction of the new variables, thresholds and models, and also the methods for evaluation of these last ones are described; in Section 3 the results presentation, the Lisbon district mortality of 1981, 1991 and 2003 summers are compared with the proposed temperatures thresholds, the model adopted for the ÍCARO surveillance system since 1999 is presented for reference and for the evaluation of the new studied models, new model parameters and the evaluation their quality are presented; in Section 4 the main conclusions are presented and discussed.

2. Material and methods

2.1. Data

In Lisbon district, the daily number of deaths and daily 3-hourly, average and maximum temperatures, from May to September, in the years of 1980 to 2003, were used. The mortality data from 1981 to 2002 were obtained from the Portuguese National Statistics Institute mortality databases.

The mortality from June to August of 2003 was supplied by the Portuguese Health General Directorate.

The data of Lisbon maximum air temperatures were supplied by the Portuguese Meteorology Institute.

2.2. Data and statistical analysis

For the daily mortality data series modelling Multiple Linear Regression Analysis was used, where functions of the air maximum temperature and indicator variables of the year were used as independent variables. Year indicator variables were used to model eventual alterations in the yearly mean mortality level.

The 'Stepwise' procedure was used to select the models, using the usual criteria, 5% significance for the variables to enter the model and 10% of significance to exit.

The validity of the models was made using the multiple correlation coefficients – R^2Adjusted – and the analysis models residual normality using the Kolmogorov-Smirnov test.

The data analysis was performed using the statistical program SPSS 13.0. For the preparation of some dependent variable the program Microsoft Excel 2002 was used.

2.3. Evaluation of the models

The evaluation of the capability of the models to correctly forecast days with increased mortality was made using the concepts of: sensitivity, specificity, positive and negative predictive values, percentage of false positives and negatives and percentage of correct decisions.

These concepts are defined below, based on the observed distributions of mortality from May to September (This is put generically and applies to any set of years chosen):

OCCURRENCE 1 – Day with high mortality – defined as 1 day where observed mortality exceeds the limit **L1 = Q3 + 1.5 * [Q3 − Q1]**;
OCCURRENCE 2 – Day with very high mortality – defined as one day where observed mortality exceeds the limit **L2 = Q3 + 3 * [Q3 − Q1]**.

(Q1 and Q3 are the 1st and the 3rd quartile respectively of the considered mortality distribution. These are well established Tukey limits for definition of *mild* and *severe* outliers. Tukey (1977); Hoaglin *et al.* (1983, 1985).)

Analogously forecast levels can be defined:

PRED 1 – The model forecasts (predicts) 1 day with a high number of deaths for the respective process of mortality (superior to L1);
PRED 2 – The model forecasts (predicts) 1 day with a very high number of deaths for the respective process of mortality (superior to L2);

Defining then,

Sensitivity – proportion of days where a high number of deaths was forecasted, among the days where effectively a high number of deaths was observed.

$$\text{Sensitivity} = \text{P(PRED|OCCURRENCE)}$$

Specificity – proportion of days where a non-high number of deaths was forecasted, among the days where effectively a high number of deaths did not occur.

$$\text{Specificity} = \text{P}(\overline{\text{PRED}}|\overline{\text{OCCURRENCE}})$$

(Bar represents negation of the given event (the complementary event). $\overline{\text{PRED}}$ means 'not predicted [a high level of deaths]'.)

DOI: 10.1002/joc

Positive Predictive Value – proportion of days where a high number of deaths occurred, among the days where a high number of deaths was forecasted by the model.

$$\text{Positive Predictive Value} = \text{P(OCCURRENCE|PRED)}$$

Value Negative Predictive – proportion of days where a high number of deaths did not occur, among the days where a number non high number of deaths was forecasted by the model.

$$\text{Negative Predictive Value} = \text{P}(\overline{\text{OCCURRENCE}}|\overline{\text{PRED}})$$

% False Positives = 1 − Specificity
% False Negatives = 1 − Sensitivity

Assuming that the proportion of days that were observed with a high level of deaths can be used as an estimate of the true prevalence - P(OCCURRENCE) - the model % of correct decisions can be define as,

$$\begin{aligned}\text{P(Correct)} &= \text{P(OCCURRENCE)} \\ &\times \text{P(PRED|OCCURRENCE)} + \text{P}(\overline{\text{OCCURRENCE}}) \\ &\times \text{P}(\overline{\text{PRED}}|\overline{\text{OCCURRENCE}})\end{aligned}$$

$$\begin{aligned}\text{P(Correct)} &= \text{P(OCCURRENCE)} \times \text{Sensitivity} \\ &+ (1 - \text{P(OCCURRENCE)}) \times \text{Specificity}\end{aligned}$$

$$\begin{aligned}\text{P(Correct)} &= \text{P(OCCURRENCE)} \\ &\times (\text{Sensitivity} - \text{Specificity}) + \text{Specificity}\end{aligned}$$

2.4. Construction of the accumulated thermal overload variable

The variable accumulated thermal overcharge (ATO) proposed by Nunes and Canto e Castro, can be easily generalized for any level of temperature threshold τ:

$$ATO_t(\tau) = HLen_t(\tau) \times Exc_t(\tau),$$

Where:

$$HLen_t(\tau) = \begin{cases} Hlen_{t-1}(\tau) + 1 & if \quad T\max_t \geq \tau \\ 0 & if \quad T\max_t < \tau \end{cases},$$

is the number of consecutive days that the maximum air temperature (Tmax) is above τ until day t and

$$Exc_t(\tau) = \begin{cases} T\max_t - \tau & if \quad T\max_t > \tau \\ 0 & if \quad T\max_t \leq \tau \end{cases},$$

is the excess of the maximum temperature above τ on day t.

2.5. Construction of the generalized accumulated thermal overload variable

For modelling purposes of the relation of the heatwaves occurrence phenomenon and mortality, and for discussion, generalization of the ATO variable is proposed here. This generalization will allow to test whether other models assuming dynamic threshold through the summer will give a better model response to phenomena that fixed thresholds seem not to correctly detect, as is the case of excess mortality occurring in the month of May with temperatures above 30 °C but below 32 °C.

Thus the GATO variable is defined as follows:

$$GATO_t(\tau) = GHLen_t(\tau) \times Exc_t(\tau),$$

where:

$$GHLen_t(\tau) = \begin{cases} GHLen_{t-1}(\tau) + 1 & if \quad T\max_t \geq \tau \\ GHLen_{t-1}(\tau) - 1 & if \quad T\max_t < \tau \wedge GHLen_{t-1}(\tau) > 0 \\ 0 & if \quad T\max_t < \tau \wedge GHLen_{t-1}(\tau) = 0 \end{cases}$$

is a balance (weighting) of the occurrence of several very close days where the maximum air temperature is above $\boldsymbol{\tau}$ until day t and

$$Exc_t(\tau)$$

is the excess of the maximum temperature above $Exc_t(\tau)$ in day t, as defined in 2.4.

The rational of the $GHLen_t(\tau)$ variable is to allow that the effect of the occurrence of several days of high temperatures (above of the threshold $\boldsymbol{\tau}$) can keep latent during some time, so that if another peak of heat occurs very near, the effect of the previous peak can still be included in the model forecast. In this way, if in the middle a big heatwave occurs one or two days with temperature below the respective threshold, the modelling process does not make a reset only makes a "cooling down", reducing the latent load of heat excess exposition.

2.6. Model for the heat-mortality relation

The assayed generic model in this work is given by:

$$\begin{aligned}Y_t = C + \alpha \times IYear + \beta \times GATO_{t-1}(\tau) + \gamma \\ \times GATO_t(\tau) + \delta \times Exc_t(\tau) + \varepsilon\end{aligned}$$

Where,

Y_t represents the number of deaths in day t;
$Iyear$ is an indicating variable or set of indicator variables of the year (1981, 1991 or 2003);
C is the regression constant,

DOI: 10.1002/joc

$\alpha, \beta, \gamma, \delta$ are regression parameters and
ε is a random process (white noise).

2.7. Definition of temperature thresholds

In this study, four different models were defined based on the daily maximum temperature variation. The temperature in each model was defined as:

Threshold I: $\boldsymbol{\tau} = \mathbf{32}$ °C from May to September

$$tau1 = 32 \quad if \quad t \geq week18$$

This fixed threshold corresponds to the established by Nunes *et al.* (1999). It is supported in the literature and it was observed that it corresponds approximately to 97.5 percentile of Lisbon's maximum air temperature during the summer period.

Threshold II: $\boldsymbol{\tau} = \mathbf{32}$ °C in May and June, $\boldsymbol{\tau} = \mathbf{35}$ °C from July to September

$$tau2 = \begin{cases} 32 & if \quad week18 \leq t < week28 \\ 35 & if \quad t \geq week28 \end{cases}$$

This threshold corresponds to a model, defined in some literature, which indicates that the impact of heatwaves tend to be greater when occurrence is precocious in the summer (Smoyer *et al.*, 2000; WHO 2004). The threshold 32 °C was obtained from the literature (CDC, 1995, 1996, 1997; Smoyer-Tomic *et al.*, 2003). The threshold 35 °C was obtained from the study of the weekly distribution of the temperatures during the broadened period of summer, from May to September, corresponding to the distribution free (non-parametric) estimate of the summer highest weekly 97.5 percentile (see 3.2).

Threshold III: $\boldsymbol{\tau}$ varies accordingly the weekly superior limits of the maximum air temperatures distribution starting in $\boldsymbol{\tau} = \mathbf{29}$ °C in May (until week 22 – end of May/beginning of June) increasing 1 °C per week until week 28 remaining at $\tau = 35$ °C until week 33, starting decreasing later slightly less than 1 °C per week.

$$tau3 = \begin{cases} 29 & if \quad t < week22 \\ 29 + (t - 22) & if \quad week22 \leq t < week28 \\ 35 & if \quad week28 \leq t < week34 \\ 35 - \frac{3}{4}(t - 33) & if \quad t \geq week34 \end{cases}$$

This threshold is a simplification of the weekly evolution of 97.5 percentiles of Lisbon air temperatures (see 3.2), corresponding to the usual evolution of these temperatures.

Threshold IV: $\boldsymbol{\tau}$ varies accordingly the weekly superior limits of the maximum air temperatures distribution starting in $\boldsymbol{\tau} = \mathbf{29}$ °C in May (until week 22 – end of May/beginning of June) increasing 1 °C per week until week 28 (2nd week of July), remaining $\boldsymbol{\tau} = \mathbf{35}$ °C until the end of September.

$$tau4 = \begin{cases} 29 + (t - 22) & if \quad week22 \leq t < week2 \\ 35 & if \quad week28 \end{cases}$$

This threshold is a mixture of thresholds II and III; therefore it makes the synthesis of the two ideas. One that the eventual occurrence corresponds to the adaptation to increasing temperatures that normally occur along the weeks, and another one that after the adaptation to temperatures up to the threshold of 35 °C there is no clear reason this adaptation to vanish as quickly as highest observed temperatures decrease at the end of summer.

3. Models nomenclature

For the purpose of a better understanding of models studied here are named as models I, II, III and IV. This numeration corresponds to the application of the model enunciated in 2.5. using the respective Accumulated Thermal Overload variable (ATO or GATO) with threshold I, II, III or IV in agreement to the defined in 2.6. Figure 1 shows the different thresholds considered.

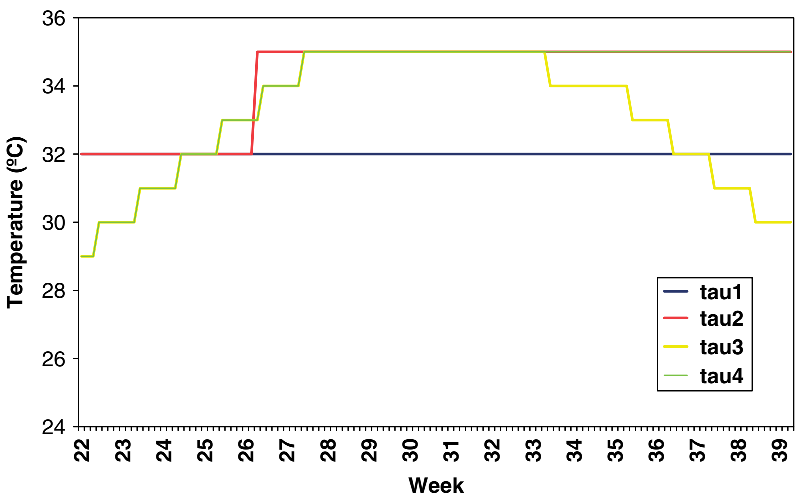

Figure 1. Weekly temperature thresholds values distribution for the 4 models. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

4. Results

4.1. The ICARO surveillance system

In Portugal a Heat Health Watch Warning system exists since 1999 – it is called the íCARO surveillance system – for heatwaves with potential harm effects on Portuguese population health. The system history is described in detail in Nogueira (2005).

This system was based on simple model knowledge described on Section 3.4 that showed a strong association between heat occurrence and raised levels of mortality for the population of Lisbon's district.

This knowledge generated a technical partnership between the Portuguese National Health and Meteorology institutes. This, with joint efforts, could generate up to 3 days advanced information on heatwave occurrence.

These two institutions, by their nature could not take mitigation actions to the population in general and to the health sector. Therefore, since the first moment, other institutions that fulfilled that gap were asked to be part of the surveillance system and agreed to do so, namely, the Portuguese National Health Directorate and the Portuguese National Service of Firemen and Civil Protection.

As an entity that crossed several institutions, the íCARO surveillance system, encountered some problems in articulating all involved partners. Sometimes it seemed that the technical information was not having adequate attention by the acting partners; and it is also easy to imagine that the acting partners sometimes felt annoyed by so much information and sometimes it surprisingly changes due to volatility of weather predictions. Prior to the big heatwave of 2003, predicted heatwave warnings were not always taken seriously but within the íCARO surveillance system's involved partners always intervened trying to mitigate eventual damages to the population's health.

The heatwave episode of 2003 showed that the íCARO model and surveillance system objectives and existence were completely justified. As can be seen in Figures 2 and 3, 2003 observed mortality and predicted íCARO indexes (reverted to expected mortality) for the Portuguese mainland and for the district of Lisbon are shown, several suppositions of the surveillance system were met:

1. The correct flagging of excess mortality episodes based solely on air temperatures from June to end of August;
2. The model based on Lisbon district showed overall correspondence in all of the mainland country;
3. The model was calibrated for heatwaves in June and July only and it showed remarkable results in this period;

Some other remarkable aspects were not that much expected, for example,

(a) Model resisted well to several heat replicates and was able to follow replicated peaks in mortality;
(b) Model gave a good response during the month of August where prior heatwave knowledge was not available;

Overall, the íCARO model seemed to overestimate mortality during the big heatwave, aspect that is not bad for a surveillance system, but it would be excessive to think that it was a flaw of the system. Since it was predicting a not previously seen long heatwave, in a time period also not previously experienced and probably reflecting heatwave warnings and interventions generated in the recent past by the system itself and even within this big heatwave.

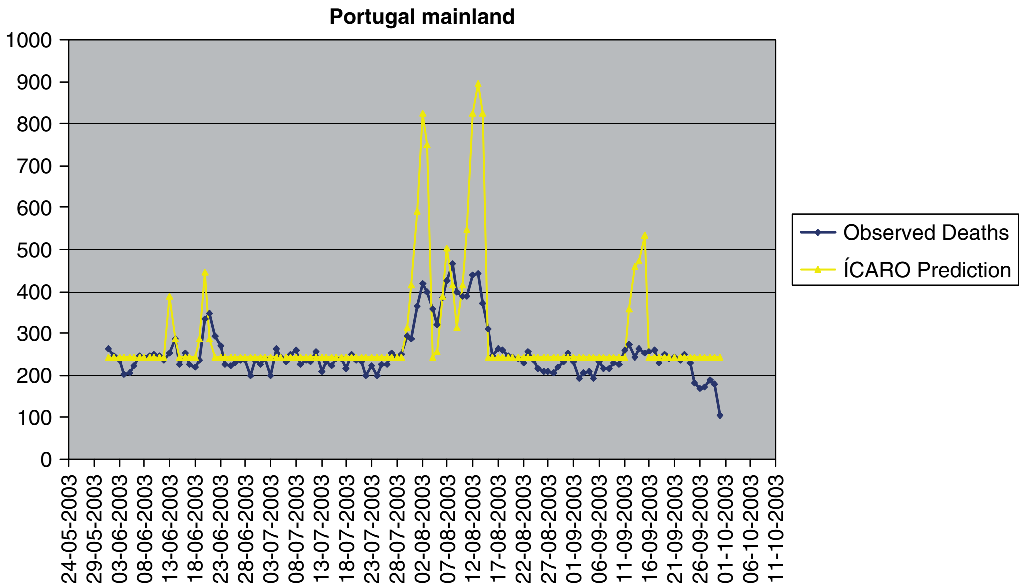

Figure 2. Daily ICARO surveillance system mortality risk predicted and observed daily mortality in the mainland of Portugal from 1st June to 21st September 2003. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

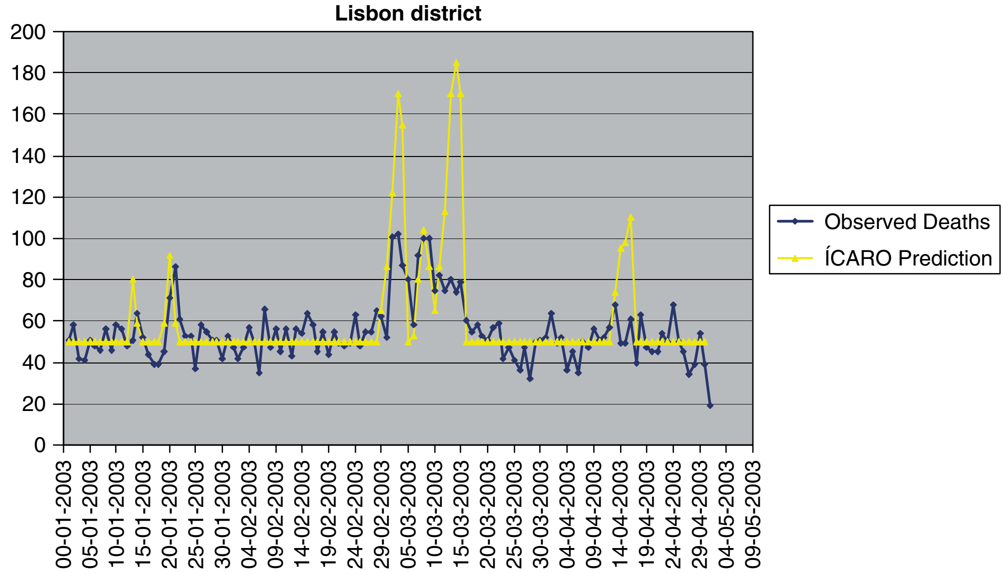

Figure 3. Daily ICARO surveillance system mortality risk predicted and observed daily mortality in Lisbon district from 1st June to 21st September 2003. This figure is available in colour online at www.interscience.wiley.com/ijoc

Table I. Summary of air temperatures distribution per week in Lisbon (1981 to 2000).

	3-hourly distributed air temperatures in Lisbon (1981–2000)			Daily mean air temperatures in Lisbon (1981–2000)			Daily maximum air temperature in Lisbon (1981–2000)		
Week	Percentile 25	MEDIAN	Percentile 75	Percentile 25	MEDIAN	Percentile 75	Percentile 25	MEDIAN	Percentile 75
18	13.4	15.6	18	14.1	15.7	18.3	18.0	19.9	24.3
19	13.4	15.4	17.8	14.3	15.7	16.8	18.3	19.9	21.7
20	13.8	15.7	18.3	14.6	16.3	17.8	18.2	20.6	23.4
21	14.6	16.6	19.6	15.8	17.2	18.5	19.7	21.6	24.2
22	15.4	17.4	20	16.6	17.5	19.3	20.6	22.4	25.5
23	15.6	18.0	20.6	17.0	18.2	19.5	21.1	23.1	25.7
24	16.8	19.8	23.8	18.1	19.7	23.1	23.0	26.1	30.7
25	17.2	19.5	22.6	18.6	19.6	21.0	23.2	25.0	27.8
26	17.4	19.7	23.2	18.9	20.2	21.8	23.6	25.6	29.4
27	17.4	20.0	23.2	19.0	20.3	21.5	23.6	25.7	28.6
28	18.4	21.4	25.4	20.3	21.4	24.5	25.8	27.6	31.7
29	18.8	22.0	25.7	20.1	21.7	24.8	25.6	28.5	32.5
30	18.7	21.6	25	20.4	21.6	23.8	25.4	27.8	31.8
31	18.8	21.6	25.1	20.4	21.8	23.6	25.7	28.1	31.2
32	19	22.0	25.4	20.6	21.8	24.1	26.0	27.7	31.3
33	18.7	21.1	24.8	20.6	21.5	22.9	26.3	27.8	30.2
34	18.8	21.0	24.2	20.4	21.4	22.8	25.5	27.2	30.2
35	18.8	20.8	24.2	20.3	21.5	22.7	25.3	27.4	30.6
36	19	21.2	24.8	20.7	21.5	23.3	26.1	28.3	30.9
37	18.4	20.7	23.6	19.2	21.2	22.7	24.7	26.5	30.4
38	17.7	19.7	22.2	18.6	20.1	21.4	22.8	24.8	27.5
39	17.2	19.2	21.6	17.9	19.4	21.0	22.0	24.3	27.4
40	16.4	18.3	20.9	17.4	18.6	20.1	21.1	24.0	26.5
Mean	17.1	19.5	22.6	18.4	19.7	21.5	23.1	25.2	28.4
SD	1.88	2.12	2.54	2.15	2.06	2.29	2.74	2.79	3.08

4.2. Distributions of air temperatures in Lisbon and in Portugal

4.2.1. Lisbon's temperatures

From observation of air temperatures distribution per week in Lisbon during the years of 1981 to 2000 (Table I) it is clear that during the heat season, temperatures have a well defined pattern. Daily temperatures increase steadily up the 32nd week of the year (approximately begin-mid August) and decrease again after that.

Table I shows air temperatures distribution using 3-hourly registered observations, daily mean and maximum

DOI: 10.1002/joc

temperatures median and quartiles. It was expectable that 3-hourly distributed temperatures would show a wider variability when compared with daily average temperatures (when measured in interquartile ranges per week). And it is remarkable that there is a great gap between an observed maximum daily temperature and mean temperature that the individual was exposed to. As an example, consider a median day in the considered period, a maximum temperature of 25.2 °C corresponds to a daily average exposure to 19.7 °C.

Table II presents results of smoothed non-parametric extreme threshold values for the considered air temperatures distributions. These extreme values are the standard for the determination of *mild* (above L1) and *severe* outliers (above L2) in exploratory data analysis.

Threshold values estimates take into account the variability within weeks, and maximum variability occurs around the 30th week, which ends up giving week extreme expected values around the 30th week instead of the 32nd.

Table II. Estimated air temperature theoretical thresholds.

	3-hourly distributed air temperatures in Lisbon (1981–2000)		Daily mean air temperatures in Lisbon (1981–2000)		Daily maximum air temperature in Lisbon (1981–2000)	
Week	L1	L2	L1	L2	L1	L2
18	24.0	30.6	22.0	27.0	29.6	36.9
19	24.8	31.6	22.2	27.0	30.2	37.4
20	25.7	32.6	22.7	27.3	31.1	38.4
21	26.3	33.4	22.5	26.6	30.8	37.6
22	28.3	36.1	24.4	29.3	34.0	42.0
23	29.4	37.5	24.9	29.5	34.7	42.5
24	30.4	38.7	25.6	30.3	36.1	44.4
25	31.4	40.1	26.0	30.5	36.7	45.0
26	32.9	42.2	27.5	32.5	38.3	47.0
27	33.3	42.5	27.7	32.7	38.4	46.8
28	34.0	43.5	28.6	33.9	39.7	48.7
29	34.5	44.2	29.0	34.4	40.1	49.0
30	35.2	45.0	29.9	35.6	40.7	49.7
31	34.8	44.4	29.0	34.2	39.8	48.3
32	34.1	43.2	27.9	32.4	38.7	46.5
33	33.6	42.5	27.4	31.5	38.1	45.6
34	33.4	42.1	27.1	31.1	37.8	45.0
35	32.7	41.1	26.8	30.8	37.8	45.1
36	31.7	39.6	26.7	30.7	37.5	45.0
37	30.9	38.5	26.5	30.8	37.1	44.9
38	29.9	37.2	26.1	30.5	36.3	44.1
39	28.6	35.6	25.5	30.1	35.2	43.0
40	27.3	33.8	24.5	28.8	33.2	40.4
Mean	30.7	39.0	26.1	30.8	36.2	44.1
SD	3.41	4.43	2.27	2.47	3.31	3.79

L1 and L2 are Tukey limits for mild and severe outliers in observed data
L1 = Q3 + 1.5 * [Q3 − Q1]
L2 = Q3 + 3 * [Q3 Q1]
Q1 and Q3 are the 1st and 3rd quartile of respective temperature distribution
These limits were smoothed using a 5 weeks moving average centred on the given week.

It is also noteworthy that *mild outliers* threshold (L1) estimated for 3-hourly distributed air temperatures is very similar to *extreme* outliers threshold (L2) for daily mean temperatures. This means that a very extreme temperature in a limited time period does not relate necessarily to a wide exposure to extreme temperatures.

On the other end, as expected, within considered weeks, 3-hourly distributed temperatures' *extreme outliers* thresholds are always within the *mild and severe outlier* thresholds of daily maximum temperatures.

4.2.2. Portugal mainland's temperatures

Figures 4 and 5 Show the distributions of air temperatures' 97.5% percentile during the months of May to September across the 18 districts of Portugal mainland. It is seen that mainland's temperatures show north–south and coast–interior gradients.

It is not difficult to accept that Lisbon's temperature may be a fair estimate for all the districts average temperature.

4.3. Construction of the models

In Figures 6–8 are presented, observed daily maximum temperature values for the district of Lisbon from June to September for the years of 1981, 1991 and 2003. These summers correspond to the occurrence of well known heatwave phenomena in Portugal and, in particular, also in the Lisbon district.

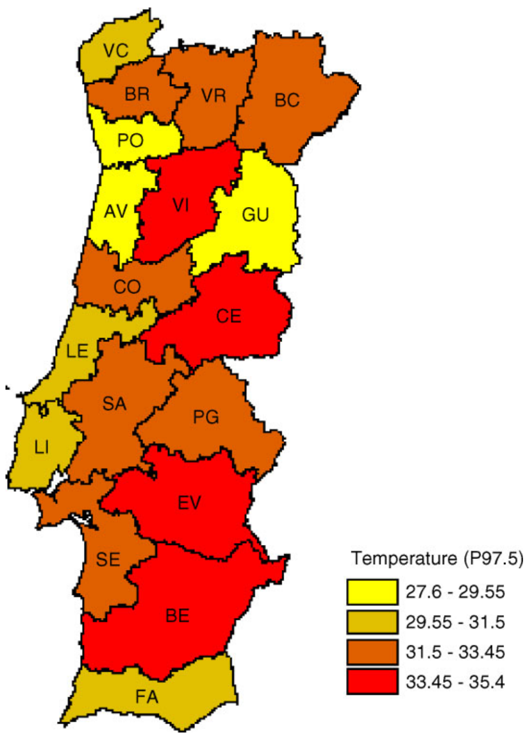

Figure 4. 3-hourly air temperature 97.5% percentile (p97.5) May to September 1981–2000 per district of Portugal mainland. This figure is available in colour online at www.interscience.wiley.com/ijoc

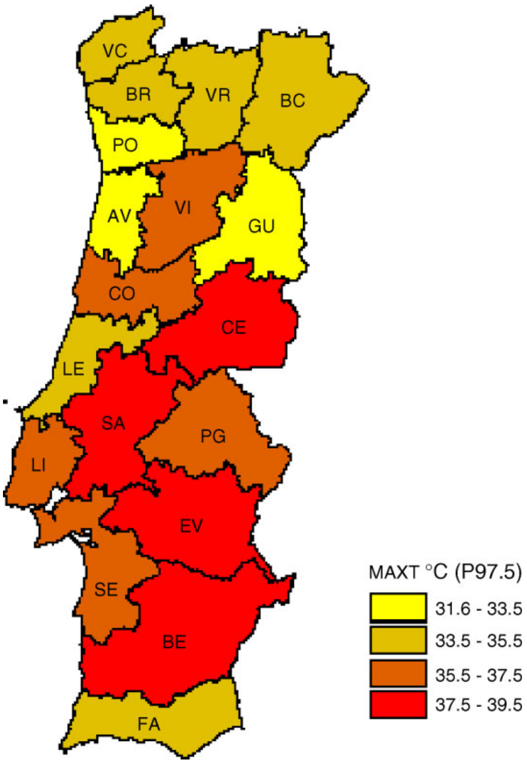

Figure 5. Daily maximum temperature 97.5% percentile (p97.5) May to September 1981–2000. This figure is available in colour online at www.interscience.wiley.com/ijoc

The proposed dynamic thresholds, besides being relatively intuitive and being in agreement with the literature that refer to the decreasing impacts in later summer, are also remarkably in accordance with the respective mortality lines.

4.4. The model adopted for the Portuguese surveillance system

The Portuguese Heat Health Watch Warning system for heatwaves with potential impact in the morbidity and mortality (The íCARO surveillance system), is active since 1999 (Nogueira *et al*., 1999; Nogueira, 2005) has proved itself as being based on a good rational and efficiently functioning in heatwaves situations (Nogueira, 2005).

The model used in the íCARO surveillance system since 1999, resulted from the adaptation of a time series model established by Nunes *et al*. (1999) for Lisbon district mortality from 1980 to 1982 and later adapted to include 1990 to 1992 data (Nunes *et al*., 1999).

This adaptation consisted in using only summer data, from July to August, and application of the linear regression model defined in 2.5 (with the ATO variable defined in 2.3. and not its generalized version GATO – defined in 2.4). This model presents an adjustment of about 51%. The only variable that was statistically significant for this model construction was the ATO of the previous day $ATO_{t-1}(\tau)$. The model was simply given by $Y_t = 44.3 + 1.6ATO_{t-1}(\tau)$. This model evaluation measures are presented in Tables III and IV.

Although from the statistical point of view the model did not show the ideal characteristics (it does not verify the residuals normality assumption), its choice was justified because this was the best known time series model.

From the validity point of view, as defined in 2.3., the model shows high specificity, excellent predictive values and globally very good capacity of correct decisions. However it also showed a low sensitivity and subsequent high percentages of false negatives (Table II). The apparently contradictory qualities and defects of this model are due to the very low prevalence of days with high level of mortality. In fact, very low prevalence of hot and very hot days produces estimates based on very small samples. And in years with particularly hot summers the occurrence of 3 consecutive days with very high temperatures can be quite common but not always generating peaks of mortality. Therefore inclusion of several years summer's data for the sensibility estimation can lead to an increase of its values, as high temperatures get more rare the higher probability exists that high mortality levels relate to them.

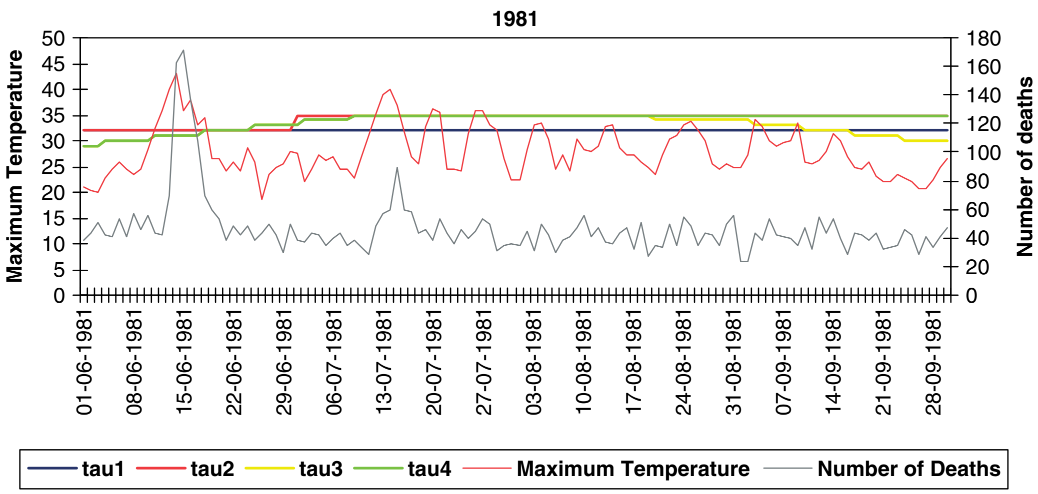

Figure 6. Lisbon's observed daily maximum temperature distribution values and number of deaths from June to September of 1981 and distribution of the 4 thresholds of temperature. This figure is available in colour online at www.interscience.wiley.com/ijoc

DOI: 10.1002/joc

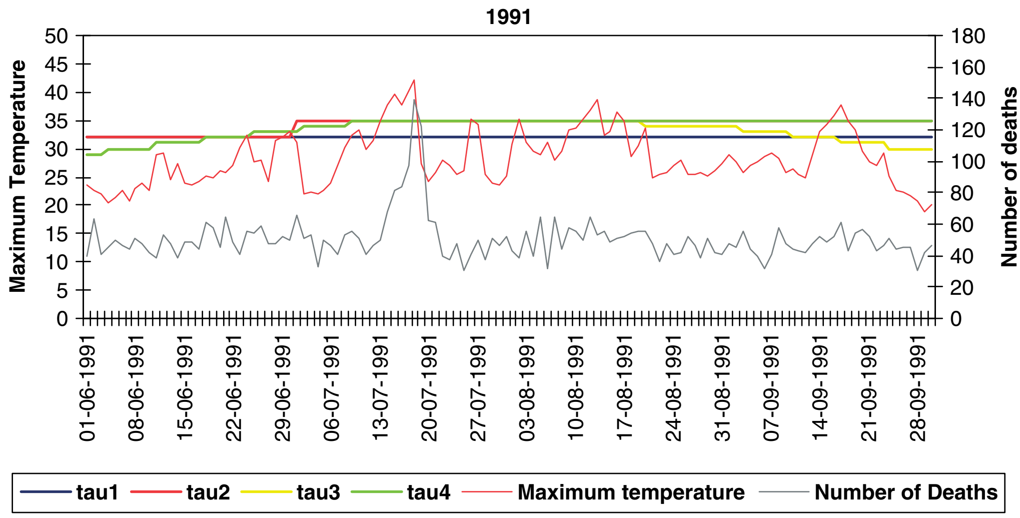

Figure 7. Observed daily maximum temperature distribution values and number of deaths from June to September of 1991 and distribution of the 4 thresholds of temperature. This figure is available in colour online at www.interscience.wiley.com/ijoc

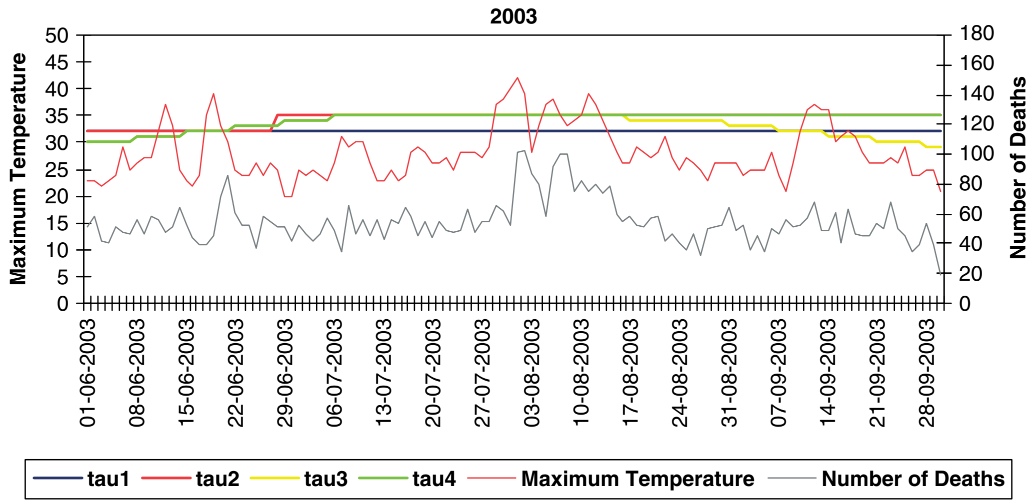

Figure 8. Observed daily maximum temperature distribution values and number of deaths from June to September of 2003 and distribution of the 4 thresholds of temperature. This figure is available in colour online at www.interscience.wiley.com/ijoc

Table III. ÍCARO model (1999) evaluation for the total of deaths in the months from June to August.

Definition	ÍCARO	Sensitivity	Specificity	False +	False −	Positive preditive values	Negative preditive values	Correct decision prob.
OCCURRENCE 1 (L1 = 75.5)	81 + 91	0.300	1.000	0.000	0.700	1.000	0.961	0.962
	81 + 2003	0.313	1.000	0.000	0.688	1.000	0.939	0.940
	1981–2003	0.212	0.999	0.001	0.788	0.778	0.988	0.987
OCCURRENCE 2 (L2 = 95)	81 + 91	0.000	1.000	0.000	1.000	–	0.962	0.962
	81 + 2003	0.125	0.994	0.006	0.875	0.500	0.962	0.957
	1981–2003	0.091	1.000	0.000	0.909	0.500	0.995	0.995

81 + 91 – results are based on 1981 and 1991 summer's data (only these two years are considered because of model's history);
81 + 2003 – The results are based on 1981 and 2003 summer's data;
1981–2003 – the results are based on 1981 to 2003 summer's data.

Table IV. ÍCARO model evaluation (that supports the implemented ICARO surveillance system) for the total of deaths in the months from May to September.

Definition	Years	Sensitivity	Specificity	False +	False −	Positive preditive values	Negative preditive values	Correct decision prob.
OCCURRENCE 1 (L1 = 71.5)	81 + 91	0.250	1.000	0.000	0.750	1.000	0.970	0.971
	81 + 2003	0.316	1.000	0.000	0.684	1.000	0.952	0.953
	1981–2003	0.133	0.999	0.001	0.867	0.800	0.985	0.985
OCCURRENCE 2 (L2 = 88)	81 + 91	0.000	1.000	0.000	1.000	–	0.974	0.974
	81 + 2003	0.100	0.996	0.004	0.900	0.500	0.967	0.964
	1981–2003	0.077	1.000	0.000	0.923	0.500	0.997	0.996

81 + 91 – results are based on 1981 and 1991 summer's data (only these 2 years are considered because of model's history);
81 + 2003 – The results are based on 1981 and 2003 summer's data;
1981–2003 – the results are based on 1981 to 2003 summer's data.

3.5. The difference between ATO and GATO variables

The big heatwave of 2003 showed that heatwaves do not occur with only a single peak of temperature and that consecutive peaks of high temperatures can occur with only one or two days of interval. The existing model would reset all gained impact to zero and would restart from zero. But although the previous model could flag heat occurrence it could not predict correctly its magnitude. In fact it overestimated it (Nogueira, 2005). In an attempt to generalize the prolonged overheat impact it is proposed to consider a weighting of all recent heat days. To show the difference between the variables Thermal Overload Accumulated and its generalization let us see the following example.

Table V attempts to show the difference between the ATO and GATO variables. An artificial example is presented, with temperature observed for 6 days, assuming $\tau = 32\,^{\circ}\mathrm{C}$. Hence, the last two columns, up to day 4 when the temperature suddenly drops to 29 °C, the two variables are equal, and it is after day 4 that differences occurs. In fact using GATO effects are expected to be higher since it is still considering potential remaining damages from the previous heat peaks that can be severed by the new heat peak.

3.6. Models for Lisbon's district total mortality

The four models fit (adjustment) results to Lisbon's district daily mortality from May to September of years 1981, 1991 and 2003 and respective measure values are presented in Table VI.

Of the four models assayed with the different temperatures thresholds, the model that statistically seems to have a better adjustment quality is Model IV (68.9%), which also presents normally distributed residuals (K-S: p = 0.163). Model II also showed very similar results with an adjusted determination coefficient of 68.7%. Also in this last model the residuals do not show departure from a white noise process (K-S: p = 0.175).

In models I and III, where adjusted determination coefficients of 50 and 58,7% were observed, respectively, the residuals were not normally distributed.

Table V. Example of the variables $HLen_t(\tau)$, $GHLen_t(\tau)$, $\mathrm{EXC}_t(\tau)$, $ATO_t(\tau)$ and $\mathrm{GATO}_t(\tau)$ functioning.

Observed maximum temperature	$HLen_t$ (τ)	$GHLen_t$ (τ)	EXC_t (τ)	ATO_t (τ)	GATO_t (τ)
33 °C	1	1	1	1	1
35 °C	2	2	2	4	4
38 °C	3	3	6	18	18
29°	0	2	0	0	0
33 °C	1	3	1	1	**3**
34 °C	2	4	2	4	**8**

$\mathrm{EXC}_t(\tau)$ – Excess of the maximum temperature above (τ) in day t
$\mathrm{ATO}_t(\tau)$ – Accumulated thermal overcharge where maximum air temperatures is above (τ) until in day t
$\mathrm{GATO}_t(\tau)$ – Generalized accumulated thermal overcharge where maximum air temperatures is above (τ) until in day t
$HLen_t(\tau)$ – Number of consecutive days that maximum air temperature is above (τ) in day t
$GHLen_t(\tau)$ – Balance (weighting) of occurrence of several very close days with maximum air temperature is above (τ) in day t.

In all assayed models with exception of Model I, the indicator variable of year 1991 and the $\mathrm{GATO_t}(\tau)$ were statistically significant. On the other hand the variable $EXC_t(\tau)$ only revealed to be significant in Model I, and was removed from all the other models. The year indicator variable for 2003 and $GATO_{t-1}(\tau)$, were significant in the four models although having distinct values for its parameters. The presence of the year indicator variables controls for the increase of the average number of daily deaths throughout the time span considered (years 1981, 1991 and 2003).

On Figure 9 the adjustment of the four assayed models to the Portuguese big heatwaves can be observed. All the models seem to forecast the excess of deaths in the three heatwaves (1981, 1991 and 2003). There seems to be a differentiation between model I and the remaining models.

- Model I seems to adjust itself well to the heatwave of 2003 and to adapt itself relatively poorly to the peaks of mortality of 1981 and 1991.

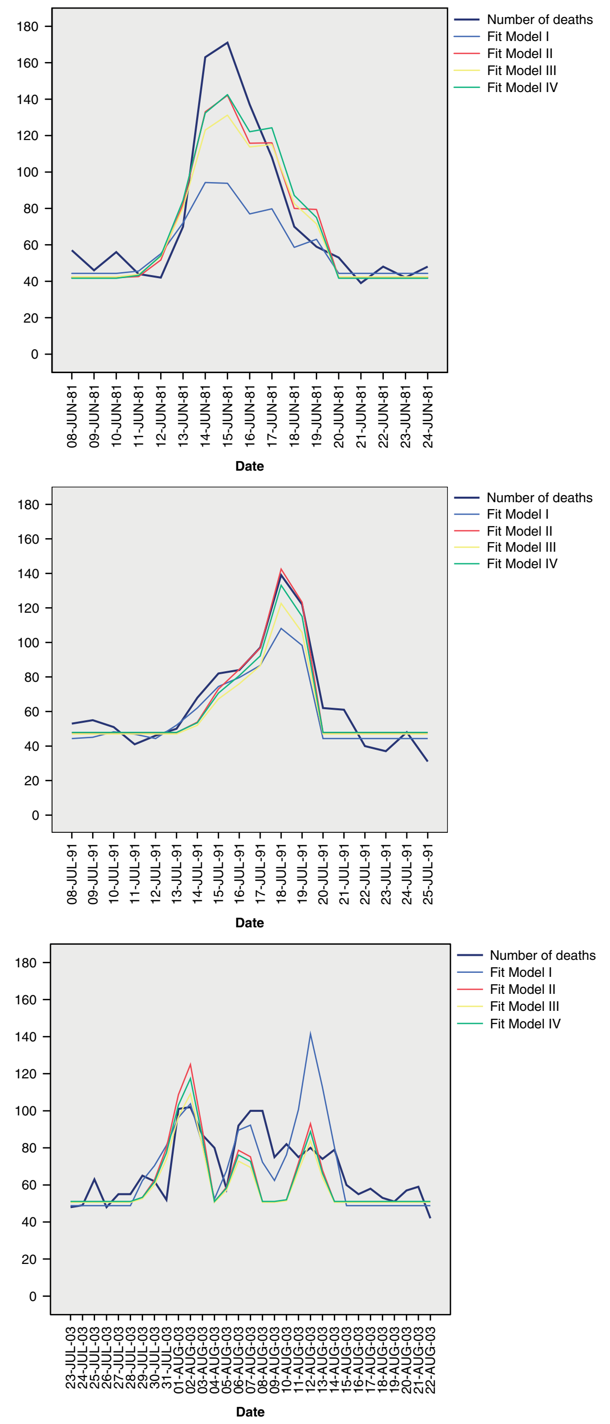

Figure 9. Models I, II, III and IV Fit to the observed data of mortality in the district of Lisbon, for periods of the big heatwaves of 1981, 1991 and 2003. This figure is available in colour online at www.interscience.wiley.com/ijoc

- Models II, III and IV seem to be much more sensible to the beginning of the heatwaves occurrences, but they adapt relatively poorly to mortality occurred during the long heatwave of 2003.

In Tables VI and VII the assayed models validity results for the two established limits of mortality are presented. These data are comparable with the ones presented in 3.1 for used íCARO model.

Here comparing with the íCARO model, higher sensitivity values for these new models are visible. The values of specificity are consistently high. Positive predictive values are not ideal but they present appreciable values for a phenomenon with a low prevalence. The Positive predictive values are consistently high, above 98%. And the correct decision probabilities for both the limits (L1 = 71.5 and L2 = 88 deaths) are also consistently high.

3.7. Models for Lisbon's district mortality in individuals with 65 years or older

In Table IX, the four models fit to the mortality data for individuals aged 65 years and older from May to September of 1981, 1991 and 2003 and respective validity measures are presented.

In the four assayed models considering as dependent variable the mortality of the individuals with aged 65 years and older, the model that seemed to have a better fit (adjustment) quality was Model II (74,1%), also presenting normally distributed residuals (K-S: $p = 0.283$). Model IV also presented very similar results, with an adjusted determination coefficient of 73.9% and residuals following a Normal distribution (K-S: $p = 0.141$).

Adjusted determination coefficients of 58.1% and 67.8% were obtained for models I and III, respectively. The residuals showed non-normality only in model I with $p = 0.005$; model III Kolmogorov–Smirnov had the value of $p = 0.067$.

In all the assayed models with exception of Model I, variable $GATO_t(\tau)$ was statistically significant, correctly quantifying an increase of the average number of daily deaths in 1981, 1991 and 2003. On the other hand variable $\mathrm{EXC}_t(\tau)$, only showed to be statistically significant in Model I and II, having been removed of all the other models.

From the observation of the regults, a good fit (adjustment) of the four assayed models to the Portuguese big heatwaves mortality is noticeable (Figure 10), as well as the to the one of the broadened summer period 2003 (Figure 10) in the age group of 65 years and older individuals. All the models forecasted the death excess of the heatwaves of 1981, 1991 and 2003. As it happened for the all age groups mortality above there seems to be a differentiation between model I and the remaining models. However, for the mortality of this particular aged population all the mortality peaks are well estimated by all models.

Globally, also for this age group, it is seen that:

- Model I seems to fit itself better to the heatwave of 2003 and to fit itself relatively poorly to the only peaks of mortality of 1981 and 1991.
- Models II, III and IV, seem to be much more sensible to the beginning of the heatwaves occurrences, but they fit relatively poorly to mortality occurred during the long heatwave of 2003.

Table VI. Models adjustment results for data of 1981, 1991 and 2003: estimates of the parameters for the total deaths, with generalized accumulated thermal overload.

	Constant	I1991	I2003	$GATO_{t-1}(\tau)$	$GATO_t(\tau)$	$Exc_t(\tau)$	
Model I	44.280	–	4.526	0.858	–	2.604	$R^2{}_{adj} = 0.500$
	$p < 0.001$		$p < 0.001$	$p < 0.001$		$p = 0.001$	K-S: $p = 0.003$
Model II	41.968	5.687	8.918	1.771	1.115	–	$R^2{}_{adjt} = 0.687$
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$		K-S: $p = 0.175$
Model III	42.441	4.594	8.156	1.386	0.867	–	$R^2{}_{adjt} = 0.587$
	$p < 0.001$	$p = 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$		K-S: $p = 0.04$
Model IV	41.691	6.230	9.385	1.639	1.036	–	$R^2{}_{adj} = 0.689$
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$		K-S: $p = 0.163$

Constant – linear regression constant;
I1991 – indicator variable for the year 1991;
I2003 – indicator variable for the year 2003;
$GATO_t(\tau)$ – generalized accumulated thermal overcharge where maximum air temperatures is above (τ) until in day t
$EXC_t(\boldsymbol{\tau})$ – excess of the maximum temperature above $(\boldsymbol{\tau})$ in day t;
p – significance level for regression coefficients;
K-S – significance level for normality Kolmogorov-Smirnov test;
$R^2{}_{adj}$ – adjusted R Square.

Table VII. Models evaluation of for the total mortality from May to September, considering the threshold value of 71,5.

	Years	Sensitivity	Specificity	False +	False −	Positive preditive values	Negative preditive Values	Correct decision prob.
Model I	81 + 91 + 2003	0.808	0.983	0.017	0.192	0.750	0.988	0.972
	1981 + 2003	0.842	0.988	0.012	0.158	0.842	0.988	0.978
	1981–2003	0.433	0.992	0.008	0.567	0.500	0.990	0.983
Model II	81 + 91 + 2003	0.654	0.983	0.017	0.346	0.708	0.978	0.963
	1981 + 2003	0.632	0.973	0.027	0.368	0.632	0.973	0.949
	1981–2003	0.300	0.996	0.004	0.700	0.563	0.988	0.984
Model III	81 + 91 + 2003	0.538	0.965	0.035	0.462	0.500	0.970	0.939
	1981 + 2003	0.474	0.969	0.031	0.526	0.529	0.961	0.935
	1981–2003	0.283	0.990	0.010	0.717	0.333	0.987	0.978
Model IV	81 + 91 + 2003	0.615	0.978	0.022	0.385	0.640	0.975	0.956
	1981 + 2003	0.579	0.977	0.023	0.421	0.647	0.969	0.949
	1981–2003	0.300	0.994	0.006	0.700	0.450	0.988	0.982

81 + 91 + 2003 – the results are based on 1981, 1991 and 2003 summer's data;
81 + 2003 – the results are based on 1981 and 2003 summer's data;
1981–2003 – the results are based on summer's data from 1981 to 2003.

notice that the variables present in all the four mortality models are the same ones, it is curious to observe, in Figure 11, that models III and IV are better to estimate the impact of the first mortality wave, associate to a small peak of heat that occurs in June of 2003.

In Tables X and XI the results of validity for the assayed models for the two established limits of mortality are presented. These data are comparable with íCARO's model ones presented in 3.1 and total mortality models' in 3.4.

Models sensitivity values are visibly higher than the íCARO model ones, but the results are completely similar to the ones obtained for the total daily mortality models. The specificity values are consistently high. The positive predictive values are not ideal but they present appreciable values for a phenomenon with a low prevalence. The negative predictive values are consistently high, above of 98%. The correct decision probability for both the limits (L1 = 56.5 and L2 = 73 deaths) is also consistently high. Perhaps noteworthy is the fact that models II, III and IV sensitivity is slightly higher than in models for all ages mortality.

4. Discussion and conclusions

The results obtained with the four models are very similar and all of them had a good fit to data. However, the best determination coefficient for the total of deaths was obtained by model IV (68,9%), with Model II having a very close result (68,7%). Model I was the one that presented the lower determination coefficient (50%).

DOI: 10.1002/joc

Table VIII. Models Evaluation of for the total mortality from May to September, considering the threshold value of 88.

	Years	Sensitivity	Specificity	False +	False −	Positive preditive values	Negative preditive values	Correct decision prob.
Model I	81 + 91 + 2003	0.615	0.993	0.007	0.385	0.727	0.988	0.981
	1981 + 2003	0.600	0.989	0.011	0.400	0.667	0.985	0.975
	1981–2003	0.615	0.997	0.003	0.385	0.444	0.999	0.996
Model II	81 + 91 + 2003	0.692	0.995	0.005	0.308	0.818	0.990	0.986
	1981 + 2003	0.600	0.992	0.008	0.400	0.750	0.985	0.978
	1981–2003	0.692	0.998	0.002	0.308	0.529	0.999	0.997
Model III	81 + 91 + 2003	0.615	0.995	0.005	0.385	0.800	0.988	0.984
	1981 + 2003	0.600	1.000	0.000	0.400	1.000	0.985	0.985
	1981–2003	0.615	0.997	0.003	0.385	0.471	0.999	0.996
Model IV	81 + 91 + 2003	0.692	0.993	0.007	0.308	0.750	0.990	0.984
	1981 + 2003	0.600	0.996	0.004	0.400	0.857	0.985	0.982
	1981–2003	0.692	0.997	0.003	0.308	0.429	0.999	0.995

81 + 91 + 2003 – the results are based on 1981, 1991 and 2003 summer's data;
81 + 2003 – the results are based on 1981 and 2003 summer's data;
1981–2003 – the results are based on summer's data from 1981 to 2003.

Table IX. Models using Generalized Accumulated Thermal Overload fit (adjustment) to mortality data in individuals aged 65 years and older data of 1981, 1991 and 2003 results and validity measures values.

	Constant	I1991	I2003	$\mathrm{GATO_{t-1}}(\tau)$	$\mathrm{GATO_t}(\tau)$	$\mathrm{Exc_t}(\tau)$	
Model I	27.506	4.102	10.556	0.779	–	2.223	$R^2{}_{adj} = 0.581$
	$p < 0.001$	$p = 0.002$	$p < 0.001$	$p < 0.001$	–	$p < 0.001$	K-S: $p = 0.005$
Model II	27.042	6.383	13.076	1.533	1.333	−1.750	$R^2{}_{adjt} = 0.741$
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.047$	K-S: $p = 0.283$
Model III	27.078	5.493	13.398	1.320	0.752	–	$R^2{}_{adj} = 0.678$
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	–	K-S: $p = 0.067$
Model IV	26.680	6.865	13.369	1.459	0.842	–	$R^2{}_{adjt} = 0.739$
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	–	K-S: $p = 0.141$

Constant – linear regression constant;
I1991 – indicator variable for the year 1991;
I2003 – indicator variable for the year 2003;
$\mathrm{GATO}_t(\boldsymbol{\tau})$ – generalized accumulated thermal overcharge where maximum air temperatures is above $(\boldsymbol{\tau})$ until in day t
$\mathrm{EXC}_t(\boldsymbol{\tau})$ – excess of the maximum temperature above $(\boldsymbol{\tau})$ in day t;
p – significance level for regression coefficients;
K-S – significance level for normality Kolmogorov-Smirnov test;
$R^2{}_{adj}$ - adjusted R Square.

This means that model I explains less of the observed data which reveals that the fixed threshold model is in disadvantage to the dynamic threshold models. A relevant difference between model I (fixed threshold) and the others (models II, III and IV – dynamic threshold models) is that in all models all mortality predictions depend mainly on the previous day information and secondarily on the given day itself. Model I depends on the day itself through the excess of temperature above the fixed threshold while the remaining models this dependence is through the accumulated thermal overload (ATO or GATO).

The analysis of the models validity, using sensitivity, specificity, predictive values and correct decision percentage notions, does not allow the clear distinction of the models.

The results jointly with the graphical visualization of the goodness of fit of the models forecasts to the observed mortality, allowed dividing the models in two groups. Model I, with the fixed reference threshold of 32 °C and the remaining models (models II, III and IV).

The model I corresponds to the classic model adopted by the íCARO surveillance system, showed some misalignment with heatwaves that were more distant in time (1981 and 1991) and a noteworthy capability to adjust the mortality peaks replicates occurred during the long heatwave of 2003 heat. However it seemed to overestimate the impacts of these peaks. Despite the slight difference of model I considered here, to the ICARO system implemented one, these results agree with the forecasted impacts during the heatwave of 2003 where it already widely overestimated the impacts

DOI: 10.1002/joc

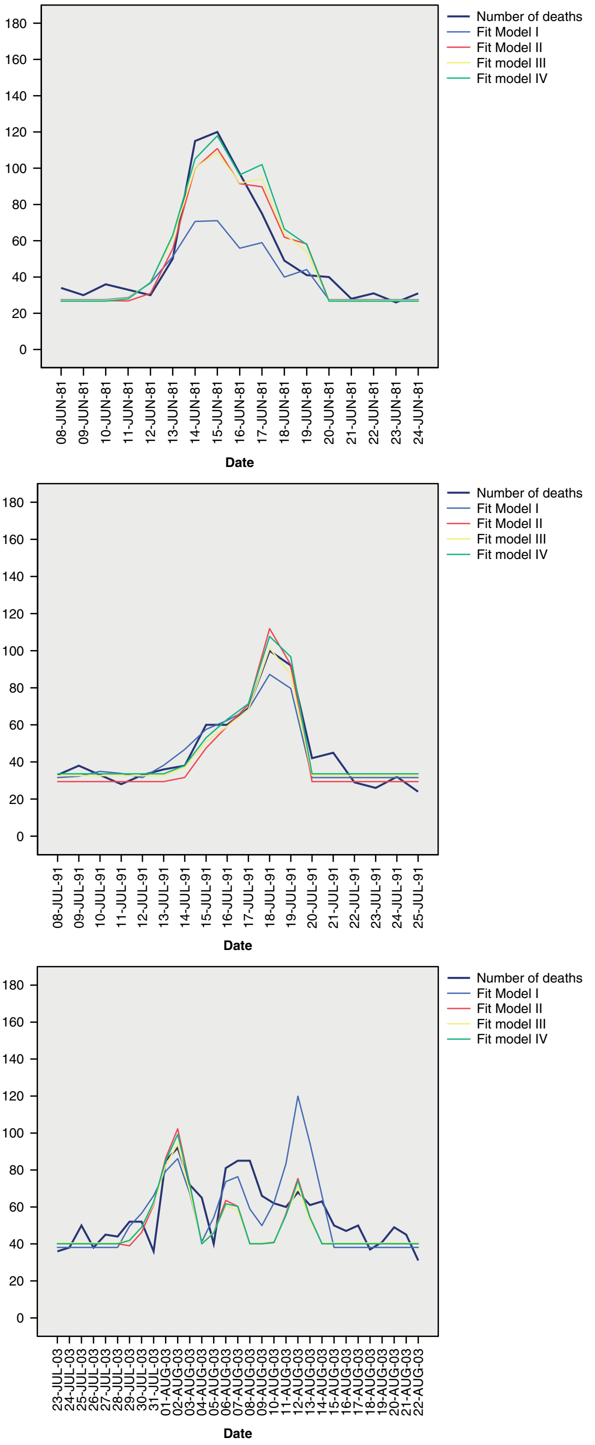

Figure 10. Models I, II, III and IV fit to the Lisbon district observed mortality data for the individuals aged 65 years and older, in the periods of the big heatwaves of 1981, 1991 and 2003. This figure is available in colour online at www.interscience.wiley.com/ijoc

of these heat peaks replicates seen in 3.1 (Nogueira, 2005).

Models II, III and IV showed opposite virtues to model I, they are particularly good accounting for the effect of mortality isolated peaks or the initial peak of mortality of the great heatwave of 2003, but they fit relatively poorly to the following peaks of mortality (replicates). However, they show the capacity to follow the same peaks, although globally underestimating their effects, i.e. when heat replicates generate increase mortality, models II, III and IV also increase their magnitude, but underestimating the real observed magnitude.

The models demonstrated a higher goodness of fit to the data of individuals aged 65 years or older, which is accordingly to what was expectable. On the other hand, of the study of the Portuguese big heatwaves episodes, there seems to be an evolution of lower impact of the heatwaves in the younger age groups along the years (Calado *et al*., 2004a,b). This eventual improvement can be justified by, health systems improvements, better adaptation of the families to the phenomenon of the heatwaves, improvement of the quality of housing conditions and places of children permanence and greater use of refrigeration systems.

It can be argued that the relative lack of fit of models II, III and IV, in the heatwave July – August of 2003, may result of the intervention that occurred during its unexpected long duration. In fact, interventions from health and civil protection authorities occurred during the heatwave and breaking news in social media of impacts (high mortality) also occurred during the 17 days that the heatwave lasted.

- In these circumstances model I could be suggesting that, for its good fit to the 2003 heatwave, more mortality was expected.
- The remaining models, for their good adjustments to initial impacts of heat, can be suggested that a greater initial impact and later lower impact was expected. Effects that previous interventions, early in that summer and in previous summers, may have attenuated but were not able to prevent.

Observation and discussion of diagnostic statistics (sensitivity, specificity, predictive values and correct decision probabilities – Tables VII, VIII, X and XI) presented some problems because there is a big variation across models and sets of years considered. Nevertheless, two distinct patterns emerged:

1. For mortality threshold 1–Moderate high levels of mortality (L1 – Tables VII and X)

Model I has higher values of sensitivity, predictive (both positive and negative) and correct decision probabilities, when compared to all other models.

2. For mortality threshold 2 – Very high levels of mortality (L2 – Tables VIII and XI)

Models II, III e IV show higher values of sensitivity, predictive (both positive and negative) and correct decision probabilities, when compared to Model I.

And among these several models they are almost indistinguishable in terms of diagnostic characteristics.

Using diagnostic values in all years (1981 to 2003) as criteria model II has advantage only in positive predictive value (slightly higher).

These results hold for both all ages and 65 years and older age groups mortalities.

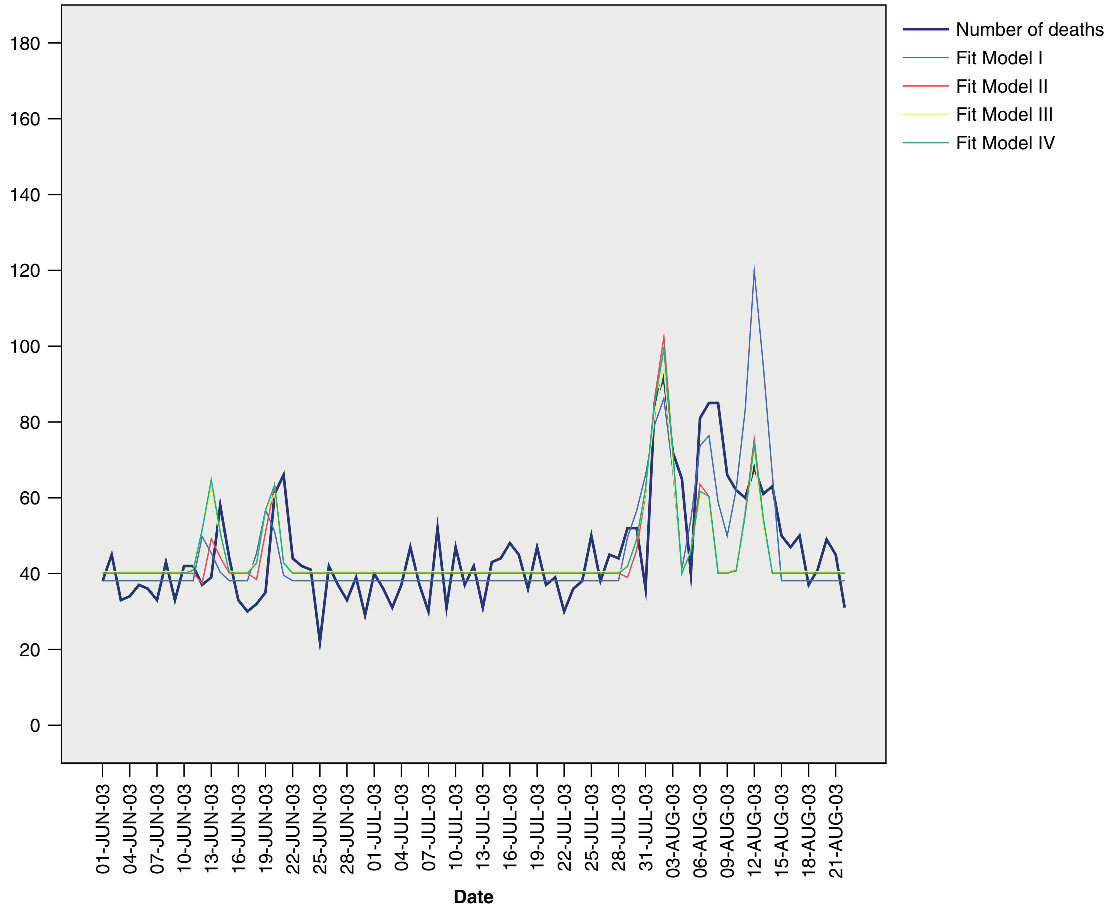

Figure 11. Models I, II, III and IV fit to the Lisbon's district observed mortality data for individuals aged 65 and older, in the period from 1st of June to 22nd of August 2003. This figure is available in colour online at www.interscience.wiley.com/ijoc

Table X. Evaluation of models for mortality in individuals aged 65 years and older, considering the months from May to September and the Threshold of value 56.5.

	Years	Sensitivity	Specificity	False +	False −	Positive preditive values	Negative preditive values	Correct decision prob.
Model I	81 + 91 + 2003	0.760	0.983	0.017	0.240	0.731	0.985	0.970
	1981 + 2003	0.700	0.984	0.016	0.300	0.778	0.977	0.964
	1981–2003	0.541	0.996	0.004	0.459	0.571	0.995	0.991
Model II	81 + 91 + 2003	0.600	0.993	0.007	0.400	0.833	0.976	0.970
	1981 + 2003	0.550	0.988	0.012	0.450	0.786	0.966	0.956
	1981–2003	0.405	0.997	0.003	0.595	0.625	0.994	0.991
Model III	81 + 91 + 2003	0.600	0.965	0.035	0.400	0.517	0.975	0.944
	1981 + 2003	0.550	0.973	0.027	0.450	0.611	0.965	0.942
	1981–2003	0.432	0.990	0.010	0.568	0.314	0.994	0.984
Model IV	81 + 91 + 2003	0.600	0.975	0.025	0.400	0.600	0.975	0.953
	1981 + 2003	0.550	0.976	0.024	0.450	0.647	0.965	0.945
	1981–2003	0.432	0.993	0.007	0.568	0.400	0.994	0.987

81 + 91 + 2003 – The results are based on 1981, 1991 and 2003 summer's data;
81 + 2003 – The results are based on 1981 and 2003 summer's data;
1981–2003 – the results are based on summer's data from 1981 to 2003.

Models III and IV, for both all age groups mortality and mortality of individuals aged 65 years and more, showed indications of having a better validity for the mortality of May and September. This fact was expected because the design of the respective reference thresholds was particularly established for these months. The fact that the Model III does not show a better fit, points out that it is not necessary to increase heat sensitivity

Table XI. Evaluation of models for the mortality in individuals aged 65 years considering the months from May to September and the threshold of value 73.

	Years	Sensitivity	Specificity	False +	False −	Positive preditive values	Negative preditive values	Correct decision prob.
Model I	81 + 91 + 2003	0.545	0.993	0.007	0.455	0.667	0.988	0.981
	1981 + 2003	0.444	0.989	0.011	0.556	0.571	0.981	0.971
	1981–2003	0.545	0.997	0.003	0.455	0.400	0.999	0.996
Model II	81 + 91 + 2003	0.727	0.998	0.002	0.273	0.889	0.993	0.991
	1981 + 2003	0.667	0.996	0.004	0.333	0.857	0.989	0.985
	1981–2003	0.727	0.998	0.002	0.273	0.533	0.999	0.997
Model III	81 + 91 + 2003	0.727	0.995	0.005	0.273	0.800	0.993	0.988
	1981 + 2003	0.667	1.000	0.000	0.333	1.000	0.989	0.989
	1981–2003	0.727	0.997	0.003	0.273	0.471	0.999	0.997
Model IV	81 + 91 + 2003	0.727	0.995	0.005	0.273	0.800	0.993	0.988
	1981 + 2003	0.667	0.996	0.004	0.333	0.857	0.989	0.985
	1981–2003	0.727	0.997	0.003	0.273	0.471	0.999	0.997

81 + 91 + 2003 – The results are based on 1981, 1991 and 2003 summer's data;
1981 + 2003 – The results are based on 1981 and 2003 summer's data to the of;
1981–2003 – the results are based on summer's data from 1981 to 2003.

at the summer's end. On the other hand, this fact is also due to the lack of anomalous mortality episodes in September. The two models rationales can only be distinguished when such anomalous episodes will occur.

The results of this work do not contradict the hypothesis that the heatwaves impacts are greater (more severe) when they occur early in summer. Better goodness of fit to the small known peaks of mortality were obtained with models II, III and IV that used dynamic thresholds with initial growing limits, with apparent success.

The relatively low value of models sensitivity, are not surprising values when we bear in mind that daily data of 23 summers is being used. These low values can occur due to the adopted case definition. Days of high or very high mortality can occur without that such phenomenon corresponding to a heatwave. External causes, road accidents or similar occurrences can increase mortality in one isolated day. Any such occurrence diminishes sensitivity and increases the percentage of false negatives without this affecting the quality of the models.

The occurrence of heatwaves is associated with heat persistence during some time and this is one of the implications of the quality of the models studied here. Therefore, the days where anomalous excesses of mortality occur that are not associated with heat happen sporadically and they do not associate in small clusters. Thus considering a probability of occurrence of a false negative of 0.2 in a 3 days forecast, the probability of missing the forecast in the 3 days is only 0.008; if eventually that probability of error goes up to 0.4 the probability of error in 3 days is 0.064. Thus, the relatively reduced values of sensitivity and false negatives are not really problematic.

In fact, the high values of specificity are much more important and the consequent very low values of false positives. If we will have 3 consecutive days where it is forecasted mortality excess, and the percentage of false positives is 0.005 (as it happens in the models studied here) the probability of error in all the 3 days is 0.000000125.

This work had as main motivation the update of the used model for the íCARO surveillance system generated for data of waves of heat of 1981 and 1991: on one hand new data referring to a great heatwave occurred in 2003 existed; on the other hand, the acquired knowledge the heatwaves of more moderate impact in mortality had evolved sufficiently in the scope of íCARO project.

The discussion of models done here was meant to evolve the knowledge on the mechanism of the mortality associated with the occurrence of extreme heat and to choose the best model for the Lisbon heatwaves surveillance system. In terms of the model for heatwaves occurrence forecasting it is intended that the respective impacts are detected, mainly the initials ones. Apparently any of models II, III or IV is adjusted for this main requirement. From the discussion made, it was considered that model IV had the conditions to be the elected model. This because,

It has the best statistical properties;

It synthesizes the greatest capability of correctly forecasting heatwaves of with impacts mortality in precocious in the period of summer.

Acknowledgements

Fundação para a Ciência e Tecnologia (FCT - Portuguese Foundation for Science and Technology) and its *Programa Operacional 'Ciência, Tecnologia e Inopvação'* (POCTI) and co-funded by European Community Fund FEDER. (POCT1/ESP/39679/2001). *'Impactos na Saúde em Portugal de Eventos Extremos:*

Passado, Presente e Futuro (ImpactE)', funded by Fundação Calouste Gulbenkian.

References

Brucker G. 2005. Vulnerable populations: lessons learnt from the summer 2003 heat waves in Europe. *Euro Surveillance* **10**(7): 147 (available at http://www.eurosurveillance.org/em/v10n07/1007-221.asp, last accessed 11-11-2005).

Calado R, Botelho J, Catarino J, Carreira M, Nogueira PJ, Paixão EJ, Falcão JM. 2004a. Onda de calor de Agosto de 2003: os seus efeitos sobre a mortalidade da população portuguesa. Relatório científico. Observatório Nacional de Saúde – Instituto Nacional de Saúde Dr. Ricardo Jorge e Direcção de Serviços de Informação e Análise – Direcção-Geral da Saúde 2004.Lisboa.

Calado R, Botelho J, Catarino J, Carreira M, Nogueira PJ, Paixão EJ, Falcão JM. 2004b. Onda de calor de Agosto de 2003: os seus efeitos sobre a mortalidade da população portuguesa. Revista Nacional de Saúde Pública. Julho/Dezembro de 2004. Vol. 2; 7–20.

Clarke JF. 1972. Some climatological aspects of heat waves in the contiguous United States. *Environmental Health* **5**: 76–84.

Centers for Disease Control and Prevention (CDC). 1995. Heat-related illness and deaths-United States, 1994–1995. *MMWR Morbidity and Mortality Weekly Report* **44**(25): 465–468.

Centers for Disease Control and Prevention (CDC). 1996. Heat-wave related mortality – milwaukee, Wisconsin, July 1995. *JAMA* **276**(4): 275.

Centers for Disease Control and Prevention (CDC). 1997. Heat-related deaths-Dallas, Wichita, and Cook Counties, Texas, and United Stattes, 1996. *MMWR Morbidity and Mortality Weekly Report* **46**(23): 528–531.

Centers for Disease Control and Prevention (CDC). 2002. Heat related deaths – four states, July-August 2001, and United states, 1979–1999. Technical Report 51. *Morbidity and Mortality Weekly Report* 567–570.

Conti S, Meli P, Minelli G, Solimini R, Toccaceli V, Vichi M, Beltrano C, Perini L. 2005. Epidemiologic study of mortality during the summer 2003 heat wave in Italy. *Environment Research* **98**: 390–399.

Dessai S. 2002. Heat stress and mortality in Lisbon Part I: model construction and validation. *International Journal of Biometeorology* **47**: 6–12.

Diaz J, Jordan A, Garcia R, Lopez C, Alberdi JC, Hernandez E, Otero A. 2002. Heat waves in Madrid 1986–1997: effects on the health of the elderly. *International Archives of Occupational and Environmental Health* **75**(3): 163–170.

Falcão JM, Castro MJ, Falcão MLM. 1988. Efeitos de uma onda de calor na mortalidade da população de Lisboa. *Saúde em Números* **3**(2): 10–12.

Falcão JM, Paixão EJ, Nogueira PJ, Nunes B. 2004. *Onda de calor de Julho de 2004: Efeitos Sobre a Mortalidade na região do Algarve, Agosto de 2004*. Observatório Nacional de Saúde – Instituto Nacional de Saúde Dr. Ricardo Jorge: Lisboa.

Garcia AC, Nogueira PJ, Falcão JM. 1999. *Onda de calor de Junho de 1981 em Portugal: Efeitos na Mortalidade*. Revista Portuguesa de Saúde Pública; Escola Nacional de Saúde Publica: Lisbon, Portugal Volume Temático I; 67–77.

Garcia-Herrera R, Díaz J, Trigo RM, Hernandéz E. 2005. Extreme summer temperatures in Ibéria: health impacts and associated synoptic conditions. *Annales Geographycae* **23**: 239–251.

Henschel A, Burton LL, Margolies L, Smith JE. 1969. An analysis of the heath deaths in St. Louis during July, 1966. *American Journal of Public Health* **59**(12): 2232–2242.

Hoaglin DC, Mosteller F, Tukey JW (eds). 1983. *Understanding Robust and Exploratory Data Analysis* John Wiley & Sons: New York, NY, USA.

Hoaglin DC, Mosteller F, Tukey JW (eds). 1985. *Exploring Data Tables, Trends and Shapes* John Wiley & Sons: New York, NY, USA.

Kilbourne EM. 1999. The spectrum of illness during heat waves. *American Journal of Preventive Medicine* **16**(4): 359–360.

Kunst AE, Looman CWN, Mackenbach JP. 1993. Outdoor air temperature and mortality in the Neatherlands: a Time series analysis. *American Journal of Epidemiology* **137**(3): 331–341.

Nogueira PJ. 1999. Distribuição da temperatura do ar em Lisboa, [Não publicado.].

Nogueira PJ. 2005. Examples of Heat Health Warning Systems: Lisbon's ICARO's surveillance system, summer of 2003. In *Extreme Weather Events and Public Health Responses*, Kirch W, Menne B, Bertollini R (eds). World Health Organization. Printed in European Union. ISBN 3-540-24417-4 Springer: Berlin, Heidelberg, New York.

Nogueira PJ, Nunes B, Dias CM, Falcão JM. 1999. *Um sistema de vigilância e alerta de ondas de calor com efeitos na mortalidade: o índice ícaro*. Revista Portuguesa de Saúde Pública: Escola Nacional de Saúde Publica: Lisbon, Portugal Volume Temático: I; 79–84.

Nunes B, Canto e Castro L, Falcão JM. 1999. *Efeitos do Excesso do calor na Mortalidade: Modelos de Previsão. Relatório Cientifico*. Observatório Nacional de Saúde – Instituto Nacional de Saúde Dr. Ricardo Jorge: Lisboa.

Nunes B, Canto e Castro L, Nogueira PJ. 2001. *The Effect of Excessive Heat on Mortality. Observatório Nacional de Saúde*. Instituto Nacional de Saúde Dr. Ricardo Jorge: Not Published.

Paixão EJ, Nogueira PJ. 2002. *Estudo da Onda de calor de Julho de 1991 em Portugal: Efeitos na Mortalidade; Relatório Cientifico. Outubro2002*. Observatório Nacional de Saúde – Instituto Nacional de saúde Dr. Ricardo Jorge: Lisboa.

Paixão EJ, Nogueira PJ. 2003. *Efeitos de uma onda de calor na mortalidade*, Vol. 21. Revista Portuguesa de Saúde Pública: Escola Nacional de Saúde Publica: Lisbon, Portugal N.°1; 41–54.

Saez M, Sunyer J, Castellsague J, Murillo C, Anto JM. 1995. Relationship between weather temperature and mortality: a time series analysis approach in Barcelona. *International Journal of Epidemiology* **24**(3): 576–582.

Schuman SH, Arbor A, Anderson CP, Oliver J. 1964. Epidemiology of Successive Heat Waves in Michigan in 1962 and 1963. *JAMA* **189**(10): 733–738.

Semenza JC, McCullough JE, Flanders WD, McGeehin MA, Lumpkin JR. 1999. Excess hospital admissions during the July 1995 heat wave in Chicago. *American Journal of Preventive Medicine* **16**(4): 269–277.

Smoyer KE, Rainham DG, Hewko JN. 2000. Heat-stress-related mortality in five cities in Southern Ontario: 1980–1996. *International Journal of Biometeorology* **44**(4): 190–197.

Smoyer-Tomic KE, Kuhn R, Hudson A. 2003. Heat Wave Hazards: an overview of heat waves impacts in Canada. *Natural Hazards* **28**: 463–485.

Trigo RM, Garcia-Herrera R, Diaz J, Trigo IF, Valente MA. 2005. How exceptional was the early August 2003 heatwave in France. *Geographical Research Letters* **32**: L10701.

Tukey JW. 1977. *Exploratory Data Analysis* Addison-Wesley: New York, NY, USA ISBN 0-201-07616-0.

World Health Organization (WHO) 2004. Heat waves: risk and responses. *Health and Global Environmental Series No. 2*. World Health Organization.